



**HAL**  
open science

# Well-posedness and propagation of chaos for McKean-Vlasov equations with jumps and locally Lipschitz coefficients

Xavier Erny

► **To cite this version:**

Xavier Erny. Well-posedness and propagation of chaos for McKean-Vlasov equations with jumps and locally Lipschitz coefficients. 2021. hal-03139922v1

**HAL Id: hal-03139922**

**<https://hal.science/hal-03139922v1>**

Preprint submitted on 12 Feb 2021 (v1), last revised 1 Mar 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Well-posedness and propagation of chaos for McKean-Vlasov equations with jumps and locally Lipschitz coefficients

Xavier Erny

*Université Paris-Saclay, CNRS, Univ Evry, Laboratoire de Mathématiques et Modélisation d'Evry, 91037, Evry, France*

**Abstract:** We study McKean-Vlasov equations where the coefficients are locally Lipschitz continuous. We prove the strong well-posedness and a propagation of chaos property in this framework. These questions can be treated with classical arguments under the assumptions that the coefficients are globally Lipschitz continuous. In the locally Lipschitz case, we use truncation arguments and Osgood's lemma instead of Grönwall's lemma. This approach entails technical difficulties in the proofs, in particular for the existence of solution of the McKean-Vlasov equations that are considered. This proof relies on a Picard iteration scheme that is not guaranteed to converge in an  $L^1$ -sense because the coefficients are not Lipschitz continuous. However, we still manage to prove its convergence in distribution, and the (strong) well-posedness of the equation using a generalization of Yamada and Watanabe results.

**MSC2020 subject classifications:** 60J60, 60K35.

**Keywords and phrases:** McKean-Vlasov equations, Mean field interaction, Interacting particle systems, Propagation of chaos.

## 1. Introduction

The aim of this paper is to prove the strong well-posedness and a propagation of chaos property for McKean-Vlasov equations. These are SDEs where the coefficients depend on the solution of the equation and on the law of this solution. This type of equation arises naturally in the framework of  $N$ -particle systems where the particles interact in a mean field way: this phenomenon can be seen as a law of large numbers. Indeed, in examples where the dynamic of the  $N$ -particle system is directed by an SDE, the mean field interactions can be expressed as a dependency of the coefficients on the empirical measure of the system. And as  $N$  goes to infinity, this empirical measure converges to the law of any particle of the limit system. This entails natural dependencies of the coefficients on the law of the solution of the limit SDE. For instance, see [De Masi et al. \(2015\)](#) and [Fournier and Löcherbach \(2016\)](#) for examples in neural network modeling, [Fischer and Livieri \(2016\)](#) for an example in portfolio modeling, and [Carmona, Delarue and Lacker \(2016\)](#) for an application in mean field games.

The weak well-posedness and the propagation of chaos are classical for the McKean-Vlasov equations without jump term, even without assuming that the coefficients are Lipschitz continuous. [Gärtner \(1988\)](#) treats both questions in this frame. We can also mention more recent work on the well-posedness of McKean-Vlasov equations without jump term as [Mishura and Veretennikov \(2020\)](#) and [Chaudru de Raynal \(2020\)](#), with different assumptions on the smoothness of the coefficients, and [Lacker \(2018\)](#) that investigates the well-posedness and the propagation of chaos.

In this paper, we consider McKean-Vlasov equations with jumps. The questions about the strong well-posedness and the propagation of chaos have also been studied in this framework under globally

Lipschitz assumptions on the coefficients: see [Graham \(1992\)](#) for the well-posedness and [Andreis, Dai Pra and Fischer \(2018\)](#) for the propagation of chaos. Note that in Section 4 of [Andreis, Dai Pra and Fischer \(2018\)](#), these questions are treated in a multi-dimensional case, where the drift coefficient is of the form  $-\nabla b_1(x) + b_2(x, m)$ , where  $b_1$  is  $C^1$  and convex, and  $b_2$ , as well as the jump coefficient and the volatility coefficient, are globally Lipschitz.

The novelty of our results is to work on McKean-Vlasov equations with jumps and with generic locally Lipschitz assumptions (see [Assumption 1](#) for a precise and complete statement of the hypothesis).

The first main result is the strong well-posedness, under this locally Lipschitz assumption, of the following McKean-Vlasov equation

$$dX_t = b(X_t, \mu_t)dt + \sigma(X_t, \mu_t)dW_t + \int_{\mathbb{R}_+ \times E} \Phi(X_{t-}, \mu_{t-}, u) \mathbb{1}_{\{z \leq f(X_{t-}, \mu_{t-})\}} d\pi(t, z, u),$$

where  $\mu_t$  is the distribution of  $X_t$ ,  $W$  a Brownian motion,  $\pi$  a Poisson measure and  $E$  some measurable space (see the beginning of [Section 2](#) for details on the notation). To prove the well-posedness when the coefficients are locally Lipschitz continuous, we adapt the computations of the proofs in the globally Lipschitz continuous case. We use a truncation argument to handle the dependency of the local Lipschitz constant w.r.t. to the variables. On the contrary of the globally Lipschitz case, Grönwall's lemma does not allow to conclude immediately. In the locally Lipschitz case, we have to use a generalization of this lemma: Osgood's lemma (see [Lemma 4.1](#)). The uniqueness of solution of the McKean-Vlasov equation follows rather quickly from Osgood's lemma and the truncation argument, but other difficulties emerge in the proof of the existence of solution. We construct a weak solution of the equation using a Picard iteration scheme, but the fact that the coefficients are only locally Lipschitz continuous does not allow to prove that this scheme converges in an  $L^1$ -sense. Instead, we prove that a subsequence converges in distribution to some limit that is shown to be a solution of the equation. Some technical difficulties emerge in this part of the proof for two reasons. The first one is that the Picard scheme is not shown to converge but only to have a converging subsequence. This implies that we need to control the variation between two consecutive steps of the scheme. The second one is that we only prove a convergence in distribution, thus it is not straightforward that the limit of the Picard scheme is solution to the equation. It is shown studying its semimartingale characteristics.

The second main result is a propagation of chaos property of McKean-Vlasov particle systems under the same locally Lipschitz assumptions. More precisely, it is the convergence of the following  $N$ -particle system

$$\begin{aligned} dX_t^{N,i} &= b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)dW_t^i + \int_{\mathbb{R}_+ \times F^{N*}} \Psi(X_{t-}^{N,i}, \mu_{t-}^N, v^i) \mathbb{1}_{\{z \leq f(X_{t-}^{N,i}, \mu_{t-}^N)\}} d\pi^i(t, z, v) \\ &+ \frac{1}{N} \sum_{j=1}^N \int_{\mathbb{R}_+ \times F^{N*}} \Theta(X_{t-}^{N,j}, X_{t-}^{N,i}, \mu_{t-}^N, v^j, v^i) \mathbb{1}_{\{z \leq f(X_{t-}^{N,j}, \mu_{t-}^N)\}} d\pi^j(t, z, v), \end{aligned}$$

where  $\mu_t^N := N^{-1} \sum_{j=1}^N \delta_{X_t^{N,j}}$ , to the infinite system

$$\begin{aligned} d\bar{X}_t^i &= b(\bar{X}_t^i, \bar{\mu}_t)dt + \sigma(\bar{X}_t^i, \bar{\mu}_t)dW_t^i + \int_{\mathbb{R}_+ \times F^{N*}} \Psi(\bar{X}_{t-}^i, \bar{\mu}_{t-}, v^i) \mathbb{1}_{\{z \leq f(\bar{X}_{t-}^i, \bar{\mu}_{t-})\}} d\pi^i(t, z, v) \\ &+ \int_{\mathbb{R}} \int_{F^{N*}} \Theta(x, \bar{X}_t^i, \bar{\mu}_t, v^1, v^2) f(x, \bar{\mu}_t) d\nu(v) d\bar{\mu}_t(x), \end{aligned}$$

where  $\bar{\mu}_t := \mathcal{L}(\bar{X}_t)$ , as  $N$  goes to infinity. The  $W^i$  ( $i \geq 1$ ) are independent standard Brownian motions, the  $\pi^i$  ( $i \geq 1$ ) are independent Poisson measures,  $F$  is some measurable space and  $\mathbb{N}^*$  denotes the set of the positive integers (see Section 3 for details on the notation). The proof of this second main result relies on a similar reasoning as the one used to prove the uniqueness of the McKean-Vlasov equation: a truncation argument and Osgood's lemma.

Let us note that this propagation of chaos property has already been proven under different hypothesis. Indeed, the  $N$ -particle system and the limit system above are the same as in Andreis, Dai Pra and Fischer (2018). Note also that, in this model, for each  $N \in \mathbb{N}^*$ , the particles  $X_t^{N,i}$  ( $1 \leq i \leq N$ ) do not only interact through the empirical measure  $\mu_t^N$ , but also through the simultaneous jumps term.

Let us mention the other conditions in Assumption 1. We assume that some boundedness conditions on the coefficients hold true. To the best of our knowledge, in order to prove a priori estimates for solutions of McKean-Vlasov equations, one needs to assume that the coefficients are bounded w.r.t. the measure variable. In this paper, we need the coefficients to be bounded w.r.t. to both variables, because for the truncation arguments mentioned above, we need a priori estimates on the exponential moments of the solutions of the McKean-Vlasov equation. This is also the reason why we need conditions on exponential moments.

Let us finally mention that, in this paper, we chose to work in dimension one to simplify the notation, but the results still hold in finite dimensions.

**Organization.** In Section 2, we state and prove our first main result: the well-posedness of the McKean-Vlasov equation (1) with locally Lipschitz coefficients. Section 3 is devoted to our second main result, the propagation of chaos in the same framework.

### 1.1. Notation

Let us introduce some notation we use throughout the paper:

- If  $X$  is random variable, we note  $\mathcal{L}(X)$  its distribution.
- If  $X$  and  $X_n$  ( $n \in \mathbb{N}^*$ ) are random variables, we note  $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$  for " $(X_n)_n$  converges in distribution to  $X$ ".
- $\mathcal{P}_1(\mathbb{R})$  is the space of probability measures on  $\mathbb{R}$  with finite first moment. This space will always be endowed with the first-order Wasserstein metric  $W_1$  defined by: for  $m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$ ,

$$W_1(m_1, m_2) = \inf_{X_1 \sim m_1, X_2 \sim m_2} \mathbb{E}[d(X_1, X_2)] = \sup_{f \in Lip_1} \int_{\mathbb{R}} f(x) dm_1(x) - \int_{\mathbb{R}} f(x) dm_2(x),$$

with  $Lip_1$  the space of Lipschitz continuous functions w.r.t. the metric  $d$  with Lipschitz constant non-greater than one. Let us note that characterizations of this convergence are given in Theorem 6.9 and Definition 6.8 of Villani (2008).

- For  $T > 0$  and  $(G, d)$  a Polish space,  $D([0, T], G)$  (resp.  $D(\mathbb{R}_+, G)$ ) denotes the space of càdlàg  $G$ -valued functions defined on  $[0, T]$  (resp.  $\mathbb{R}_+$ ) endowed with Skorohod topology, whence this space is Polish. Let us recall that the convergence of a sequence  $(x_n)_n$  of  $D([0, T], G)$  to some  $x$  in Skorohod topology is equivalent to the existence of continuous increasing functions  $\lambda_n$  satisfying  $\lambda_n(0) = 0$ ,  $\lambda_n(T) = T$  and both

$$\sup_{0 \leq t \leq T} |\lambda_n(t) - t| \text{ and } \sup_{0 \leq t \leq T} d(x(\lambda_n(t)), x_n(t))$$

vanish as  $n$  goes to infinity. In the following, we call such a sequence  $(\lambda_n)_n$  a sequence of time-changes.

- $L$  denotes the Lipschitz constant of the coefficients (see Assumption 1).
- $C$  denotes any arbitrary positive constant, whose value can change from line to line in an equation. If the constant depends on some parameter  $\theta$ , we write  $C_\theta$  instead.

## 2. Well-posedness of McKean-Vlasov equations

This section is dedicated to prove the well-posedness of the following McKean-Vlasov equation.

$$dX_t = b(X_t, \mu_t)dt + \sigma(X_t, \mu_t)dW_t + \int_{\mathbb{R}_+ \times E} \Phi(X_{t-}, \mu_{t-}, u) \mathbb{1}_{\{z \leq f(X_{t-}, \mu_{t-})\}} d\pi(t, z, u), \quad (1)$$

with  $\mu_t = \mathcal{L}(X_t)$ ,  $W$  a standard one-dimensional Brownian motion,  $\pi$  a Poisson measure on  $\mathbb{R}_+ \times \mathbb{R}_+ \times E$  having intensity  $dt \cdot dz \cdot d\rho(u)$ , where  $(E, \mathcal{E}, \rho)$  is a  $\sigma$ -finite measure space. The assumptions on the coefficients are specified in Assumption 1 below. Let us note here that  $f$  is assumed to be non-negative.

### Assumption 1.

1. *Locally Lipschitz conditions: there exists a  $a > 0$ , such that for all  $x_1, x_2 \in \mathbb{R}, m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$ ,*

$$\begin{aligned} & |b(x_1, m_1) - b(x_2, m_2)| + \int_E \int_{\mathbb{R}_+} |\Phi(x_1, m_1, u) \mathbb{1}_{\{z \leq f(x_1, m_1)\}} - \Phi(x_2, m_2, u) \mathbb{1}_{\{z \leq f(x_2, m_2)\}}| dz d\rho(u) \\ & \leq L \left( 1 + |x_1| + |x_2| + \int_{\mathbb{R}} e^{a|x|} dm_1(x) + \int_{\mathbb{R}} e^{a|x|} dm_2(x) \right) (|x_1 - x_2| + W_1(m_1, m_2)). \end{aligned}$$

2. *Globally Lipschitz condition for  $\sigma$ : for all  $x_1, x_2 \in \mathbb{R}, m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$ ,*

$$|\sigma(x_1, m_1) - \sigma(x_2, m_2)| \leq L (|x_1 - x_2| + W_1(m_1, m_2)).$$

3. *Boundedness conditions: the functions  $b, \sigma$  and  $f$  are bounded (uniformly w.r.t. all the variables), and for the same constant  $a > 0$  as in Item 1.,*

$$\sup_{x \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})} \int_E e^{a|\Phi(x, m, u)|} d\rho(u) < \infty.$$

4. *Initial condition: for the same  $a > 0$  as in Items 1. and 3.,*

$$\mathbb{E} \left[ e^{a|X_0|} \right] < \infty.$$

**Remark 2.1.** *If we consider equation (1) without the jump term (that is  $\Phi \equiv 0$ ), then, we can adapt the proof of Theorem 2.3 to the case where  $\sigma$  is also locally Lipschitz continuous. More precisely, we can replace the two first Items of Assumption 1 by: for all  $x_1, x_2 \in \mathbb{R}, m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$ ,*

$$\begin{aligned} & |b(x_1, m_1) - b(x_2, m_2)| + |\sigma(x_1, m_1) - \sigma(x_2, m_2)| \\ & \leq L \left( 1 + \sqrt{|x_1|} + \sqrt{|x_2|} + \sqrt{\int_{\mathbb{R}} e^{a|x|} dm_1(x)} + \sqrt{\int_{\mathbb{R}} e^{a|x|} dm_2(x)} \right) (|x_1 - x_2| + W_1(m_1, m_2)). \end{aligned}$$

See Remark 2.6 for more details on the adaptation of the proof.

Note that Item 3 of Assumption 1 implies that, for all  $n \in \mathbb{N}$ ,

$$\sup_{x \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})} \int_E |\Phi(x, m, u)|^n d\nu(u) < \infty.$$

**Remark 2.2.** Under the boundedness conditions of Assumption 1, a sufficient condition to obtain the locally Lipschitz condition of the jump term in Assumption 1 is given by the following direct conditions on the functions  $f$  and  $\Phi$ : for all  $x_1, x_2 \in \mathbb{R}, m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$ ,

$$\begin{aligned} & |f(x_1, m_1) - f(x_2, m_2)| + \int_E |\Phi(x_1, m_1, u) - \Phi(x_2, m_2, u)| d\rho(u) \\ & \leq L \left( 1 + |x_1| + |x_2| + \int_{\mathbb{R}} e^{a|x|} dm_1(x) + \int_{\mathbb{R}} e^{a|x|} dm_2(x) \right) (|x_1 - x_2| + W_1(m_1, m_2)). \end{aligned}$$

**Example 1.** A natural form of the coefficient of a McKean-Vlasov equation is given by the so-called "true McKean-Vlasov" case. For simplicity, we only give the form for  $b$ , but a similar form can be considered for the other coefficient.

$$b(x, m) = \int_{\mathbb{R}} \tilde{b}(x, y) dm(y),$$

with  $\tilde{b} : \mathbb{R}^2 \rightarrow \mathbb{R}$ .

For  $b$  to satisfy the conditions of Assumption 1 in this example, it is sufficient to assume that  $\tilde{b}$  is bounded and that: for all  $x, x', y, y' \in \mathbb{R}$ ,

$$|\tilde{b}(x, y) - \tilde{b}(x', y')| \leq C(1 + |x| + |x'|)(|x - x'| + |y - y'|).$$

Indeed, for any  $x, x' \in \mathbb{R}, m, m' \in \mathcal{P}_1(\mathbb{R})$ ,

$$\begin{aligned} |b(x, m) - b(x', m')| & \leq |b(x, m) - b(x', m)| + |b(x', m) - b(x', m')| \\ & \leq \int_{\mathbb{R}} |\tilde{b}(x, y) - \tilde{b}(x', y)| dm(y) + \left| \int_{\mathbb{R}} \tilde{b}(x', y) dm(y) - \int_{\mathbb{R}} \tilde{b}(x', y) dm'(y) \right| \\ & \leq C(1 + |x| + |x'|)|x - x'| + C(1 + 2|x'|)W_1(m, m'), \end{aligned}$$

where the second quantity of the last line has been obtained using Kantorovich-Rubinstein duality (see Remark 6.5 of Villani (2008)) and the fact that, for a fixed  $x'$ , the function  $y \mapsto \tilde{b}(x', y)$  is Lipschitz continuous with Lipschitz constant  $C(1 + 2|x'|)$ .

**Theorem 2.3.** Under Assumption 1, there exists a unique strong solution of (1).

The rest of this section is dedicated to prove Theorem 2.3.

### 2.1. A priori estimates for equation (1)

In this section, we prove the following a priori estimates for the solutions of the SDE (1).

**Lemma 2.4.** Grant the boundedness conditions and the initial condition of Assumption 1. Any solution  $(X_t)_{t \geq 0}$  of (1) satisfies for all  $t > 0$ ,

- (i)  $\sup_{0 \leq s \leq t} \mathbb{E} [e^{a|X_s|}] < \infty$ , with  $a > 0$  the same as constant as in Assumption 1,  
(ii)  $\mathbb{E} \left[ \sup_{0 \leq s \leq t} |X_s| \right] < \infty$ .

*Proof.* Let us prove (i). It is sufficient to prove

$$\sup_{0 \leq s \leq t} \mathbb{E} [e^{aX_s}] < \infty, \quad (2)$$

and

$$\sup_{0 \leq s \leq t} \mathbb{E} [e^{-aX_s}] < \infty. \quad (3)$$

To prove (2), let us apply Ito's formula

$$\begin{aligned} e^{aX_t} &= e^{aX_0} + a \int_0^t e^{aX_s} b(X_s, \mu_s) ds + a \int_0^t e^{aX_s} \sigma(X_s, \mu_s) dW_s + \frac{a^2}{2} \int_0^t e^{aX_s} \sigma(X_s, \mu_s)^2 ds \\ &\quad + \int_{[0,t] \times \mathbb{R}_+ \times E} \left[ e^{a(X_{s-} + \Phi(X_{s-}, \mu_{s-}, u))} - e^{aX_{s-}} \right] \mathbb{1}_{\{z \leq f(X_{s-}, \mu_{s-})\}} d\pi(s, z, u). \end{aligned}$$

Introducing, for  $M > 0$ , the stopping time  $\tau_M := \inf\{t > 0 : |X_t| > M\}$ , we have

$$\begin{aligned} \mathbb{E} [e^{aX_{t \wedge \tau_M}}] &\leq \mathbb{E} [e^{aX_0}] + a \|b\|_\infty \int_0^t \mathbb{E} [e^{aX_{s \wedge \tau_M}}] ds + \frac{1}{2} a^2 \|\sigma\|_\infty^2 \int_0^t \mathbb{E} [e^{aX_{s \wedge \tau_M}}] ds \\ &\quad + \int_0^t \int_E e^{aX_{s \wedge \tau_M}} \left[ e^{\Phi(X_{s \wedge \tau_M}, \mu_{s \wedge \tau_M}, u)} - 1 \right] f(X_{s \wedge \tau_M}, \mu_{s \wedge \tau_M}) d\rho(u) ds. \end{aligned}$$

Then, introducing  $u_t^M := \mathbb{E} [e^{aX_{t \wedge \tau_M}}]$ , and using the boundedness condition of  $\Phi$  from Assumption 1, we obtain, for all  $t > 0$ ,

$$u_t^M \leq \mathbb{E} [e^{aX_0}] + K \int_0^t u_s^M ds,$$

with

$$K := a \|b\|_\infty + \frac{1}{2} a^2 \|\sigma\|_\infty^2 + \|f\|_\infty \sup_{x \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})} \int_E e^{a|\Phi(x, m, u)|} d\rho(u) < \infty.$$

Consequently, Grönwall's lemma implies

$$\sup_{0 \leq s \leq t} u_s^M \leq \mathbb{E} [e^{aX_0}] e^{Kt}.$$

As the bound above does not depend on  $M$ , it implies that  $\tau_M$  goes to infinity almost surely as  $M$  goes to infinity. Fatou's lemma then implies

$$\sup_{0 \leq s \leq t} \mathbb{E} [e^{aX_t}] \leq \mathbb{E} [e^{aX_0}] e^{Kt}.$$

This proves (2), and with the same reasoning we can prove

$$\sup_{0 \leq s \leq t} \mathbb{E} [e^{-aX_t}] \leq \mathbb{E} [e^{-aX_0}] e^{Kt}.$$

which proves (3), whence the point (i) of the lemma.

To prove the point (ii), let us use the following bound, which is a direct consequence of the form of the SDE (1),

$$\mathbb{E} \left[ \sup_{0 \leq s \leq t} |X_s| \right] \leq \mathbb{E} [|X_0|] + \|b\|_\infty t + \mathbb{E} \left[ \sup_{0 \leq s \leq t} \left| \int_0^s \sigma(X_t, \mu_r) dW_r \right| \right] + \|f\|_\infty \int_0^t \mathbb{E} \left[ \int_E |\Phi(X_s, \mu_s, u)| d\rho(u) \right] ds.$$

Then the result follows from Burkholder-Davis-Gundy's inequality and the boundedness conditions of  $\sigma$  and  $\Phi$  from Assumption 1.  $\square$

## 2.2. Pathwise uniqueness for equation (1)

**Proposition 2.5.** *Grant Assumption 1. The pathwise uniqueness property holds true for (1).*

*Proof.* Let  $(\hat{X}_t)_{t \geq 0}$  and  $(\check{X}_t)_{t \geq 0}$  be two solutions of (1) defined w.r.t. the same initial condition  $X_0$ , the same Brownian motion  $W$  and the same Poisson measure  $\pi$ . The proof consists in showing that the function

$$u(t) := \mathbb{E} \left[ \sup_{0 \leq s \leq t} |\hat{X}_s - \check{X}_s| \right]$$

is zero. This choice of function  $u$  is inspired of the proof of Theorem 2.1 of [Graham \(1992\)](#). It allows to treat equations with both a jump term and a Brownian term.

We know that, for all  $t \geq 0$ ,  $u_t < \infty$  by Lemma 2.4.(ii).

Writing  $\hat{\mu}_t := \mathcal{L}(\hat{X}_t)$  and  $\check{\mu}_t := \mathcal{L}(\check{X}_t)$ , we have

$$\begin{aligned} |\hat{X}_s - \check{X}_s| &\leq \int_0^s |b(\hat{X}_r, \hat{\mu}_r) - b(\check{X}_r, \check{\mu}_r)| dr + \left| \int_0^s (\sigma(\hat{X}_r, \hat{\mu}_r) - \sigma(\check{X}_r, \check{\mu}_r)) dW_r \right| \\ &\quad + \int_{[0, s] \times \mathbb{R}_+ \times \mathbb{R}} \left| \Phi(\hat{X}_{r-}, \hat{\mu}_{r-}, u) \mathbb{1}_{\{z \leq f(\hat{X}_{r-}, \hat{\mu}_{r-})\}} - \Phi(\check{X}_{r-}, \check{\mu}_{r-}, u) \mathbb{1}_{\{z \leq f(\check{X}_{r-}, \check{\mu}_{r-})\}} \right| d\pi(r, z, u). \end{aligned}$$

This implies that

$$\begin{aligned} u(t) &\leq L \mathbb{E} \left[ \left( \int_0^t (|\hat{X}_s - \check{X}_s| + W_1(\hat{\mu}_s, \check{\mu}_s))^2 ds \right)^{1/2} \right] \\ &\quad + 2L \int_0^t \mathbb{E} \left[ \left( 1 + |\hat{X}_s| + |\check{X}_s| + \int_{\mathbb{R}} e^{a|x|} d\hat{\mu}_s(x) + \int_{\mathbb{R}} e^{a|x|} d\check{\mu}_s(x) \right) (|\hat{X}_s - \check{X}_s| + W_1(\hat{\mu}_s, \check{\mu}_s)) \right] ds \end{aligned} \tag{4}$$

where we have used Burkholder-Davis-Gundy's inequality to deal with the Brownian term that corresponds to the term at the first line above. The term at the first line corresponds to the controls of the drift term and the jump term.

By Lemma 2.4.(i), we have for all  $t \geq 0$ ,

$$\sup_{0 \leq s \leq t} \int_{\mathbb{R}} e^{a|x|} d\hat{\mu}_s(x) + \sup_{0 \leq s \leq t} \int_{\mathbb{R}} e^{a|x|} d\check{\mu}_s(x) \leq C_t < \infty.$$



And, from the definition of  $W_1$ , we have the following bound

$$W_1(\hat{\mu}_s, \check{\mu}_s) \leq \mathbb{E} \left[ \left| \hat{X}_s - \check{X}_s \right| \right] \leq u(s).$$

Then, (4) and Lemma 2.4 imply that, for all  $0 \leq t \leq T$ ,

$$\begin{aligned} u(t) &\leq \int_0^t \mathbb{E} \left[ \left( 1 + |\hat{X}_s| + |\check{X}_s| + C_T \right) \left( |\hat{X}_s - \check{X}_s| + u(s) \right) \right] ds \\ &\quad + L \mathbb{E} \left[ \left( \int_0^t \left( |\hat{X}_s - \check{X}_s| + u(s) \right)^2 ds \right)^{1/2} \right] \\ &\leq C_T \int_0^t \mathbb{E} \left[ \left( 1 + |\hat{X}_s| + |\check{X}_s| \right) \left( |\hat{X}_s - \check{X}_s| + u(s) \right) \right] ds + 2L\sqrt{t}u(t) \\ &\leq C_T \int_0^t \mathbb{E} \left[ \left( 1 + |\hat{X}_s| + |\check{X}_s| \right) \left( |\hat{X}_s - \check{X}_s| \right) \right] ds + C_T \int_0^t u(s) ds + 2L\sqrt{t}u(t), \end{aligned}$$

where we have bounded the second integral of the RHS of the first inequality above by  $t$  times the supremum of the integrand. Note that the value of  $C_T$  changes from line to line.

Now, to end the proof, we have to control a term of the type  $(1 + |x| + |y|)|x - y|$ . To do so, we use a truncation argument based on the following inequality: for all  $x, y \in \mathbb{R}, R > 0$ ,

$$(1 + |x| + |y|)|x - y| \leq (1 + 2R)|x - y| + (1 + |x| + |y|)|x - y| (\mathbf{1}_{\{|x| > R\}} + \mathbf{1}_{\{|y| > R\}}).$$

Let  $R : s \mapsto R_s > 0$  be the truncation function whose values will be chosen later. By Lemma 2.4, for any  $0 \leq s \leq T$ ,

$$\mathbb{E} \left[ \left( 1 + |\hat{X}_s| + |\check{X}_s| \right) \left| \hat{X}_s - \check{X}_s \right| \right] \leq (1 + 2R_s)u(s) + C_T \sqrt{\mathbb{P} \left( |\hat{X}_s| > R_s \right)} + C_T \sqrt{\mathbb{P} \left( |\check{X}_s| > R_s \right)}.$$

The exponential moments proven in Lemma 2.4.(i) are used to control the two last term above. Indeed, by Markov's inequality

$$\mathbb{P} \left( |\hat{X}_s| > R_s \right) + \mathbb{P} \left( |\check{X}_s| > R_s \right) \leq C_T e^{-aR_s}.$$

Consequently, defining  $r_s := aR_s/2$ , for any  $0 \leq t \leq T$ ,

$$u(t) \leq C_T \int_0^t \left[ (1 + r_s)u(s) + e^{-r_s} \right] ds + 2L\sqrt{t}u(t).$$

Now, let  $T = 1/(16L^2)$  such that  $2L\sqrt{T} \leq 1/2$ . Then, we can rewrite the above inequality as, for all  $t \in [0, T]$ ,

$$u(t) \leq C_T \int_0^t \left[ (1 + r_s)u(s) + e^{-r_s} \right] ds.$$

Let us prove by contradiction that, for all  $t \leq T$ ,  $u(t) = 0$ . To do so, let  $t_0 := \inf\{t > 0 : u(t) > 0\}$  and assume that  $t_0 < T$ . Notice that, as  $u$  is non-decreasing, this implies that, for all  $t \in [0, t_0]$ ,  $u(t) = 0$ . In particular, for all  $t \in [t_0, T]$ ,

$$u(t) \leq C_T \int_{t_0}^t \left[ (1 + r_s)u(s) + e^{-r_s} \right] ds.$$

Besides  $u(t)$  is finite and bounded (see Lemma 2.4.(ii)) on  $[0, T]$ , say by a constant  $D > 1$ . Let  $v(t) := u(t)/(De^2) < e^{-2}$ . Obviously  $v$  satisfies the same inequality as  $u$  above. Now we define  $r_s := -\ln v(s)$ , so that, for all  $t_0 < t \leq T$ ,

$$v(t) \leq C_T \int_{t_0}^t (2 - \ln v(s))v(s)ds \leq -2C_T \int_{t_0}^t v(s) \ln v(s)ds,$$

where we have used that  $2 - \ln v(s) \leq -2 \ln v(s)$  since  $-\ln v(s) \geq 2$ .

In particular, for any  $c \in ]0, e^{-2}[$ , for all  $t_0 \leq t \leq T$ ,

$$v(t) \leq c - 2C_T \int_{t_0}^t v(s) \ln v(s)ds.$$

Then, introducing  $M : x \in ]0, e^{-2}[ \mapsto \int_x^{e^{-2}} \frac{1}{-s \ln s} ds$ , we may apply Osgood's lemma (see Lemma 4.1) with  $\gamma \equiv 2C_T$  and  $\mu(v) = (-\ln v)v$  to obtain that

$$-M(v(T)) + M(c) \leq \int_{t_0}^T 2C_T ds = 2C_T(T - t_0)$$

or equivalently,

$$M(c) \leq M(v(T)) + 2C_T T.$$

Recalling that we assumed  $v(T) > 0$  such that the right hand side of the above equality is finite, if we let  $c$  tend to 0, we obtain

$$M(0) = \int_0^{e^{-2}} \frac{1}{-s \ln s} ds \leq \int_{v(T)}^{e^{-2}} \frac{1}{-s \ln s} ds + 2C_T T < \infty,$$

which is absurd since  $M(0) = \infty$ .

A consequence of the above considerations is that for all  $t \in [0, T]$ ,  $u(t) = 0$ . Recalling the definition of  $u$ , we have proven that the processes  $\hat{X}$  and  $\check{X}$  are equal on  $[0, T]$ .

We can repeat this argument on the interval  $[T, 2T]$  and iterate up to any finite time interval  $[0, T_0]$  since  $T = 1/(16L^2)$  does only depend on the coefficients of the system but not on the initial condition. This proves the pathwise-uniqueness property for the McKean-Vlasov equation (1).  $\square$

Let us complete our previous Remark 2.1.

**Remark 2.6.** *The adaptation suggested in Remark 2.1 is the following: in the proof of Proposition 2.5 above, one has to replace the distance  $\mathbb{E} \left[ \sup_{s \leq t} |\hat{X}_s - \check{X}_s| \right]$  by  $\mathbb{E} \left[ (\hat{X}_t - \check{X}_t)^2 \right]$ , and one has to do similar changes in the proof of Proposition 2.11 below.*

### 2.3. Existence of a weak solution of equation (1)

Before proving the existence of solution of (1), let us state some elementary lemmas about series and Skorohod topology, whose proofs are postponed to the Appendix.

**Lemma 2.7.** *Let  $(u_n)_{n \geq 0}$  be a sequence of non-negative real numbers, and  $S_n = \sum_{k=0}^n u_k$  ( $n \in \mathbb{N}$ ). If there exists  $0 < \varepsilon < 1$ , such that for all  $n \in \mathbb{N}$ ,  $S_n \leq Cn^{1-\varepsilon}$ , then there exists a subsequence of  $(u_n)_{n \geq 0}$  that converges to 0.*

**Lemma 2.8.** *Let  $T > 0$ , and  $(x_n)_n$  be a sequence of càdlàg functions converging to some càdlàg function  $x$  in  $D([0, T], \mathbb{R})$ . In addition, if  $(y_n)_n$  is a sequence of càdlàg functions that satisfies*

$$\sup_{0 \leq t \leq T} |x_n(t) - y_n(t)| \xrightarrow{n \rightarrow \infty} 0,$$

*then, the sequence  $(x_n, y_n)_n$  converges to  $(x, x)$  in  $D([0, T], \mathbb{R}^2)$ .*

**Lemma 2.9.** *Let  $T > 0$ , and  $x$  and  $x_n$  ( $n \in \mathbb{N}$ ) be càdlàg functions. Let  $\lambda_n$  ( $n \in \mathbb{N}$ ) be continuous, increasing functions satisfying  $\lambda_n(0) = 0$ ,  $\lambda_n(T) = T$ , and that both*

$$\sup_{0 \leq t \leq T} |x_n(t) - x(\lambda_n(t))| \quad \text{and} \quad \sup_{0 \leq t \leq T} |t - \lambda_n(t)|$$

*vanish as  $n$  goes to infinity. Then,*

$$\sup_{0 \leq t \leq T} \left| \int_0^t x_n(s) ds - \int_0^{\lambda_n(t)} x(s) ds \right| \xrightarrow{n \rightarrow \infty} 0.$$

**Remark 2.10.** *An interesting consequence of the previous lemma is that, if  $(x_n, y_n)_n$  converges to  $(x, y)$  in  $D([0, T], \mathbb{R}^2)$ , then  $(x_n, \int_0^{\cdot} y_n(s) ds)_n$  converges to  $(x, \int_0^{\cdot} y(s) ds)$  in  $D([0, T], \mathbb{R}^2)$ . Note that it is important to have convergence in  $D([0, T], \mathbb{R}^2)$  instead than in  $D([0, T], \mathbb{R})^2$ . The difference between these two topologies is that, the convergence in  $D([0, T], \mathbb{R}^2)$  means that the two coordinates have to share the same sequence of time-changes, whereas for the convergence in  $D([0, T], \mathbb{R})^2$  each of the coordinates has its own sequence.*

The aim of this section is to construct a weak solution of the McKean-Vlasov equation (1), using a Picard iteration. The idea of the proof is to show that this scheme converges to a solution of (1). However, because of our locally Lipschitz conditions, we cannot prove it directly. Instead, we prove that a subsequence converges in distribution by tightness. That is why, in a first time, we only construct a weak solution.

**Proposition 2.11.** *Grant Assumption 1. There exists a weak solution of (1) on  $[0, T]$ , with  $T = 1/(16L^2)$ .*

*Proof.* As in the proof of the pathwise uniqueness of Section 2.2, we work on a time interval  $[0, T]$  where  $T > 0$  is a number whose value can be fixed at  $1/(16L^2)$ .

*Step 1.* In this first step, we introduce the iteration scheme, and state its basic properties at (5). Let  $X_t^{[0]} := X_0$ , and define the process  $X^{[n+1]}$  from  $X^{[n]}$  and  $\mu_t^{[n]} := \mathcal{L}(X_t^{[n]})$  by

$$\begin{aligned} X_t^{[n+1]} := & X_0 + \int_0^t b(X_s^{[n]}, \mu_s^{[n]}) ds + \int_0^t \sigma(X_s^{[n]}, \mu_s^{[n]}) dW_s \\ & + \int_{[0, t] \times \mathbb{R}_+ \times \mathbb{R}} \Phi(X_{s-}^{[n]}, \mu_{s-}^{[n]}, u) \mathbf{1}_{\{z \leq f(X_{s-}^{[n]}, \mu_{s-}^{[n]})\}} d\pi(s, z, u) \end{aligned}$$

Note that, thanks to the boundedness conditions of Assumption 1, using the same computations as in the proof of Lemma 2.4, we can prove that, for all  $t \geq 0$ ,

$$\sup_{n \in \mathbb{N}} \sup_{0 \leq s \leq t} \mathbb{E} \left[ e^{a|X_s^{[n]}|} \right] < \infty \quad \text{and} \quad \sup_{n \in \mathbb{N}} \mathbb{E} \left[ \sup_{0 \leq s \leq t} |X_s^{[n]}| \right] < \infty. \quad (5)$$

*Step 2.* Now let us show that  $(X^{[n]}, X^{[n+1]})_n$  has a converging subsequence in distribution in  $D([0, T], \mathbb{R}^2)$ , by showing that it satisfies Aldous' tightness criterion:

- (a) for all  $\varepsilon > 0$ ,  $\lim_{\delta \downarrow 0} \limsup_{N \rightarrow \infty} \sup_{(S, S') \in A_{\delta, T}} P(|X_{S'}^{[n]} - X_S^{[n]}| + |X_{S'}^{[n+1]} - X_S^{[n+1]}| > \varepsilon) = 0$ , where  $A_{\delta, T}$  is the set of all pairs of stopping times  $(S, S')$  such that  $0 \leq S \leq S' \leq S + \delta \leq T$  a.s.,
- (b)  $\lim_{K \uparrow \infty} \sup_n \mathbb{P}(\sup_{t \in [0, T]} |X_t^{[n]}| + |X_t^{[n+1]}| \geq K) = 0$ .

Assertion (b) is a straightforward consequence of (5) and Markov's inequality. To check assertion (a), notice that, for any  $(S, S') \in A_{\delta, T}$ , by BDG inequality,

$$\mathbb{E} \left[ \left| X_{S'}^{[n+1]} - X_S^{[n+1]} \right| \right] \leq \|b\|_{\infty} \delta + \|\sigma\|_{\infty} \sqrt{\delta} + \delta \|f\|_{\infty} \sup_{0 \leq s \leq T} \mathbb{E} \left[ \int_E |\Phi(X_s^{[n]}, \mu_s^{[n]}, u)| d\rho(u) \right]. \quad (6)$$

Then, by tightness, there exists a subsequence of  $(X^{[n]}, X^{[n+1]})_n$  that converges in distribution to some  $(X, Y)$  in  $D([0, T], \mathbb{R}^2)$ . In the rest of the proof, we work on this subsequence without writing it explicitly for the sake of notation.

*Step 3.* In this step, we show that  $X = Y$  almost surely. Note that, since we work on a subsequence, this is not obvious. It is for this part of the proof that we need to restrict our processes to a time interval of the form  $[0, T]$ . By Lemma 2.8, it is sufficient to prove that, for a subsequence,

$$\mathbb{E} \left[ \sup_{0 \leq s \leq T} \left| X_s^{[n+1]} - X_s^{[n]} \right| \right] \quad (7)$$

vanishes as  $n$  goes to infinity. Indeed, (7) implies that, for another subsequence,  $\sup_{s \leq T} |X_s^{[n+1]} - X_s^{[n]}|$  converges to zero almost surely. Then, we can apply Skorohod representation theorem (see Theorem 6.7 of Billingsley (1999)) to the following sequence

$$\left( X^{[n]}, X^{[n+1]} \right)_n$$

that converges in distribution in  $D([0, T], \mathbb{R}^2)$  to  $(X, Y)$ . Thus we can consider, for  $n \in \mathbb{N}$ , random variables  $(\tilde{X}^{[n]}, \tilde{X}^{[n+1]})$  (resp.  $(\tilde{X}, \tilde{Y})$ ) having the same distribution as  $(X^{[n]}, X^{[n+1]})$  (resp.  $(X, Y)$ ) for which the previous convergence is almost sure. In particular, we also know that  $\sup_{s \leq T} |\tilde{X}_s^{[n+1]} - \tilde{X}_s^{[n]}|$  vanishes almost surely. Hence, by Lemma 2.8, the representing r.v.  $(\tilde{X}^{[n]}, \tilde{X}^{[n+1]})$  converges a.s. to the representing r.v.  $(\tilde{X}, \tilde{Y})$  in  $D([0, T], \mathbb{R}^2)$ . As a consequence  $\tilde{X} = \tilde{Y}$  almost surely, and so  $X = Y$  almost surely.

Now let us prove (7). Let

$$u^{[n]}(t) := \mathbb{E} \left[ \sup_{0 \leq s \leq t} \left| X_s^{[n+1]} - X_s^{[n]} \right| \right].$$

By (5),

$$\sup_{n \in \mathbb{N}} u^{[n]}(t) < \infty.$$

Let us fix some  $n \in \mathbb{N}^*$  and consider a truncation function  $r_t^{[n]} > 0$  whose values will be fixed later. The same truncation argument used in Section 2.2 allows to prove that, for all  $0 \leq k \leq n-1, t \leq T$ ,

$$u^{[k+1]}(t) \leq C_T \int_0^t \left[ (1 + r_s^{[n]}) u^{[k]}(s) + e^{-r_s^{[n]}} \right] ds + 2L\sqrt{T} u^{[k]}(t)$$

$$\leq C_T \int_0^t \left[ (1 + r_s^{[n]}) u^{[k]}(s) + e^{-r_s^{[n]}} \right] ds + \frac{1}{2} u^{[k]}(t).$$

where  $C_T > 0$  does not depend on  $n$  thanks to (5). The second inequality above comes from the fact that we fix the value of  $T > 0$  such that  $L\sqrt{T} < 1/4$ .

Now, introducing  $S_n(t) := \sum_{k=0}^n u_t^{[k]}$  and summing the above inequality from  $k = 0$  to  $k = n - 1$ , we have, for all  $t \leq T$ ,

$$S_n(t) \leq C_T + C_T \int_0^t \left[ (1 + r_s^{[n]}) S_n(s) ds + n e^{-r_s^{[n]}} \right] ds + \frac{1}{2} S_n(t),$$

where we have used that  $u_t^{[0]} \leq C_T$  and  $S_{n-1}(t) \leq S_n(t)$ . This implies

$$S_n(t) \leq C_T + C_T \int_0^t \left[ (1 + r_s^{[n]}) S_n(s) ds + n e^{-r_s^{[n]}} \right] ds.$$

Let  $D_T := \max(\sup_{k \geq 0} \sup_{s \leq T} |u_s^{[k]}|, C_T, 1) < \infty$ , and introduce

$$R_n(t) := \frac{S_n(t)}{(n+1)D_T e^2} \leq e^{-2}.$$

Consequently, for all  $t \leq T$ ,

$$R_n(t) \leq \frac{1}{n+1} + C_T \int_0^t \left[ (1 + r_s^{[n]}) R_n(s) + e^{-r_s^{[n]}} \right] ds.$$

Finally we choose  $r_t^{[n]} := -\ln R_n(t) \geq 2$  and obtain for all  $t \leq T$ ,

$$R_n(t) \leq \frac{1}{n+1} + C_T \int_0^t (2 - \ln R_n(s)) R_n(s) ds \leq \frac{1}{n+1} - C_T \int_0^t R_n(s) \ln R_n(s) ds.$$

As before we apply Osgood's lemma. Let  $M(x) := \int_x^{e^{-2}} \frac{1}{-s \ln s} ds = \ln(-\ln x) - \ln 2$ . Then

$$-M(R_n(T)) + M(1/(n+1)) \leq C_T T$$

or equivalently

$$R_n(T) \leq (n+1)^{-e^{-C_T T}} \text{ such that } S_n(T) \leq C_T n^{1-e^{-C_T T}}.$$

Lemma 2.7 above then implies that there exists a subsequence of  $(u_T^{[n]})_n$  that converges to 0 as  $n$  goes to infinity. This proves (7).

*Step 4.* Let us prove that a subsequence of  $(\mu^{[n]})_n$  converges to some limit  $\mu : t \mapsto \mu_t$  in the following sense

$$\sup_{0 \leq t \leq T} W_1(\mu_t^{[n]}, \mu_t) \xrightarrow{n \rightarrow \infty} 0,$$

where  $\mu_t := \mathcal{L}(X_t)$  for a.e.  $t \leq T$ .

We prove this point by proving that the sequence of functions  $\mu^{[n]} : t \mapsto \mu_t^{[n]} = \mathcal{L}(X_t^{[n]}) \in \mathcal{P}_1(\mathbb{R})$  is relatively compact, using Arzelà-Ascoli's theorem.

To begin with, the definition of  $W_1$  and the same computation as the one used to obtain (6) allows to prove that, for all  $s, t \leq T$ , for all  $n \in \mathbb{N}$ ,

$$W_1(\mu_t^{[n]}, \mu_s^{[n]}) \leq \mathbb{E} \left[ \left| X_t^{[n]} - X_s^{[n]} \right| \right] \leq C \left( |t - s| + \sqrt{|t - s|} \right), \quad (8)$$

for a constant  $C > 0$  independent of  $n$ .

This implies that the sequence  $\mu^{[n]} : t \mapsto \mu_t^{[n]}$  is equicontinuous. In addition, by (5) we know that, for every  $t \leq T$ , the set  $(\mu_t^{[n]})_n$  is tight, and whence relatively compact (in the topology of the weak convergence, but not in  $\mathcal{P}_1(\mathbb{R})$  a priori). Indeed, for any  $\varepsilon > 0$ , considering  $M_\varepsilon := \sup_n \mathbb{E} \left[ |X_t^{[n]}| \right] / \varepsilon$ , we have, for all  $n$ ,

$$\mu_t^{[n]}(\mathbb{R} \setminus [-M_\varepsilon, M_\varepsilon]) = \mathbb{P} \left( |X_t^{[n]}| > M_\varepsilon \right) \leq \frac{1}{M_\varepsilon} \mathbb{E} \left[ |X_t^{[n]}| \right] \leq \varepsilon.$$

In particular, for every  $t \leq T$ , we can consider a subsequence of  $(\mu_t^{[n]})_n$  that converges weakly. To prove that this convergence holds for the metric  $W_1$ , we rely on the characterization (iii) of  $W_1$  given in Definition 6.8, and Theorem 6.9 of Villani (2008). According to this result, the convergence of the same subsequence of  $(\mu_t^{[n]})_n$  for  $W_1$  follows from (5), Markov's inequality, Cauchy-Schwarz's inequality and the fact that,

$$\mathbb{E} \left[ |X_t^{[n]}| \mathbb{1}_{\{|X_t^{[n]}| > R\}} \right] \leq \frac{1}{R} \sup_{k \in \mathbb{N}} \mathbb{E} \left[ |X_t^{[k]}| \right] \mathbb{E} \left[ (X_t^{[k]})^2 \right]^{1/2} \xrightarrow{R \rightarrow \infty} 0.$$

We can then conclude that, for all  $t \leq T$ , the sequence  $(\mu_t^{[n]})_n$  is also relatively compact on  $\mathcal{P}_1(\mathbb{R})$ .

Then, thanks to (8), Arzelà-Ascoli's theorem implies that the sequence  $(\mu^{[n]})_n$  is relatively compact. As a consequence, there exists a subsequence of  $(\mu^{[n]})_n$  (as previously, we do not write this subsequence explicitly in the notation) that converges to some  $\mu : t \mapsto \mu_t \in \mathcal{P}_1(\mathbb{R})$  in the following sense

$$\sup_{0 \leq t \leq T} W_1(\mu_t^{[n]}, \mu_t) \xrightarrow{n \rightarrow \infty} 0.$$

The last thing to show in this step is that  $\mu_t = \mathcal{L}(X_t)$  for a.e.  $t \leq T$ . By construction,  $\mu_t$  is the limit of  $\mu_t^{[n]} := \mathcal{L}(X_t^{[n]})$ . Recalling that  $X^{[n]}$  converges to  $X$  in distribution in Skorohod topology, we know that for all continuity point  $t$  of  $s \mapsto \mathcal{L}(X_s)$ ,  $\mu_t = \mathcal{L}(X_t)$ .

*Step 5.* Recall that, for a subsequence,  $(X^{[n]}, X^{[n+1]})_n$  converges to  $(X, X)$  in distribution in  $D([0, T], \mathbb{R}^2)$ , and  $(\mu^{[n]})_n$  (which is a sequence of deterministic and continuous functions from  $\mathbb{R}_+$  to  $\mathcal{P}_1(\mathbb{R})$ ) converges uniformly to  $\mu$  on  $[0, T]$ . The aim of this step is to prove that  $(X^{[n]}, X^{[n+1]}, \mu^{[n]})_n$  converges to  $(X, X, \mu)$  in distribution in  $D([0, T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$ . We consider  $\mu^{[n]}$  in the previous distribution even though it is deterministic, because the important point in the convergence we want to prove is that  $\mu^{[n]}$  must converge w.r.t. the same sequence of time-changes as the one of  $(X^{[n]}, X^{[n+1]})$  to be able to apply Lemma 2.9 almost surely. In particular, it is important to have convergence in the topology of  $D([0, T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$  rather than in the weaker topology  $D([0, T], \mathbb{R}^2) \times \mathcal{P}_1(\mathbb{R})$ .

As  $\mu$  is continuous, we have that, for any sequence of time-changes  $(\lambda_n)_n$  the convergence

$$\sup_{0 \leq t \leq T} W_1(\mu_t^{[n]}, \mu_{\lambda_n(t)}) \xrightarrow{n \rightarrow \infty} 0. \quad (9)$$

By Skorohod's representation theorem, we can assume that some representative r.v.  $(\tilde{X}^{[n]}, \tilde{X}^{[n+1]})$  of  $(X^{[n]}, X^{[n+1]})$  converges a.s. to representative r.v.  $(\tilde{X}, \tilde{X})$  of  $(X, X)$  in  $D([0, T], \mathbb{R}^2)$ . This implies that, almost surely, there exists a sequence of time-changes  $(\lambda_n)_n$  such that

$$\sup_{0 \leq t \leq T} \left| \tilde{X}_t^{[n]} - \tilde{X}_{\lambda_n(t)} \right| \text{ and } \sup_{0 \leq t \leq T} \left| \tilde{X}_t^{[n+1]} - \tilde{X}_{\lambda_n(t)} \right|$$

vanish as  $n$  goes to infinity. So, by (9), almost surely, there exists a sequence of time-changes  $(\lambda_n)_n$  such that

$$\sup_{0 \leq t \leq T} d \left[ \left( \tilde{X}_t^{[n]}, \tilde{X}_t^{[n+1]}, \mu^{[n]} \right)_t, \left( \tilde{X}_{\lambda_n(t)}, \tilde{X}_{\lambda_n(t)}, \mu_{\lambda_n(t)} \right) \right] \xrightarrow{n \rightarrow \infty} 0,$$

with  $d[(x, y, m), (x', y', m')] = |x - x'| + |y - y'| + W_1(m, m')$ . In particular, we know that the sequence  $(\tilde{X}^{[n]}, \tilde{X}^{[n+1]}, \mu^{[n]})_n$  converges to  $(\tilde{X}, \tilde{X}, \mu)$  almost surely in  $D([0, T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$ . This implies that  $(X^{[n]}, X^{[n+1]}, \mu^{[n]})_n$  converges to  $(X, X, \mu)$  in distribution in  $D([0, T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$ .

*Step 6.* This step concludes the proof, showing that  $X$  is solution to (1). In order to prove that  $X$  is solution to (1), we use the fact that, using the notation of Definitions II.2.6 and II.2.16 of [Jacod and Shiryaev \(2003\)](#),  $X^{[n+1]}$  is a semimartingale with characteristics  $(B^{[n+1]}, C^{[n+1]}, \nu^{[n+1]})$  given by

$$\begin{aligned} B_t^{[n+1]} &= \int_0^t b(X_s^{[n]}, \mu_s^{[n]}) ds, \\ C_t^{[n+1]} &= \int_0^t \sigma(X_s^{[n]}, \mu_s^{[n]})^2 ds, \\ \nu^{[n+1]}(dt, dx) &= f(X_t^{[n]}, \mu_t^{[n]}) dt \int_E \delta_{\Phi(X_t^{[n]}, \mu_t^{[n]}, u)}(dx) d\rho(u). \end{aligned}$$

Let us note that, above, we have chosen as truncation function  $h = 0$ , hence the modified second characteristics  $\tilde{C}^{[n+1]}$  is the same as  $C^{[n+1]}$ .

Recall that, in *Step 5*, we have shown that, for a subsequence,  $(X^{[n]}, X^{[n+1]}, \mu^{[n]})_n$  converges in distribution in  $D([0, T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$  to  $(X, X, \mu)$ . Using once again Skorohod's representation theorem, we can consider representative r.v. for which the previous convergence is almost sure. Whence, by Lemma 2.9, for all  $g \in C_b(\mathbb{R})$ , the following convergences hold almost surely for the representative r.v. and hence in distribution:

$$\begin{aligned} \left( X^{[n+1]}, B^{[n+1]}, C^{[n+1]} \right) &\xrightarrow{n \rightarrow +\infty} \left( X, \int_0^\cdot b(X_s, \mu_s) ds, \int_0^\cdot \sigma(X_s, \mu_s)^2 ds \right), \\ \left( X^{[n+1]}, \int_{[0, \cdot] \times \mathbb{R}} g(x) \nu^{[n+1]}(ds, dx) \right) &\xrightarrow{n \rightarrow +\infty} \left( X, \int_0^\cdot \int_E g(\Phi(X_s, \mu_s, u)) f(X_s, \mu_s) d\rho(u) ds \right), \end{aligned}$$

where the convergences hold respectively in the spaces  $D(\mathbb{R}_+, \mathbb{R}^3)$  and  $D(\mathbb{R}_+, \mathbb{R}^2)$ .

Then, Theorem IX.2.4 of [Jacod and Shiryaev \(2003\)](#) implies that  $X$  is a semimartingale with characteristics  $(B, C, \nu)$  given by

$$\begin{aligned} B_t &= \int_0^t b(X_s, \mu_s) ds, \\ C_t &= \int_0^t \sigma(X_s, \mu_s)^2 ds, \end{aligned}$$

$$\nu(dt, dx) = f(X_t, \mu_t) dt \int_E \delta_{\Phi(X_t, \mu_t, u)}(dx) d\rho(u).$$

Then, we can use the canonical representation of  $X$  (see Theorem II.2.34 of [Jacod and Shiryaev \(2003\)](#)):  $X = X_0 + B + M^c + Id * \mu^X$ , where  $M^c$  is a continuous locale martingale,  $\mu^X = \sum_s \mathbb{1}_{\{\Delta Y_s \neq 0\}} \delta_{(s, X_s)}$  is the jump measure of  $X$  (let us recall that we chose the truncation function  $h = 0$ ) and  $(Id * \mu^X)_t := \int_0^t \int_{\mathbb{R}} x d\mu^X(s, x)$ . By definition of the characteristics,  $\langle M^c \rangle_t = C_t$ . Whence, by Theorem II.7.1 of [Ikeda and Watanabe \(1989\)](#), there exists a Brownian motion  $W$  such that

$$M_t^c = \int_0^t \sigma(X_s, \mu_s) dW_s. \tag{10}$$

In addition, we know that  $\nu$  is the compensator of  $\mu^X$ . We rely on Theorem II.7.4 of [Ikeda and Watanabe \(1989\)](#). Using the notation therein, we introduce  $Z = \mathbb{R}_+ \times E$ ,  $m(dz, du) = dz\rho(du)$  and

$$\theta(t, z, u) := \Phi(X_{t-}, \mu_{t-}, u) \mathbb{1}_{\{z \leq f(X_{t-}, \mu_{t-})\}}.$$

According to Theorem II.7.4 of [Ikeda and Watanabe \(1989\)](#), there exists a Poisson measure  $\pi$  on  $\mathbb{R}_+ \times \mathbb{R}_+ \times E$  having intensity  $dt \cdot dz \cdot d\rho(u)$  such that, for all  $A \in \mathcal{B}(\mathbb{R})$ ,

$$\mu^X([0, t] \times A) = \int_0^t \int_0^\infty \int_E \mathbb{1}_{\{\theta(s, z, u) \in A\}} d\pi(s, z, u).$$

This implies that

$$(Id * \mu^X)_t = \int_{[0, t] \times \mathbb{R}_+ \times E} \Phi(X_{s-}, \mu_{s-}, u) \mathbb{1}_{\{z \leq f(X_{s-}, \mu_{s-})\}} d\pi(s, z, u). \tag{11}$$

Finally, recalling that  $X = X_0 + B + M^c + Id * \mu^X$ , (10) and (11), we have just shown that  $X$  is a weak solution to (1) on  $[0, T]$ .  $\square$

#### 2.4. Proof of Theorem 2.3

In Section 2.2, we have proven the (global) pathwise uniqueness of solutions of (1), and, in Section 2.3, the existence of a weak solution of (1) on  $[0, T]$ , with  $T = 1/(16L^2)$ .

Then, generalizations of Yamada-Watanabe results allows to construct a strong solution on  $[0, T]$ : it is a consequence of Theorem 1.5 and Lemma 2.10 of [Kurtz \(2014\)](#) (see the discussion before Lemma 2.10 or Example 2.14 for more details).

More precisely, given a Brownian motion  $W$ , a Poisson random measure  $\pi$  and an initial condition  $X_0$ , there exists a strong solution  $(X_t)_{0 \leq t \leq T}$  defined w.r.t. these  $W, \pi, X_0$ . Then, one can construct a strong solution  $(X_t)_{T \leq t \leq 2T}$  on  $[T, 2T]$  defined w.r.t. the Brownian motion  $(W_{T+t} - W_T)_{t \geq 0}$ , the Poisson measure  $\pi_T$  defined by

$$\pi_T(A \times B) = \pi(\{T + x : x \in A\} \times B),$$

and the initial condition  $X_T$ . Iterating this reasoning, we can construct a strong solution of (1) on  $[0, kT]$  for any  $k \in \mathbb{N}^*$ , with  $T = 1/(16L^2) > 0$ . Hence, there exists a (global) strong solution of (1). This proves Theorem 2.3.



### 3. Propagation of chaos

In this section, we prove a propagation of chaos for McKean-Vlasov systems: Theorem 3.3. This property in the globally Lipschitz case has been proven in Proposition 3.1 of Andreis, Dai Pra and Fischer (2018). Let us introduce the  $N$ -particle system  $(X_t^{N,i})_{1 \leq i \leq N}$

$$\begin{aligned} dX_t^{N,i} = & b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)dW_t^i + \int_{\mathbb{R}_+ \times F^{\mathbb{N}^*}} \Psi(X_{t-}^{N,i}, \mu_{t-}^N, v^i) \mathbb{1}_{\{z \leq f(X_{t-}^{N,i}, \mu_{t-}^N)\}} d\pi^i(t, z, v) \\ & + \frac{1}{N} \sum_{j=1}^N \int_{\mathbb{R}_+ \times F^{\mathbb{N}^*}} \Theta(X_{t-}^{N,j}, X_{t-}^{N,i}, \mu_{t-}^N, v^j, v^i) \mathbb{1}_{\{z \leq f(X_{t-}^{N,j}, \mu_{t-}^N)\}} d\pi^j(t, z, v), \end{aligned} \quad (12)$$

with  $\mu^N := N^{-1} \sum_{j=1}^N \delta_{X^{N,j}}$ ,  $W^i$  ( $i \geq 1$ ) independent standard one-dimensional Brownian motions, and  $\pi^i$  ( $i \geq 1$ ) independent Poisson measures on  $\mathbb{R}_+^2 \times F^{\mathbb{N}^*}$  with intensity  $dt \cdot dz \cdot d\nu(v)$ , where  $F$  is a measurable space, and  $\nu$  is a  $\sigma$ -finite symmetric measure on  $F^{\mathbb{N}^*}$  (i.e.  $\nu$  is invariant under finite permutations).

In the following, we assume that  $b, \sigma$  and  $f$  satisfy the same conditions as in Assumption 1, and that  $\Psi$  satisfies the same as  $\Phi$  for some constant  $a > 0$ , with  $E = F^{\mathbb{N}^*}$  and  $\rho = \nu$ . We also assume that  $\Theta$  satisfies similar conditions: for all  $x_1, x'_1, x_2, x'_2 \in \mathbb{R}$ ,  $m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$ ,

$$\begin{aligned} & \int_{F^{\mathbb{N}^*}} \int_{\mathbb{R}_+} |\Theta(x_1, x'_1, m_1, v^1, v^2) \mathbb{1}_{\{z \leq f(x_1, m_1)\}} - \Theta(x_2, x'_2, m_2, v^1, v^2) \mathbb{1}_{\{z \leq f(x_2, m_2)\}}| dz d\nu(v) \\ & \leq L \left( 1 + |x_1| + |x'_1| + |x_2| + |x'_2| + \int_{\mathbb{R}} e^{a|x|} dm_1(x) + \int_{\mathbb{R}} e^{a|x|} dm_2(x) \right) \\ & \quad (|x_1 - x_2| + |x'_1 - x'_2| + W_1(m_1, m_2)), \end{aligned}$$

and

$$\sup_{x, x' \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})} \int_{F^{\mathbb{N}^*}} e^{a|\Theta(x, x', m, v^1, v^2)|} d\nu(v) < \infty.$$

In addition, we assume that each  $X_0^{N,i}$  ( $i \geq 1, N \geq 1$ ) satisfies the initial condition of Assumption 1, and that, for every  $N \in \mathbb{N}^*$ , the system  $(X_0^{N,i})_{1 \leq i \leq N}$  is i.i.d.

We prove that these  $N$ -particles systems converge as  $N$  goes to infinity to the following limit system.

$$\begin{aligned} d\bar{X}_t^i = & b(\bar{X}_t^i, \bar{\mu}_t)dt + \sigma(\bar{X}_t^i, \bar{\mu}_t)dW_t^i + \int_{\mathbb{R}_+ \times F^{\mathbb{N}^*}} \Psi(\bar{X}_{t-}^i, \bar{\mu}_{t-}, v^i) \mathbb{1}_{\{z \leq f(\bar{X}_{t-}^i, \bar{\mu}_{t-})\}} d\pi^i(t, z, v) \\ & + \int_{\mathbb{R}} \int_{F^{\mathbb{N}^*}} \Theta(x, \bar{X}_t^i, \bar{\mu}_t, v^1, v^2) f(x, \bar{\mu}_t) d\nu(v) d\bar{\mu}_t(x), \end{aligned} \quad (13)$$

where  $\bar{\mu}_t = \mathcal{L}(\bar{X}_t)$ . We assume that the variables  $\bar{X}_0^i$  ( $i \geq 1$ ) are i.i.d. and satisfy the initial condition of Assumption 1.

Let us remark that the (strong) well-posedness of equation (13) is a consequence of Theorem 2.3 for the same  $\sigma, \Psi = \Phi$  and for the drift

$$b(x, m) + \int_{\mathbb{R}} \int_{F^{\mathbb{N}^*}} \Theta(y, x, m, v^1, v^2) f(y, m) d\nu(v) dm(y).$$

One can also prove the (strong) well-posedness of equation (12) using a similar reasoning as the one used in the proof of Theorem 2.3. The only difference is for the *Step 4* of the proof of Proposition 2.11, since, for (12) the measure  $\mu^N$  is not deterministic. Instead of proving that the sequence of measures  $(\mu^{[n]})_n$  constructed in the Picard scheme is relatively compact by Arzelà-Ascoli's theorem, we rely exclusively on the following lemma whose proof is postponed to Appendix.

**Lemma 3.1.** *Let  $N \in \mathbb{N}^*$ ,  $T > 0$ , and  $(x^k)_{1 \leq k \leq N}$  and  $(x_n^k)_{1 \leq k \leq N}$  ( $n \in \mathbb{N}$ ) be càdlàg functions. Define*

$$\mu_n(t) := N^{-1} \sum_{k=1}^N \delta_{x_n^k(t)} \quad \text{and} \quad \mu(t) := \sum_{k=1}^N \delta_{x^k(t)}.$$

*Let  $\lambda_n$  ( $n \in \mathbb{N}$ ) be continuous, increasing functions satisfying  $\lambda_n(0) = 0$ ,  $\lambda_n(T) = T$ , and that, for any  $1 \leq k \leq N$ ,*

$$\sup_{0 \leq t \leq T} |x_n^k(t) - x^k(\lambda_n(t))| \quad \text{and} \quad \sup_{0 \leq t \leq T} |t - \lambda_n(t)|$$

*vanish as  $n$  goes to infinity. Then,*

$$\sup_{0 \leq t \leq T} W_1(\mu_n(t), \mu(\lambda_n(t))) \xrightarrow{n \rightarrow \infty} 0.$$

**Remark 3.2.** *This lemma allows to prove that, if  $(x_n, y_n)_n$  converges to  $(x, y)$  in  $D([0, T], (\mathbb{R}^N)^2)$ , then, the sequence  $(x_n, y_n, \mu_n)_n$  converges to  $(x, y, \mu)$  in  $D([0, T], (\mathbb{R}^N)^2 \times \mathcal{P}_1(\mathbb{R}))$ .*

In the following, we assume that  $(X^{N,i})_{1 \leq i \leq N}$  and  $(\bar{X}^i)_{i \geq 1}$  are strong solutions of respectively (12) and (13) defined w.r.t. the same Brownian motions  $W^i$  ( $i \geq 1$ ) and the same Poisson measures  $\pi^i$  ( $i \geq 1$ ) such that all the systems are defined on the same space. In addition, we assume that the following condition holds true

$$\varepsilon_0^N := \mathbb{E} \left[ \left| X_0^{N,1} - \bar{X}_0^1 \right| \right] \xrightarrow{N \rightarrow \infty} 0. \tag{14}$$

Now let us state the main result of this section: the propagation of chaos of the  $N$ -particle systems, that is, the convergence of the systems  $(X^{N,i})_{1 \leq i \leq N}$  to the i.i.d. system  $(\bar{X}^i)_{i \geq 1}$  as  $N$  goes to infinity. We comment the convergence speed in Remark 3.5.

**Theorem 3.3.** *We have, for all  $T_0 > 0$ ,*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T_0} \left| X_t^{N,1} - \bar{X}_t^1 \right| \right] \xrightarrow{N \rightarrow \infty} 0.$$

*Consequently, for all  $k \geq 1$ , the following weak convergence holds true:*

$$\mathcal{L}(X^{N,1}, X^{N,2}, \dots, X^{N,k}) \xrightarrow{N \rightarrow \infty} \mathcal{L}(\bar{X}^1) \otimes \mathcal{L}(\bar{X}^2) \otimes \dots \otimes \mathcal{L}(\bar{X}^k),$$

*in the product topology of the topology of the uniform convergence on every compact set.*

**Remark 3.4.** *We just state the result of Theorem 3.3 for the first coordinate because both systems  $(X^{N,i})_{1 \leq i \leq N}$  and  $(\bar{X}^i)_{i \geq 1}$  are exchangeable.*

**Remark 3.5.** In the proof of Theorem 3.3, we obtain a convergence speed for

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T_0} |X_t^{N,1} - \bar{X}_t^1| \right] \xrightarrow{N \rightarrow \infty} 0$$

that depends on  $T_0$ . Indeed, if  $T := 1/(16L^2)$ , the formula (16) below gives a convergence speed for

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^{N,1} - \bar{X}_t^1| \right]$$

of the form

$$S_0^N := C^1 \left( \varepsilon_0^N + N^{-1/2} \right)^{C^2},$$

for some positive constants  $C^1, C^2$ , where  $\varepsilon_0^N$  is given at (14). And, for all  $k \in \mathbb{N}^*$ , the convergence speed  $S_k^N$  of

$$\mathbb{E} \left[ \sup_{kT \leq t \leq (k+1)T} |X_t^{N,1} - \bar{X}_t^1| \right]$$

can be obtained inductively by

$$S_k^N = C^1 \left( S_{k-1}^N + N^{-1/2} \right)^{C^2}$$

for the same constants  $C^1, C^2$  for all  $k$ .

Before proving Theorem 3.3, let us state a lemma about some a priori estimates of the process  $(X_t^{N,1})_{t \geq 0}$ .

**Lemma 3.6.** For every  $N \in \mathbb{N}^*$ , let  $(X^{N,i})_{1 \leq i \leq N}$  be the solution of (12). For any  $t \geq 0$ ,

- (i)  $\sup_{N \in \mathbb{N}^*} \sup_{0 \leq s \leq t} \mathbb{E} \left[ e^{a|X_s^{N,1}|} \right] < \infty$  and  $\sup_{0 \leq s \leq t} \mathbb{E} \left[ e^{a|\bar{X}_s^1|} \right] < \infty$ ,
- (ii)  $\sup_{N \in \mathbb{N}^*} \mathbb{E} \left[ \sup_{0 \leq s \leq t} |X_s^{N,1}| \right] < \infty$  and  $\mathbb{E} \left[ \sup_{0 \leq s \leq t} |\bar{X}_s^1| \right] < \infty$ .

*Sketch of proof of Lemma 3.6.* We only give the main steps of the proof of the first part of the point (i) since the main part of the proof relies on the same computations as in the proof of Lemma 2.4. The only new property of the lemma is the fact that the bounds has to be uniform in  $N$ . This property is easy to prove for the point (ii) of the lemma in this model because of the normalization in  $N^{-1}$  of the jump term  $\Theta$  of (12). For the point (i), as in the proof of Lemma 2.4, we apply Itô's formula to both functions  $x \mapsto e^{ax}$  and  $x \mapsto e^{-ax}$ , and then, one has to note that, by exchangeability,

$$\begin{aligned} \mathbb{E} \left[ e^{aX_t^{N,1}} \right] &\leq \mathbb{E} \left[ e^{aX_0^{N,1}} \right] + a\|b\|_\infty \int_0^t \mathbb{E} \left[ e^{aX_s^{N,1}} \right] ds + \frac{1}{2} a^2 \|\sigma\|_\infty^2 \int_0^t \mathbb{E} \left[ e^{aX_s^{N,1}} \right] ds \\ &\quad + \|f\|_\infty \int_0^t \mathbb{E} \left[ e^{aX_s^{N,1}} \int_{F^{\mathbb{N}^*}} \left( e^{a\Psi(X_s^{N,1}, \mu_s^N, v^1)} - 1 \right) d\nu(v) \right] ds \\ &\quad + (N-1)\|f\|_\infty \int_0^t \mathbb{E} \left[ e^{aX_s^{N,1}} \int_{F^{\mathbb{N}^*}} \left( e^{aN^{-1}\Theta(X_s^{N,2}, X_s^{N,1}, \mu_s^N, v^1, v^2)} - 1 \right) d\nu(v) \right] ds. \end{aligned}$$

So, the only term that does depend on  $N$  is the one of the last line that is bounded by

$$\begin{aligned} \|f\|_\infty \int_0^t \mathbb{E} \left[ e^{aX_s^{N,1}} \int_{F^{\mathbb{N}^*}} N \left( e^{aN^{-1}|\Theta(X_s^{N,2}, X_s^{N,1}, \mu_s^N, v^1, v^2)|} - 1 \right) d\nu(v) \right] ds \\ \leq \|f\|_\infty \int_0^t \mathbb{E} \left[ e^{aX_s^{N,1}} \right] ds \sup_{x, x' \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})} \int_{F^{\mathbb{N}^*}} e^{a|\Theta(x, x', m, v^1, v^2)|} d\nu(v), \end{aligned}$$

where we have used the fact that, for all  $x \geq 0, n \in \mathbb{N}^*, n(e^{x/n} - 1) \leq e^x$ .

Finally, Grönwall's lemma allows to obtain a bound of  $\mathbb{E} \left[ e^{aX_s^{N,1}} \right]$  that is uniform in  $N$ . With the same reasoning, we can obtain the same bound for  $\mathbb{E} \left[ e^{-aX_s^{N,1}} \right]$ . This allows to have a bound for  $\mathbb{E} \left[ e^{a|X_s^{N,1}|} \right]$  uniform in  $N$ .  $\square$

**Remark 3.7.** For the previous proof to be formal, we should have introduced stopping times as in the proof of Lemma 2.4, do the computations above for the stopped processes, and then apply Fatou's lemma to prove that the control obtained for the stopped processes still hold for the real process.

*Proof of Theorem 3.3.* We follow the ideas of Proposition 3.1 of Andreis, Dai Pra and Fischer (2018). Instead of introducing an auxiliary system, we rewrite artificially the SDE (12) as

$$\begin{aligned} dX_t^{N,i} = & b(X_t^{N,i}, \mu_t^N) dt + \sigma(X_t^{N,i}, \mu_t^N) dW_t^i + \int_{\mathbb{R}_+ \times F^{\mathbb{N}^*}} \Psi(X_{t-}^{N,i}, \mu_{t-}^N, v^i) \mathbb{1}_{\{z \leq f(X_{t-}^{N,i}, \mu_{t-}^N)\}} d\pi^i(t, z, v) \\ & + \int_{\mathbb{R}} \int_{F^{\mathbb{N}^*}} \Theta(x, X_t^{N,i}, \mu_t^N, v^1, v^2) f(x, \mu_t^N) d\nu(v) d\mu_t^N(x) + dG_t^N, \end{aligned}$$

where

$$\begin{aligned} G_t^N = & \frac{1}{N} \sum_{j=1}^N \left[ \int_{[0,t] \times \mathbb{R}_+ \times F^{\mathbb{N}^*}} \Theta(X_{s-}^{N,j}, X_{s-}^{N,i}, \mu_{s-}^N, v^j, v^i) \mathbb{1}_{\{z \leq f(X_{s-}^{N,j}, \mu_{s-}^N)\}} d\pi^j(s, z, v) \right. \\ & \left. - \int_0^t \int_{F^{\mathbb{N}^*}} \Theta(X_s^{N,j}, X_s^{N,i}, \mu_s^N, v^1, v^2) f(X_s^{N,j}, \mu_s^N) d\nu(v) ds \right]. \end{aligned}$$

Let us define

$$u^N(t) := \mathbb{E} \left[ \sup_{0 \leq s \leq t} |X_s^{N,1} - \bar{X}_s^1| \right].$$

Let us note  $\bar{\mu}_t^N := N^{-1} \sum_{j=1}^N \delta_{\bar{X}_t^j}$ . By triangle inequality and since for all  $m, m' \in \mathcal{P}_1(\mathbb{R}), W_1(m, m') \leq W_2(m, m')$ , we have for all  $t \geq 0$ ,

$$W_1(\mu_t^N, \bar{\mu}_t) \leq W_1(\mu_t^N, \bar{\mu}_t^N) + W_1(\bar{\mu}_t^N, \bar{\mu}_t) \leq \frac{1}{N} \sum_{j=1}^N |X_t^{N,j} - \bar{X}_t^j| + W_2(\bar{\mu}_t^N, \bar{\mu}_t).$$

Besides, using Theorem 1 of Fournier and Guillin (2015) with  $d = 1, p = 2$  and any  $q > 4$  (and using Lemma 3.6), we have

$$\mathbb{E} [W_2(\bar{\mu}_t^N, \bar{\mu}_t)] \leq C \mathbb{E} [|\bar{X}_t|^q]^{p/q} N^{-1/2} \leq C \left( 1 + \mathbb{E} \left[ e^{a|\bar{X}_t|} \right] \right)^{p/q} N^{-1/2} \leq C_t N^{-1/2}.$$

As a consequence, and thanks to Lemma 3.6, for all  $t \geq 0$ ,

$$\mathbb{E} \left[ \sup_{0 \leq s \leq t} W_1(\mu_s^N, \mu_s) \right] \leq u^N(t) + C_t N^{-1/2}.$$

Whence the same truncation arguments as the ones used in Sections 2.2 and 2.3 allows to prove that, for all  $t \leq T$ ,

$$u^N(t) \leq u^N(0) + C_T \int_0^t \left[ (1 + r_s^N) u^N(s) + e^{-r_s^N} \right] ds + 2L\sqrt{T} u^N(t) + \mathbb{E} \left[ \sup_{0 \leq s \leq T} |G_s^N| \right] + C_T N^{-1/2},$$

where  $C_T$  does not depend on  $N$  thanks to Lemma 3.6.

Now, fixing  $T = 1/(16L^2)$  (such that  $2L\sqrt{T} \leq 1/2$ ), we obtain

$$u^N(t) \leq 2u^N(0) + C_T \int_0^t \left[ (1 + r_s^N) u^N(s) + e^{-r_s^N} \right] ds + 2\mathbb{E} \left[ \sup_{0 \leq s \leq T} |G_s^N| \right] + C_t N^{-1/2}. \quad (15)$$

To control the term  $G_s^N$ , we use BDG inequality,

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |G_t^N|^2 \right] \leq \frac{C}{N^2} \sum_{j=1}^N \mathbb{E} \left[ \int_0^t \int_{F^{\mathbb{N}^*}} \Theta(X_s^{N,j}, X_s^{N,i}, \mu_s^N, v^j, v^i)^2 f(X_s^{N,j}, \mu_s^N) d\nu(v) ds \right] \leq \frac{C_T}{N},$$

where we have used the boundedness conditions of the functions  $f$  and  $\Theta$ , and the fact that the Poisson measures  $\pi^j$  ( $1 \leq j \leq N$ ) are independent.

Hence, by Cauchy-Schwarz's inequality,

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |G_t^N| \right] \leq C_T N^{-1/2}.$$

Then we can rewrite (15) as

$$u^N(t) \leq 2u^N(0) + C_T \int_0^t \left[ (1 + r_s^N) u^N(s) + e^{-r_s^N} \right] ds + C_T N^{-1/2}.$$

Now, let  $D_T := \max(C_T, \sup_N \sup_{s \leq T} u^N(s), 1)$  which is finite by Lemma 3.6.(ii), and define

$$v^N(t) := \frac{u^N(t)}{D_T e^2} \leq e^{-2}.$$

Choosing  $r_t^N := -\ln v^N(t)$ , we have for all  $0 \leq t \leq T$ ,

$$\begin{aligned} v^N(t) &\leq 2v^N(0) + C_T \int_0^t (2 - \ln v^N(s)) v^N(s) ds + \frac{C_T}{\sqrt{N}} \\ &\leq 2v^N(0) - 2C_T \int_0^t v^N(s) \ln v^N(s) ds + \frac{C_T}{\sqrt{N}}, \end{aligned}$$

and Osgood's lemma allows to conclude that

$$-M(v(T)) + M(2v^N(0) + C_T N^{-1/2}) \leq 2C_T T,$$

where  $M(x) = \int_x^{e^{-2}} \frac{1}{-s \ln s} ds = \ln(-\ln x) - \ln 2$ . This implies that

$$\ln(-\ln(2v^N(0) + C_T^1 N^{-1/2})) - \ln(-\ln v^N(t)) \leq C_T^2,$$

for some constants  $C_T^1, C_T^2 > 0$ , where we distinguish  $C_T^1$  and  $C_T^2$  for clarity. This implies that, for all  $0 \leq t \leq T$ ,

$$v^N(t) \leq \left(2v^N(0) + C_T^1 N^{-1/2}\right)^{\exp(-C_T^2)}. \quad (16)$$

Hence

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^{N,1} - \bar{X}_t^1| \right] \xrightarrow{N \rightarrow \infty} 0,$$

for a  $T > 0$  sufficiently small that does not depend on the initial conditions (recalling that we have taken  $T = 1/(16L^2)$ ). Then, iterating this reasoning on  $[T, 2T]$ , we can prove

$$\mathbb{E} \left[ \sup_{T \leq t \leq 2T} |X_t^{N,1} - \bar{X}_t^1| \right] \xrightarrow{N \rightarrow \infty} 0,$$

noticing that the "initial conditions" on  $[T, 2T]$  satisfy the same condition as (14):

$$\varepsilon_T^N := \mathbb{E} \left[ |X_T^{N,1} - \bar{X}_T^1| \right] \xrightarrow{N \rightarrow \infty} 0.$$

Finally, by induction, we can prove that for all  $k \in \mathbb{N}^*$ ,

$$\mathbb{E} \left[ \sup_{0 \leq t \leq kT} |X_t^{N,1} - \bar{X}_t^1| \right] \xrightarrow{N \rightarrow \infty} 0,$$

which proves the result. □

## 4. Appendix

### 4.1. Osgood's lemma

We have used many times a generalization of Grönwall's lemma, which is Osgood's lemma. Let us write it explicitly for self-containedness (see e.g. Lemma 3.4 of Bahouri, Chemin and Danchin (2011)).

**Lemma 4.1.** *Let  $\varrho$  be a measurable function from  $[t_0, T]$  to  $[0, b]$ ,  $\gamma$  a locally integrable function from  $[t_0, T]$  to  $\mathbb{R}_+$ , and  $\mu$  a continuous and non-decreasing function from  $[0, b]$  to  $\mathbb{R}_+$ . Suppose that for all  $t \in [t_0, T]$  and for some  $c \in ]0, b[$ ,*

$$\varrho(t) \leq c + \int_{t_0}^t \gamma(s) \mu(\varrho(s)) ds.$$

Then, with  $M(x) := \int_x^b \frac{ds}{\mu(s)}$ ,

$$-M(\varrho(t)) + M(c) \leq \int_{t_0}^t \gamma(s) ds.$$

#### 4.2. Proof of Lemma 2.7

We construct an increasing sequence of positive integers  $(n_k)_{k \geq 1}$  such that for all  $k \in \mathbb{N}^*$ ,

$$u_{n_k} \leq C n_k^{-\varepsilon/2}. \quad (17)$$

Let  $n_1 = 1$ . By hypothesis,  $u_1 = S_1 \leq C$ . Let  $k \in \mathbb{N}^*$ , assume that the  $k$  first terms of the sequence  $n_1, \dots, n_k$  are defined such that (17) holds. Then we put  $n_{k+1} := \min\{n > n_k : u_n \leq C n^{-\varepsilon/2}\}$ . We only have to prove that the above set is not empty. This is shown by contradiction. If the set was empty, this would imply that for all  $n > n_k$ ,  $u_n > C n^{-\varepsilon/2}$ . Consequently, for all  $n > n_k$ ,

$$\begin{aligned} S_n &= S_{n_k} + \sum_{i=n_k+1}^n u_i > S_{n_k} + C \sum_{i=n_k+1}^n i^{-\varepsilon/2} \geq S_{n_k} + C \int_{n_k+1}^{n+1} \frac{dt}{t^{\varepsilon/2}} \\ &= S_{n_k} + \frac{C}{1-\varepsilon/2} (n+1)^{1-\varepsilon/2} - \frac{C}{1-\varepsilon/2} (n_k+1)^{1-\varepsilon/2}. \end{aligned}$$

Letting  $n \rightarrow \infty$  in the above inequality we see that it is in contradiction with the hypothesis that for all  $n \in \mathbb{N}^*$ ,  $S_n \leq C n^{1-\varepsilon}$ .

#### 4.3. Proof of Lemma 2.8

By hypothesis, there exists a sequence of increasing and continuous functions  $\lambda_n$  such that  $\lambda_n(0) = 0$ ,  $\lambda_n(T) = T$ , and that both

$$\sup_{0 \leq t \leq T} |x_n(t) - x(\lambda_n(t))| \quad \text{and} \quad \sup_{0 \leq t \leq T} |t - \lambda_n(t)|$$

vanish as  $n$  goes to infinity.

To prove the lemma, it is sufficient to prove that  $(y_n)_n$  converges to  $x$  in the same sense as above with the same sequence of time-changes  $(\lambda_n)_n$ . It is a direct consequence of the fact that, for all  $0 \leq t \leq T$ ,

$$|y_n(t) - x(\lambda_n(t))| \leq |y_n(t) - x_n(t)| + |x_n(t) - x(\lambda_n(t))|.$$

#### 4.4. Proof of Lemma 2.9

The proof relies on the following inequalities: for any  $0 \leq t \leq T$

$$\begin{aligned} &\left| \int_0^t x_n(s) ds - \int_0^{\lambda_n(t)} x(s) ds \right| \\ &\leq \int_0^t |x_n(s) - x(\lambda_n(s))| ds + \int_0^t |x(\lambda_n(s)) - x(s)| ds + \left| \int_0^t x(s) ds - \int_0^{\lambda_n(t)} x(s) ds \right| \\ &\leq T \sup_{0 \leq s \leq T} |x_n(s) - x(\lambda_n(s))| + \int_0^T |x(\lambda_n(s)) - x(s)| ds + \sup_{0 \leq s \leq T} |x(s)| \cdot \sup_{0 \leq s \leq T} |s - \lambda_n(s)|. \end{aligned}$$

To conclude the proof, one has to notice that the three terms above do not depend on  $t$  and vanish as  $n$  goes to infinity.

#### 4.5. Proof of Lemma 3.1

The proof consists in noticing that, for all  $t \leq T$ ,

$$W_1(\mu_n(t), \mu(\lambda_n(t))) \leq \frac{1}{N} \sum_{k=1}^N |x_n^k(t) - x^k(\lambda_n(t))|.$$

#### Acknowledgements

The author would like to thank his thesis supervisors Eva Löcherbach and Dasha Loukianova for their support and for fruitful discussion about Osgood's lemma and the truncation arguments used in the paper.

#### Bibliography

- ANDREIS, L., DAI PRA, P. and FISCHER, M. (2018). McKean–Vlasov limit for interacting systems with simultaneous jumps. *Stochastic Analysis and Applications* **36** 960–995. Publisher: Taylor & Francis \_eprint: <https://doi.org/10.1080/07362994.2018.1486202>.
- BAHOURI, H., CHEMIN, J.-Y. and DANCHIN, R. (2011). *Fourier Analysis and Nonlinear Partial Differential Equations. Grundlehren der mathematischen Wissenschaften* **343**. Springer.
- BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, Second ed. Wiley Series In Probability And Statistics.
- CARMONA, R., DELARUE, F. and LACKER, D. (2016). Mean field games with common noise. *The Annals of Probability* **44** 3740–3803.
- CHAUDRU DE RAYNAL, P. E. (2020). Strong well posedness of McKean–Vlasov stochastic differential equations with Hölder drift. *Stochastic Processes and their Applications* **130** 79–107.
- DE MASI, A., GALVES, A., LÖCHERBACH, E. and PRESUTTI, E. (2015). Hydrodynamic Limit for Interacting Neurons. *Journal of Statistical Physics* **158** 866–902.
- FISCHER, M. and LIVIERI, G. (2016). Continuous time mean-variance portfolio optimization through the mean field approach. *ESAIM: Probability and Statistics* **20** 30–44. Publisher: EDP Sciences.
- FOURNIER, N. and GUILLIN, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields* **162** 707–738.
- FOURNIER, N. and LÖCHERBACH, E. (2016). On a toy model of interacting neurons. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques* **52** 1844–1876.
- GÄRTNER, J. (1988). On the McKean-Vlasov Limit for Interacting Diffusions. *Mathematische Nachrichten* **137** 197–248. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mana.19881370116>.
- GRAHAM, C. (1992). McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Processes and their Applications* **40** 69–82.
- IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, Second ed. North-Holland Publishing Company.
- JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, Second ed. Springer-Verlag BerlinHeidelberg NewYork.
- KURTZ, T. (2014). Weak and strong solutions of general stochastic models. *Electronic Communications in Probability* **19**.



- LACKER, D. (2018). On a strong form of propagation of chaos for McKean-Vlasov equations. *Electronic Communications in Probability* **23**. Publisher: The Institute of Mathematical Statistics and the Bernoulli Society. [MR3841406](#)
- MISHURA, Y. S. and VERETENNIKOV, A. Y. (2020). Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Theory of Probability and Mathematical Statistics*. Publisher: American Mathematical Society.
- VILLANI, C. (2008). *Optimal transport, old and new*. Springer.