



**HAL**  
open science

# Path-dependent McKean-Vlasov equation: strong well-posedness and convergence of an interpolated Euler scheme

Armand Bernou, Yating Liu

## ► To cite this version:

Armand Bernou, Yating Liu. Path-dependent McKean-Vlasov equation: strong well-posedness and convergence of an interpolated Euler scheme. 2022. hal-03843062v1

**HAL Id: hal-03843062**

**<https://hal.science/hal-03843062v1>**

Preprint submitted on 7 Nov 2022 (v1), last revised 4 Mar 2023 (v3)

**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.

# Path-dependent McKean-Vlasov equation: strong well-posedness and convergence of an interpolated Euler scheme

Armand Bernou <sup>\*</sup>                      Yating Liu <sup>†</sup>

7th November, 2022

## Abstract

We consider the path-dependent McKean-Vlasov equation, in which both the drift and the diffusion coefficients are allowed to depend on the whole trajectory of the process up to the current time  $t$ , and depend on the corresponding marginal distributions. We prove the strong well-posedness of the equation in the  $L^p$  setting,  $p \geq 2$ , locally in time. Then, we introduce an interpolated Euler scheme, a key object to simulate numerically the process, and we prove the convergence of this scheme towards the strong solution in the  $L^p$  norm. Our result is quantitative and provides an explicit rate. As applications we give results for two mean-field equations arising in biology and neuroscience.

*Keywords:* path-dependent McKean-Vlasov equation, interpolated Euler scheme, well-posedness of non-linear SDEs, convergence rate of numerical scheme.

## 1 Introduction

In this paper, we consider the path-dependent McKean-Vlasov equation in  $\mathbb{R}^d$  of the form

$$\begin{cases} dX_t = b(t, X_{\cdot \wedge t}, \mu_{\cdot \wedge t})dt + \sigma(t, X_{\cdot \wedge t}, \mu_{\cdot \wedge t})dB_t, \\ X_0 : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \text{ random variable} \end{cases} \quad (1.1)$$

where the terms  $X_{\cdot \wedge t}$  and  $\mu_{\cdot \wedge t}$  in the coefficients  $b$  and  $\sigma$  keep track of the whole trajectory of  $X$  and its marginal distribution  $\mu$  between 0 and  $t > 0$  (see below (1.2) and (1.3) for the precise definitions) and  $(B_t)_{t \geq 0}$  is an  $\mathbb{R}^q$ -valued Brownian motion independent of  $X_0$ .

The equation (1.1) can be seen as the generalization of the classical McKean-Vlasov equation  $dX_t = b(t, X_t, \mu_t)dt + \sigma(t, X_t, \mu_t)dB_t$ , first introduced by McKean in [McK67] as a stochastic model naturally associated to a class of non-linear PDEs. See also [Szn91] for a systematic presentation of the McKean-Vlasov equation.

Originally used for the study of plasma physics, the McKean-Vlasov equation has since then been popularized for applications in opinion dynamics [HK02], finance (for instance through the rank-based model, see [KF09] and the references therein) and neurosciences [CCP11, CPSS15, DIRT15]. It also plays a key role in the theory of mean-field games, with applications in biological models on animal competition, road traffic engineering and dynamic economic models, see Huang-Malhamé-Caines [CHM06] and the references within.

---

<sup>\*</sup>Dipartimento di Matematica, Università di Roma La Sapienza, P.le Aldo Moro 2, 00185 Roma, Italy. E-mail address: armand.bernou@uniroma1.it

<sup>†</sup>CEREMADE, CNRS, UMR 7534, Université. Paris-Dauphine, PSL University, 75016 Paris, France. E-mail address: liu@ceremade.dauphine.fr

The path-dependent version (1.1) studied in this paper is the underlying dynamics in models of optimal control investigated (mainly for financial applications) by Cosso et al. [CGK<sup>+</sup>20]. These path-dependent McKean-Vlasov dynamics also appear in the recent work of Tomašević on the 2d parabolic-parabolic Keller-Segel equation [Tom21, Equation (1.2)] and in models of cortical signals, such as the Jansen-Rit model [JR95, FTC09]. In this paper, our first focus is on deriving a proof of existence and uniqueness for the strong solution of (1.1) with finite  $L^p$  norm. In a second part, we focus on an interpolated Euler scheme, a key tool that gives a clear road-map for the numerical simulation of (1.1). We prove its convergence towards the unique strong solution, and derive the rate of this convergence in  $L^p$ .

## 1.1 Model

We place ourselves in a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  satisfying the usual condition. Let  $T > 0$  be fixed. For  $p \geq 1$ , we write  $\mathcal{P}_p(\mathbb{R}^d)$  for the set of probability distributions on  $\mathbb{R}^d$  admitting a finite moment of order  $p$ ,  $\mathcal{C}([0, T], E)$  for the set of continuous maps from  $[0, T]$  to some topological space  $E$ . Write  $\mathcal{B}(\mathbb{R}^d)$  for the Borel  $\sigma$ -algebra of  $\mathbb{R}^d$ . We introduce the definitions of  $X_{\cdot \wedge t}$ ,  $\mu_{\cdot \wedge t}$  and give more precision on the form of our drift and diffusion coefficients in (1.1). Let  $\alpha = (\alpha_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathbb{R}^d)$  and let  $(\nu_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ . For a fixed  $t_0 \in [0, T]$ , we define  $\alpha_{\cdot \wedge t_0} = (\alpha_{t \wedge t_0})_{t \in [0, T]}$  by

$$\alpha_{t \wedge t_0} := \begin{cases} \alpha_t & \text{if } t \in [0, t_0], \\ \alpha_{t_0} & \text{if } t \in (t_0, T]. \end{cases} \quad (1.2)$$

Then it is obvious that  $\alpha_{\cdot \wedge t_0} \in \mathcal{C}([0, T], \mathbb{R}^d)$ . Similarly, we define  $\nu_{\cdot \wedge t_0} = (\nu_{t \wedge t_0})_{t \in [0, T]}$  by

$$\nu_{t \wedge t_0} := \begin{cases} \nu_t & \text{if } t \in [0, t_0], \\ \nu_{t_0} & \text{if } t \in (t_0, T], \end{cases} \quad (1.3)$$

and it is straightforward to see that  $\nu_{\cdot \wedge t_0} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ .

Let  $\mathbb{M}_{d,q}(\mathbb{R})$  denote the space of matrices of size  $d \times q$ , equipped with the operator norm  $\|\cdot\|$ . Our path-dependent McKean-Vlasov equation writes

$$X_t = X_0 + \int_0^t b(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) ds + \int_0^t \sigma(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) dB_s \quad (1.4)$$

where

- $X_0 : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  is a random vector in  $L^p(\mathbb{P})$ ,
- $b : [0, T] \times \mathcal{C}([0, T], \mathbb{R}^d) \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \rightarrow \mathbb{R}^d$ ,
- $\sigma : [0, T] \times \mathcal{C}([0, T], \mathbb{R}^d) \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \rightarrow \mathbb{M}_{d,q}(\mathbb{R})$ ,
- $(B_t)_{t \in [0, T]}$  is an  $(\mathcal{F}_t)$ -standard Brownian motion valued in  $\mathbb{R}^q$ , independent of  $X_0$ .
- $\mu_{\cdot \wedge t}$  denotes the marginal distributions of the process  $X_{\cdot \wedge t}$ , that is, for every  $s \in [0, T]$ ,

$$\mu_{s \wedge t} = \mathbb{P} \circ X_{s \wedge t}^{-1}.$$

Our goal is to prove the well-posedness of the path-dependent McKean-Vlasov equation (1.4) and to prove the convergence of an interpolated Euler scheme defined further in (1.12), with a quantitative estimate for the convergence rate. The interpolated Euler scheme presented here is

of strong importance, as it gives a road-map to simulate the equation. Our estimate quantifies the corresponding error, which is also key in practice and will be the starting point of a future work devoted to the implementable particle method for the numerical simulation. This line of reasoning is inspired by the numerical analysis for the classical McKean-Vlasov equation (see e.g. [BT97, Equation (2.3)], [Liu19, Section 7.1], [AKH02, Section 3] and [HL22]). Besides, the Euler scheme also plays a major role in the study of the convex order (see e.g. [LP20] and [LP22] for the classical McKean-Vlasov equation) which we also plan to study in a forthcoming work.

## 1.2 Assumptions and main results

We shall work with two sets of assumptions, both depending on an index  $p \geq 2$ . The first one is required to derive our proof of strong well-posedness in  $L^p$  spaces. The second one is needed to obtain the convergence of our interpolated Euler scheme. Remark that the time horizon  $T > 0$  is fixed. We recall the definition of the Wasserstein distance  $\mathcal{W}_p$  on  $\mathcal{P}_p(\mathbb{R}^d)$ :

$$\begin{aligned} \mathcal{W}_p(\mu, \nu) &:= \inf_{\pi \in \Pi(\mu, \nu)} \left( \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \pi(dx, dy) \right)^{\frac{1}{p}} \\ &= \inf \left\{ \mathbb{E} [|X - Y|^p]^{\frac{1}{p}}, X, Y : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \text{ with } \mathbb{P} \circ X^{-1} = \mu, \mathbb{P} \circ Y^{-1} = \nu \right\} \end{aligned} \quad (1.5)$$

where  $\Pi(\mu, \nu)$  denotes the set of probability measures on  $(\mathbb{R}^d \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)^{\otimes 2})$  with marginals  $\mu$  and  $\nu$ .

**Assumption (I).** There exists  $p \geq 2$  such that

1.  $X_0 \in L^p(\mathbb{P})$ .
2. The coefficient functions  $b, \sigma$  are continuous in  $t$  and Lipschitz continuous in  $\alpha$  and in  $(\mu_t)_{t \in [0, T]}$  with respect to the sup norm  $\|\cdot\|_{\text{sup}}$  and the Wasserstein distance  $\mathcal{W}_p$  uniformly in  $t$ , i.e. there exists  $L > 0$  s.t.

$$\begin{aligned} \forall t \in [0, T], \forall \alpha, \beta \in \mathcal{C}([0, T], \mathbb{R}^d) \text{ and } \forall (\mu_t)_{t \in [0, T]}, (\nu_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)), \\ \left| b(t, \alpha, (\mu_t)_{t \in [0, T]}) - b(t, \beta, (\nu_t)_{t \in [0, T]}) \right| \leq L \left[ \|\alpha - \beta\|_{\text{sup}} + d_p((\mu_t)_{t \in [0, T]}, (\nu_t)_{t \in [0, T]}) \right], \\ \left\| \left\| \sigma(t, \alpha, (\mu_t)_{t \in [0, T]}) - \sigma(t, \beta, (\nu_t)_{t \in [0, T]}) \right\| \right\| \leq L \left[ \|\alpha - \beta\|_{\text{sup}} + d_p((\mu_t)_{t \in [0, T]}, (\nu_t)_{t \in [0, T]}) \right], \end{aligned} \quad (1.6)$$

where  $d_p$  is defined by

$$d_p((\mu_t)_{t \in [0, T]}, (\nu_t)_{t \in [0, T]}) := \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \nu_t). \quad (1.7)$$

**Theorem 1.** *Under Assumption (I), there exists a unique strong solution  $X = (X_t)_{t \in [0, T]}$  from  $(\Omega, \mathcal{F}, \mathbb{P})$  to  $(\mathcal{C}([0, T], \mathbb{R}^d), \|\cdot\|_{\text{sup}})$  of the path-dependent McKean-Vlasov equation (1.4). Moreover, there exists a constant  $\Gamma > 0$  depending only on  $b, \sigma, L, T, d, p$  such that*

$$\left\| \sup_{t \in [0, T]} |X_t| \right\|_p \leq \Gamma \left( 1 + \|X_0\|_p \right). \quad (1.8)$$

We then turn to the quantitative result regarding the convergence of our interpolated Euler scheme. In the following definition,  $M \in \mathbb{N}^*$  should be thought as the temporal discretization number, while  $h := \frac{T}{M}$  is the time step. We propose an interpolated Euler scheme, in which we only need to consider a discrete sequence of random variables and a discrete sequence of

probability measures as the inputs of each step. To simplify the notation, we will denote by  $x_{0:m} := (x_0, \dots, x_m)$ ,  $\mu_{0:m} := (\mu_0, \dots, \mu_m)$ . Our discretization scheme uses the following interpolator.

**Definition 2** (Interpolator). (a) For every  $m = 1, \dots, M$ , we define a piecewise affine interpolator  $i_m$  on  $m + 1$  points in  $\mathbb{R}^d$  by

$$x_{0:m} \in (\mathbb{R}^d)^{m+1} \longmapsto i_m(x_{0:m}) = (\bar{x}_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathbb{R}^d), \quad (1.9)$$

where for every  $t \in [0, T]$ ,  $\bar{x}_t$  is defined by

$$\begin{aligned} \forall k = 0, \dots, m-1, \forall t \in [t_k, t_{k+1}), \quad \bar{x}_t &= \frac{1}{h}(t_{k+1} - t)x_k + \frac{1}{h}(t - t_k)x_{k+1}, \\ \forall t \in [t_m, T], \quad \bar{x}_t &= x_m. \end{aligned}$$

By convention, we define, for every  $t \in [0, T]$ ,  $i_0(x_0)_t := x_0$ .

(b) Let  $p \geq 1$ . For every  $m = 1, \dots, M$ , we define a piecewise affine interpolator for  $m + 1$  probability measures in  $\mathcal{P}_p(\mathbb{R}^d)$ , still denoted by  $i_m$  with a slight abuse of notation, by

$$\mu_{0:m} \in (\mathcal{P}_p(\mathbb{R}^d))^{m+1} \longmapsto i_m(\mu_{0:m}) = (\bar{\mu}_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)), \quad (1.10)$$

where for every  $t \in [0, T]$ ,  $\bar{\mu}_t$  is defined by

$$\begin{aligned} \forall k = 0, \dots, m-1, \forall t \in [t_k, t_{k+1}), \quad \bar{\mu}_t &= \frac{1}{h}(t_{k+1} - t)\mu_k + \frac{1}{h}(t - t_k)\mu_{k+1}, \\ \forall t \in [t_m, T], \quad \bar{\mu}_t &= \mu_m. \end{aligned} \quad (1.11)$$

By convention, we define, for every  $t \in [0, T]$ ,  $i_0(\mu_0)_t := \mu_0$ .

With this at hand, we define our interpolated Euler scheme in which we use the short-hand notation  $Y_{t_0:t_m}$  (respectively,  $\nu_{t_0:t_m}$ ) to denote  $(Y_{t_0}, \dots, Y_{t_m})$  (resp.  $(\nu_{t_0}, \dots, \nu_{t_m})$ ).

**Definition 3.** Let  $M \in \mathbb{N}^*$ ,  $h = \frac{T}{M}$ . For every  $m = 0, \dots, M$ , we set  $t_m = mh$ . Given a Brownian motion  $(B_t)_{t \in [0, T]}$  and  $X_0$ , the discretization scheme  $(\tilde{X}_{t_m}^M)_{0 \leq m \leq M}$  of the path-dependent McKean-Vlasov equation (1.4) is defined as follows :

1.  $\tilde{X}_0^M = X_0$ ;
2. for all  $m \in \{0, \dots, M-1\}$ ,

$$\tilde{X}_{t_{m+1}}^M = \tilde{X}_{t_m}^M + h b_m(t_m, \tilde{X}_{t_0:t_m}^M, \tilde{\mu}_{t_0:t_m}^M) + \sqrt{h} \sigma_m(t_m, \tilde{X}_{t_0:t_m}^M, \tilde{\mu}_{t_0:t_m}^M) Z_{m+1}, \quad (1.12)$$

where, for all  $k \in \{0, \dots, M\}$ ,  $\tilde{\mu}_{t_k}^M$  is the probability distribution of  $\tilde{X}_{t_k}^M$ , where, for  $m = 0, \dots, M-1$ ,  $Z_{m+1} = \frac{1}{\sqrt{h}}(B_{t_{m+1}} - B_{t_m})$ , and where the applications

$$\begin{aligned} b_m &: [0, T] \times (\mathbb{R}^d)^{m+1} \times (\mathcal{P}_p(\mathbb{R}^d))^{m+1} \longrightarrow \mathbb{R}^d, \\ \sigma_m &: [0, T] \times (\mathbb{R}^d)^{m+1} \times (\mathcal{P}_p(\mathbb{R}^d))^{m+1} \longrightarrow \mathbb{M}_{d,q}(\mathbb{R}) \end{aligned}$$

are defined as follows

$$\begin{aligned} \forall t \in [0, T], \quad x_{0:m} \in (\mathbb{R}^d)^{m+1}, \quad \mu_{0:m} \in (\mathcal{P}_p(\mathbb{R}^d))^{m+1} \\ b_m(t, x_{0:m}, \mu_{0:m}) := b(t, i_m(x_{0:m}), i_m(\mu_{0:m})), \quad \sigma_m(t, x_{0:m}, \mu_{0:m}) := \sigma(t, i_m(x_{0:m}), i_m(\mu_{0:m})). \end{aligned} \quad (1.13)$$

The notations  $b_m$  and  $\sigma_m$  can often discretize computations from a numerical point of view. For instance, if

$$b(t, (X_s)_{s \in [0, T]}, (\mu_s)_{s \in [0, T]}) := \int_0^t \mathbb{E} [\phi(X_s)] ds$$

with a bounded function  $\phi$ , then

$$b_m(t_m, \tilde{X}_{t_0:t_m}^M, \tilde{\mu}_{t_0:t_m}^M) = \frac{h}{2} \left( \mathbb{E} [\phi(\tilde{X}_{t_0}^M)] + \mathbb{E} [\phi(\tilde{X}_{t_m}^M)] \right) + h \sum_{k=1}^{m-1} \mathbb{E} [\phi(\tilde{X}_{t_k}^M)].$$

In order to prove the convergence of the interpolated Euler scheme (1.12) to the unique strong solution of (1.4), we will first assume Assumption (I), guaranteeing the uniqueness of the latter, but also some additional regularity on the coefficients.

**Assumption (II).** The coefficient functions  $b, \sigma$  are  $\gamma$ -Hölder in  $t$ ,  $0 < \gamma \leq 1$ , uniformly in  $\alpha$  and in  $(\mu_t)_{t \in [0, T]}$ , i.e. there exists  $L > 0$  s.t.

$$\begin{aligned} & \forall t, s \in [0, T], \forall \alpha \in \mathcal{C}([0, T], \mathbb{R}^d) \text{ and } \forall (\mu_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)), \\ & \left| b(t, \alpha, (\mu_t)_{t \in [0, T]}) - b(s, \alpha, (\mu_t)_{t \in [0, T]}) \right| \vee \left\| \sigma(t, \alpha, (\mu_t)_{t \in [0, T]}) - \sigma(s, \alpha, (\mu_t)_{t \in [0, T]}) \right\| \\ & \leq L \left( 1 + \|\alpha\|_{\text{sup}} + \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \delta_0) \right) |t - s|^\gamma, \end{aligned} \quad (1.14)$$

where  $\delta_0$  is the Dirac measure at 0.

With this at hand, we state our second main result, that regards the discretization scheme.

**Theorem 4** (Convergence rate of the interpolated Euler Scheme). *Under Assumptions (I) and (II), for  $(B_t)_{t \in [0, T]}$  an  $\mathbb{R}^q$ -valued Brownian motion, for  $(X_t)_{t \in [0, T]}$  the unique strong solution to (1.4) given by Theorem 1, for  $(\tilde{X}_{t_m}^M)_{0 \leq m \leq M}$  the Euler scheme from Definition 3 with parameter  $M$  large enough, for  $h = \frac{T}{M}$ , one has,*

$$\left\| \sup_{0 \leq m \leq M} \left| X_{t_m} - \tilde{X}_{t_m}^M \right| \right\|_p \leq \tilde{C} \left( h^\gamma + (h |\ln(h)|)^{\frac{1}{2}} \right), \quad (1.15)$$

where  $\tilde{C} > 0$  is a constant depending on  $L, p, d, \|X_0\|_p, T$  and  $\gamma$ .

From Definition 3, we can define a continuous version of  $(\tilde{X}_{t_m}^M)_{0 \leq m \leq M}$ , denoted by  $\widehat{X}^M = (\widehat{X}_t^M)_{t \in [0, T]}$  and defined by  $\widehat{X}^M := i_M(\tilde{X}_{t_0:t_m}^M)$ . Then we have the following convergence.

**Corollary 5.** *Under Assumptions (I) and (II), for  $M$  large enough, one has*

$$\left\| \sup_{t \in [0, T]} \left| X_t - \widehat{X}_t^M \right| \right\|_p \leq \tilde{C} \left( h^\gamma + (h |\ln(h)|)^{\frac{1}{2}} \right), \quad (1.16)$$

where  $\tilde{C} > 0$  is a constant depending on  $L, p, d, \|X_0\|_p, T$  and  $\gamma$ .

*Remark 6.* Theorems 1 and 4 in this paper can be generalized to the following equation having both path-dependent and marginal part in the coefficient functions

$$dX_t = b(t, X_t, \mu_t, X_{\cdot \wedge t}, \mu_{\cdot \wedge t}) dt + \sigma(t, X_t, \mu_t, X_{\cdot \wedge t}, \mu_{\cdot \wedge t}) dB_t, \quad (1.17)$$

when

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

$$\sigma : [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \times \mathcal{C}([0, T], \mathbb{R}^d) \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \rightarrow \mathbb{M}_{d,q}(\mathbb{R})$$

satisfy some adapted conditions. More precisely, for the well-posedness result, we require that for some  $p \geq 2$ ,

1.  $X_0 \in L^p(\mathbb{P})$ .
2. The coefficient functions  $b, \sigma$  are continuous in  $t$  and Lipschitz continuous in the parameters  $(x, \nu, \alpha, (\mu_t)_{t \in [0, T]})$  in the following sense: there exists  $L > 0$  s.t.

$$\begin{aligned} & \forall t \in [0, T], \forall x, x' \in \mathbb{R}^d, \forall \nu, \nu' \in \mathcal{P}_p(\mathbb{R}^d), \\ & \quad \forall \alpha, \alpha' \in \mathcal{C}([0, T], \mathbb{R}^d) \text{ and } \forall (\mu_t)_{t \in [0, T]}, (\mu'_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)), \\ & \left| b(t, x, \nu, \alpha, (\mu_t)_{t \in [0, T]}) - b(t, x', \nu', \alpha', (\mu'_t)_{t \in [0, T]}) \right| \\ & \quad \leq L \left[ |x - x'| + \mathcal{W}_p(\nu, \nu') + \|\alpha - \alpha'\|_{\text{sup}} + d_p((\mu_t)_{t \in [0, T]}, (\mu'_t)_{t \in [0, T]}) \right], \\ & \left\| \sigma(t, x, \nu, \alpha, (\mu_t)_{t \in [0, T]}) - \sigma(t, x', \nu', \alpha', (\mu'_t)_{t \in [0, T]}) \right\| \\ & \quad \leq L \left[ |x - x'| + \mathcal{W}_p(\nu, \nu') + \|\alpha - \alpha'\|_{\text{sup}} + d_p((\mu_t)_{t \in [0, T]}, (\mu'_t)_{t \in [0, T]}) \right]. \end{aligned}$$

To extend Theorem 4 to this larger setting for the following interpolated Euler scheme with the same time discretization as in Definition 3

$$\begin{aligned} \widetilde{X}_{t_0}^M &= X_0, \\ \widetilde{X}_{t_{m+1}}^M &= \widetilde{X}_{t_m}^M + h b(t_m, \widetilde{X}_{t_m}^M, \widetilde{\mu}_{t_m}^M, i_m(\widetilde{X}_{t_0:t_m}^M), i_m(\widetilde{\mu}_{t_0:t_m}^M)) \\ & \quad + \sqrt{h} \sigma(t_m, \widetilde{X}_{t_m}^M, \widetilde{\mu}_{t_m}^M, i_m(\widetilde{X}_{t_0:t_m}^M), i_m(\widetilde{\mu}_{t_0:t_m}^M)) Z_{m+1}, \end{aligned} \quad (1.18)$$

we require in addition the following Hölder continuity assumption: the coefficient functions  $b, \sigma$  are  $\gamma$ -Hölder in  $t$ ,  $0 < \gamma \leq 1$ , uniformly in  $x, \alpha, \nu$  and  $(\mu_t)_{t \in [0, T]}$ , i.e. there exists  $L > 0$  s.t.

$$\begin{aligned} & \forall t, s \in [0, T], \forall x \in \mathbb{R}^d, \forall \nu \in \mathcal{P}_p(\mathbb{R}^d), \forall \alpha \in \mathcal{C}([0, T], \mathbb{R}^d) \text{ and } \forall (\mu_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)), \\ & \left| b(t, x, \nu, \alpha, (\mu_t)_{t \in [0, T]}) - b(s, x, \nu, \alpha, (\mu_t)_{t \in [0, T]}) \right| \\ & \quad \leq L \left( 1 + |x| + \mathcal{W}_p(\nu, \delta_0) + \|\alpha\|_{\text{sup}} + \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \delta_0) \right) |t - s|^\gamma, \\ & \left\| \sigma(t, x, \nu, \alpha, (\mu_t)_{t \in [0, T]}) - \sigma(s, x, \nu, \alpha, (\mu_t)_{t \in [0, T]}) \right\| \\ & \quad \leq L \left( 1 + |x| + \mathcal{W}_p(\nu, \delta_0) + \|\alpha\|_{\text{sup}} + \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \delta_0) \right) |t - s|^\gamma. \end{aligned}$$

This extension is particularly useful in the case where the equation investigated is of the form

$$dX_t = b(t, X_t, \mu_{\cdot \wedge t}) dt + \sigma(t, X_t, \mu_{\cdot \wedge t}) dB_t \quad \text{or} \quad dX_t = b(t, X_{\cdot \wedge t}, \mu_t) dt + \sigma(t, X_{\cdot \wedge t}, \mu_t) dB_t$$

in which case the corresponding assumptions are weaker than Assumptions (I) and (II).

### 1.3 Strategy and plan of the paper

The strategy to establish the well-posedness (Section 3) is largely inspired from Bouleau [Bou88], see also [Liu19, Chapter 5] where the second author derived similar results for the classical McKean-Vlasov equation (without path-dependency), and Lacker [Lac18]. While the norms used in this method are more involved than in earlier works, the main idea of the proof

is reminiscent of the one of Sznitman [Szn91] based on earlier work by Dobrushin [Dob70]. By considering appropriate trajectorial spaces, and by introducing norms depending on well-chosen parameters, we are able to perform a classical fixed-point argument. We introduce first the space

$$\mathcal{H}_{p,C,T} := \left\{ Y \in L^p_{\mathcal{C}([0,T],\mathbb{R}^d)}(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{P}) \text{ s.t. } Y \text{ is } (\mathcal{F}_t)_{t \in [0,T]} \text{-adapted.} \right\},$$

that we endow with a suitable norm  $\|\cdot\|_{p,C,T}$  with parameter  $C > 0$ , so that  $\|Y\|_{p,C,T} \rightarrow 0$  when  $C \rightarrow \infty$  for all  $Y \in \mathcal{H}_{p,C,T}$ . Then, we identify every probability measure on  $\mathcal{C}([0,T], \mathbb{R}^d)$  admitting  $p$  moments with a continuous map from  $[0,T]$  to  $\mathcal{P}_p(\mathbb{R}^d)$ . The main argument is then to endow the Banach product space  $\mathcal{H}_{p,C,T} \times \mathcal{C}([0,T], \mathcal{P}_p(\mathbb{R}^d))$  with a suitable distance, denoted  $d_{\mathcal{H} \times \mathcal{P}}$  and to show that, roughly, the map

$$\begin{aligned} \Phi_C : \mathcal{H}_{p,C,T} \times \mathcal{C}([0,T], \mathcal{P}_p(\mathbb{R}^d)) &\rightarrow \mathcal{H}_{p,C,T} \times \mathcal{C}([0,T], \mathcal{P}_p(\mathbb{R}^d)) & (1.19) \\ (Y, (\nu_t)_{t \in [0,T]}) &\mapsto \left( \underbrace{\left( X_0 + \int_0^t b(s, Y_{\cdot \wedge s}, \nu_{\cdot \wedge s}) ds + \int_0^t \sigma(s, Y_{\cdot \wedge s}, \nu_{\cdot \wedge s}) dB_s \right)_{t \in [0,T]}}_{=: \Phi_C^{(1)}(Y, (\nu_t)_{t \in [0,T]})}, \iota(P_{\Phi_C^{(1)}(Y, (\nu_t)_{t \in [0,T]})}) \right) \end{aligned}$$

is Lipschitz continuous, with a Lipschitz constant strictly smaller than one for  $C$  large enough, turning  $\Phi_C$  into a contraction mapping. Here  $\iota(P_{\Phi_C^{(1)}(Y, (\nu_t)_{t \in [0,T]})})$  in (1.19) denotes the marginal distributions of  $\Phi_C^{(1)}(Y, (\nu_t)_{t \in [0,T]})$ . This allows to perform a fixed-point argument to obtain the well-posedness of (1.4).

Section 4 is devoted to the study of the convergence of the interpolated Euler scheme. We start by giving the definition of the theoretical continuous extension  $(\tilde{X}_t)_{t \in [0,T]}$  of  $(\tilde{X}_{t_m})_{0 \leq m \leq M}$  in (1.12). The objective of Section 4 is to prove the convergence of  $\tilde{X} = (\tilde{X}_t)_{t \in [0,T]}$  to the unique solution  $X$  in  $L^p$ -norm, which directly implies Theorem 4. To do this, we first study the properties of the interpolator  $i_m$  and link the uniform norm of the interpolated process and the interpolated marginal distributions with the underlying collection of points. In a second part, we prove that the sup norm of  $(\tilde{X}_t)_{t \in [0,T]}$  is bounded in  $L^p$  and study the  $L^p$ -norm of a specific modulus of continuity of  $(\tilde{X}_t)_{t \in [0,T]}$  adapted to our temporal discretization. The proof relies on a combination of functional inequalities with Lévy's modulus of continuity theorem for the control of the diffusive component. The use of the latter is the key point limiting our rate of convergence in the final result. Finally, we obtain Theorem 4 and Corollary 5 by combining the properties of the interpolated Euler scheme and its continuous extension with our assumptions on the drift and diffusion coefficients.

## 1.4 Notations

We place ourselves in a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  satisfying the usual condition. The law of any random variable  $X$  is denoted by  $P_X = \mathbb{P} \circ X^{-1}$ . Sometimes we write  $X \sim \mu$  to indicate that  $X$  has distribution  $\mu$ . In the whole paper,  $(B_t)_{t \in [0,T]}$  denotes a  $(\mathcal{F}_t)_{t \in [0,T]}$  Brownian motion valued in  $\mathbb{R}^q$ ,  $q \in \mathbb{N}^*$ . We denote by  $\mathcal{N}(0, \mathbf{I}_q)$  the  $\mathbb{R}^q$  standard normal distribution, where  $\mathbf{I}_q$  is the  $q \times q$  identity matrix. The  $L^p$  norm is denoted  $\|\cdot\|_p$  for  $p \in (0, \infty]$ . We write  $|\cdot|$  for the Euclidean norm on  $\mathbb{R}^d$ ,  $\delta_x$  for the Dirac measure at  $x$ ,  $\|\cdot\|$  for the operator norm on  $\mathbb{M}_{d,q}(\mathbb{R})$ , the space of matrices of dimensions  $(d, q)$ . We recall that, for  $A \in \mathbb{M}_{d,q}(\mathbb{R})$ ,

$$\|A\| := \sup_{z \in \mathbb{R}^q, |z| \leq 1} |Az|.$$



We write  $(E, \|\cdot\|_E)$  for the Banach space  $E$  endowed with the norm  $\|\cdot\|_E$ . Let  $\mathcal{P}(E)$  denote the set of probability distributions on  $E$ , while  $\mathcal{P}_p(E)$  denotes the set of probability distributions with  $p$ -th finite moment. We write  $\text{supp}(\mu)$  for the support of a probability distribution  $\mu$ . The Wasserstein distance on  $\mathcal{P}_p(\mathbb{R}^d)$  is denoted by  $\mathcal{W}_p$ , defined by (1.5). We shall use repeatedly the space of  $\mathbb{R}^d$ -valued continuous applications, denoted  $\mathcal{C}([0, T], \mathbb{R}^d)$ , that is,

$$\mathcal{C}([0, T], \mathbb{R}^d) := \left\{ \alpha = (\alpha_t)_{t \in [0, T]} \text{ such that } t \in [0, T] \mapsto \alpha_t \text{ is continuous} \right\}.$$

We endow this space with the supremum norm  $\|\alpha\|_{\text{sup}} = \sup_{t \in [0, T]} |\alpha_t|$ . The projection  $\pi_t : \mathcal{C}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}^d$  is defined by  $\pi_t(\alpha) = \alpha_t$  for all  $(\alpha_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathbb{R}^d)$ .

A handful of further spaces will be introduced throughout the text, that we gather here for clarity. The space  $L^p_{\mathcal{C}([0, T], \mathbb{R}^d)}(\Omega, \mathcal{F}, \mathbb{P})$  is the  $L^p$ -space of random variables defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  with values in  $\mathcal{C}([0, T], \mathbb{R}^d)$ . We endow this space with the norm  $\|\cdot\|_{p, \mathcal{C}, T}$  defined in (3.1). We further consider the space  $\mathcal{H}_{p, \mathcal{C}, T}$  of  $(\mathcal{F}_t)_{t \geq 0}$ -adapted processes in  $L^p_{\mathcal{C}([0, T], \mathbb{R}^d)}(\Omega, \mathcal{F}, \mathbb{P})$ , and the space  $\mathcal{P}_p(\mathcal{C}([0, T], \mathbb{R}^d))$  of probability distributions  $\mu$  on  $\mathcal{C}([0, T], \mathbb{R}^d)$  such that

$$\int_{\mathcal{C}([0, T], \mathbb{R}^d)} \|\xi\|_{\text{sup}}^p \mu(d\xi) < \infty.$$

The Wasserstein distance on  $\mathcal{P}_p(\mathcal{C}([0, T], \mathbb{R}^d))$  is denoted  $\mathbb{W}_p$ , see (3.4). For any two probability distributions  $\mu, \nu$ , the set of all probability distributions with marginals  $\mu$  and  $\nu$  is denoted  $\Pi(\mu, \nu)$ . We also introduce  $\mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  the space of probability distributions  $(\mu_t)_{t \in [0, T]}$  such that  $t \in [0, T] \mapsto \mu_t \in \mathcal{P}_p(\mathbb{R}^d)$  is continuous with respect to the distance  $\mathcal{W}_p$  (see (1.5)). We endow this space with the distance

$$d_p\left((\mu_t)_{t \in [0, T]}, (\nu_t)_{t \in [0, T]}\right) := \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \nu_t).$$

We also use the distance  $d_{\mathcal{H} \times \mathcal{P}}$  defined on  $\mathcal{H}_{p, \mathcal{C}, T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  in (3.12). At last, the application  $\iota$  is key to the fixed-point argument performed in Section 3, and sends elements from  $\mathcal{P}_p(\mathcal{C}([0, T], \mathbb{R}^d))$  to  $\mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  by  $\iota(\mu) = (\mu_t)_{t \in [0, T]}$ , see Lemma 11.

## 2 Applications

### 2.1 The mean-field Jansen-Rit model for multi-population neural networks

We consider the mean-field equations arising from Jansen and Rit's model [JR95], in the form of the equations given by Faugeras-Touboul-Cessac [FTC09]. This model includes three different neurons population and is used to get a deeper understanding of cortical signals, more specifically of the emergence of oscillations in the electrical activity of the brain registered by an electroencephalogram after a stimulation of a sensory pathway. The three populations are organised as follows: the pyramidal population, thereafter numbered 1, the excitatory feedback population, indexed by 2, and the inhibitory interneuron population, indexed by 3. More details on the model can be found in [FTC09], see in particular Figure 2 for a graphical representation.

At the level of the particle system, given a number  $N_j \in \mathbb{N}^*$  of neurons in population  $j$ , the equations for the potential of the neuron  $i$  in population  $j$  takes the following form

$$V_{j,i}(t) = V_{j,i}(0) + \int_0^t g_j(t-s) \left( \sum_{k=1}^3 \sum_{\ell=1}^{N_k} \bar{J}_{j,k} S(V_{k,\ell}(s)) + I_j(s) \right) ds + \int_0^t f_j(s) dW_s^{j,i}, \quad (2.1)$$

where, for  $t > 0$ ,  $g_j(t) = K_j e^{-\frac{t}{\tau_j}}$  represents the so-called  $g$ -shape of the postsynaptic potentials of population  $j$ , ensuring a modulation of the exchanges with time, while  $J_{k,j}$  is the strength of the postsynaptic potentials elicited by neurons from population  $j$  on neurons of population  $k$ . We assume that

$$K_1 = K_2 > 0, \quad K_3 > 0, \quad \tau_1 = \tau_2 > 0, \quad \tau_3 > 0, \quad \bar{J}_{i,j} = \frac{J_{i,j}}{N_j}$$

with

$$J = \begin{pmatrix} 0 & J_{1,2} & J_{1,3} \\ J_{2,1} & 0 & 0 \\ J_{3,1} & 0 & 0 \end{pmatrix}$$

where  $J_{1,2}, J_{1,3}, J_{2,1}$  and  $J_{3,1}$  are non-zero constants.

The functions  $I_j, f_j$  from  $\mathbb{R}_+$  to  $\mathbb{R}$  are assumed to be Lipschitz. Finally, the Brownian motions  $(W_t^{i,j})_{t \geq 0}$  for  $\{(i,j) : i \in \{1,2,3\}, j \in \{1, \dots, N_i\}\}$  are assumed to be mutually independent, and the function  $S$  is given on  $\mathbb{R}$  by

$$S(v) = \frac{v_m}{1 + e^{r(v_0 - v)}},$$

with  $r > 0$  and  $0 < v_0 < v_m$ .

As the number of particles in each population grows to infinity, it is natural to expect the system to be described by the following system of three McKean-Vlasov equations. Write  $\mu_t^j$  for the distribution of the potential of population  $j \in \{1,2,3\}$  at time  $t > 0$ . The mean-field limit of (2.1) is given by the system

$$\bar{V}_j(t) = \bar{V}_j(0) + \int_0^t g_j(t-s) \sum_{k=1}^3 \int_{\mathbb{R}} J_{j,k} S(x) \mu_s^k(dx) ds + \int_0^t g_j(t-s) I_j(s) ds + \int_0^t f_j(s) dW_s^j, \quad (2.2)$$

where  $(W^1, W^2, W^3)$  are three independent Brownian motions.

One can see clearly from the assumptions made on (2.1) that (2.2) fits our framework:

**Proposition 7.** *Under the previous hypotheses on  $g_j, J, S, I_j$  and  $f_j, j \in \{1,2,3\}$  the system (2.2) satisfies Assumptions (I) and (II).*

This provides an additional proof of well-posedness on finite time  $[0, T]$  for any  $T > 0$  for the model with  $\bar{V}_j(0) = 0$  for all  $j \in \{1,2,3\}$  as treated in [FTC09]. In addition, Proposition 7 induces a moment propagation result, in the sense that if  $\bar{V}_j(0) \in L^p, p \geq 2$  for  $j \in \{1,2,3\}$ , then  $\bar{V}_j(t) \in L^p$  at all time  $t \in [0, T]$ .

While the numerics of the limiting equation are investigated in [FTC09], the fact that (2.2) satisfies Assumption (II) provides a different path towards the simulation of the corresponding dynamics. In particular our interpolated Euler scheme is a first step towards the numerical investigation not only of the limiting equation, but also of the particle system and its convergence. For models of neuron masses, one key aspect is the understanding of the steady states and of their dependency with regards to the parameters of the system. As the parameters values change, different qualitative behaviors might appear, thus a bifurcation analysis must be conducted to get a clear picture of the possible outcomes. A numerical study of (2.2) and of the corresponding particle system could shed a new light on this matter: the non-Markovian property of the limiting equation could lead to new behaviors compared to classical model for neuron masses, see Faugeras [FTC09].

## 2.2 A regularized equation for the 2-dimensional parabolic-parabolic Keller-Segel model

In [Tom21], Tomašević provides a stochastic interpretation, based on earlier work of Talay-Tomašević [TT20], of the parabolic-parabolic Keller-Segel model via a stochastic representation which falls into the framework of (1.17). In particular, the drift of the corresponding process depends on the past of its law. The Keller-Segel equation describes the time evolution of the density  $\rho_t$  of a cell population, and of the concentration  $c_t$  of a chemical attractant. The term parabolic-parabolic refers to the fact that the chemical attractant itself is not constant in time, as opposed to alternative Keller-Segel models which involve time-dependency solely for the cell population. We refer to Horstmann [Hor03, Hor04] for a review of the standard Keller-Segel model and its variations.

In the study of the stochastic representation, Tomašević introduced a regularized problem, [Tom21, Equation (3.6)] on which we will focus. Let  $T > 0$  be fixed. The equation, set in  $[0, T] \times \mathbb{R}^2$ , takes the following form

$$\begin{cases} dX_t^\epsilon = b_0^\epsilon(t, X_t^\epsilon) + \chi \int_0^t \left[ e^{-\lambda(t-s)} (K_{t-s}^\epsilon * \mu_s^\epsilon)(X_t^\epsilon) ds \right] dt + dW_t, \\ \mu_s^\epsilon = P_{X_s^\epsilon}, \end{cases} \quad (2.3)$$

where  $\epsilon > 0$  is the regularization parameter,  $\lambda, \chi$  are positive constants, and

$$b_0^\epsilon(t, x) := \chi e^{-\lambda t} (\nabla c_0 * g_t^\epsilon)(x), \quad g_t^\epsilon(x) := \frac{1}{2\pi(t+\epsilon)} e^{-\frac{|x|^2}{2t}}, \quad K_t^\epsilon(x) := -\frac{x}{2\pi(t+\epsilon)^2} e^{-\frac{|x|^2}{2t}},$$

where  $c_0$  belongs to  $H^1(\mathbb{R}^2)$ , the usual Sobolev space.

The equation (2.3) is key to the argument in [Tom21] because drift and density estimates can be obtained for the regularized process, those being uniform in the regularization parameter  $\epsilon$  when a condition on the size of  $\chi$  is fulfilled.

A proof of well-posedness in  $L^1(\mathbb{R}^d)$  for the problem (2.3) is provided in [Tom21, Theorem A.1] with the initial condition that  $X_0$  is an  $\mathcal{F}_0$ -measurable random variable, where  $(\mathcal{F}_t)_{t \geq 0}$  is the filtration associated to the Brownian motion  $(W_t)_{t \geq 0}$ . A straightforward application of our result gives a similar well-posedness in the  $L^p$  framework for  $p \geq 2$ , implying also propagation of moments in finite time.

**Proposition 8.** *Let  $p \geq 2$ . Assume that  $T > 0$  and  $X_0 \in L^p$ . Then there exists a unique process  $(X_t)_{t \geq 0}$  solving (2.3) continuous in time such that  $\|\sup_{s \in [0, T]} |X_s|\|_p < \infty$ .*

The proof of Proposition 8, as well as the following Proposition 9, is postponed to Appendix A. We now turn to the issue of simulating the Keller-Segel parabolic-parabolic equation in dimension 2. As noticed in [Tom21], the existence of the particle system corresponding to the nonregularized particle system and its propagation of chaos is a difficult problem, tackled with the introduction of a Markovian (enriched) particle system in [FT22]. This gives a first approach to the simulation of the problem via the particle system.

As a first step towards a different approach, based on a direct Euler scheme for a non-markovian particle system associated to the regularized path-dependent equation thanks to the process from Definition 3, we prove that the interpolated Euler scheme corresponding to (2.3) converges, in  $L^p$  norm, to the desired solution, in the case where  $c_0$  also belongs to  $W^{1, \infty}(\mathbb{R}^d)$ .

**Proposition 9.** *Assume that  $c_0 \in H^1(\mathbb{R}^d) \cap W^{1, \infty}(\mathbb{R}^d)$ . Let  $(X_t)_{t \in [0, T]}$  be the unique strong solution to (2.3) given by Proposition 8. Let  $(\tilde{X}_{t_m}^M)_{0 \leq m \leq M}$  be the interpolated Euler scheme defined by (1.18). Then, for  $M$  large enough, for some  $\tilde{C} > 0$  independent of  $M$ , for  $h = \frac{T}{M}$ , we*

have

$$\left\| \sup_{0 \leq m \leq M} \left| X_{t_m} - \widetilde{X}_{t_m}^M \right| \right\|_p \leq \tilde{C} \left( h^\gamma + (h |\ln(h)|)^{\frac{1}{2}} \right).$$

### 3 Strong well-posedness

In this section, we prove Theorem 1, following the strategy sketched in Section 1.3. Assumption (I) is supposed to hold throughout the rest of this paper.

Let  $L^p_{\mathcal{C}([0,T],\mathbb{R}^d)}(\Omega, \mathcal{F}, \mathbb{P})$  denote the space of  $\mathcal{C}([0,T], \mathbb{R}^d)$ -valued r.v.  $Y = (Y_t)_{t \in [0,T]}$  having an  $L^p$ -norm  $\|Y\|_p := [\mathbb{E} \|Y\|_{\text{sup}}^p]^{1/p} = [\mathbb{E} \sup_{t \in [0,T]} |Y_t|^p]^{1/p} < +\infty$ . For a fixed constant  $C > 0$ , we define another norm  $\|\cdot\|_{p,C,T}$  on  $L^p_{\mathcal{C}([0,T],\mathbb{R}^d)}(\Omega, \mathcal{F}, \mathbb{P})$  by

$$\|Y\|_{p,C,T} := \sup_{t \in [0,T]} e^{-Ct} \left\| \sup_{0 \leq s \leq t} |Y_s| \right\|_p. \quad (3.1)$$

It is obvious that  $\|\cdot\|_{p,C,T}$  and  $\|\cdot\|_p$  are equivalent since

$$\forall Y \in L^p_{\mathcal{C}([0,T],\mathbb{R}^d)}(\Omega, \mathcal{F}, \mathbb{P}), \quad e^{-CT} \|Y\|_p \leq \|Y\|_{p,C,T} \leq \|Y\|_p. \quad (3.2)$$

We define

$$\mathcal{H}_{p,C,T} := \left\{ Y \in L^p_{\mathcal{C}([0,T],\mathbb{R}^d)}(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{P}) \text{ s.t. } Y \text{ is } (\mathcal{F}_t)_{t \in [0,T]} \text{ - adapted.} \right\}. \quad (3.3)$$

The next lemma shows that  $\mathcal{H}_{p,C,T}$  endowed with the norm  $\|\cdot\|_{p,C,T}$  is a Banach space. For simplicity we skip its proof, which can be found in [Liu19, Lemma 5.1.1].

**Lemma 10.** *The space  $\mathcal{H}_{p,C,T}$  equipped with  $\|\cdot\|_{p,C,T}$  is a complete space.*

For any random variable  $Y \in L^p_{\mathcal{C}([0,T],\mathbb{R}^d)}(\Omega, \mathcal{F}, \mathbb{P})$ , its probability distribution  $P_Y$  naturally lies in

$$\begin{aligned} & \mathcal{P}_p(\mathcal{C}([0,T], \mathbb{R}^d)) \\ & := \left\{ \mu \text{ probability distribution on } \mathcal{C}([0,T], \mathbb{R}^d) \text{ s.t. } \int_{\mathcal{C}([0,T], \mathbb{R}^d)} \|\alpha\|_{\text{sup}}^p \mu(d\alpha) < +\infty \right\}. \end{aligned}$$

We also define an  $L^p$ -Wasserstein distance  $\mathbb{W}_p$  on  $\mathcal{P}_p(\mathcal{C}([0,T], \mathbb{R}^d))$  by

$$\begin{aligned} & \forall \mu, \nu \in \mathcal{P}_p(\mathcal{C}([0,T], \mathbb{R}^d)), \\ & \mathbb{W}_p(\mu, \nu) := \left[ \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{C}([0,T], \mathbb{R}^d) \times \mathcal{C}([0,T], \mathbb{R}^d)} \|x - y\|_{\text{sup}}^p \pi(dx, dy) \right]^{\frac{1}{p}}, \end{aligned} \quad (3.4)$$

where  $\Pi(\mu, \nu)$  denote the set of probability measures on  $\mathcal{C}([0,T], \mathbb{R}^d) \times \mathcal{C}([0,T], \mathbb{R}^d)$  with respective marginals  $\mu$  and  $\nu$ . The space  $\mathcal{P}_p(\mathcal{C}([0,T], \mathbb{R}^d))$  equipped with  $\mathbb{W}_p$  is complete and separable since  $(\mathcal{C}([0,T], \mathbb{R}^d), \|\cdot\|_{\text{sup}})$  is a Polish space (see [Bol08]).

Let us consider now

$$\mathcal{C}([0,T], \mathcal{P}_p(\mathbb{R}^d))$$

$$:= \left\{ (\mu_t)_{t \in [0, T]} \text{ s.t. } t \mapsto \mu_t \text{ is a continuous application from } [0, T] \text{ to } (\mathcal{P}_p(\mathbb{R}^d), \mathcal{W}_p) \right\}$$

equipped with the distance

$$d_p((\mu_t)_{t \in [0, T]}, (\nu_t)_{t \in [0, T]}) := \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \nu_t). \quad (3.5)$$

As  $(\mathcal{P}_p(\mathbb{R}^d), \mathcal{W}_p)$  is a complete space (see [Bol08]),  $\mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  equipped with the uniform distance  $d_p$  is also a complete space.

For any  $t \in [0, T]$ , we define  $\pi_t : \mathcal{C}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}^d$  by  $\alpha \mapsto \pi_t(\alpha) = \alpha_t$ . The following lemma, and its proof, can be found in [Liu19, Lemma 5.1.2].

**Lemma 11.** *The application  $\iota : \mathcal{P}_p(\mathcal{C}([0, T], \mathbb{R}^d)) \rightarrow \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  defined by*

$$\mu \mapsto \iota(\mu) = (\mu \circ \pi_t^{-1})_{t \in [0, T]} = (\mu_t)_{t \in [0, T]}$$

*is well-defined.*

**Lemma 12.** *Under Assumption (I), the coefficient functions  $b$  and  $\sigma$  have a linear growth in  $\alpha$  and  $(\mu_t)_{t \in [0, T]}$  in the sense that there exists a constant  $C_{b, \sigma, L, T}$  s.t. for every  $t \in [0, T]$ ,  $\alpha \in \mathcal{C}([0, T], \mathbb{R}^d)$ ,  $(\mu_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ ,*

$$|b(t, \alpha, (\mu_t)_{t \in [0, T]})| \vee \|\sigma(t, \alpha, (\mu_t)_{t \in [0, T]})\| \leq C_{b, \sigma, L, T} \left( 1 + \|\alpha\|_{\text{sup}} + \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \delta_0) \right). \quad (3.6)$$

*Proof of Lemma 12.* Let  $\delta_{0, [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  be such that  $\forall t \in [0, T]$ ,  $\delta_{0, [0, T]}(t) = \delta_0$  and let  $\mathbf{0} \in \mathcal{C}([0, T], \mathbb{R}^d)$  be such that  $\forall t \in [0, T]$ ,  $\mathbf{0}(t) = 0$ . Then

$$\begin{aligned} & \left| b(t, \alpha, (\mu_t)_{t \in [0, T]}) \right| - \left| b(t, \mathbf{0}, \delta_{0, [0, T]}) \right| \leq \left| b(t, \alpha, (\mu_t)_{t \in [0, T]}) - b(t, \mathbf{0}, \delta_{0, [0, T]}) \right| \\ & \leq L \left( \|\alpha - \mathbf{0}\|_{\text{sup}} + d_p((\mu_t)_{t \in [0, T]}, \delta_{0, [0, T]}) \right) = L \left( \|\alpha\|_{\text{sup}} + \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \delta_0) \right). \end{aligned}$$

Consequently,

$$\left| b(t, \alpha, (\mu_t)_{t \in [0, T]}) \right| \leq \left( \sup_{t \in [0, T]} \left| b(t, \mathbf{0}, \delta_{0, [0, T]}) \right| \vee L \right) \left( \|\alpha\|_{\text{sup}} + \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \delta_0) + 1 \right). \quad (3.7)$$

Similarly, we have

$$\|\sigma(t, \alpha, (\mu_t)_{t \in [0, T]})\| \leq \left( \sup_{t \in [0, T]} \|\sigma(t, \mathbf{0}, \delta_{0, [0, T]})\| \vee L \right) \left( \|\alpha\|_{\text{sup}} + \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t, \delta_0) + 1 \right) \quad (3.8)$$

so that one can take  $C_{b, \sigma, L, T} := \sup_{t \in [0, T]} \left| b(t, \mathbf{0}, \delta_{0, [0, T]}) \right| \vee \sup_{t \in [0, T]} \|\sigma(t, \mathbf{0}, \delta_{0, [0, T]})\| \vee L$  to obtain (3.6).  $\square$

Before proving that the McKean-Vlasov equation (1.4) has a unique strong solution under Assumption (I), we first recall two important technical tools used throughout the proof: the generalized Minkowski Inequality and the Burkölder-Davis-Gundy Inequality. For the proof of these two inequalities, we refer to [Pag18, Section 7.8] among other references.

**Lemma 13** (The Generalized Minkowski Inequality). *For any (bi-measurable) process  $X =$*

$(X_t)_{t \geq 0}$ , for every  $p \in [1, \infty)$  and for every  $T \in [0, +\infty]$ ,

$$\left\| \int_0^T X_t dt \right\|_p \leq \int_0^T \|X_t\|_p dt.$$

**Lemma 14** (Burkholder-Davis-Gundy Inequality (continuous time)). *For every  $p \in (0, +\infty)$ , there exists two real constants  $c_p^{BDG} > 0$  and  $C_p^{BDG} > 0$  such that, for every continuous local martingale  $(X_t)_{t \in [0, T]}$  null at 0, denoting  $(\langle X \rangle_t)_{t \in [0, T]}$  its total variation process,*

$$c_p^{BDG} \left\| \sqrt{\langle X \rangle_T} \right\|_p \leq \left\| \sup_{t \in [0, T]} |X_t| \right\|_p \leq C_p^{BDG} \left\| \sqrt{\langle X \rangle_T} \right\|_p.$$

Note that under Assumption (I),  $t \rightarrow \sigma(t, X_{\cdot \wedge t}, \mu_{\cdot \wedge t})$  is adapted and continuous, hence progressively measurable. Recall also that  $p \geq 2$ . A direct application of those two inequalities provides the following lemma.

**Lemma 15.** *Let  $(B_t)_{t \in [0, T]}$  be a  $(\mathcal{F}_t)_{t \in [0, T]}$  standard Brownian motion, and  $(H_t)_{t \in [0, T]}$  be an  $(\mathcal{F}_t)_{t \in [0, T]}$  progressively measurable process having values in  $\mathbb{M}_{d, q}(\mathbb{R})$  such that  $\int_0^T \|H_t\|^2 dt < \infty$ ,  $\mathbb{P}$ -a.s.. Then, for all  $t \in [0, T]$ ,*

$$\left\| \sup_{s \in [0, t]} \left| \int_0^s H_u dB_u \right| \right\|_p \leq C_{d, p}^{BDG} \left[ \int_0^t \|H_u\|^2 du \right]^{\frac{1}{2}}.$$

*Proof.* Notice first that it follows from Lemma 14 that  $\int_0^t H_s dB_s$  is a  $d$ -dimensional local martingale satisfying

$$\left\| \sup_{s \in [0, t]} \left| \int_0^s H_u dB_u \right| \right\|_p \leq C_{d, p}^{BDG} \left\| \sqrt{\int_0^t \|H_u\|^2 du} \right\|_p. \quad (3.9)$$

Applying this, and using that when  $U \geq 0$ ,  $\|\sqrt{U}\|_p = \|U\|_{\frac{p}{2}}^{\frac{1}{2}}$ , we obtain

$$\left\| \sup_{s \in [0, t]} \left| \int_0^s H_u dB_u \right| \right\|_p \leq C_{d, p}^{BDG} \left\| \int_0^t \|H_u\|^2 du \right\|_{\frac{p}{2}}^{\frac{1}{2}} \leq C_{d, p}^{BDG} \left[ \int_0^t \|H_u\|^2 du \right]_{\frac{p}{2}}^{\frac{1}{2}}$$

where we used Minkowski's inequality (recall that  $p \geq 2$ ) to obtain the last inequality. The proof follows by noticing that  $\| |U|^2 \|_{\frac{p}{2}} = \|U\|_p^2$ .  $\square$

**Lemma 16.** *Under Assumption (I), for any  $(X, (\mu_t)_{t \in [0, T]})$ ,  $(Y, (\nu_t)_{t \in [0, T]})$  in the space  $\mathcal{H}_{p, C, T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  and for any  $t \in [0, T]$ , one has*

$$\begin{aligned} & \left\| \sup_{s \in [0, t]} \left| \int_0^s [b(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})] du \right| \right\|_p \\ & \leq L \int_0^t \left[ \left\| \sup_{s \in [0, u]} |X_s - Y_s| \right\|_p + \sup_{s \in [0, u]} \mathcal{W}_p(\mu_s, \nu_s) \right] du, \end{aligned}$$

and

$$\left\| \sup_{s \in [0, t]} \left| \int_0^s [\sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})] dB_u \right| \right\|_p$$

$$\leq C_{d,p,L} \left\{ \int_0^t \left[ \left\| \sup_{s \in [0,u]} |X_s - Y_s| \right\|_p^2 + \sup_{s \in [0,u]} \mathcal{W}_p^2(\mu_s, \nu_s) \right] du \right\}^{\frac{1}{2}},$$

where  $C_{d,p,L}$  is a positive constant only depending on  $d, p, L$ .

*Proof.* For any  $(X, (\mu_t)_{t \in [0,T]}), (Y, (\nu_t)_{t \in [0,T]}) \in \mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ , for any  $t \in [0, T]$ , we have

$$\begin{aligned} & \left\| \sup_{s \in [0,t]} \left| \int_0^s [b(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})] du \right| \right\|_p \\ & \leq \left\| \int_0^t |b(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})| du \right\|_p \\ & \leq \int_0^t \|b(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})\|_p du \quad (\text{by Lemma 13}) \\ & \leq \int_0^t \left\| L \left[ \|X_{\cdot \wedge u} - Y_{\cdot \wedge u}\|_{\text{sup}} + d_p((\mu_{v \wedge u})_{v \in [0,T]}, (\nu_{v \wedge u})_{v \in [0,T]}) \right] \right\|_p du \\ & \leq \int_0^t \left\| L \left[ \sup_{s \in [0,u]} |X_s - Y_s| + \sup_{s \in [0,u]} \mathcal{W}_p(\mu_s, \nu_s) \right] \right\|_p du \\ & \quad (\text{by Assumption (I) and by definitions (1.2) and (1.3)}) \tag{3.10} \\ & \leq L \int_0^t \left[ \left\| \sup_{s \in [0,u]} |X_s - Y_s| \right\|_p + \sup_{s \in [0,u]} \mathcal{W}_p(\mu_s, \nu_s) \right] du \end{aligned}$$

and, by applying Lemma 15,

$$\begin{aligned} & \left\| \sup_{s \in [0,t]} \left| \int_0^s [\sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})] dB_u \right| \right\|_p \\ & \leq C_{d,p}^{BDG} \left[ \int_0^t \left\| \|\sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})\| \right\|_p^2 du \right]^{\frac{1}{2}}. \end{aligned}$$

By Assumption (I) and by definition of  $\alpha_{\cdot \wedge u}$  and  $\mu_{\cdot \wedge u}$  in (1.2) and (1.3), we get

$$\begin{aligned} & \left\| \sup_{s \in [0,t]} \left| \int_0^s [\sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})] dB_u \right| \right\|_p \\ & \leq C_{d,p}^{BDG} \left[ \int_0^t \left\| L \left[ \sup_{s \in [0,u]} |X_s - Y_s| + \sup_{s \in [0,u]} \mathcal{W}_p(\mu_s, \nu_s) \right] \right\|_p^2 du \right]^{\frac{1}{2}} \\ & \leq C_{d,p}^{BDG} L \left[ \int_0^t \left[ \left\| \sup_{s \in [0,u]} |X_s - Y_s| \right\|_p + \sup_{s \in [0,u]} \mathcal{W}_p(\mu_s, \nu_s) \right]^2 du \right]^{\frac{1}{2}} \\ & \leq \sqrt{2} C_{d,p}^{BDG} L \left\{ \int_0^t \left[ \left\| \sup_{s \in [0,u]} |X_s - Y_s| \right\|_p^2 + \sup_{s \in [0,u]} \mathcal{W}_p^2(\mu_s, \nu_s) \right] du \right\}^{\frac{1}{2}}. \end{aligned}$$

The conclusion follows by letting  $C_{d,p,L} = \sqrt{2} C_{d,p}^{BDG} L$ .  $\square$

The idea of our proof of well-posedness follows from Feyel's approach, originally developed for the existence and uniqueness of a strong solution to the SDE  $dX_t = b(X_t)dt + \sigma(X_t)dB_t$ , see [Bou88, Section 7].

We define a distance  $d_{p,C,T}$  on  $\mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  as follows:

$$\forall (\mu_t)_{t \in [0,T]}, (\nu_t)_{t \in [0,T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)),$$

$$d_{p,C,T}((\mu_t)_{t \in [0,T]}, (\nu_t)_{t \in [0,T]}) := \sup_{t \in [0,T]} e^{-Ct} \mathcal{W}_p(\mu_t, \nu_t). \quad (3.11)$$

We also define a distance  $d_{\mathcal{H} \times \mathcal{P}}$  on  $\mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  as follows:

$$\begin{aligned} \forall (X, (\mu_t)_{t \in [0,T]}), (Y, (\nu_t)_{t \in [0,T]}) \in \mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)), \\ d_{\mathcal{H} \times \mathcal{P}}\left((X, (\mu_t)_{t \in [0,T]}), (Y, (\nu_t)_{t \in [0,T]})\right) = \|X - Y\|_{p,C,T} + \sup_{t \in [0,T]} e^{-Ct} \mathcal{W}_p(\mu_t, \nu_t). \end{aligned} \quad (3.12)$$

Recall that  $X_0 \in L^p(\mathbb{R}^d)$  is given by Assumption (I). We define an application<sup>1</sup>

$$\Phi_C : \mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \rightarrow \mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$$

by

$$\begin{aligned} \forall (Y, (\nu_t)_{t \in [0,T]}) \in \mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)), \\ \Phi_C(Y, (\nu_t)_{t \in [0,T]}) \\ = \left( \underbrace{\left( X_0 + \int_0^t b(s, Y_{\cdot \wedge s}, \nu_{\cdot \wedge s}) ds + \int_0^t \sigma(s, Y_{\cdot \wedge s}, \nu_{\cdot \wedge s}) dB_s \right)}_{=: \Phi_C^{(1)}(Y, (\nu_t)_{t \in [0,T]})}, \iota(P_{\Phi_C^{(1)}(Y, (\nu_t)_{t \in [0,T]})}) \right). \end{aligned}$$

The application  $\Phi_C$  has the following properties.

**Proposition 17.** (i) Under Assumption (I), the map  $\Phi_C$  is well-defined, and there holds

$$\begin{aligned} \left\| \sup_{s \in [0,t]} \left| \Phi_C^{(1)}(Y, \iota(P_Y))_s \right| \right\|_p \leq \|X_0\|_p + C_{b,\sigma,L,T}(2T + C_{d,p}^{BDG} \sqrt{2T}) + 2C_{b,\sigma,L,T} \int_0^t \left\| \sup_{s \in [0,u]} |Y_s| \right\|_p du \\ + 2\sqrt{2} C_{d,p}^{BDG} C_{b,\sigma,L,T} \left( \int_0^t \left\| \sup_{s \in [0,u]} |Y_s| \right\|_p^2 du \right)^{\frac{1}{2}}. \end{aligned} \quad (3.13)$$

(ii) Under Assumption (I),  $\Phi_C$  is Lipschitz continuous in the sense that: for any  $(X, \iota(P_X))$  and  $(Y, \iota(P_Y))$  in  $\mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ ,

$$d_{\mathcal{H} \times \mathcal{P}}\left(\Phi_C(X, \iota(P_X)), \Phi_C(Y, \iota(P_Y))\right) \leq \left( \frac{K_1}{C} + \frac{K_2}{\sqrt{C}} \right) d_{\mathcal{H} \times \mathcal{P}}\left((X, \iota(P_X)), (Y, \iota(P_Y))\right),$$

where  $K_1, K_2$  are real positive constants which do not depend on the constant  $C$ .

*Proof.* (i) It follows from Lemma 11 that for every  $Y \in \mathcal{H}_{p,C,T}$ ,  $\iota(P_Y) \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ .

Let  $\nu = P_Y$ . We only need to prove  $\Phi_C^{(1)}(Y, \iota(\nu)) \in \mathcal{H}_{p,C,T}$ . For any  $t \in [0, T]$ ,

$$\begin{aligned} \left\| \sup_{s \in [0,t]} \left| \Phi_C^{(1)}(Y, \iota(\nu))_s \right| \right\|_p &= \left\| \sup_{s \in [0,t]} \left| X_0 + \int_0^s b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u}) du + \int_0^s \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u}) dB_u \right| \right\|_p \\ &\leq \left\| |X_0| + \int_0^t |b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})| du + \sup_{s \in [0,t]} \left| \int_0^s \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u}) dB_u \right| \right\|_p \\ &\leq \|X_0\|_p + \left\| \int_0^t |b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})| du \right\|_p + \left\| \sup_{s \in [0,t]} \left| \int_0^s \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u}) dB_u \right| \right\|_p. \end{aligned} \quad (3.14)$$

<sup>1</sup>The  $C$  in the subscript of  $\Phi_C$  is the same constant  $C$  as in  $(\mathcal{H}_{p,C,T}, \|\cdot\|_{p,C,T})$ . We carry this notation throughout this section.



Owing to Assumption (I), we have  $\|X_0\|_p < +\infty$ . For the second part of (3.14), it follows from Lemma 12 that

$$\begin{aligned}
\left\| \int_0^t b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u}) du \right\|_p &\leq \int_0^t \|b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})\|_p du \\
&\leq \int_0^t \left\| C_{b,\sigma,L,T} (1 + \|Y_{\cdot \wedge u}\|_{\sup} + \sup_{s \in [0,T]} \mathcal{W}_p(\nu_{s \wedge u}, \delta_0)) \right\|_p du \\
&\leq \int_0^t C_{b,\sigma,L,T} (1 + \left\| \sup_{s \in [0,T]} |Y_{s \wedge u}| \right\|_p + \sup_{s \in [0,T]} \|Y_{s \wedge u}\|_p) du \\
&\leq 2C_{b,\sigma,L,T} \int_0^t \left( 1 + \left\| \sup_{s \in [0,T]} |Y_{s \wedge u}| \right\|_p \right) du \\
&\leq 2C_{b,\sigma,L,T} \int_0^t \left( 1 + \left\| \sup_{s \in [0,u]} |Y_s| \right\|_p \right) du
\end{aligned} \tag{3.15}$$

where we used that  $\sup_{s \in [0,T]} \|Y_{s \wedge u}\|_p \leq \left\| \sup_{s \in [0,T]} |Y_{s \wedge u}| \right\|_p$  and

$$\left\| \sup_{s \in [0,T]} |Y_{s \wedge u}| \right\|_p = \left\| \sup_{s \in [0,u]} |Y_s| \right\|_p.$$

By definition of  $\|\cdot\|_{p,C,T}$ , we deduce

$$\left\| \int_0^t b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u}) du \right\|_p \leq 2C_{b,\sigma,L,T} \int_0^t (1 + e^{CT} \|Y\|_{p,C,T}) du < +\infty,$$

where we used (3.2) in the following way

$$\left\| \sup_{s \in [0,u]} |Y_s| \right\|_p \leq \left\| \sup_{s \in [0,T]} |Y_s| \right\|_p \leq e^{CT} \|Y\|_{p,C,T} < +\infty. \tag{3.16}$$

On the other hand, by using Lemma 15,

$$\left\| \sup_{s \in [0,t]} \left| \int_0^s \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u}) dB_u \right| \right\|_p \leq C_{d,p}^{BDG} \left[ \int_0^t \|\sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})\|_p^2 du \right]^{\frac{1}{2}}.$$

As before, we then invoke Lemma 12 to obtain

$$\begin{aligned}
&\left\| \sup_{s \in [0,t]} \left| \int_0^s \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u}) dB_u \right| \right\|_p \\
&\leq C_{d,p}^{BDG} \left\{ \int_0^t \left\| C_{b,\sigma,L,T} (1 + \|Y_{\cdot \wedge u}\|_{\sup} + \sup_{s \in [0,T]} \mathcal{W}_p(\nu_{s \wedge u}, \delta_0)) \right\|_p^2 du \right\}^{\frac{1}{2}} \\
&\leq C_{d,p}^{BDG} C_{b,\sigma,L,T} \left\{ \int_0^t \left[ 1 + \left\| \sup_{s \in [0,T]} |Y_{s \wedge u}| \right\|_p + \sup_{s \in [0,T]} \mathcal{W}_p(\nu_{s \wedge u}, \delta_0) \right]^2 du \right\}^{\frac{1}{2}} \\
&\leq C_{d,p}^{BDG} C_{b,\sigma,L,T} \left\{ \int_0^t \left[ 1 + \left\| \sup_{s \in [0,T]} |Y_{s \wedge u}| \right\|_p + \sup_{s \in [0,T]} \|Y_{s \wedge u}\|_p \right]^2 du \right\}^{\frac{1}{2}} \\
&\leq C_{d,p}^{BDG} C_{b,\sigma,L,T} \left\{ \int_0^t \left[ 1 + 2 \left\| \sup_{s \in [0,T]} |Y_{s \wedge u}| \right\|_p \right]^2 du \right\}^{\frac{1}{2}},
\end{aligned}$$

where we used again the convex inequality  $\sup_{s \in [0,T]} \|Y_{s \wedge u}\|_p \leq \left\| \sup_{s \in [0,T]} |Y_{s \wedge u}| \right\|_p$ . Using again

$\left\| \sup_{s \in [0, T]} |Y_{s \wedge u}| \right\|_p = \left\| \sup_{s \in [0, u]} |Y_s| \right\|_p$ , we get

$$\begin{aligned} \left\| \sup_{s \in [0, t]} \left| \int_0^s \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u}) dB_u \right| \right\|_p &\leq C_{d,p}^{BDG} C_{b,\sigma,L,T} \left\{ 2T + 8 \int_0^t \left\| \sup_{s \in [0, u]} |Y_s| \right\|_p^2 du \right\}^{\frac{1}{2}} \\ &\leq C_{d,p}^{BDG} C_{b,\sigma,L,T} \left( \sqrt{2T} + 2\sqrt{2} \left\{ \int_0^t \left\| \sup_{s \in [0, u]} |Y_s| \right\|_p^2 du \right\}^{\frac{1}{2}} \right) \\ &< +\infty, \end{aligned} \quad (3.17)$$

where the last inequality of the above formula is due to (3.16), and where we used  $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$  for  $a, b \geq 0$ .

Hence for every  $t \in [0, T]$ ,  $\left\| \sup_{s \in [0, t]} |\Phi_C^{(1)}(Y, \iota(\nu))_s| \right\|_p < +\infty$ , which directly implies

$$\left\| \Phi_C^{(1)}(Y, \iota(\nu)) \right\|_{p,C,T} = \sup_{t \in [0, T]} e^{-Ct} \left\| \sup_{s \in [0, t]} |\Phi_C^{(1)}(Y, \iota(\nu))_s| \right\|_p < +\infty.$$

Thus  $\Phi_C^{(1)}(Y, \iota(\nu)) \in \mathcal{H}_{p,C,T}$ . The inequality (3.13) follows by injecting (3.15) and (3.17) into (3.14).

(ii) We split the proof of this inequality into three steps.

**Step 1.** We first prove that for any  $X, Y \in \mathcal{H}_{p,C,T}$ ,  $d_{p,C,T}(\iota(P_X), \iota(P_Y)) \leq \|X - Y\|_{p,C,T}$ . In fact

$$\begin{aligned} d_{p,C,T}(\iota(P_X), \iota(P_Y)) &= \sup_{t \in [0, T]} e^{-Ct} \mathcal{W}_p(P_X \circ \pi_t^{-1}, P_Y \circ \pi_t^{-1}) \leq \sup_{t \in [0, T]} e^{-Ct} \|X_t - Y_t\|_p \\ &\leq \sup_{t \in [0, T]} e^{-Ct} \left\| \sup_{s \in [0, t]} |X_s - Y_s| \right\|_p = \|X - Y\|_{p,C,T}. \end{aligned}$$

**Step 2.** We prove that  $\Phi_C^{(1)}$  is Lipschitz continuous, in the sense that

$$\left\| \Phi_C^{(1)}(X, \iota(\mu)) - \Phi_C^{(1)}(Y, \iota(\nu)) \right\|_{p,C,T} \leq \left( \frac{2L}{C} + \frac{C_{d,p,L}}{\sqrt{C}} \right) \|X - Y\|_{p,C,T}, \quad (3.18)$$

where  $C_{d,p,L} > 0$  is the constant given by Lemma 16 and is independent of the parameter  $C$  of the application  $\Phi_C$ . For any  $X, Y \in \mathcal{H}_{p,C,T}$ , set  $\mu = P_X$  and  $\nu = P_Y$ . Then

$$\begin{aligned} &\left\| \Phi_C^{(1)}(X, \iota(\mu)) - \Phi_C^{(1)}(Y, \iota(\nu)) \right\|_{p,C,T} \\ &= \left\| \int_0^\cdot (b(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})) du + \int_0^\cdot (\sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})) dB_u \right\|_{p,C,T} \\ &\leq \left\| \int_0^\cdot (b(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})) du \right\|_{p,C,T} \\ &\quad + \left\| \int_0^\cdot (\sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})) dB_u \right\|_{p,C,T} \\ &= \sup_{t \in [0, T]} e^{-Ct} \left\| \sup_{s \in [0, t]} \left| \int_0^s [b(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})] du \right| \right\|_p \end{aligned} \quad (3.19)$$

$$+ \sup_{t \in [0, T]} e^{-Ct} \left\| \sup_{s \in [0, t]} \left| \int_0^s [\sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})] dBu \right| \right\|_p.$$

We treat the two terms in (3.19) separately. Owing to Lemma 16, we first have

$$\begin{aligned} & \sup_{t \in [0, T]} e^{-Ct} \left\| \sup_{s \in [0, t]} \left| \int_0^s [b(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - b(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})] du \right| \right\|_p \\ & \leq L \sup_{t \in [0, T]} e^{-Ct} \int_0^t \left[ \left\| \sup_{s \in [0, u]} |X_s - Y_s| \right\|_p + \sup_{s \in [0, u]} \mathcal{W}_p(\mu_s, \nu_s) \right] du \\ & \leq L \sup_{t \in [0, T]} e^{-Ct} \int_0^t \left[ \left\| \sup_{s \in [0, u]} |X_s - Y_s| \right\|_p + \sup_{s \in [0, u]} \|X_s - Y_s\|_p \right] du \\ & \quad (\text{since } \mathcal{W}_p(\mu_s, \nu_s) \leq \|X_s - Y_s\|_p) \\ & \leq 2L \sup_{t \in [0, T]} e^{-Ct} \int_0^t e^{Cu} \left( e^{-Cu} \left\| \sup_{s \in [0, u]} |X_s - Y_s| \right\|_p \right) du \\ & \leq 2L \sup_{t \in [0, T]} e^{-Ct} \int_0^t e^{Cu} du \|X - Y\|_{p, C, T} \quad (\text{by the definition of } \|\cdot\|_{p, C, T} \text{ in (3.1)}) \\ & = 2L \sup_{t \in [0, T]} e^{-Ct} \frac{e^{Ct} - 1}{C} \|X - Y\|_{p, C, T} \\ & \leq \frac{2L}{C} \|X - Y\|_{p, C, T}. \end{aligned}$$

On the other hand

$$\begin{aligned} & \sup_{t \in [0, T]} e^{-Ct} \left\| \sup_{s \in [0, t]} \left| \int_0^s [\sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(u, Y_{\cdot \wedge u}, \nu_{\cdot \wedge u})] dBu \right| \right\|_p \\ & \leq \sup_{t \in [0, T]} e^{-Ct} C_{d, p, L} \left\{ \int_0^t \left[ \left\| \sup_{s \in [0, u]} |X_s - Y_s| \right\|_p^2 + \sup_{s \in [0, u]} \mathcal{W}_p^2(\mu_s, \nu_s) \right] du \right\}^{\frac{1}{2}} \quad (\text{by Lemma 16}) \\ & \leq \sup_{t \in [0, T]} e^{-Ct} C_{d, p, L} \left\{ \int_0^t \left[ \left\| \sup_{s \in [0, u]} |X_s - Y_s| \right\|_p^2 + \sup_{s \in [0, u]} \|X_s - Y_s\|_p^2 \right] du \right\}^{\frac{1}{2}} \\ & \leq \sqrt{2} C_{d, p, L} \sup_{t \in [0, T]} e^{-Ct} \left\{ \int_0^t e^{2Cu} \left( e^{-Cu} \left\| \sup_{s \in [0, u]} |X_s - Y_s| \right\|_p \right)^2 du \right\}^{\frac{1}{2}} \\ & \leq \sqrt{2} C_{d, p, L} \|X - Y\|_{p, C, T} \sup_{t \in [0, T]} e^{-Ct} \left\{ \int_0^t e^{2Cu} du \right\}^{\frac{1}{2}} \\ & \leq \sqrt{2} C_{d, p, L} \|X - Y\|_{p, C, T} \sup_{t \in [0, T]} e^{-Ct} \left[ \frac{e^{2Ct} - 1}{2C} \right]^{\frac{1}{2}} \\ & \leq \frac{C_{d, p, L}}{\sqrt{C}} \|X - Y\|_{p, C, T}, \end{aligned}$$

since  $\sup_{t \in [0, T]} e^{-Ct} \left[ \frac{e^{2Ct} - 1}{2C} \right]^{\frac{1}{2}} \leq \frac{1}{\sqrt{2C}}$ . Injecting those two results in (3.19), we obtain (3.18).

**Step 3.** We combine the results of the previous two steps to conclude. We have, from the definition of  $d_{\mathcal{H} \times \mathcal{P}}$  (3.12),

$$\begin{aligned} & d_{\mathcal{H} \times \mathcal{P}} \left( \Phi_C(X, \iota(\mu)), \Phi_C(Y, \iota(\nu)) \right) \\ & = \left\| \Phi_C^{(1)}(X, \iota(\mu)) - \Phi_C^{(1)}(Y, \iota(\nu)) \right\|_{p, C, T} + d_{p, C, T} \left( P_{\Phi_C^{(1)}(X, \iota(\mu))}, P_{\Phi_C^{(1)}(Y, \iota(\nu))} \right) \end{aligned}$$

and using Step 1, and (3.18),

$$\begin{aligned} d_{\mathcal{H} \times \mathcal{P}}\left(\Phi_C(X, \iota(\mu)), \Phi_C(Y, \iota(\nu))\right) &\leq 2\left(\frac{2L}{C} + \frac{C_{d,p,L}}{\sqrt{C}}\right) \|X - Y\|_{p,C,T} \\ &\leq 2\left(\frac{2L}{C} + \frac{C_{d,p,L}}{\sqrt{C}}\right) d_{\mathcal{H} \times \mathcal{P}}\left((X, \mu), (Y, \nu)\right). \end{aligned}$$

The proof follows by letting  $K_1 = 4L$  and  $K_2 = 2C_{d,p,L}$ .  $\square$

To obtain the precise description of the upper bound of  $\|\sup_{s \in [0, T]} |X_s|\|_p$ , we will need the following version of Gronwall's lemma. We refer to [Pag18, Lemma 7.3] for a proof (among many others).

**Lemma 18** (“À la Gronwall” Lemma). *Let  $f : [0, T] \rightarrow \mathbb{R}_+$  be a Borel, locally bounded and non-decreasing function and let  $\psi : [0, T] \rightarrow \mathbb{R}_+$  be a non-negative non-decreasing function satisfying*

$$\forall t \in [0, T], \quad f(t) \leq A \int_0^t f(s) ds + B \left( \int_0^t f^2(s) ds \right)^{\frac{1}{2}} + \psi(t),$$

where  $A, B$  are two positive real constants. Then, for any  $t \in [0, T]$ ,

$$f(t) \leq 2e^{(2A+B^2)t} \psi(t).$$

Proposition 17 directly implies the existence and uniqueness of a strong solution of the McKean-Vlasov equation (1.4) as shown below.

*Proof of Theorem 1.* Proposition 17 implies that  $\Phi_C$  is a Lipschitz continuous function. Thus,  $F_C := \Phi_C(\mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d)))$  is a closed set in  $\mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ . Moreover, for a large enough constant  $C$ , we have  $(\frac{K_1}{C} + \frac{K_2}{\sqrt{C}}) < 1$ , so that  $\Phi_C$  is a contraction mapping. Therefore,  $\Phi_C$  has a unique fixed point  $(H, \iota(P_H)) \in F_C \subset \mathcal{H}_{p,C,T} \times \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  and this process  $H$  is the unique strong solution of the McKean-Vlasov equation (1.4).

We turn to the proof of (1.8). Let  $(X, P_X)$  be the unique strong solution of (1.4). Then,  $\Phi_C^{(1)}(X, \iota(P_X)) = X$  since  $X$  is a fixed point of the application  $\Phi_C$ . Therefore, (3.13) takes the following form

$$\begin{aligned} \left\| \sup_{s \in [0, t]} |X_s| \right\|_p &\leq \|X_0\|_p + C_{b,\sigma,L,T}(2T + C_{d,p}^{BDG} \sqrt{2T}) + 2C_{b,\sigma,L,T} \int_0^t \left\| \sup_{s \in [0, u]} |X_s| \right\|_p du \\ &\quad + 2\sqrt{2} C_{d,p}^{BDG} C_{b,\sigma,L,T} \left( \int_0^t \left\| \sup_{s \in [0, u]} |X_s| \right\|_p^2 du \right)^{\frac{1}{2}}. \end{aligned}$$

We let  $f(t) := \left\| \sup_{s \in [0, t]} |X_s| \right\|_p$ , and apply Lemma 18 to get

$$\left\| \sup_{s \in [0, t]} |X_s| \right\|_p \leq C_{p,d,b,\sigma,L,T} e^{C_{p,d,b,\sigma,L,T} t} (1 + \|X_0\|_p),$$

with the constant  $C_{p,d,b,\sigma,L,T} > 0$  defined by

$$C_{p,d,b,\sigma,L,T} = (4C_{b,\sigma,L,T} + 8(C_{d,p}^{BDG} C_{b,\sigma,L,T})^2) \vee 2(1 \vee C_{b,\sigma,L,T} T + \sqrt{2T} C_{d,p}^{BDG} C_{b,\sigma,L,T}). \quad (3.20)$$

The conclusion follows by choosing  $t = T$  and  $\Gamma = C_{p,d,b,\sigma,L,T} e^{C_{p,d,b,\sigma,L,T} T}$ .  $\square$

## 4 Interpolated Euler scheme and associated convergences

This section is devoted to the proof of Theorem 4. For this purpose, we first define the associated theoretical continuous Euler scheme  $(\widetilde{X}_t^M)_{t \in [0, T]}$  of  $(\widetilde{X}_{t_m}^M)_{0 \leq m \leq M}$  from (1.12). We use the same temporal discretization as Definition 3: let  $M \in \mathbb{N}^*$ ,  $h = \frac{T}{M}$ . For every  $m = 0, \dots, M$ , we set  $t_m = mh$ . As the size of the discretization parameter  $M$  will sometimes play a role, we write  $(\widetilde{X}_t^M)_{t \geq 0}$  when we wish to emphasize the dependency of the process in  $M$  and omit this superscript when it is clear from context.

**Definition 19.** Given a Brownian motion  $(B_t)_{t \in [0, T]}$ ,  $X_0 \in L^p(\mathbb{R}^d)$  and the discretize scheme  $(\widetilde{X}_{t_0:t_m}^M)$  with the associated probability distributions  $(\widetilde{\mu}_{t_0:t_m}^M)$  from Definition 3, using the same notations  $b_m, \sigma_m$  as in (1.13), we define the continuous Euler scheme,  $(\widetilde{X}_t^M)_{t \in [0, T]}$  by setting, for all  $t \in (t_m, t_{m+1}]$ ,

$$\begin{cases} \widetilde{X}_0^M &= X_0, \\ \widetilde{X}_t^M &= \widetilde{X}_{t_m}^M + (t - t_m) b_m(t_m, \widetilde{X}_{t_0:t_m}^M, \widetilde{\mu}_{t_0:t_m}^M) + \sigma_m(t_m, \widetilde{X}_{t_0:t_m}^M, \widetilde{\mu}_{t_0:t_m}^M)(B_t - B_{t_m}). \end{cases} \quad (4.1)$$

According to the definition of  $b_m$  and  $\sigma_m$  in (1.13), the continuous Euler scheme (4.1) writes, for  $t \in (t_m, t_{m+1}]$ ,  $m \in \{0, \dots, M-1\}$ ,

$$\widetilde{X}_t^M = \widetilde{X}_{t_m}^M + (t - t_m) b\left(t_m, i_m(\widetilde{X}_{t_0:t_m}^M), i_m(\widetilde{\mu}_{t_0:t_m}^M)\right) + \sigma_m\left(t_m, i_m(\widetilde{X}_{t_0:t_m}^M), i_m(\widetilde{\mu}_{t_0:t_m}^M)\right)(B_t - B_{t_m}).$$

In order to compare this with equation (1.4), we write, for all  $t \in [0, T]$ ,  $\widetilde{\mu}_t^M$  for the distribution of  $\widetilde{X}_t^M$ , and for all  $m \in \{0, \dots, M-1\}$  we set

$$\underline{t} := t_m, \quad [\underline{t}] := m \quad \text{if } t \in [t_m, t_{m+1}). \quad (4.2)$$

With this at hand, the process  $(\widetilde{X}_t^M)_{t \in [0, T]}$  satisfies

$$\widetilde{X}_t^M = \widetilde{X}_0^M + \int_0^t b\left(\underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}^M), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}^M)\right) ds + \int_0^t \sigma\left(\underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}^M), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}^M)\right) dB_s. \quad (4.3)$$

Theorem 4 is a direct result of the following proposition.

**Proposition 20.** Under Assumptions (I) and (II), for  $(B_t)_{t \in [0, T]}$  an  $\mathbb{R}^q$ -valued Brownian motion, for  $(X_t)_{t \in [0, T]}$  the unique strong solution to (1.4) given by Theorem 1, for  $(\widetilde{X}_t^M)_{t \in [0, T]}$  the Euler scheme from Definition 19 with parameter  $M$  large enough, for  $h = \frac{T}{M}$ , one has,

$$\left\| \sup_{t \in [0, T]} \left| X_t - \widetilde{X}_t^M \right| \right\|_p \leq \tilde{C} \left( h^\gamma + (h |\ln(h)|)^{\frac{1}{2}} \right), \quad (4.4)$$

where  $\tilde{C} > 0$  is a constant depending on  $L, p, d, \|X_0\|_p, T$  and  $\gamma$ .

This section is organized as follows. Section 4.1 shows several preliminary results for the interpolator  $i_m$  that will be used for the proof of Proposition 20. Section 4.2 gathers several properties of the process  $(\widetilde{X}_t^M)_{t \geq 0}$  from Definition 19. Finally, in Section 4.3, we prove Proposition 20, Theorem 4 and Corollary 5.

*Remark 21.* We might define the classical continuous Euler scheme  $(\bar{X}_t)_{t \geq 0}$  by setting

1.  $\bar{X}_0 = X_0$ ;

2. for all  $m \in \{0, \dots, M-1\}$ ,  $t \in (t_m, t_{m+1}]$

$$\bar{X}_t = \bar{X}_{t_m} + (t - t_m) b(t_m, \bar{X}_{\cdot \wedge t_m}, \bar{\mu}_{\cdot \wedge t_m}) + \sigma(t_m, \bar{X}_{\cdot \wedge t_m}, \bar{\mu}_{\cdot \wedge t_m}) (B_t - B_{t_m}).$$

The convergence of this non-implementable continuous Euler scheme towards the solution of (1.4) can then be proved using similar arguments as those developed in this section.

#### 4.1 Preliminary results

We gather several properties that will be used for the proof of Proposition 20. For two probability measures  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$  and for  $\lambda \in [0, 1]$ , we define  $\lambda\mu + (1 - \lambda)\nu$  by

$$\forall B \in \mathcal{B}(\mathbb{R}^d), \quad (\lambda\mu + (1 - \lambda)\nu)(B) := \lambda\mu(B) + (1 - \lambda)\nu(B). \quad (4.5)$$

It is easy to check that  $\lambda\mu + (1 - \lambda)\nu \in \mathcal{P}_p(\mathbb{R}^d)$ .

**Lemma 22.** *Let  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$  with  $p \geq 1$ .*

(a) *The application  $\tau : \lambda \in [0, 1] \mapsto \tau(\lambda) = \lambda\mu + (1 - \lambda)\nu \in \mathcal{P}_p(\mathbb{R}^d)$  is  $\frac{1}{p}$ -Hölder continuous with respect to the Wasserstein distance  $\mathcal{W}_p$ . Moreover, for every  $\lambda_1, \lambda_2 \in [0, 1]$ , we have*

$$\mathcal{W}_p(\tau(\lambda_1), \tau(\lambda_2)) \leq |\lambda_1 - \lambda_2|^{\frac{1}{p}} \mathcal{W}_p(\mu, \nu).$$

(b) *Let  $\delta_0$  denote the Dirac measure on  $0 \in \mathbb{R}^d$ . Then*

$$\sup_{\lambda \in [0, 1]} \mathcal{W}_p(\tau(\lambda), \delta_0) \leq \mathcal{W}_p(\mu, \delta_0) \vee \mathcal{W}_p(\nu, \delta_0).$$

Remark that Lemma 22 implies that the interpolator  $i_m$  defined by (1.10) and (1.11) is well defined.

*Proof of Lemma 22.* Let  $X, Y$  be such that  $P_X = \mu$ ,  $P_Y = \nu$  and consider another random variable  $U$  having uniform distribution on  $[0, 1]$ , independent of  $(X, Y)$ . One can easily check that for all  $\lambda \in [0, 1]$ ,

$$1_{\{U \leq \lambda\}} X + 1_{\{U > \lambda\}} Y \sim \tau(\lambda).$$

(a) Let  $\lambda_1, \lambda_2 \in [0, 1]$ . We assume without loss of generality that  $\lambda_1 < \lambda_2$ . We have

$$\begin{aligned} & \mathcal{W}_p^p(\tau(\lambda_1), \tau(\lambda_2)) \\ & \leq \mathbb{E} \left[ \left| 1_{\{U \leq \lambda_1\}} X + 1_{\{U > \lambda_1\}} Y - 1_{\{U \leq \lambda_2\}} X - 1_{\{U > \lambda_2\}} Y \right|^p \right] \\ & = \mathbb{E} \left[ \left| -1_{\{\lambda_1 < U \leq \lambda_2\}} X + 1_{\{\lambda_1 < U \leq \lambda_2\}} Y \right|^p \right] \\ & = \mathbb{E} \left[ 1_{\{\lambda_1 < U \leq \lambda_2\}} |X - Y|^p \right] = \mathbb{P}(\lambda_1 < U \leq \lambda_2) \mathbb{E} [|X - Y|^p] \quad (\text{as } U \perp (X, Y)) \\ & = (\lambda_2 - \lambda_1) \mathbb{E} [|X - Y|^p]. \end{aligned} \quad (4.6)$$

Taking the infimum over  $(X, Y) \in \Pi(\mu, \nu)$ , we find

$$\mathcal{W}_p(\tau(\lambda_1), \tau(\lambda_2)) \leq (\lambda_2 - \lambda_1)^{\frac{1}{p}} \mathcal{W}_p(\mu, \nu),$$

where  $\mathcal{W}_p(\mu, \nu)$  is finite since  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ . This concludes the proof of (a).

(b) For every fixed  $\lambda \in [0, 1]$ ,

$$\begin{aligned}
\mathcal{W}_p^p(\tau(\lambda), \delta_0) &= \mathbb{E} \left[ \left| X1_{\{U \leq \lambda\}} + Y1_{\{U > \lambda\}} \right|^p \right] \\
&= \mathbb{E} \left[ \left| X1_{\{U \leq \lambda\}} + Y1_{\{U > \lambda\}} \right|^p 1_{\{U \leq \lambda\}} \right] + \mathbb{E} \left[ \left| X1_{\{U \leq \lambda\}} + Y1_{\{U > \lambda\}} \right|^p 1_{\{U > \lambda\}} \right] \\
&= \mathbb{E} \left[ |X|^p 1_{\{U \leq \lambda\}} \right] + \mathbb{E} \left[ |Y|^p 1_{\{U > \lambda\}} \right] = \lambda \mathbb{E} \left[ |X|^p \right] + (1 - \lambda) \mathbb{E} \left[ |Y|^p \right] \\
&\leq \lambda \mathcal{W}_p^p(\mu, \delta_0) + (1 - \lambda) \mathcal{W}_p^p(\nu, \delta_0) \leq \mathcal{W}_p^p(\mu, \delta_0) \vee \mathcal{W}_p^p(\nu, \delta_0).
\end{aligned}$$

Then we can conclude since the previous inequality is true for every  $\lambda \in [0, 1]$ .  $\square$

**Lemma 23** (Properties of the interpolator  $i_m$ ). *Let  $m \in \{1, \dots, M\}$ .*

(a) For every  $x_{0:m} \in (\mathbb{R}^d)^{m+1}$ ,  $\|i_m(x_{0:m})\|_{\sup} = \sup_{0 \leq k \leq m} |x_k|$ .

(b) For every  $\mu_{0:m} \in (\mathcal{P}_p(\mathbb{R}^d))^{m+1}$ ,  $\sup_{t \in [0, T]} \mathcal{W}_p(i_m(\mu_{0:m})_t, \delta_0) = \sup_{0 \leq k \leq m} \mathcal{W}_p(\mu_k, \delta_0)$ .

*Proof of Lemma 23.* (a) First, it is obvious that  $\sup_{0 \leq k \leq m} |x_k| \leq \|i_m(x_{0:m})\|_{\sup}$  by the definition of  $i_m$ . For every  $k \in \{0, \dots, m-1\}$ , for every  $t \in [t_k, t_{k+1}]$ , we have

$$|i_m(x_{0:m})_t| \leq |x_k| \vee |x_{k+1}| \leq \sup_{0 \leq k \leq m} |x_k| \quad (4.7)$$

and for every  $t \in [t_m, T]$ , we have  $|i_m(x_{0:m})_t| = x_m \leq \sup_{0 \leq k \leq m} |x_k|$ . Then we can conclude  $\sup_{0 \leq k \leq m} |x_k| = \|i_m(x_{0:m})\|_{\sup}$ .

(b) First, it is obvious that  $\sup_{t \in [0, T]} \mathcal{W}_p(i_m(\mu_{0:m})_t, \delta_0) \geq \sup_{0 \leq k \leq m} \mathcal{W}_p(\mu_k, \delta_0)$  by the definition of  $i_m$ . For every  $k \in \{0, \dots, m-1\}$ , we have

$$\begin{aligned}
\sup_{t \in [t_k, t_{k+1}]} \mathcal{W}_p(i_m(\mu_{0:m})_t, \delta_0) &\leq \mathcal{W}_p(\mu_k, \delta_0) \vee \mathcal{W}_p(\mu_{k+1}, \delta_0) \quad (\text{by Lemma 22-(b)}) \\
&\leq \sup_{0 \leq k \leq m} \mathcal{W}_p(\mu_k, \delta_0)
\end{aligned}$$

and

$$\sup_{t \in [t_m, T]} \mathcal{W}_p(i_m(\mu_{0:m})_t, \delta_0) = \mathcal{W}_p(\mu_m, \delta_0) \leq \sup_{0 \leq k \leq m} \mathcal{W}_p(\mu_k, \delta_0).$$

Then we can conclude that  $\sup_{t \in [0, T]} \mathcal{W}_p(i_m(\mu_{0:m})_t, \delta_0) = \sup_{0 \leq k \leq m} \mathcal{W}_p(\mu_k, \delta_0)$ .  $\square$

## 4.2 Properties of the discretization scheme

We gather here several properties of the process  $(\widetilde{X}_t^M)_{t \geq 0}$  from Definition 19.

**Proposition 24.** *For all  $M \in \mathbb{N}^*$ , write  $(\widetilde{X}_t^M)_{t \in [0, T]}$  for the process from Definition 19 with parameter  $M$ . Then under Assumptions (I), we have*

(a) For every  $M \in \mathbb{N}^*$ ,  $\left\| \sup_{t \in [0, T]} |\widetilde{X}_t^M| \right\|_p \leq \Gamma(1 + \|X_0\|_p)$  with the same constant  $\Gamma$  from Theorem 1.

(b) There exists a constant  $\kappa$  depending on  $L, b, \sigma, \|X_0\|_p, p, d, T$  such that for  $M \in \mathbb{N}^*$  large enough, there holds

$$\left\| \sup_{0 \leq m \leq M-1} \sup_{v \in [t_m, t_{m+1}]} \left| \widetilde{X}_v^M - \widetilde{X}_{t_m}^M \right| \right\|_p \leq \kappa (h |\ln(h)|)^{\frac{1}{2}}.$$

Proposition 24 directly implies the following result.

**Corollary 25.** *Under Assumptions (I), we have, for a large enough time discretization number  $M \in \mathbb{N}^*$ ,*

$$\begin{aligned} & \left\| \left\| \tilde{X}^M - i_M(\tilde{X}_{t_0:t_M}^M) \right\|_{\sup} \right\|_p \leq 2\kappa (h |\ln(h)|)^{\frac{1}{2}} \\ \text{and } d_p\left(\left(\tilde{\mu}_t\right)_{t \in [0, T]}, i_M\left(\tilde{\mu}_{t_0:t_M}\right)\right) & \leq 3\kappa (h |\ln(h)|)^{\frac{1}{2}}. \end{aligned} \quad (4.8)$$

*Proof of Corollary 25.* Let  $M$  be fixed and large enough. We drop the superscript in  $\tilde{X}^M$  for simplicity. It is obvious that

$$\begin{aligned} & \left\| \tilde{X} - i_M(\tilde{X}_{t_0:t_M}) \right\|_{\sup} = \sup_{0 \leq m \leq M-1} \sup_{t \in [t_m, t_{m+1}]} \left| \tilde{X}_t - i_M(\tilde{X}_{t_0:t_M})_t \right| \\ & \leq \sup_{0 \leq m \leq M-1} \sup_{t \in [t_m, t_{m+1}]} \left[ \left| \tilde{X}_t - \tilde{X}_{t_m} \right| + \left| i_M(\tilde{X}_{t_0:t_M})_t - \tilde{X}_{t_m} \right| \right] \\ & \leq \sup_{0 \leq m \leq M-1} \sup_{t \in [t_m, t_{m+1}]} \left[ \left| \tilde{X}_t - \tilde{X}_{t_m} \right| + \left| \tilde{X}_{t_{m+1}} - \tilde{X}_{t_m} \right| \right] \\ & \leq 2 \sup_{0 \leq m \leq M-1} \sup_{t \in [t_m, t_{m+1}]} \left| \tilde{X}_t - \tilde{X}_{t_m} \right|. \end{aligned} \quad (4.9)$$

Then we conclude by applying Proposition 24-(b).

Consider now random variables  $(U_m)_{0 \leq m \leq M}$  i.i.d. having the uniform distribution on  $[0, 1]$  and independent of the process  $(\tilde{X}_t)_{t \in [0, T]}$ . For every  $m \in \{0, \dots, M-1\}$  and for every  $t \in [t_m, t_{m+1}]$ ,

$$1_{\{U_m > \frac{t-t_m}{h}\}} \tilde{X}_{t_m} + 1_{\{U_m \leq \frac{t-t_m}{h}\}} \tilde{X}_{t_{m+1}} \sim i_M(\tilde{X}_{t_0:t_M})_t.$$

Then

$$\begin{aligned} d_p\left(\left(\tilde{\mu}_t\right)_{t \in [0, T]}, i_M\left(\tilde{\mu}_{t_0:t_M}\right)\right) & = \sup_{t \in [0, T]} \mathcal{W}_p\left(\tilde{\mu}_t, i_M\left(\tilde{\mu}_{t_0:t_M}\right)_t\right) \\ & = \sup_{0 \leq m \leq M-1} \sup_{t \in [t_m, t_{m+1}]} \mathcal{W}_p\left(\tilde{\mu}_t, i_M\left(\tilde{\mu}_{t_0:t_M}\right)_t\right) \\ & \leq \sup_{0 \leq m \leq M-1} \sup_{t \in [t_m, t_{m+1}]} \left\| \tilde{X}_t - 1_{\{U_m > \frac{t-t_m}{h}\}} \tilde{X}_{t_m} - 1_{\{U_m \leq \frac{t-t_m}{h}\}} \tilde{X}_{t_{m+1}} \right\|_p \\ & \leq \sup_{0 \leq m \leq M-1} \sup_{t \in [t_m, t_{m+1}]} \left( \left\| (\tilde{X}_t - \tilde{X}_{t_m}) 1_{\{U_m > \frac{t-t_m}{h}\}} \right\|_p + \left\| (\tilde{X}_t - \tilde{X}_{t_{m+1}}) 1_{\{U_m \leq \frac{t-t_m}{h}\}} \right\|_p \right) \\ & \leq 3 \sup_{0 \leq m \leq M-1} \sup_{t \in [t_m, t_{m+1}]} \left\| \tilde{X}_t - \tilde{X}_{t_m} \right\|_p \leq 3\kappa (h |\ln(h)|)^{\frac{1}{2}}, \end{aligned} \quad (4.10)$$

where the last inequality comes from Proposition 24-(b).  $\square$

*Proof of Proposition 24.* (a) **Step 1.** In this first step, we prove that for every fixed  $M \in \mathbb{N}^*$

$$\left\| \sup_{0 \leq k \leq M} |\tilde{X}_{t_k}| \right\|_p < +\infty \quad (4.11)$$

by induction. First,  $\|\tilde{X}_{t_0}\|_p = \|X_0\|_p < +\infty$  by Assumption (I). Now assume that, for some  $l \geq 0$ ,  $\left\| \sup_{0 \leq k \leq l} |\tilde{X}_{t_k}| \right\|_p < +\infty$ . It follows that

$$\left\| \sup_{0 \leq k \leq l+1} |\tilde{X}_{t_k}| \right\|_p \leq \left\| \sup_{0 \leq k \leq l} |\tilde{X}_{t_k}| + \left( |\tilde{X}_{t_{l+1}}| - \sup_{0 \leq k \leq l} |\tilde{X}_{t_k}| \right)_+ \right\|_p$$



$$\begin{aligned}
&\leq \left\| \sup_{0 \leq k \leq l} |\widetilde{X}_{t_k}| \right\|_p + \left\| \left( |\widetilde{X}_{t_{l+1}}| - \sup_{0 \leq k \leq l} |\widetilde{X}_{t_k}| \right)_+ \right\|_p \quad (\text{by the Minkowski inequality}) \\
&\leq \left\| \sup_{0 \leq k \leq l} |\widetilde{X}_{t_k}| \right\|_p + \left\| \left( |\widetilde{X}_{t_{l+1}}| - |\widetilde{X}_{t_l}| \right)_+ \right\|_p \leq \left\| \sup_{0 \leq k \leq l} |\widetilde{X}_{t_k}| \right\|_p + \left\| |\widetilde{X}_{t_{l+1}}| - |\widetilde{X}_{t_l}| \right\|_p \\
&\leq \left\| \sup_{0 \leq k \leq l} |\widetilde{X}_{t_k}| \right\|_p + \left\| \widetilde{X}_{t_{l+1}} - \widetilde{X}_{t_l} \right\|_p. \tag{4.12}
\end{aligned}$$

Moreover,

$$\begin{aligned}
\left\| \widetilde{X}_{t_{l+1}} - \widetilde{X}_{t_l} \right\|_p &= \left\| h b_l(t_l, \widetilde{X}_{t_0:t_l}, \widetilde{\mu}_{t_0:t_l}) + \sqrt{h} \sigma_l(t_l, \widetilde{X}_{t_0:t_l}, \widetilde{\mu}_{t_0:t_l}) Z_{l+1} \right\|_p \\
&\leq h \left\| b \left( t_l, i_l(\widetilde{X}_{t_0:t_l}), i_l(\widetilde{\mu}_{t_0:t_l}) \right) \right\|_p + \sqrt{h} \left\| \left\| \sigma \left( t_l, i_l(\widetilde{X}_{t_0:t_l}), i_l(\widetilde{\mu}_{t_0:t_l}) \right) \right\| \right\|_p \left\| Z_{l+1} \right\|_p \\
&\quad (\text{by the definition of } b_l \text{ and } \sigma_l \text{ in Definition 2 and as } Z_{l+1} \perp \sigma(\mathcal{F}_{t_l})) \\
&\leq \left( h + \sqrt{h} C_{p,q} \right) \left\| C_{b,\sigma,L,T} \left( 1 + \|i_l(\widetilde{X}_{t_0:t_l})\|_{\text{sup}} + \sup_{t \in [0,T]} \mathcal{W}_p(i_l(\widetilde{\mu}_{t_0:t_l})_t, \delta_0) \right) \right\|_p,
\end{aligned}$$

where we used Lemma 12, where  $C_{p,q} = \|Z_{l+1}\|_p < +\infty$  as  $Z_{l+1} \sim \mathcal{N}(0, \mathbf{I}_q)$  is a constant depending only on  $p$  and  $q$ . We now invoke Lemma 23 to obtain

$$\begin{aligned}
\left\| \widetilde{X}_{t_{l+1}} - \widetilde{X}_{t_l} \right\|_p &\leq \left( h + \sqrt{h} C_{p,q} \right) \left\| C_{b,\sigma,L,T} \left( 1 + \sup_{0 \leq k \leq l} |\widetilde{X}_{t_k}| + \sup_{0 \leq k \leq l} \mathcal{W}_p(\widetilde{\mu}_{t_k}, \delta_0) \right) \right\|_p \tag{4.13} \\
&\leq C_{b,\sigma,L,T} \left( h + \sqrt{h} C_{p,q} \right) \left( 1 + \left\| \sup_{0 \leq k \leq l} |\widetilde{X}_{t_k}| \right\|_p + \sup_{0 \leq k \leq l} \left\| \widetilde{X}_{t_k} \right\|_p \right) \\
&\leq C_{b,\sigma,L,T} \left( h + \sqrt{h} C_{p,q} \right) \left( 1 + 2 \left\| \sup_{0 \leq k \leq l} |\widetilde{X}_{t_k}| \right\|_p \right) \\
&< +\infty \tag{4.14}
\end{aligned}$$

where we used the induction hypothesis to obtain the last inequality. Thus  $\left\| \sup_{0 \leq k \leq l+1} |\widetilde{X}_{t_k}| \right\|_p < +\infty$  which concludes the proof of (4.11) by induction.

**Step 2.** We prove that  $\left\| \sup_{t \in [0,T]} |\widetilde{X}_t| \right\|_p < +\infty$ . First, by (4.3) we get for every  $t \in [0, T]$ ,

$$\begin{aligned}
&\left\| \sup_{u \in [0,t]} |\widetilde{X}_u| \right\|_p \\
&= \left\| \sup_{u \in [0,t]} \left| X_0 + \int_0^u b \left( \underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}) \right) ds \right. \right. \\
&\quad \left. \left. + \int_0^u \sigma \left( \underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}) \right) dB_s \right| \right\|_p \tag{4.15} \\
&\leq \|X_0\|_p + \left\| \int_0^t \left| b \left( \underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}) \right) \right| ds \right\|_p \\
&\quad + \left\| \sup_{u \in [0,t]} \left| \int_0^u \sigma \left( \underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}) \right) dB_s \right| \right\|_p, \tag{4.16}
\end{aligned}$$

where we used Minkowski's inequality to obtain the second inequality. The second term in (4.15) can be upper bounded as follows: using Lemma 13,

$$\begin{aligned}
&\left\| \int_0^t \left| b \left( \underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}) \right) \right| ds \right\|_p \leq \int_0^t \left\| b \left( \underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}) \right) \right\|_p ds \\
&\leq \int_0^t \left\| C_{b,\sigma,L,T} \left( 1 + \|i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}})\|_{\text{sup}} + \sup_{u \in [0,T]} \mathcal{W}_p(i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}})_u, \delta_0) \right) \right\|_p ds
\end{aligned}$$

$$\begin{aligned}
&= \int_0^t \left\| C_{b,\sigma,L,T} \left( 1 + \sup_{0 \leq k \leq [\underline{s}]} |\widetilde{X}_{t_k}| + \sup_{0 \leq k \leq [\underline{s}]} \mathcal{W}_p(\widetilde{\mu}_{t_k}, \delta_0) \right) \right\|_p ds \\
&= \int_0^t C_{b,\sigma,L,T} \left( 1 + 2 \left\| \sup_{0 \leq k \leq [\underline{s}]} |\widetilde{X}_{t_k}| \right\|_p \right) ds \\
&\leq T C_{b,\sigma,L,T} + 2 C_{b,\sigma,L,T} \int_0^t \left\| \sup_{0 \leq k \leq [\underline{s}]} |\widetilde{X}_{t_k}| \right\|_p ds
\end{aligned} \tag{4.17}$$

which is finite by (4.11). We used Lemma 12 to obtain the second inequality, and Lemma 23 to deduce the next equality.

Moreover, combining Lemmas 15 and 12, the third term in (4.15) can be upper bounded as follows

$$\begin{aligned}
&\left\| \sup_{u \in [0,t]} \left| \int_0^u \sigma \left( \underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}) \right) dB_s \right| \right\|_p \\
&\leq C_{d,p}^{BDG} \left\{ \int_0^t \left\| C_{b,\sigma,L,T} \left( 1 + \|i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]})\|_{\text{sup}} + \sup_{u \in [0,T]} \mathcal{W}_p(i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]})_u, \delta_0) \right) \right\|_p^2 ds \right\}^{\frac{1}{2}} \\
&= C_{d,p}^{BDG} \left\{ \int_0^t \left\| C_{b,\sigma,L,T} \left( 1 + \sup_{0 \leq k \leq [\underline{s}]} |\widetilde{X}_{t_k}| + \sup_{0 \leq k \leq [\underline{s}]} \mathcal{W}_p(\widetilde{\mu}_{t_k}, \delta_0) \right) \right\|_p^2 ds \right\}^{\frac{1}{2}} \\
&\leq \sqrt{2T} C_{d,p}^{BDG} C_{b,\sigma,L,T} + 2 C_{d,p}^{BDG} C_{b,\sigma,L,T} \left\{ \int_0^t \left\| \sup_{0 \leq k \leq [\underline{s}]} |\widetilde{X}_{t_k}| \right\|_p^2 ds \right\}^{\frac{1}{2}}
\end{aligned} \tag{4.18}$$

which is again finite by (4.11). We used again Lemma 23 to get the third line. We conclude that  $\left\| \sup_{t \in [0,T]} |\widetilde{X}_t| \right\|_p < +\infty$ .

**Step 3.** We conclude the proof of (a). Using that

$$\left\| \sup_{0 \leq k \leq [\underline{s}]} |\widetilde{X}_{t_k}| \right\|_p^2 \leq \left\| \sup_{u \in [0,s]} |\widetilde{X}_u| \right\|_p^2$$

by the definition of  $[\underline{s}]$ , see (4.2), the inequalities (4.15), (4.17) and (4.18) in the previous step imply that for every  $t \in [0, T]$

$$\begin{aligned}
&\left\| \sup_{u \in [0,t]} |\widetilde{X}_u| \right\|_p \\
&\leq \|X_0\|_p + T C_{b,\sigma,L,T} + 2 C_{b,\sigma,L,T} \int_0^t \left\| \sup_{u \in [0,s]} |\widetilde{X}_u| \right\|_p ds \\
&\quad + \sqrt{2T} C_{d,p}^{BDG} C_{b,\sigma,L,T} + 2 C_{d,p}^{BDG} C_{b,\sigma,L,T} \left\{ \int_0^t \left\| \sup_{u \in [0,s]} |\widetilde{X}_u| \right\|_p^2 ds \right\}^{\frac{1}{2}}.
\end{aligned}$$

Hence, by applying Lemma 18 with  $f(t) := \left\| \sup_{u \in [0,t]} |\widetilde{X}_u| \right\|_p$ , we obtain

$$\left\| \sup_{u \in [0,t]} |\widetilde{X}_u| \right\|_p \leq C_{p,d,b,\sigma,L,T} e^{C_{p,d,b,\sigma,L,T} t} (1 + \|X_0\|_p)$$

with the constant  $C_{p,d,b,\sigma,L,T} > 0$  defined by (3.20). Then

$$\left\| \sup_{u \in [0,T]} |\widetilde{X}_u| \right\|_p \leq C_{p,d,b,\sigma,L,T} e^{C_{p,d,b,\sigma,L,T} T} (1 + \|X_0\|_p),$$

and we conclude by recognizing  $\Gamma = C_{p,d,b,\sigma,L,T} e^{C_{p,d,b,\sigma,L,T} T}$  from Theorem 1.

(b) By hypothesis,  $M$  is large enough so that  $h = \frac{T}{M} \leq \frac{1}{2}$ . We have

$$\begin{aligned}
& \left\| \sup_{0 \leq m \leq M-1} \sup_{v \in [t_m, t_{m+1}]} \left| \widetilde{X}_v - \widetilde{X}_{t_m} \right| \right\|_p \\
& \leq \left\| \sup_{0 \leq m \leq M-1} \sup_{v \in [t_m, t_{m+1}]} \left| (v - t_m) b_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) + \sigma_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m})(B_v - B_{t_m}) \right| \right\|_p \\
& \leq \left\| \sup_{0 \leq m \leq M-1} \sup_{v \in [t_m, t_{m+1}]} \left[ \left| (v - t_m) b_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) \right| + \left| \sigma_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m})(B_v - B_{t_m}) \right| \right] \right\|_p \\
& \leq \left\| \sup_{0 \leq m \leq M-1} \left[ h \left| b_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) \right| + \left\| \sigma_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) \right\| \sup_{v \in [t_m, t_{m+1}]} \left| B_v - B_{t_m} \right| \right] \right\|_p \\
& \leq h \left\| \sup_{0 \leq m \leq M-1} \left| b_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) \right| \right\|_p \\
& \quad + \left\| \sup_{0 \leq m \leq M-1} \left[ \left\| \sigma_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) \right\| \sup_{v \in [t_m, t_{m+1}]} \left| B_v - B_{t_m} \right| \right] \right\|_p
\end{aligned}$$

where we used that  $|t_{m+1} - t_m| = h$  and Minkowski's inequality. We now apply Lévy's modulus of continuity theorem, see e.g. [CR81, Theorem 1.1.1] to handle the Brownian component in the last inequality: there exists  $M_0 \geq 0$  such that for  $M \geq M_0$ ,

$$\begin{aligned}
\left\| \sup_{0 \leq m \leq M-1} \sup_{v \in [t_m, t_{m+1}]} \left| \widetilde{X}_v - \widetilde{X}_{t_m} \right| \right\|_p & \leq h \left\| \sup_{0 \leq m \leq M-1} \left| b_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) \right| \right\|_p \\
& \quad + 2(h |\ln(h)|)^{\frac{1}{2}} \left\| \sup_{0 \leq m \leq M-1} \left\| \sigma_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) \right\| \right\|_p.
\end{aligned}$$

We now treat the two terms on the right-hand-side of this inequality. First, by definition of  $b_m$ ,

$$\begin{aligned}
& \left\| \sup_{0 \leq m \leq M-1} \left| b_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) \right| \right\|_p = \left\| \sup_{0 \leq m \leq M-1} \left| b(t_m, i_m(\widetilde{X}_{t_0:t_m}), i_m(\widetilde{\mu}_{t_0:t_m})) \right| \right\|_p \\
& \leq \left\| \sup_{0 \leq m \leq M-1} C_{b,\sigma,L,T} \left( 1 + \|i_m(\widetilde{X}_{t_0:t_m})\|_{\sup} + \sup_{u \in [0, T]} \mathcal{W}_p(i_m(\widetilde{\mu}_{t_0:t_m})_u, \delta_0) \right) \right\|_p \\
& \quad \text{(by Lemma 12)} \\
& \leq \left\| \sup_{0 \leq m \leq M-1} C_{b,\sigma,L,T} \left( 1 + \sup_{0 \leq k \leq m} |\widetilde{X}_k| + \sup_{0 \leq k \leq m} \mathcal{W}_p(\widetilde{\mu}_{t_k}, \delta_0) \right) \right\|_p \quad \text{(by Lemma 23)} \\
& \leq \left\| C_{b,\sigma,L,T} \left( 1 + \sup_{0 \leq k \leq M} |\widetilde{X}_k| + \sup_{0 \leq k \leq M} \mathcal{W}_p(\widetilde{\mu}_{t_k}, \delta_0) \right) \right\|_p \\
& \leq C_{b,\sigma,L,T} \left( 1 + 2 \left\| \sup_{0 \leq k \leq M} |\widetilde{X}_k| \right\|_p \right) \leq C_{b,\sigma,L,T} \left( 1 + 2\Gamma(1 + \|X_0\|_p) \right) < +\infty. \quad (4.19)
\end{aligned}$$

Let  $C_\star := C_{b,\sigma,L,T} \left( 1 + 2\Gamma(1 + \|X_0\|_p) \right)$ , where we recall that  $\Gamma$  is given in Theorem 1. By a

similar computation, we obtain

$$\left\| \sup_{0 \leq m \leq M-1} \left\| \sigma_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m}) \right\| \right\|_p \leq C_\star.$$

Then, using that for  $h \in [0, \frac{1}{2}]$ ,  $h \leq (h |\ln(h)|)^{\frac{1}{2}}$ ,

$$\left\| \sup_{0 \leq m \leq M-1} \sup_{v \in [t_m, t_{m+1}]} \left| \widetilde{X}_v - \widetilde{X}_{t_m} \right| \right\|_p \leq 3C_\star (h |\ln(h)|)^{\frac{1}{2}} \quad (4.20)$$

and we can conclude by letting  $\kappa := 3C_\star$ .  $\square$

### 4.3 Proof of Proposition 20, Theorem 4 and Corollary 5

Before turning to the proof of Proposition 20, we briefly prove Theorem 4 and Corollary 5, which are two easy consequences of Proposition 20.

*Proof of Theorem 4.* The proof is straightforward since

$$\left\| \sup_{0 \leq m \leq M} \left| X_{t_m} - \widetilde{X}_{t_m}^M \right| \right\|_p \leq \left\| \sup_{t \in [0, T]} \left| X_t - \widetilde{X}_t^M \right| \right\|_p$$

by the definition of  $(\widetilde{X}_{t_m}^M)_{0 \leq m \leq M}$  and  $(\widetilde{X}_t^M)_{t \in [0, T]}$  in Definition 3 and Definition 19.  $\square$

*Proof of Corollary 5.* Corollary 25 implies that  $\| \|\widehat{X} - \widetilde{X}\|_{\text{sup}} \|_p \leq 2\kappa (h |\ln(h)|)^{\frac{1}{2}}$ . Then the result is a direct application of Theorem 4.  $\square$

*Proof of Proposition 20.* For every  $s \in [0, T]$ , we have

$$\begin{aligned} X_s - \widetilde{X}_s &= \int_0^s \left[ b(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - b(\underline{u}, i_{[\underline{u}]}(\widetilde{X}_{t_0:t_{[\underline{u}]}}), i_{[\underline{u}]}(\widetilde{\mu}_{t_0:t_{[\underline{u}]}})) \right] du \\ &\quad + \int_0^s \left[ \sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(\underline{u}, i_{[\underline{u}]}(\widetilde{X}_{t_0:t_{[\underline{u}]}}), i_{[\underline{u}]}(\widetilde{\mu}_{t_0:t_{[\underline{u}]}})) \right] dB_u, \end{aligned}$$

and we set

$$f(t) := \left\| \sup_{s \in [0, t]} \left| X_s - \widetilde{X}_s \right| \right\|_p.$$

It follows from Proposition 24-(a) that  $\widetilde{X} = (\widetilde{X}_t)_{t \in [0, T]} \in L^p_{\mathcal{C}([0, T], \mathbb{R}^d)}(\Omega, \mathcal{F}, \mathbb{P})$ . Consequently,  $\widetilde{\mu} \in \mathcal{P}_p(\mathcal{C}([0, T], \mathbb{R}^d))$  and  $\iota(\widetilde{\mu}) = (\widetilde{\mu}_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  by applying Lemma 11. Hence,

$$\begin{aligned} f(t) &= \left\| \sup_{s \in [0, t]} \left| X_s - \widetilde{X}_s \right| \right\|_p \\ &\leq \left\| \int_0^t \left| b(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - b(\underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}})) \right| ds \right. \\ &\quad \left. + \sup_{s \in [0, t]} \left| \int_0^s \left[ \sigma(u, X_{\cdot \wedge u}, \mu_{\cdot \wedge u}) - \sigma(\underline{u}, i_{[\underline{u}]}(\widetilde{X}_{t_0:t_{[\underline{u}]}}), i_{[\underline{u}]}(\widetilde{\mu}_{t_0:t_{[\underline{u}]}})) \right] dB_u \right| \right\|_p \\ &\leq \int_0^t \left\| b(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - b(\underline{s}, i_{[\underline{s}]}(\widetilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\widetilde{\mu}_{t_0:t_{[\underline{s}]}})) \right\|_p ds \end{aligned}$$

$$+ C_{d,p}^{BDG} \left[ \int_0^t \left\| \left\| \sigma(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - \sigma(\underline{u}, i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})) \right\| \right\|_p^2 ds \right]^{\frac{1}{2}} \quad (4.21)$$

using Lemma 15. The first term in (4.21) can be upper bounded by

$$\begin{aligned} & \int_0^t \left\| b(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - b(\underline{s}, i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})) \right\|_p ds \\ & \leq \int_0^t \|b(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - b(\underline{s}, X_{\cdot \wedge s}, \mu_{\cdot \wedge s})\|_p ds \\ & \quad + \int_0^t \left\| b(\underline{s}, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - b(\underline{s}, i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})) \right\|_p ds. \end{aligned} \quad (4.22)$$

For the first term in (4.22), we use Assumption (II) to obtain

$$\begin{aligned} & \int_0^t \|b(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - b(\underline{s}, X_{\cdot \wedge s}, \mu_{\cdot \wedge s})\|_p ds \\ & \leq \int_0^t L \left\| 1 + \|X_{\cdot \wedge s}\|_{\text{sup}} + \sup_{u \in [0, T]} W_p(\mu_{u \wedge s}, \delta_0) \right\|_p |s - \underline{s}|^\gamma ds \\ & \leq (LT + 2LT \sup_{t \in [0, T]} \|X_t\|_p) h^\gamma \\ & \leq h^\gamma 2LT \Gamma(1 + \|X_0\|_p), \end{aligned} \quad (4.23)$$

where we used (1.8) to obtain the last inequality. For the second term of (4.22), we have

$$\begin{aligned} & \int_0^t \left\| b(\underline{s}, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - b(\underline{s}, i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})) \right\|_p ds \\ & \leq \int_0^t \left\| L \left[ \|X_{\cdot \wedge s} - i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}})\|_{\text{sup}} + d_p((\mu_{v \wedge s})_{v \in [0, T]}, i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})) \right] \right\|_p ds \\ & \leq L \int_0^t \left\| \|X_{\cdot \wedge s} - i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}})\|_{\text{sup}} \right\|_p ds + L \int_0^t d_p((\mu_{v \wedge s})_{v \in [0, T]}, i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})) ds \\ & \leq L \int_0^t \left\| \|X_{\cdot \wedge s} - \tilde{X}_{\cdot \wedge s}\|_{\text{sup}} \right\|_p ds + L \int_0^t \left\| \|\tilde{X}_{\cdot \wedge s} - i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}})\|_{\text{sup}} \right\|_p ds \\ & \quad + L \int_0^t d_p((\mu_{v \wedge s})_{v \in [0, T]}, (\tilde{\mu}_{v \wedge s})_{v \in [0, T]}) ds + L \int_0^t d_p((\tilde{\mu}_{v \wedge s})_{v \in [0, T]}, i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})) ds \\ & \leq L \int_0^t \left\| \sup_{v \in [0, s]} \|X_v - \tilde{X}_v\|_p \right\|_p ds + 2LT \kappa(h |\ln(h)|)^{\frac{1}{2}} \\ & \quad + L \int_0^t \sup_{v \in [0, s]} \mathcal{W}_p(\mu_v, \tilde{\mu}_v) ds + 3LT \kappa(h |\ln(h)|)^{\frac{1}{2}} \\ & \leq L \int_0^t f(s) ds + LT 5\kappa(h |\ln(h)|)^{\frac{1}{2}} + L \int_0^t \sup_{v \in [0, s]} \|X_v - \tilde{X}_v\|_p ds \\ & \leq 2L \int_0^t f(s) ds + 5LT \kappa(h |\ln(h)|)^{\frac{1}{2}}, \end{aligned} \quad (4.24)$$

where we used Corollary 25 to obtain the fourth inequality. Now we consider the second term of (4.21). It follows by applying Lemma 15 and norm inequalities that

$$C_{d,p}^{BDG} \left[ \int_0^t \left\| \left\| \sigma(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - \sigma(\underline{s}, i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})) \right\| \right\|_p^2 ds \right]^{\frac{1}{2}}$$

$$\begin{aligned}
&\leq \sqrt{2}C_{d,p}^{BDG} \left[ \int_0^t \left\| \left\| \sigma(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - \sigma(\underline{s}, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) \right\| \right\|_p^2 ds \right]^{\frac{1}{2}} \\
&\quad + \sqrt{2}C_{d,p}^{BDG} \left[ \int_0^t \left\| \left\| \sigma(\underline{s}, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - \sigma\left(\underline{s}, i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})\right) \right\| \right\|_p^2 ds \right]^{\frac{1}{2}}. \quad (4.25)
\end{aligned}$$

For the first term in (4.25), we use the same argument as the one giving (4.23) to get

$$\left[ \int_0^t \left\| \left\| \sigma(s, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - \sigma(\underline{s}, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) \right\| \right\|_p^2 ds \right]^{\frac{1}{2}} \leq h^\gamma (\sqrt{2T} + 2\sqrt{T}\Gamma_2(1 + \|X_0\|_p)) \quad (4.26)$$

for some constant  $\Gamma_2 > 0$  depending explicitly on  $\kappa$  from (1.8) and the constants of Lemma 13 and Assumptions (I) and (II). The second term of (4.25) can be upper bounded as follows

$$\begin{aligned}
&\sqrt{2}C_{d,p}^{BDG} \left[ \int_0^t \left\| \left\| \sigma(\underline{s}, X_{\cdot \wedge s}, \mu_{\cdot \wedge s}) - \sigma\left(\underline{s}, i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}}), i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})\right) \right\| \right\|_p^2 ds \right]^{\frac{1}{2}} \\
&\leq \sqrt{2}C_{d,p}^{BDG} \left[ \int_0^t \left\| L \left[ \|X_{\cdot \wedge s} - i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}})\|_{\text{sup}} + d_p\left((\mu_{v \wedge s})_{v \in [0, T]}, i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})\right) \right] \right\|_p^2 ds \right]^{\frac{1}{2}} \\
&\leq 2LC_{d,p}^{BDG} \left[ \int_0^t \left\| \|X_{\cdot \wedge s} - i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}})\|_{\text{sup}} \right\|_p^2 ds \right]^{\frac{1}{2}} \\
&\quad + 2LC_{d,p}^{BDG} \left[ \int_0^t d_p\left((\mu_{v \wedge s})_{v \in [0, T]}, i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})\right)^2 ds \right]^{\frac{1}{2}} \\
&\leq 2\sqrt{2}LC_{d,p}^{BDG} \left[ \int_0^t \left\| \|X_{\cdot \wedge s} - \tilde{X}_{\cdot \wedge s}\|_{\text{sup}} \right\|_p^2 ds \right]^{\frac{1}{2}} \\
&\quad + 2\sqrt{2}LC_{d,p}^{BDG} \left[ \int_0^t \left\| \|\tilde{X}_{\cdot \wedge s} - i_{[\underline{s}]}(\tilde{X}_{t_0:t_{[\underline{s}]}})\|_{\text{sup}} \right\|_p^2 ds \right]^{\frac{1}{2}} \\
&\quad + 2\sqrt{2}LC_{d,p}^{BDG} \left[ \int_0^t d_p\left((\mu_{v \wedge s})_{v \in [0, T]}, (\tilde{\mu}_{v \wedge s})_{v \in [0, T]}\right)^2 ds \right]^{\frac{1}{2}} \\
&\quad + 2\sqrt{2}LC_{d,p}^{BDG} \left[ \int_0^t d_p\left((\tilde{\mu}_{v \wedge s})_{v \in [0, T]}, i_{[\underline{s}]}(\tilde{\mu}_{t_0:t_{[\underline{s}]}})\right)^2 ds \right]^{\frac{1}{2}} \\
&\leq 4\sqrt{2}LC_{d,p}^{BDG} \left[ \int_0^t f(s)^2 ds \right]^{\frac{1}{2}} + 2\sqrt{2}LC_{d,p}^{BDG} \sqrt{T} 5\kappa(h|\ln(h)|)^{\frac{1}{2}} \quad (4.27)
\end{aligned}$$

by a similar reasoning as the one leading to (4.24). Bringing those inequalities together, we find

$$\begin{aligned}
f(t) &\leq L 2h^\gamma T \Gamma_2 (1 + \|X_0\|_p) + 2L \int_0^t f(s) ds + 5LT\kappa(h|\ln(h)|)^{\frac{1}{2}} \\
&\quad + h^\gamma (\sqrt{2T} + 2\sqrt{T}\Gamma_2(1 + \|X_0\|_p)) + 4\sqrt{2}LC_{d,p}^{BDG} \left[ \int_0^t f(s)^2 ds \right]^{\frac{1}{2}} \\
&\quad + 10\sqrt{2}LC_{d,p}^{BDG} \sqrt{T} \kappa(h|\ln(h)|)^{\frac{1}{2}}. \quad (4.28)
\end{aligned}$$

The conclusion follows by applying Lemma 18.  $\square$

## Acknowledgment

A.B. acknowledges partial funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement n° 864066).

## A Proofs for Subsection 2.2

We provide in this appendix the proofs of the results from Subsection 2.2.

*Proof of Proposition 8.* It suffices to show that Assumption (I) is satisfied. Since the diffusion matrix is the identity, we only need to focus on the drift coefficient. We recall from [Tom21, Proof of Proposition 3.9] that there exists  $C_\epsilon > 0$  such that for all  $t \in [0, T]$  and  $x, y \in \mathbb{R}^2$ ,

$$|b_0^\epsilon(t, x) - b_0^\epsilon(t, y)| + |K_t^\epsilon(x) - K_t^\epsilon(y)| \leq C_\epsilon |x - y|, \quad (\text{A.1})$$

where for  $t = 0$ , we consider the natural extension of  $b_0^\epsilon$  and  $K_t^\epsilon$ , namely,

$$b_0^\epsilon(0, x) = 0 \quad \text{and} \quad K_0^\epsilon(x) = 0. \quad (\text{A.2})$$

Using this, one sees easily that  $b_0^\epsilon$  satisfy the Lipschitz condition in the sense of Assumption (I), the continuity in time being straightforward from its definition. The continuity in time of the second drift term

$$A(t, x, (\mu_s)_{s \in [0, T]}) := \chi \int_0^t e^{-\lambda(t-s)} \left[ \int_{\mathbb{R}^d} K_{t-s}^\epsilon(x - y) \mu_s(dy) \right] ds,$$

is deduced from the form of  $K_t^\epsilon(x)$ . We focus now on the Lipschitz condition for this term. We have, for all  $t \in [0, T]$ ,  $x_1, x_2 \in \mathbb{R}^d$  and  $(\mu_t)_{t \in [0, T]}, (\nu_t)_{t \in [0, T]}$  in  $C([0, T], \mathcal{P}_p(\mathbb{R}^d))$ , by triangle inequality,

$$\begin{aligned} & \left| A(t, x_1, (\mu_s)_{s \in [0, T]}) - A(t, x_2, (\nu_s)_{s \in [0, T]}) \right| \\ & \leq \chi \left| \int_0^t e^{-\lambda(t-s)} \left[ \int_{\mathbb{R}^d} \left( K_{t-s}^\epsilon(x_1 - y) - K_{t-s}^\epsilon(x_2 - y) \right) \mu_s(dy) \right] ds \right| \\ & \quad + \chi \left| \int_0^t e^{-\lambda(t-s)} \int_{\mathbb{R}^d} K_{t-s}^\epsilon(x_2 - y) (\mu_s - \nu_s)(dy) ds \right| \\ & := B_t^1 + B_t^2, \end{aligned}$$

the last equality standing for definitions of  $B_t^1$  and  $B_t^2$ . Using Jensen's inequality and (A.1), we obtain

$$B_t^1 \leq \frac{\chi}{\lambda} C_\epsilon |x_1 - x_2| (1 - e^{-\lambda t}) \leq \frac{\chi}{\lambda} C_\epsilon |x_1 - x_2|.$$

For the second term, we recall the dual representation of the Wasserstein distance  $\mathcal{W}_1$  (see e.g. [Edw11], [Vil09, Remark 6.5]), namely, for every  $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$ ,

$$\mathcal{W}_1(\mu, \nu) = \sup \left\{ \int_{\mathbb{R}^d} \varphi d\mu - \int_{\mathbb{R}^d} \varphi d\nu \mid \varphi : \mathbb{R}^d \rightarrow \mathbb{R} \text{ Lipschitz continuous} \right. \\ \left. \text{with Lipschitz constant } [\varphi]_{\text{Lip}} \leq 1 \right\}$$

and the fact that for every  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ ,  $p \geq 1$ ,  $\mathcal{W}_1(\mu, \nu) \leq \mathcal{W}_p(\mu, \nu)$  (see e.g. [Vil09, Remark 6.6]). This implies

$$B_t^2 \leq \chi \int_0^t e^{-\lambda(t-s)} C_\epsilon \mathcal{W}_p(\mu_s, \nu_s) ds \leq \frac{\chi}{\lambda} C_\epsilon d_p((\mu_s)_{s \in [0, T]}, (\nu_s)_{s \in [0, T]}),$$

which concludes the proof that the drift coefficient satisfies Assumption (I), and the proposition follows by applying Theorem 1.  $\square$

*Proof of Proposition 9.* We treat the two terms of the drift separately. Again from [Tom21, Proof of Proposition 3.9], there exists  $C_\epsilon > 0$  such that for all  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$ ,

$$|K_t^\epsilon(x)| \leq C_\epsilon, \quad (\text{A.3})$$

where for  $t = 0$ , we consider the same extension of  $K_t^\epsilon$  as in (A.2).

**Step 1: Term involving  $b_0^\epsilon$ .** Note that, for  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$ ,

$$b_0^\epsilon(t, x) = \chi e^{-\lambda t} \int_{\mathbb{R}^d} \nabla c_0(y) \frac{1}{t+\epsilon} \frac{1}{2\pi} e^{-\frac{|x-y|^2}{2t}} dy.$$

Hence

$$\begin{aligned} \partial_t b_0^\epsilon(t, x) &= -\lambda b_0^\epsilon(t, x) - \chi e^{-\lambda t} \int_{\mathbb{R}^d} \nabla c_0(y) \frac{1}{(t+\epsilon)^2} \frac{1}{2\pi} e^{-\frac{|x-y|^2}{2t}} dy \\ &\quad - \chi e^{-\lambda t} \int_{\mathbb{R}^d} \nabla c_0(y) \frac{1}{t+\epsilon} \frac{|x-y|^2}{4\pi t^2} e^{-\frac{|x-y|^2}{2t}} dy. \end{aligned}$$

From (A.3), the only singularity at  $t = 0$  is on the last term on the right-hand side, and easily handled by considering the change of variable  $y \rightarrow (x - y)/\sqrt{t}$  from  $\mathbb{R}^2$  to  $\mathbb{R}^2$ . Using that  $c_0 \in H^1(\mathbb{R}^d) \cap W^{1,\infty}(\mathbb{R}^d)$ , we deduce that for all  $t \geq 0$ ,  $x \in \mathbb{R}^d$ ,

$$|\partial_t b_0^\epsilon(t, x)| \leq C$$

for some constant  $C > 0$  independent of  $t$ .

**Step 2: Term in A.** With the same notations as before, we consider, for  $0 \leq u \leq t$ ,  $x \in \mathbb{R}^d$ ,  $(\mu_t)_{t \in [0, T]} \in \mathcal{C}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ ,

$$\begin{aligned} &A(t, x, (\mu_s)_{0 \leq s \leq T}) - A(u, x, (\mu_s)_{0 \leq s \leq T}) \\ &= \chi \left\{ \int_u^t e^{-\lambda(t-s)} \int_{\mathbb{R}^d} K_{t-s}^\epsilon(x-y) \mu_s(dy) ds \right. \\ &\quad \left. + \int_0^u \left[ e^{-\lambda(t-s)} \int_{\mathbb{R}^d} K_{t-s}^\epsilon(x-y) \mu_s(dy) - e^{-\lambda(u-s)} \int_{\mathbb{R}^d} K_{u-s}^\epsilon(x-y) \mu_s(dy) \right] ds. \right\} \end{aligned} \quad (\text{A.4})$$

Using (A.3) leads to

$$\left| \int_u^t e^{-\lambda(t-s)} \int_{\mathbb{R}^d} K_{t-s}^\epsilon(x-y) \mu_s(dy) ds \right| \leq C_2 |t - u|$$

for some constant  $C_2 > 0$ . For the second term on the right-hand-side of (A.4), we notice that the function defined for  $t \geq 0$ ,  $x \in \mathbb{R}^d$

$$g(t, x) = e^{-\lambda t} \frac{x}{(t+\epsilon)^2} e^{-\frac{|x|^2}{2t}}$$

is such that

$$\partial_t g(t, x) = \left( -\lambda - \frac{2}{t+\epsilon} \right) g(t, x) - e^{-\lambda t} \frac{x|x|^2}{2t^2(t+\epsilon)^2} e^{-\frac{|x|^2}{2t}}$$

so that

$$|\partial_t g(t, x)| \leq C_2$$

hence  $g$  is Lipschitz in time. We conclude that Assumption (II) holds with  $\gamma = 1$ , and applying Theorem 4, the result follows.  $\square$



## References

- [AKH02] F. Antonelli and A. Kohatsu-Higa. Rate of convergence of a particle method to the solution of the McKean-Vlasov equation. *The Annals of Applied Probability*, 12(2):423–476, 2002.
- [Bol08] F. Bolley. Separability and completeness for the Wasserstein distance. In *Séminaire de probabilités XLI*, pages 371–377. Springer, 2008.
- [Bou88] N. Bouleau. *Processus stochastiques et applications*. Hermann Paris, 1988.
- [BT97] M. Bossy and D. Talay. A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.*, 66(217):157–192, 1997.
- [CCP11] M. J. Cáceres, J. A. Carrillo, and B. Perthame. Analysis of Nonlinear Noisy Integrate& Fire Neuron Models: blow-up and steady states. *The Journal of Mathematical Neuroscience*, 1(1):7, 2011.
- [CGK<sup>+</sup>20] A. Cosso, F. Gozzi, I. Kharroubi, H. Pham, and M. Rosestolato. Optimal control of path-dependent McKean-Vlasov SDEs in infinite dimension. *arXiv: 2012.14772*, December 2020.
- [CHM06] P. E. Caines, M. Huang, and R. P. Malhamé. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information and Systems*, 6(3):221–252, 2006.
- [CPSS15] J. A. Carrillo, B. Perthame, D. Salort, and D. Smets. Qualitative properties of solutions for the noisy integrate and fire model in computational neuroscience. *Nonlinearity*, 28(9):3365–3388, August 2015.
- [CR81] M. Csörgö and P. Révész. *Strong Approximations in Probability and Statistics*. Elsevier Science & Techn., July 1981.
- [DIRT15] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Global solvability of a networked integrate-and-fire model of McKean-Vlasov type. *The Annals of Applied Probability*, 25(4), August 2015.
- [Dob70] R. L. Dobrushin. Prescribing a System of Random Variables by Conditional Distributions. *Theory of Probability & Its Applications*, 15(3):458–486, jan 1970.
- [Edw11] D.A. Edwards. On the Kantorovich-Rubinstein theorem. *Expo. Math.*, 29(4):387–398, 2011.
- [FT22] N. Fournier and M. Tomašević. On a particle system associated to the parabolic-parabolic Keller-Segel equation. *In preparation*, 2022.
- [FTC09] O. Faugeras, J. Touboul, and B. Cessac. A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in Computational Neuroscience*, 3, 2009.
- [HK02] R. Hegselmann and U. Krause. Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.
- [HL22] M. Hoffmann and Y. Liu. A statistical approach for simulating the density solution of a McKean-Vlasov equation. *In preparation*, 2022.

- [Hor03] D. Horstmann. From 1970 until present: the Keller–Segel model in chemotaxis and its consequences. I. *Jahresber, Deutsch. Math.-Verein*, 105:103–165, 2003.
- [Hor04] D. Horstmann. From 1970 until present: The Keller–Segel model in chemotaxis and its consequences. II. *Jahresber, Deutsch. Math.-Verein*, 106:51–69, 2004.
- [JR95] B. H. Jansen and V. G. Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73(4):357–366, September 1995.
- [KF09] I. Karatzas and R. Fernholz. Stochastic Portfolio Theory: an Overview. In *Special Volume: Mathematical Modeling and Numerical Methods in Finance*, pages 89–167. Elsevier, 2009.
- [Lac18] D. Lacker. Mean field games and interacting particle systems. *Preprint*, 2018. Available at <http://www.columbia.edu/~dl3133/MFGSpring2018.pdf>.
- [Liu19] Y. Liu. *Optimal Quantization: Limit Theorems, Clustering and Simulation of the McKean-Vlasov Equation*. PhD thesis, Sorbonne Université, 2019.
- [LP20] Y. Liu and G. Pagès. Functional convex order for the scaled McKean–Vlasov processes. *arXiv:2005.03154*, 2020.
- [LP22] Y. Liu and G. Pagès. Monotone convex order for the McKean–Vlasov processes. *Stochastic Processes and their Applications*, 152:312–338, 2022.
- [McK67] H. P. McKean. Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57, 1967.
- [Pag18] G. Pagès. *Numerical Probability: An Introduction with Applications to Finance*. Springer, 2018.
- [Szn91] A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [Tom21] M. Tomašević. A new McKean–Vlasov stochastic interpretation of the parabolic–parabolic Keller–Segel model: The two-dimensional case. *The Annals of Applied Probability*, 31(1), February 2021.
- [TT20] D. Talay and M. Tomašević. A new McKean–Vlasov stochastic interpretation of the parabolic–parabolic Keller–Segel model: The one-dimensional case. *Bernoulli*, 26(2), May 2020.
- [Vil09] C. Villani. *Optimal transport, Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.