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

Armand Bernou, Yating Liu

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

Bernou Armand[1] **and Liu Yating**[2]

[1] *Dipartimento di Matematica, Università di Roma La Sapienza, P.le Aldo Moro 2, 00185 Roma, Italy , e-mail:* armand.bernou@uniroma1.it

[2] *CEREMADE, CNRS, UMR 7534, Université Paris-Dauphine, PSL University, 75016 Paris, France, e-mail:* liu@ceremade.dauphine.fr

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 give a proof of strong well-posedness of the equation in the L^p setting, $p \geq 2$, locally in time, as well as the propagation of chaos properties. Then, we introduce an interpolated Euler scheme, a key object to simulate numerically the process, and we prove the convergence, with an explicit rate, of this scheme towards the strong solution in the L^p norm. As applications we give results for two mean-field limit equations arising in biology and neuroscience.

MSC2020 subject classifications: Primary 60G65, 60H35; secondary 60G20.
Keywords and phrases: path-dependent McKean-Vlasov equation, propagation of chaos, 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

$$\begin{cases} \mathrm{d}X_t = b(t, X_{\cdot\wedge t}, \mu_{\cdot\wedge t})\mathrm{d}t + \sigma(t, X_{\cdot\wedge t}, \mu_{\cdot\wedge t})\mathrm{d}B_t, \\ X_0 : (\Omega, \mathcal{F}, \mathbb{P}) \to \left(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\right) \text{ random variable} \end{cases} \tag{1.1}$$

where the arguments $X_{\cdot\wedge t}$ and $\mu_{\cdot\wedge t}$ of the coefficients b and σ keep track of the whole trajectory of $X_\cdot$ and its marginal distribution $\mu_\cdot$ between 0 and $t > 0$ (see below (1.3) and (1.4) 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

$$\mathrm{d}X_t = b(t, X_t, \mu_t)\mathrm{d}t + \sigma(t, X_t, \mu_t)\mathrm{d}B_t \tag{1.2}$$

first introduced by McKean in [34] as a stochastic model naturally associated to a class of non-linear PDEs. See also [37] for a systematic presentation of the McKean-Vlasov equation, including the notion of propagation of chaos.

We prove the well-posedness of the path-dependent McKean-Vlasov equation (1.1), the propagation of chaos properties for an associated particle system and the convergence of an interpolated Euler scheme defined further in (1.15), with a quantitative estimate for the convergence rate.

A more detailed presentation of the context surrounding this model and its applications, as well as a comparison of our results with the existing literature regarding Equation (1.1) is drawn in Subsection 1.3.

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. We write $\mathcal{C}([0,T], E)$ for the set of continuous maps from $[0,T]$ to some topological space E, and, for $p \geq 1$, $\mathcal{P}_p(\mathbb{R}^d)$ for the set of probability distributions on $\mathbb{R}^d$ admitting a finite moment of order p. Write $\mathcal{B}(\mathbb{R}^d)$ for the Borel σ-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

$\alpha = (\alpha_t)_{t\in[0,T]} \in \mathcal{C}\big([0,T],\mathbb{R}^d\big)$ and let $(\nu_t)_{t\in[0,T]} \in \mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$. 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

$$\alpha_{t\wedge t_0} \coloneqq \begin{cases} \alpha_t & \text{if} \quad t \in [0,t_0], \\ \alpha_{t_0} & \text{if} \quad t \in (t_0,T]. \end{cases} \tag{1.3}$$

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

$$\nu_{t\wedge t_0} \coloneqq \begin{cases} \nu_t & \text{if} \quad t \in [0,t_0], \\ \nu_{t_0} & \text{if} \quad t \in (t_0,T], \end{cases} \tag{1.4}$$

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

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

$$X_t = X_0 + \int_0^t b(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s})\,\mathrm{d}s + \int_0^t \sigma(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s})\,\mathrm{d}B_s \tag{1.5}$$

where

- $X_0 : (\Omega,\mathcal{F},\mathbb{P}) \to \big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$ is a random vector in $L^p(\mathbb{P})$,
- $b : [0,T]\times\mathcal{C}\big([0,T],\mathbb{R}^d\big)\times\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big) \to \mathbb{R}^d$,
- $\sigma : [0,T]\times\mathcal{C}\big([0,T],\mathbb{R}^d\big)\times\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big) \to \mathbb{M}_{d,q}(\mathbb{R})$,
- $(B_t)_{t\in[0,T]}$ 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 $s\in[0,T]$,

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

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 and our result regarding the propagation of chaos. The second one is needed to obtain the convergence of our interpolated Euler scheme. Remark that the time horizon $T > 0$ is fixed. For a Polish space (S, d_S), we recall the definition of the Wasserstein distance W_p on $\mathcal{P}_p(S)$

$$\begin{aligned} W_p(\mu,\nu) &:= \inf_{\pi\in\Pi(\mu,\nu)} \Big(\int_{S\times S} d_S(x,y)^p\,\pi(\mathrm{d}x,\mathrm{d}y)\Big)^{\frac{1}{p}} \\ &= \inf\Big\{\mathbb{E}\left[d_S(X,Y)^p\right]^{\frac{1}{p}},\ X,Y : (\Omega,\mathcal{F},\mathbb{P}) \to (S,\mathcal{S}) \text{with } \mathbb{P}\circ X^{-1} = \mu,\ \mathbb{P}\circ Y^{-1} = \nu\Big\} \end{aligned} \tag{1.6}$$

where $\Pi(\mu,\nu)$ denotes the set of probability measures on $(S\times S, \mathcal{S}^{\otimes 2})$ with marginals μ and ν, and $\mathcal{S}$ denotes the Borel σ-algebra on S generated by the distance d_S. We write $\mathcal{W}_p$ for the case $S=\mathbb{R}^d$ and $\mathbb{W}_p$ for the case $S = \mathcal{C}([0,T],\mathbb{R}^d)$ endowed with the sup norm $\|\cdot\|_{\sup}$.

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

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

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

where d_p is defined by

$$d_p\big((\mu_t)_{t\in[0,T]},(\nu_t)_{t\in[0,T]}\big) \coloneqq \sup_{t\in[0,T]} \mathcal{W}_p(\mu_t,\nu_t). \tag{1.7}$$

Theorem 1.1. *Under Assumption (I), there exists a unique strong solution $X = (X_t)_{t\in[0,T]}$ from $(\Omega, \mathcal{F}, \mathbb{P})$ to $\big(\mathcal{C}\big([0,T],\mathbb{R}^d\big), \|\cdot\|_{\sup}\big)$ of the path-dependent McKean-Vlasov equation (1.5). Moreover, there exists a constant $\Gamma > 0$ depending on b, σ, L, T, d, p such that*

$$\Big\| \sup_{t\in[0,T]} |X_t| \Big\|_p \le \Gamma\Big(1 + \|X_0\|_p\Big). \tag{1.8}$$

We then turn to our result regarding the propagation of chaos. Let $T > 0$ be the same time horizon as before. For $N \ge 1$, let $X_0^{1,N}, \dots, X_0^{N,N}$ be i.i.d. copies of X_0 in (1.5) and $B^i := (B_t^i)_{t\in[0,T]}$, $1 \le i \le N$ be N independent standard $\mathbb{R}^q$-valued Brownian motions, independent of the Brownian motion $(B_t)_{t\in[0,T]}$ in (1.5) and of $(X_0, X_0^{1,N}, \dots, X_0^{N,N})$. We introduce the N-particle system $(X_t^{1,N}, \dots, X_t^{N,N})_{t\in[0,T]}$ defined by

$$\begin{cases} X_t^{i,N} &= X_0^{i,N} + \int_0^t b\big(s, X_{\cdot\wedge s}^{i,N}, \mu_{\cdot\wedge s}^N\big)\mathrm{d}s + \int_0^t \sigma\big(s, X_{\cdot\wedge s}^{i,N}, \mu_{\cdot\wedge s}^N\big)\mathrm{d}B_s^i, \quad 1 \le i \le N,\ t \in [0,T], \\ \mu_t^N &:= \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}, \qquad t \in [0,T]. \end{cases} \tag{1.9}$$

We prove in Proposition 4.2 that this particle system is well-posed. Our result regarding propagation of chaos is the following.

Theorem 1.2. *Assume that Assumption (I) holds with $p \ge 2$. Let X be the unique solution of (1.5) given by Theorem 1.1 and write $\mu := \mathbb{P}\circ X^{-1}$ and $(\mu_t)_{t\in[0,T]}$ for its marginal distributions. Let $(X_t^{1,N}, \dots, X_t^{N,N})_{t\in[0,T]}$ be the processes defined by the N-particle system (1.9) and $(Y^1, \dots, Y^N)$ be N i.i.d. copies of X. Then*

1. *there holds, for some constant $C_{d,p,L,T} > 0$, for all $N \ge 1$,*

$$\Big\| \sup_{t\in[0,T]} \mathcal{W}_p\Big(\mu_t, \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}\Big)\Big\|_p \le C_{d,p,L,T}\Big\|\mathbb{W}_p(\mu,\nu^N)\Big\|_p, \tag{1.10}$$

   where $\nu^N := \frac{1}{N}\sum_{i=1}^N \delta_{Y^i}$ is the empirical measures of $(Y^1, \dots, Y^N)$. Moreover, the norm $\|\mathbb{W}_p(\mu,\nu^N)\|_p$ converges to 0 as $N \to \infty$.
2. *For a fixed $k \in \mathbb{N}^*$, we have the weak convergence:*

$$\Big(X^{1,N}, \dots, X^{k,N}\Big) \Rightarrow \Big(Y^1, \dots, Y^k\Big) \qquad \text{as} \quad N \to \infty. \tag{1.11}$$

We then focus on the interpolated Euler scheme, and give a quantitative convergence result. In the following definition, $M \in \mathbb{N}^*$ should be thought of as the temporal discretization number, while $h := \frac{T}{M}$ is the time step. For every $m = 0, \dots, M$, we set $t_m = mh$. We introduce 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 notations, we will write $x_{0:m} := (x_0, ..., x_m)$, $\mu_{0:m} := (\mu_0, ..., \mu_m)$. Our discretization scheme uses the following interpolator.

Definition 1.3 (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*

$$x_{0:m} \in (\mathbb{R}^d)^{m+1} \longmapsto i_m(x_{0:m}) = (\bar{x}_t)_{t\in[0,T]} \in \mathcal{C}\big([0,T],\mathbb{R}^d\big), \tag{1.12}$$

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

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

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

(*b*) *Let $p \ge 1$. For every $m = 1, ..., M$, 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*

$$\mu_{0:m} \in (\mathcal{P}_p(\mathbb{R}^d))^{m+1} \longmapsto i_m(\mu_{0:m}) = (\bar{\mu}_t)_{t\in[0,T]} \in \mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big), \tag{1.13}$$

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

$$\forall\, k = 0, ..., m-1, \quad \forall\, t \in [t_k, t_{k+1}), \quad \bar{\mu}_t = \frac{1}{h}(t_{k+1} - t)\,\mu_k + \frac{1}{h}(t - t_k)\,\mu_{k+1},$$
$$\forall\, t \in [t_m, T], \quad \bar{\mu}_t = \mu_m. \tag{1.14}$$

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 1.4. *Let $M \in \mathbb{N}^*$, $h = \frac{T}{M}$. For every $m = 0, ..., M$, we set $t_m = mh$. For the same Brownian motion $(B_t)_{t\in[0,T]}$ and random vector X_0 as in (1.5), the interpolated scheme $(\widetilde{X}^M_{t_m})_{0\le m\le M}$ of the path-dependent McKean-Vlasov equation (1.5) is defined as follows :*

1. $\widetilde{X}^M_0 = X_0$;
2. *for all $m \in \{0, \dots, M-1\}$,*

$$\widetilde{X}^M_{t_{m+1}} = \widetilde{X}^M_{t_m} + h\, b_m(t_m, \widetilde{X}^M_{t_0:t_m}, \widetilde{\mu}^M_{t_0:t_m}) + \sqrt{h}\, \sigma_m(t_m, \widetilde{X}^M_{t_0:t_m}, \widetilde{\mu}^M_{t_0:t_m}) Z_{m+1}, \tag{1.15}$$

where, for $k \in \{0, \dots, M\}$, $\widetilde{\mu}^M_{t_k}$ is the probability distribution of $\widetilde{X}^M_{t_k}$, 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

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

are defined as follows : $\forall\, t \in [0,T]$, $x_{0:m} \in (\mathbb{R}^d)^{m+1}$, $\mu_{0:m} \in \big(\mathcal{P}_p(\mathbb{R}^d)\big)^{m+1}$

$$\begin{cases} b_m(t, x_{0:m}, \mu_{0:m}) := b\big(t, i_m(x_{0:m}), i_m(\mu_{0:m})\big), \\ \sigma_m(t, x_{0:m}, \mu_{0:m}) := \sigma\big(t, i_m(x_{0:m}), i_m(\mu_{0:m})\big). \end{cases} \tag{1.16}$$

The applications b_m and σ_m can often discretize computations from a numerical point of view. For instance, if

$$b\big(t, (X_s)_{s\in[0,T]}, (\mu_s)_{s\in[0,T]}\big) := \int_0^t \mathbb{E}\,[\phi(X_s)]\mathrm{d}s \tag{1.17}$$

with a bounded function ϕ, then

$$b_m(t_m, \widetilde{X}^M_{t_0:t_m}, \widetilde{\mu}^M_{t_0:t_m}) = \frac{h}{2}\Big(\mathbb{E}\,[\phi(\widetilde{X}^M_{t_0})] + \mathbb{E}\,[\phi(\widetilde{X}^M_{t_m})]\Big) + h\sum_{k=1}^{m-1} \mathbb{E}\,[\phi(\widetilde{X}^M_{t_k})]. \tag{1.18}$$

Computing an integral quantity such as (1.17) is numerically more demanding than dealing with sums as in (1.18). This is especially important when one wants to use a particle system to estimate such quantities.

In order to prove the convergence of the interpolated Euler scheme (1.15) to the unique strong solution of (1.5), 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, σ are γ-Hölder in t for some $0 < \gamma \le 1$, uniformly in α and in $(\mu_t)_{t\in[0,T]}$, in the following sense : there exists $L > 0$ s.t.

$$\forall\, t, s \in [0,T],\ \forall\, \alpha \in \mathcal{C}\big([0,T], \mathbb{R}^d\big) \text{ and } \forall\, (\mu_t)_{t\in[0,T]} \in \mathcal{C}\big([0,T], \mathcal{P}_p(\mathbb{R}^d)\big),$$
$$\max\Big(\Big|b\big(t, \alpha, (\mu_t)_{t\in[0,T]}\big) - b\big(s, \alpha, (\mu_t)_{t\in[0,T]}\big)\Big|, \Big|\Big|\Big|\sigma\big(t, \alpha, (\mu_t)_{t\in[0,T]}\big) - \sigma\big(s, \alpha, (\mu_t)_{t\in[0,T]}\big)\Big|\Big|\Big|\Big)$$
$$\le L\Big(1 + \|\alpha\|_{\sup} + \sup_{t\in[0,T]} \mathcal{W}_p(\mu_t, \delta_0)\Big)|t - s|^\gamma, \tag{1.19}$$

where δ_0 is the Dirac measure at 0.

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

Theorem 1.5 (Convergence rate of the interpolated Euler scheme)**.** *Under Assumptions (I) and (II), for $(X_t)_{t\in[0,T]}$ the unique strong solution to (1.5) given by Theorem 1.1, for $(\widetilde{X}^M_{t_m})_{0\le m\le M}$ the interpolated Euler scheme from Definition 1.4 with parameter M large enough, for $h=\frac{T}{M}$, one has,*

$$\left\| \sup_{0\le m\le M} \left| X_{t_m} - \widetilde{X}^M_{t_m} \right| \right\|_p \le \tilde{C}\Big(h^\gamma + \big(h|\ln(h)|\big)^{\frac{1}{2}} \Big), \tag{1.20}$$

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

From Definition 1.4, we can introduce a continuous extension of $(\widetilde{X}^M_{t_m})_{0\le m\le M}$, denoted by $\widehat{X}^M = (\widehat{X}^M_t)_{t\in[0,T]}$ and defined by $\widehat{X}^M := i_M(\widetilde{X}^M_{t_0:t_M})$. Then we have the following convergence.

Corollary 1.6. *Under Assumptions (I) and (II), for M large enough, one has*

$$\left\| \sup_{t\in[0,T]} \left| X_t - \widehat{X}^M_t \right| \right\|_p \le \tilde{C}\Big(h^\gamma + \big(h|\ln(h)|\big)^{\frac{1}{2}} \Big), \tag{1.21}$$

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

1.3. Previous results and main contributions

1.3.1. The standard McKean-Vlasov equation

Originally used for the study of plasma physics, the standard McKean-Vlasov equation (1.2) has since then been popularized for applications in opinion dynamics [22], finance (for instance through the rank-based model, see [27] and the references therein) and neurosciences [5, 9, 15]. It also plays a key role in the theory of mean-field games [7, 8], with applications in biological models on animal competition, road traffic engineering and dynamic economic models, see Huang-Malhamé-Caines [6] and the references within. We also refer to the recent review [10, 11] for the different notions of propagation of chaos for the particle system associated to this equation and the strategies to prove them.

In this context, the results proved in this paper are known. Regarding the strong well-posedness in L^p and the form of propagation of chaos obtained in this paper, we refer to the illuminating exposition of Sznitman [37] and the recent notes of Lacker [28]. The Euler scheme of (1.2) is studied in a paper by the second author and Pagès [32], where an explicit rate of convergence was also obtained.

1.3.2. Path-dependent framework

The discussion regarding the generalized McKean-Vlasov equation with path-dependent coefficients appears in some recent works, see Cosso et al. [12], Lacker [29] and Djete-Possamaï-Tan [17]. In those papers, the dependency on the measure argument is different from (1.5). More specifically, the dynamics of [12, 17, 29] take the form

$$\mathrm{d}X_t = b\big(t, X_{\cdot\wedge t}, \mathcal{L}(X_{\cdot\wedge t})\big)\mathrm{d}t + \sigma\big(t, X_{\cdot\wedge t}, \mathcal{L}(X_{\cdot\wedge t})\big)\mathrm{d}B_t \tag{1.22}$$

where $\mathcal{L}(X_{\cdot\wedge t}) \in \mathcal{P}_p(\mathcal{C}([0,T],\mathbb{R}^d))$ is the probability distribution of the whole path $X_{\cdot\wedge t}$.

Certainly, in our definition of the path-dependent McKean-Vlasov equation, the measure argument $\mu_{\cdot\wedge t}$ made from the marginal distributions can be considered as a special case of the dependency on $\mathcal{L}(X_{\cdot\wedge t})$. The well-posedness of (1.5), given in Theorem 1.1 has been obtained by Djete-Possamaï-Tan [17]. We still include a proof in order to make this paper self-contained. We emphasize nevertheless several advantages of our setting (1.5):

(i) from the theoretical perspective, Lacker [29] studied the propagation of chaos of (1.22) in total variation distance by means of a Girsanov transform. This work is close to ours, since it focuses on the particle system associated to (1.22) and its convergence rather than on the limit equation (1.22). In contrast with [29], our diffusion coefficient σ is allowed to depend on the measure argument, a case which can not be treated via Girsanov argument. Also, our propagation of chaos result is given in Wasserstein distance;

(ii) our measure arguments are taken in $\mathcal{C}([0,T],\mathcal{P}_p(\mathbb{R}^d))$ instead of $\mathcal{P}_p\big(\mathcal{C}([0,T],\mathbb{R}^d)\big)$. This framework constitutes a trade-off between the theoretical aspects, the numerical perspectives and the applications. Indeed, our setting can be simulated more easily and with more precision in general, and, concerning applications, some path-dependent McKean-Vlasov equations, fitting (1.5), can also be found in the recent work on the 2d parabolic-parabolic Keller-Segel equation of Tomašević and Fournier-Tomašević [39, Equation (1.2)] [21].

We present the regularized equation from [21, 39] as an application, as well as a system, generalized from [20], corresponding to the microscopic dynamics of some neural mass model for the visual cortex and its mean-field limit. In the latter, a delay between the emission and the reception of the signals and a modulation of the exchanges with time create the path-dependency. The inclusion of this latter modulation is important from the modeling point of view, as it opens the possibility to tune in intrinsic excitability into the model.

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. Combining our results on the interpolated Euler scheme and the propagation of chaos for the particle system will be the starting point of a forthcoming work devoted to an implementable particle method for the numerical simulation of (1.5). This line of reasoning is inspired by the numerical analysis for the classical McKean-Vlasov equation, see in particular [1, Section 3], [3, Equation (2.3)], [23], [30, Section 7.1] and [31]. Besides, the Euler scheme also plays a major role in the study of the convex order (see e.g. [32] and [33] for the classical McKean-Vlasov equation) which we also plan to investigate in a future work.

1.4. Strategy and plan of the paper

Section 2 is devoted to our applications. For both models, we show that Assumptions (I) and (II) hold, yielding results of well-posedness in L^p, $p \geq 2$, of propagation of chaos for the associated particle system and the convergence of our interpolated Euler scheme.

In Section 3, we establish the well-posedness of (1.5). The strategy is largely inspired from Bouleau [4], who applied it to diffusions, see also [30, Chapter 5] where the second author derived similar results for the classical McKean-Vlasov equation (without path-dependency), and Lacker [28]. 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 [37] based on earlier work by Dobrushin [18]. 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

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

that we endow with a suitable norm $\|\cdot\|_{p,C,T}$ with parameter $C > 0$. Then, we identify every probability measure on $\mathcal{C}([0,T],\mathbb{R}^d)$ admitting a finite p moment 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

$$\begin{aligned}\Phi_C : \mathcal{H}_{p,C,T}\times\mathcal{C}([0,T],\mathcal{P}_p(\mathbb{R}^d)) &\to \mathcal{H}_{p,C,T}\times\mathcal{C}([0,T],\mathcal{P}_p(\mathbb{R}^d)) \\ (Y,(\nu_t)_{t\in[0,T]}) &\mapsto \Big(\Phi^{(1)}_C(Y,(\nu_t)_{t\in[0,T]}), \iota\big(P_{\Phi^{(1)}_C(Y,(\nu_t)_{t\in[0,T]})}\big)\Big) \end{aligned} \tag{1.23}$$

where

$$\Phi^{(1)}_C(Y,(\nu_t)_{t\in[0,T]}) := \Big(X_0 + \int_0^t b(s,Y_{\cdot\wedge s},\nu_{\cdot\wedge s})\mathrm{d}s + \int_0^t \sigma(s,Y_{\cdot\wedge s},\nu_{\cdot\wedge s})\mathrm{d}B_s\Big)_{t\in[0,T]}$$

is Lipschitz continuous, with a Lipschitz constant strictly smaller than one for C large enough, turning Φ_C into a contraction mapping. Here $\iota\big(P_{\Phi^{(1)}_C(Y,(\nu_t)_{t\in[0,T]})}\big)$ in (1.23) denotes the marginal distributions of $\Phi^{(1)}_C\big(Y,(\nu_t)_{t\in[0,T]}\big)$. This allows to perform a fixed-point argument to obtain the well-posedness of (1.5).

Section 4 is devoted to our proof of Theorem 1.2. We split it into two parts. Proposition 4.2 focuses on the well-posedness of the particle system (1.9) and relies on a by now fairly classical argument:

by writing the coefficients in a vectorized manner, we obtain the system (1.9) in the form of a path-dependent stochastic differential equation in some extended space, whose well-posedness is given by Theorem 1.1. In Subsection 4.2, we prove Theorem 1.2. Our proof is based on a coupling argument. Recalling that $(\mu_t)_{t\in[0,T]}$ denote the marginal distributions of X solving (1.5), we introduce the system $(Y^1,\dots,Y^N)$ such that for all $i\in\{1,\dots,N\}$,

$$Y_t^i = X_0^{i,N} + \int_0^t b\big(s, Y^i_{\cdot\wedge s}, \mu_{\cdot\wedge s}\big)\mathrm{d}s + \int_0^t \sigma\big(s, Y^i_{\cdot\wedge s}, \mu_{\cdot\wedge s}\big)\mathrm{d}B_s^i, \qquad t\in[0,T]. \tag{1.24}$$

Note that, by uniqueness in Theorem 1.1, $(Y^1,\dots,Y^N)$ are i.i.d. copies of X. Then, writing $(\nu_t^N)_{t\in[0,T]}$ for the empirical measure of the particle system (1.24) and noticing that

$$\sup_{t\in[0,T]} \mathcal{W}_p\big(\mu_t^N,\mu_t\big) \le \sup_{t\in[0,T]} \mathcal{W}_p\big(\mu_t^N,\nu_t^N\big) + \sup_{t\in[0,T]} \mathcal{W}_p\big(\nu_t^N,\mu_t\big),$$

we control, in expectation, the first term on the right-hand side by the second one using a coupling of (μ_t^N,ν_t^N) for all $t\in[0,T]$ and Assumption (I). Thus, for some constant $C_{p,d,T,L}$ depending on p, d, T and L,

$$\Big\| \sup_{t\in[0,T]} \mathcal{W}_p\big(\mu_t^N,\mu_t\big)\Big\|_p \le C_{p,d,T,L}\Big\| \sup_{t\in[0,T]} \mathcal{W}_p\big(\nu_t^N,\mu_t\big)\Big\|_p. \tag{1.25}$$

It only remains to use the convergence towards 0 of the right-hand side, which follows from general results of convergence in the Wasserstein distance for i.i.d. sampling. We note that this coupling strategy only provides a result valid for a finite time horizon $T>0$: the constant $C_{p,d,T,L}$ explodes as $T\to\infty$.

Section 5 is devoted to the study of the convergence of the interpolated Euler scheme. We start by giving the definition of the theoretical continuous extension $(\widetilde{X}_t)_{t\in[0,T]}$ of $(\widetilde{X}_{t_m})_{0\le m\le M}$ in (1.15). The objective of Section 5 is to prove the convergence of the process $\widetilde{X}=(\widetilde{X}_t)_{t\in[0,T]}$ towards the unique solution X in L^p-norm, which directly implies Theorem 1.5. 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 collections of points and measures. In a second part, we prove that the sup norm of $(\widetilde{X}_t)_{t\in[0,T]}$ is bounded in L^p and study the L^p-norm of a specific modulus of continuity of $(\widetilde{X}_t)_{t\in[0,T]}$ 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 1.5 and Corollary 1.6 by combining the properties of the interpolated Euler scheme and its continuous extension with our assumptions on the drift and diffusion coefficients.

1.5. Notations

We place ourselves in a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge0},\mathbb{P})$ satisfying the usual condition. The law of a random variable X is denoted by $P_X=\mathbb{P}\circ X^{-1}$. Sometimes we also write $X\sim\mu$ to indicate that X has distribution μ. In the whole paper, $(B_t)_{t\in[0,T]}$ denotes an $(\mathcal{F}_t)_{t\in[0,T]}$ Brownian motion valued in $\mathbb{R}^q$, $q\in\mathbb{N}^*$. The same holds for $(B_t^i)_{t\in[0,T]}$ with $i\in\mathbb{N}^*$. 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 $p\in(0,\infty]$. On $\mathbb{R}$, we write $a\wedge b:=\min(a,b)$ and $a\vee b=\max(a,b)$. We write $|\cdot|$ for the Euclidean norm on $\mathbb{R}^d$, $a\cdot b$ for the scalar product of $a,b\in\mathbb{R}^d$, δ_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 $A\in\mathbb{M}_{d,q}(\mathbb{R})$,

$$|||A||| := \sup_{z\in\mathbb{R}^q,|z|\le1} \big|Az\big|.$$

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 $\mathrm{supp}(\mu)$ for the support of a probability distribution μ. The Wasserstein distance on $\mathcal{P}_p(\mathbb{R}^d)$ is denoted by $\mathcal{W}_p$, defined by (1.6). We shall use repeatedly the space of $\mathbb{R}^d$-valued continuous applications, denoted $\mathcal{C}([0,T],\mathbb{R}^d)$, that is,

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

We endow this space with the supremum norm $\|\alpha\|_{\sup} = \sup_{t\in[0,T]} |\alpha_t|$. The projection π_t from $\mathcal{C}([0,T],\mathbb{R}^d)$ to $\mathbb{R}^d$ 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 $\|\cdot\|_{p,C,T}$ defined in (3.1). We further consider the space $\mathcal{H}_{p,C,T}$ of $(\mathcal{F}_t)_{t\in[0,T]}$-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 μ on $\mathcal{C}([0,T],\mathbb{R}^d)$ such that

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

The Wasserstein distance on $\mathcal{P}_p(\mathcal{C}([0,T],\mathbb{R}^d))$ is denoted by $\mathbb{W}_p$, see (1.6). For any two probability distributions μ,ν, the set of all probability distributions with marginals μ and ν is denoted $\Pi(\mu,\nu)$. We also introduce $\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$ 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 (1.6)). We endow this space with the distance

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

We also use the distance $d_{\mathcal{H}\times\mathcal{P}}$ defined on $\mathcal{H}_{p,C,T} \times \mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$ in (3.13). The application ι 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 $\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$ by $\iota(\mu) = (\mu_t)_{t\in[0,T]}$, see Lemma 3.2. We use the notation $C_{p_1,\dots,p_n}$ for a positive constant depending on parameters $p_1,\dots,p_n$ whose value is allowed to change from line to line.

2. Applications

2.1. A neural mass model with intrinsic excitability

Our starting point is the microscopic setting leading, in the mean-field limit, to Jansen and Rit's model [26], in the form of the equations given by Faugeras-Touboul-Cessac [20]. This neural mass model (NMM) includes three different neurons population and is used to get a deeper understanding of visual 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 find in [20], see in particular Figure 2 for a graphical representation.

At the level of the particle system, given a time horizon $T > 0$ and a number $N_j \in \mathbb{N}^*$ of neurons in population j, the equations for the potential of the neuron i in population j of [20] take the form

$$\mathrm{d}V_{j,i}(t) = -\frac{1}{\tau_j} V_{j,i}(t)\mathrm{d}t + \Big(\sum_{k=1}^{3}\sum_{\ell=1}^{N_k} \bar{J}_{j,k} S\big(V_{k,\ell}(t)\big) + I_j(t)\Big)\mathrm{d}t + f_j(t)\mathrm{d}W_t^{j,i}, \tag{2.1}$$

for $t \in [0,T]$, where the first drift term corresponds to a modulation of the exchanges with time realized through the so-called g-shape of the postsynaptic potentials of population j.

An important effect for the visual cortex is the so-called potentiation due to intrinsic excitability [14]: depending on its previous behavior, the sensibility of a neuron to incoming signals can vary. When the neuron was previously highly active, it reaches an excitability state in which incoming signals are magnified. In order to model this feature, we enrich the coefficients $\bar{J}_{j,k}$, constant in (2.1), by including a path-dependent function of the trajectory of the neuron at hand. As a second extension, we include a delay in the signal received by the neurons i from population j from the neuron ℓ from population k. To simplify, we consider the same delay $\triangle$ in each population, but we notice that our setting could easily adapt to treat a delay depending on the population (by a straightforward adaptation of the initial data).

We thus consider

$$\tau_1 = \tau_2 > 0, \quad \tau_3 > 0, \quad \bar{J}_{i,j} = \frac{J_{i,j}}{N_j} \tag{2.2}$$

with

$$J = \begin{pmatrix} 0 & J_{1,2} & J_{1,3} \\ J_{2,1} & 0 & 0 \\ J_{3,1} & 0 & 0 \end{pmatrix} \tag{2.3}$$

where $J_{1,2}, J_{1,3}, J_{2,1}$ and $J_{3,1}$ are functions of $[0,T] \times \mathcal{C}([0,T], \mathbb{R}^d)$ given by

$$J_{i,j}\Big(t, (\alpha_s)_{s\in[0,T]}\Big) = D_{i,j}\Big(1 + \varepsilon \int_0^t \varphi(\alpha_s)\mathrm{d}s\Big), \tag{2.4}$$

where for all $(i,j) \in \{(1,2),(1,3),(2,1),(3,1)\}$, $D_{i,j}$ are fixed, positive constants and ε is a small parameter modulating the rate-based plasticity [36, Section 6.6]. The function φ is assumed to be bounded and Lipschitz continuous from $\mathcal{C}([0,T], \mathbb{R}^d)$ to $\mathbb{R}$.

Note that this extends the model of [20], in which the $J_{i,j}$, modeling the postsynaptic strength between population i and j are all constant, and that our hypotheses allow for any choice of $J_{i,j}$ that are regular enough (see Assumptions I and II). The setting (2.4) should be thought of as a toy model illustrating our ability to take into account the effect of the potential trajectory on the postsynaptic strengths.

We also introduce a delay parameter $\triangle \in [0,T)$ and consider ultimately the following microscopic system, for $j \in \{1,2,3\}$, $i \in \{1,\dots,N_j\}$ and $t \in [0,T]$

$$\begin{aligned} \mathrm{d}V_{j,i}(t) = -\frac{1}{\tau_j} V_{j,i}(t)\mathrm{d}t + \Big(\sum_{k=1}^{3}\sum_{\ell=1}^{N_k} \bar{J}_{j,k}\Big(t, \big(V_{j,i}(\cdot)\big)_{\cdot\wedge t}\Big) S\big(V_{k,\ell}(t-\triangle)\big) + I_j(t)\Big)\mathrm{d}t \\ + f_j(t)\mathrm{d}W_t^{j,i} \end{aligned} \tag{2.5}$$

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

$$S(v) = \frac{v_m}{1 + e^{r(v_0 - v)}}, \tag{2.6}$$

with $r > 0$ and $0 < v_0 < v_m$. Note that this function is bounded and Lipschitz continuous with constant $v_m r$.

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 path-dependent McKean-Vlasov equations. Write μ_t^j for the distribution of the potential of population $j \in \{1,2,3\}$ at time t in $[0,T]$. In the mean-field limit, we obtain the following system set on $[0,T]$, for $j \in \{1,2,3\}$,

$$\begin{cases} \bar{V}_j(t) = \bar{V}_j(0) + \int_0^t e^{-\frac{t-s}{\tau_j}} \Big\{ \sum_{k=1}^3 D_{j,k}\Big(1 + \varepsilon \int_0^s \varphi\big(\bar{V}_j(u)\big)\mathrm{d}u\Big) \int_{\mathbb{R}} S(y)\mu_{s-\triangle}^k(\mathrm{d}y) \Big\}\mathrm{d}s \\ \qquad + \int_0^t e^{-\frac{t-s}{\tau_j}} I_j(s)\mathrm{d}s + \int_0^t e^{-\frac{t-s}{\tau_j}} f_j(s)\mathrm{d}W_s^j \\ \bar{V}_j(t) \sim \mu_t^j, \quad t \in [0,T], \end{cases} \tag{2.7}$$

where (W^1, W^2, W^3) are three independent Brownian motions. We summarize those assumptions as follow:

Assumption 2.1

1. For $s \in [-\triangle, 0]$, fix $\bar{V}_j(s) = 0$ (leading to $\mu_s^j = \delta_0$, $s \in [-\triangle, 0]$ in (2.7));
2. the functions $I_j, f_j : \mathbb{R}_+ \to \mathbb{R}$ are Lipschitz continuous;
3. the function S is given by (2.6);
4. the function φ appearing in the definition of $J_{i,j}$ in (2.4) is bounded, Lipschitz continuous from $\mathcal{C}([0,T], \mathbb{R}^d)$ to $\mathbb{R}$.

Proposition 2.1. *Under Assumption 2.1, the system (2.7) satisfies Assumptions (I) and (II).*

This provides a proof of well-posedness on finite time $[0,T]$ for any $T > 0$ for this upgraded version of the model treated in [20], where we have added possible intrinsic excitability as well as a delay

between the emission and reception of the signals. In addition, Proposition 2.1 induces a moment propagation result, in the sense that, still letting $\bar{V}_j(s) = 0$ for all $s \in [-\triangle, 0)$, 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 $t \in [0, T]$. Proposition 2.1 also provides a justification for the derivation of (2.7) from the particle system (2.1). More precisely, letting for all $s \in [-\triangle, T]$, $\tilde{\mu}_s = \mu_s^1 \otimes \mu_s^2 \otimes \mu_s^3$, we prove in Appendix A.1 the following.

Proposition 2.2. *Let $N \in \mathbb{N}^*$ and assume that $N_1 = N_2 = N_3 = N$. Assume that for all $i \in \{1, \dots, N\}$ and for all $s \in [-\triangle, 0]$, $(V_{1,i}(s), V_{2,i}(s), V_{3,i}(s)) = 0_{\mathbb{R}^3}$. Under Assumption 2.1, the particle system (2.1) is well-defined. Moreover, defining $\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{(V_{1,i}(t), V_{2,i}(t), V_{3,i}(t))}$, we have*

$$\left\| \sup_{t \in [0,T]} \mathcal{W}_p\big(\tilde{\mu}_t, \mu_t^N\big) \right\|_p \underset{N \to \infty}{\longrightarrow} 0.$$

We plan to study numerically (2.7) in a future work, as was done for the first model (2.1) in [20]. Propositions 2.1 and 2.2 give a path towards the simulation of the dynamics of (2.7). 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. A bifurcation analysis must be conducted to get a clear picture of the possible outcomes. A numerical study of (2.7) 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 with classical model for neuron masses [20]. We mention that the justification of neural mass models from the microscopic dynamics is a challenging topic in computational neuroscience, see e.g. Deschle et al. [16] and the references therein.

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

In [39], Tomašević provides a stochastic interpretation, based on earlier work of Talay-Tomašević [38], of the parabolic-parabolic Keller-Segel model via a stochastic representation which falls into the framework of (1.5). 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 ρ_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 [24, 25] for a review of the standard Keller-Segel model and its variations.

In the study of the stochastic representation, Tomašević [39, Equation (3.6)] introduced a regularized problem, on which we will focus. Let $T > 0$ be fixed. Set in $[0, T] \times \mathbb{R}^2$, the equation takes the following form

$$\begin{cases} \mathrm{d}X_t^\epsilon = b_0^\epsilon(t, X_t^\epsilon) + \chi \int_0^t \Big[e^{-\lambda(t-s)} \big(K_{t-s}^\epsilon * \mu_s^\epsilon\big)(X_t^\epsilon)\, \mathrm{d}s \Big] \mathrm{d}t + \mathrm{d}B_t, \\ \mu_s^\epsilon = P_{X_s^\epsilon}, \end{cases} \tag{2.8}$$

where $\epsilon > 0$ is the regularization parameter, λ, χ are positive constants, and

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

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

The equation (2.8) is key to the argument in [39] because drift and density estimates can be obtained for the regularized process, those being uniform in the regularization parameter ϵ when a condition on the size of χ is fulfilled.

In [39, Theorem A.1], a proof of well-posedness in $L^1(\mathbb{R}^d)$ for the problem (2.8) is provided 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 $(B_t)_{t \geq 0}$. We obtain a similar well-posedness in the L^p framework for $p \geq 2$, implying also propagation of moments in finite time, as corollaries of the following proposition.

Proposition 2.3. *Let $p \geq 2$. Assume that $T > 0$ and $X_0 \in L^p$. Then, the equation (2.8) satisfies Assumption (I).*

Corollary 2.4. *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 (2.8) continuous in time such that $\| \sup_{s\in[0,T]} |X_s| \|_p < \infty$.*

The proof of Proposition 2.3, as well as the following Proposition 2.6, is postponed to Appendix A.2. Corollary 2.4 and the following Corollary 2.5 are straightforward applications of Theorems 1.1 and 1.2.

As noticed in [39], 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 [21].

As a first step towards a different approach, we prove a propagation of chaos result for a path-dependent particle system associated to (2.8).

Corollary 2.5. *For $N \geq 1$, $p \geq 2$, let $X_0^{1,N}, \dots, X_0^{N,N}$ be i.i.d. copies of $X_0 \in L^p$ in (2.8) and $B^i := (B_t^i)_{t\in[0,T]}$, $1 \leq i \leq N$ be N independent standard $\mathbb{R}^2$-valued Brownian motions, independent of the Brownian motion $(B_t)_{t\in[0,T]}$ in (2.8) and $(X_0, X_0^{1,N}, \dots, X_0^{N,N})$. Then, under the same assumption as in Proposition 2.3, the N-particle exchangeable regularized Keller-Segel system set on $[0,T]$*

$$\begin{cases} X_t^{i,N} = X_0^{i,N} + \int_0^t b_0^\epsilon(s, X_s^{i,N})\,\mathrm{d}s + \chi \int_0^t \int_0^s \Big[e^{-\lambda(s-u)} \big(K_{s-u}^\epsilon * \mu_u^N\big)\big(X_s^\epsilon\big)\,\mathrm{d}u \Big]\mathrm{d}s + B_t^i, \\ \mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}, \qquad t \in [0,T], \end{cases}$$

is well-posed, and

$$\Big\| \sup_{t\in[0,T]} \mathcal{W}_p\Big(\mu_t^\epsilon, \mu_t^N\Big) \Big\|_p \underset{N\to\infty}{\longrightarrow} 0.$$

We now turn to the issue of simulating the Keller-Segel parabolic-parabolic equation in dimension 2. We show that the interpolated Euler scheme corresponding to (2.8) converges, in L^p norm, to the desired solution, in the case where c_0 also belongs to $W^{1,\infty}(\mathbb{R}^d)$. This gives a first approach to the simulation of the problem based on a non-markovian particle system associated to the regularized path-dependent equation thanks to the process from Definition 1.4. Again, the proof is presented in Appendix A.2.

Proposition 2.6. *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 (2.8) given by Proposition 2.3. Let $(\widetilde{X}_{t_m}^M)_{0\leq m\leq M}$ be the interpolated Euler scheme from Definition 1.4. Then, for M large enough, for some $\tilde{C} > 0$ independent of M, for $h = \frac{T}{M}$, we have*

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

3. Strong well-posedness

In this section, we prove Theorem 1.1, following the strategy sketched in Section 1.4. 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 $\|Y\|_p := \big[\mathbb{E}\,\|Y\|_{\sup}^p\big]^{1/p} = \big[\mathbb{E}\sup_{t\in[0,T]}|Y_t|^p\big]^{1/p} < +\infty$. 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

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

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

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

We define

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

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 [30, Lemma 5.1.1].

Lemma 3.1. *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 the space

$$\mathcal{P}_p\big(\mathcal{C}([0,T],\mathbb{R}^d)\big) \coloneqq \left\{\mu \text{ probability distribution on } \mathcal{C}([0,T],\mathbb{R}^d) \text{ s.t. } \int_{\mathcal{C}([0,T],\mathbb{R}^d)} \|\alpha\|^p_{\sup}\,\mu(\mathrm{d}\alpha) < +\infty\right\}.$$

Recall the definition of the Wasserstein distance on $\mathcal{P}_p(S)$ for (S,d_S) a Polish space from (1.6), and that we write $\mathbb{W}_p$ when $(S,d_S) = (\mathcal{C}([0,T],\mathbb{R}^d),\|\cdot\|_{\sup})$, $\mathcal{W}_p$ when $S=\mathbb{R}^d$. Since the underlying space $\big(\mathcal{C}([0,T],\mathbb{R}^d),\|\cdot\|_{\sup}\big)$ is Polish, the space $\mathcal{P}_p\big(\mathcal{C}([0,T],\mathbb{R}^d)\big)$ equipped with $\mathbb{W}_p$ is complete and separable (see [2]).

Let us consider now

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

equipped with the distance

$$d_p\big((\mu_t)_{t\in[0,T]},(\nu_t)_{t\in[0,T]}\big) \coloneqq \sup_{t\in[0,T]} \mathcal{W}_p(\mu_t,\nu_t). \tag{3.4}$$

As $\big(\mathcal{P}_p(\mathbb{R}^d),\mathcal{W}_p\big)$ is a complete space (see [2]), $\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$ equipped with the uniform distance d_p is also a complete space.

For any $t\in[0,T]$, we define $\pi_t : \mathcal{C}\big([0,T],\mathbb{R}^d\big)\to\mathbb{R}^d$ by

$$\alpha\mapsto\pi_t(\alpha)=\alpha_t. \tag{3.5}$$

The following lemma, and its proof, can be found in [30, Lemmas 5.1.2 & 5.1.3].

Lemma 3.2. *The application $\iota : \mathcal{P}_p\big(\mathcal{C}([0,T],\mathbb{R}^d)\big)\to\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$ defined by*

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

is well-defined and 1-Lipschitz continuous.

Lemma 3.3. *Under Assumption (I), the coefficient functions b and σ have a linear growth in α and $(\mu_t)_{t\in[0,T]}$ in the sense that there exists a constant $C_{b,\sigma,L,T}$ s.t. for every $t\in[0,T]$, $\alpha\in\mathcal{C}\big([0,T],\mathbb{R}^d\big)$, $(\mu_t)_{t\in[0,T]}\in\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$,*

$$\big|b(t,\alpha,(\mu_t)_{t\in[0,T]})\big| \vee \big|\!\big|\!\big|\sigma(t,\alpha,(\mu_t)_{t\in[0,T]})\big|\!\big|\!\big| \le C_{b,\sigma,L,T}\Big(1+\|\alpha\|_{\sup}+\sup_{t\in[0,T]}\mathcal{W}_p(\mu_t,\delta_0)\Big). \tag{3.6}$$

Proof of Lemma 3.3. Let $\delta_{0,[0,T]}\in\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$ be such that $\delta_{0,[0,T]}(t)=\delta_0$ for all $t\in[0,T]$ and let $\mathbf{0}\in\mathcal{C}\big([0,T],\mathbb{R}^d\big)$ be such that for all $t\in[0,T]$, $\mathbf{0}(t)=0$. Then

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

Consequently,

$$\Big|b\big(t,\alpha,(\mu_t)_{t\in[0,T]}\big)\Big| \le \Big(\sup_{t\in[0,T]}\big|b\big(t,\mathbf{0},\delta_{0,[0,T]}\big)\big| \vee L\Big)\Big(\|\alpha\|_{\sup}+\sup_{t\in[0,T]}\mathcal{W}_p(\mu_t,\delta_0)+1\Big). \tag{3.7}$$

Similarly, we have

$$\big|\!\big|\!\big|\sigma\big(t,\alpha,(\mu_t)_{t\in[0,T]}\big)\big|\!\big|\!\big| \le \Big(\sup_{t\in[0,T]}\big|\!\big|\!\big|\sigma\big(t,\mathbf{0},\delta_{0,[0,T]}\big)\big|\!\big|\!\big| \vee L\Big)\Big(\|\alpha\|_{\sup}+\sup_{t\in[0,T]}\mathcal{W}_p(\mu_t,\delta_0)+1\Big) \tag{3.8}$$

so that one can take $C_{b,\sigma,L,T} \coloneqq \sup_{t\in[0,T]}\big|b\big(t,\mathbf{0},\delta_{0,[0,T]}\big)\big| \vee \sup_{t\in[0,T]}\big|\!\big|\!\big|\sigma\big(t,\mathbf{0},\delta_{0,[0,T]}\big)\big|\!\big|\!\big| \vee L$ to conclude. □

Before proving that the McKean-Vlasov equation (1.5) 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 [35, Section 7.8] among other references.

Lemma 3.4 (The Generalized Minkowski Inequality)**.** *For any (bi-measurable) process $X = (X_t)_{t\geq 0}$, for every $p \in [1,\infty)$ and for every $T \in [0,+\infty]$,*

$$\left\| \int_0^T X_t \, \mathrm{d}t \right\|_p \leq \int_0^T \|X_t\|_p \, \mathrm{d}t.$$

Lemma 3.5 (Burkölder-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), $t \to \sigma(t, X_{\cdot\wedge t}, \mu_{\cdot\wedge t})$ 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 3.6. *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 $\int_0^T |||H_t|||^2 \mathrm{d}t < \infty$, $\mathbb{P}$-a.s.. Then, for all $t \in [0,T]$,*

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

Proof. Notice first that it follows from Lemma 3.5 that $\int_0^\cdot H_s \mathrm{d}B_s$ is a d-dimensional local martingale satisfying

$$\left\| \sup_{s\in[0,t]} \left| \int_0^s H_u \, \mathrm{d}B_u \right| \right\|_p \leq C_{d,p}^{BDG} \left\| \sqrt{\int_0^t |||H_u|||^2 \mathrm{d}u} \right\|_p . \tag{3.9}$$

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

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

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

Lemma 3.7. *Under Assumption (I), for any $\big(X, (\mu_t)_{t\in[0,T]}\big), \big(Y, (\nu_t)_{t\in[0,T]}\big)$ belowing to the space $\mathcal{H}_{p,C,T} \times \mathcal{C}\big([0,T], \mathcal{P}_p(\mathbb{R}^d)\big)$ and for any $t \in [0,T]$, one has*

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

and

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

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

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

$$\begin{aligned}
\Big\|\sup_{s\in[0,t]}\Big|\int_0^s\big[b(u,X_{\cdot\wedge u},\mu_{\cdot\wedge u})-b(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\big]\mathrm{d}u\Big|\Big\|_p &\le \Big\|\int_0^t\big|b(u,X_{\cdot\wedge u},\mu_{\cdot\wedge u})-b(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\big|\mathrm{d}u\Big\|_p\\
&\le \int_0^t\big\|b(u,X_{\cdot\wedge u},\mu_{\cdot\wedge u})-b(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\big\|_p\,\mathrm{d}u \quad \text{(by Lemma 3.4)}\\
&\le \int_0^t\Big\|L\big[\,\|X_{\cdot\wedge u}-Y_{\cdot\wedge u}\|_{\sup}+d_p\big((\mu_{v\wedge u})_{v\in[0,T]},(\nu_{v\wedge u})_{v\in[0,T]}\big)\big]\Big\|_p\,\mathrm{d}u\\
&\le \int_0^t\Big\|L\big[\sup_{s\in[0,u]}|X_s-Y_s|+\sup_{s\in[0,u]}\mathcal{W}_p(\mu_s,\nu_s)\big]\Big\|_p\,\mathrm{d}u\\
&\quad\text{(by Assumption (I) and by definitions (1.3) and (1.4))}\\
&\le L\int_0^t\Big[\Big\|\sup_{s\in[0,u]}|X_s-Y_s|\Big\|_p+\sup_{s\in[0,u]}\mathcal{W}_p(\mu_s,\nu_s)\Big]\mathrm{d}u
\end{aligned}\tag{3.10}$$

and, by applying Lemma 3.6,

$$\begin{aligned}
&\Big\|\sup_{s\in[0,t]}\Big|\int_0^s\big[\sigma(u,X_{\cdot\wedge u},\mu_{\cdot\wedge u})-\sigma(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\big]\mathrm{d}B_u\Big|\Big\|_p\\
&\quad\le C^{BDG}_{d,p}\Big[\int_0^t\big\|\|\sigma(u,X_{\cdot\wedge u},\mu_{\cdot\wedge u})-\sigma(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\|\big\|_p^2\mathrm{d}u\Big]^{\frac12}.
\end{aligned}$$

By Assumption (I) and by definition of $\alpha_{\cdot\wedge u}$ and $\mu_{\cdot\wedge u}$ in (1.3) and (1.4), we get

$$\begin{aligned}
&\Big\|\sup_{s\in[0,t]}\Big|\int_0^s\big[\sigma(u,X_{\cdot\wedge u},\mu_{\cdot\wedge u})-\sigma(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\big]\mathrm{d}B_u\Big|\Big\|_p\\
&\quad\le C^{BDG}_{d,p}\Big[\int_0^t\Big\|L\big[\sup_{s\in[0,u]}|X_s-Y_s|+\sup_{s\in[0,u]}\mathcal{W}_p(\mu_s,\nu_s)\big]\Big\|_p^2\mathrm{d}u\Big]^{\frac12}\\
&\quad\le C^{BDG}_{d,p}L\Big[\int_0^t\Big[\Big\|\sup_{s\in[0,u]}|X_s-Y_s|\Big\|_p+\sup_{s\in[0,u]}\mathcal{W}_p(\mu_s,\nu_s)\Big]^2\mathrm{d}u\Big]^{\frac12}\\
&\quad\le \sqrt{2}\,C^{BDG}_{d,p}L\Big\{\int_0^t\Big[\Big\|\sup_{s\in[0,u]}|X_s-Y_s|\Big\|_p^2+\sup_{s\in[0,u]}\mathcal{W}_p^2(\mu_s,\nu_s)\Big]\mathrm{d}u\Big\}^{\frac12}.
\end{aligned}\tag{3.11}$$

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

The idea of our proof of well-posedness follows from Feyel's approach, originally developed by Bouleau [4, Section 7] for the existence and uniqueness of a strong solution to the SDE $\mathrm{d}X_t=b(X_t)\mathrm{d}t+\sigma(X_t)\mathrm{d}B_t$. We define a distance $d_{p,C,T}$ on $\mathcal{C}([0,T],\mathcal{P}_p(\mathbb{R}^d))$ as follows:

$$\begin{aligned}
&\forall(\mu_t)_{t\in[0,T]},(\nu_t)_{t\in[0,T]}\in\mathcal{C}([0,T],\mathcal{P}_p(\mathbb{R}^d)),\\
&\qquad d_{p,C,T}\big((\mu_t)_{t\in[0,T]},(\nu_t)_{t\in[0,T]}\big):=\sup_{t\in[0,T]}e^{-Ct}\,\mathcal{W}_p(\mu_t,\nu_t).
\end{aligned}\tag{3.12}$$

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:

$$\begin{aligned}
&\forall\big(X,(\mu_t)_{t\in[0,T]}\big),\big(Y,(\nu_t)_{t\in[0,T]}\big)\in\mathcal{H}_{p,C,T}\times\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big),\\
&d_{\mathcal{H}\times\mathcal{P}}\Big(\big(X,(\mu_t)_{t\in[0,T]}\big),\big(Y,(\nu_t)_{t\in[0,T]}\big)\Big):=\|X-Y\|_{p,C,T}+\sup_{t\in[0,T]}e^{-Ct}\mathcal{W}_p(\mu_t,\nu_t).
\end{aligned}\tag{3.13}$$

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

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

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

by

$$\begin{aligned}
&\forall\,(Y,(\nu_t)_{t\in[0,T]})\in\mathcal{H}_{p,C,T}\times\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big),\\
&\Phi_C(Y,(\nu_t)_{t\in[0,T]})=\Big(\Phi_C^{(1)}(Y,(\nu_t)_{t\in[0,T]}),\iota(P_{\Phi_C^{(1)}(Y,(\nu_t)_{t\in[0,T]})})\Big),\qquad\text{where}\\
&\Phi_C^{(1)}(Y,(\nu_t)_{t\in[0,T]}):=\Big(X_0+\int_0^t b(s,Y_{\cdot\wedge s},\nu_{\cdot\wedge s})\mathrm{d}s+\int_0^t\sigma(s,Y_{\cdot\wedge s},\nu_{\cdot\wedge s})\mathrm{d}B_s\Big)_{t\in[0,T]}
\end{aligned}$$

The application Φ_C has the following properties.

Proposition 3.8. *(i) Under Assumption (I), the map Φ_C is well-defined, and there holds*

$$\begin{aligned}
&\Big\|\sup_{s\in[0,t]}\big|\Phi_C^{(1)}(Y,\iota(P_Y))_s\big|\Big\|_p\\
&\quad\le\|X_0\|_p+C_{b,\sigma,L,T}(2T+C_{d,p}^{BDG}\sqrt{2T})+2C_{b,\sigma,L,T}\int_0^t\Big\|\sup_{s\in[0,u]}|Y_s|\Big\|_p\mathrm{d}u\\
&\qquad+2\sqrt{2}C_{d,p}^{BDG}C_{b,\sigma,L,T}\Big(\int_0^t\Big\|\sup_{s\in[0,u]}|Y_s|\Big\|_p^2\mathrm{d}u\Big)^{\frac12}.
\end{aligned}\tag{3.14}$$

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

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

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

Proof. *(i)* It follows from Lemma 3.2 that for every $Y\in\mathcal{H}_{p,C,T}$, $\iota(P_Y)$ belongs to $\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$. Let $\nu=P_Y$. We need to prove that $\Phi_C^{(1)}\big(Y,\iota(\nu)\big)\in\mathcal{H}_{p,C,T}$. For any $t\in[0,T]$,

$$\begin{aligned}
&\Big\|\sup_{s\in[0,t]}\big|\Phi_C^{(1)}\big(Y,\iota(\nu)\big)_s\big|\Big\|_p\\
&=\Big\|\sup_{s\in[0,t]}\big|X_0+\int_0^s b(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\mathrm{d}u+\int_0^s\sigma(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\mathrm{d}B_u\big|\Big\|_p\\
&\le\Big\||X_0|+\int_0^t\big|b(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\big|\mathrm{d}u+\sup_{s\in[0,t]}\Big|\int_0^s\sigma(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\mathrm{d}B_u\Big|\Big\|_p\\
&\le\|X_0\|_p+\Big\|\int_0^t\big|b(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\big|\mathrm{d}u\Big\|_p+\Big\|\sup_{s\in[0,t]}\Big|\int_0^s\sigma(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\mathrm{d}B_u\Big|\Big\|_p.
\end{aligned}\tag{3.15}$$

Owing to Assumption (I), we have $\|X_0\|_p<+\infty$. For the second part of (3.15), it follows from Lemma 3.3 that

$$\begin{aligned}
\Big\|\int_0^t b(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\mathrm{d}u\Big\|_p&\le\int_0^t\|b(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\|_p\,\mathrm{d}u\\
&\le\int_0^t\Big\|C_{b,\sigma,L,T}\big(1+\|Y_{\cdot\wedge u}\|_{\sup}+\sup_{s\in[0,T]}\mathcal{W}_p(\nu_{s\wedge u},\delta_0)\big)\Big\|_p\mathrm{d}u\\
&\le\int_0^t C_{b,\sigma,L,T}\big(1+\Big\|\sup_{s\in[0,T]}|Y_{s\wedge u}|\Big\|_p+\sup_{s\in[0,T]}\|Y_{s\wedge u}\|_p\big)\mathrm{d}u\\
&\le 2C_{b,\sigma,L,T}\int_0^t\Big(1+\Big\|\sup_{s\in[0,T]}|Y_{s\wedge u}|\Big\|_p\Big)\mathrm{d}u\\
&\le 2C_{b,\sigma,L,T}\int_0^t\Big(1+\Big\|\sup_{s\in[0,u]}|Y_s|\Big\|_p\Big)\mathrm{d}u
\end{aligned}\tag{3.16}$$

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

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

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

$$\Big\| \int_0^t b(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u}) \mathrm{d}u \Big\|_p \le 2C_{b,\sigma,L,T} \int_0^t \big(1 + e^{CT} \|Y\|_{p,C,T}\big) \mathrm{d}u < +\infty,$$

where we used (3.2) in the following way

$$\Big\| \sup_{s\in[0,u]} |Y_s| \Big\|_p \le \Big\| \sup_{s\in[0,T]} |Y_s| \Big\|_p \le e^{CT} \|Y\|_{p,C,T} < +\infty. \tag{3.17}$$

On the other hand, by using Lemma 3.6,

$$\Big\| \sup_{s\in[0,t]} \Big| \int_0^s \sigma(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u}) \mathrm{d}B_u \Big| \Big\|_p \le C_{d,p}^{BDG} \Big[\int_0^t \big\| \|\sigma(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u})\| \big\|_p^2 \mathrm{d}u \Big]^{\frac{1}{2}}.$$

As before, we then invoke Lemma 3.3 to obtain

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

where we used again the inequality $\sup_{s\in[0,T]} \|Y_{s\wedge u}\|_p \le \| \sup_{s\in[0,T]} |Y_{s\wedge u}| \|_p$. Making again use of $\| \sup_{s\in[0,T]} |Y_{s\wedge u}| \|_p = \| \sup_{s\in[0,u]} |Y_s| \|_p$, we get

$$\begin{aligned}
\Big\| \sup_{s\in[0,t]} \Big| \int_0^s \sigma(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u}) \mathrm{d}B_u \Big| \Big\|_p &\le C_{d,p}^{BDG}\, C_{b,\sigma,L,T} \Big\{ 2T + 8 \int_0^t \Big\| \sup_{s\in[0,u]} |Y_s| \Big\|_p^2 \mathrm{d}u \Big\}^{\frac{1}{2}} \\
&\le C_{d,p}^{BDG}\, C_{b,\sigma,L,T} \Big(\sqrt{2T} + 2\sqrt{2} \Big\{ \int_0^t \Big\| \sup_{s\in[0,u]} |Y_s| \Big\|_p^2 \mathrm{d}u \Big\}^{\frac{1}{2}} \Big) \\
&< +\infty,
\end{aligned} \tag{3.18}$$

where the last inequality of the above formula is due to (3.17), and where we used the inequality $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ for $a, b \ge 0$.

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

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

Thus $\Phi_C^{(1)}\big(Y, \iota(\nu)\big) \in \mathcal{H}_{p,C,T}$. The inequality (3.14) follows by injecting (3.16) and (3.18) into (3.15).

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

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

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

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

$$\Big\|\Phi_C^{(1)}\big(X, \iota(\mu)\big) - \Phi_C^{(1)}\big(Y, \iota(\nu)\big)\Big\|_{p,C,T} \le \Big(\frac{2L}{C} + \frac{C_{d,p,L}}{\sqrt{C}}\Big)\|X - Y\|_{p,C,T}, \tag{3.19}$$

where $C_{d,p,L} > 0$ is the constant given by Lemma 3.7 and is independent of the parameter C of the application Φ_C. For any $X, Y \in \mathcal{H}_{p,C,T}$, set $\mu = P_X$ and $\nu = P_Y$. Then

$$\begin{aligned} &\Big\|\Phi_C^{(1)}\big(X, \iota(\mu)\big) - \Phi_C^{(1)}\big(Y, \iota(\nu)\big)\Big\|_{p,C,T} \\ &= \Big\| \int_0^\cdot \big(b(u, X_{\cdot\wedge u}, \mu_{\cdot\wedge u}) - b(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u})\big)\mathrm{d}u \\ &\qquad + \int_0^\cdot \big(\sigma(u, X_{\cdot\wedge u}, \mu_{\cdot\wedge u}) - \sigma(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u})\big)\mathrm{d}B_u\Big\|_{p,C,T} \qquad (3.20) \\ &\le \Big\|\int_0^\cdot \big(b(u, X_{\cdot\wedge u}, \mu_{\cdot\wedge u}) - b(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u})\big)\mathrm{d}u\Big\|_{p,C,T} \\ &\qquad + \Big\|\int_0^\cdot \big(\sigma(u, X_{\cdot\wedge u}, \mu_{\cdot\wedge u}) - \sigma(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u})\big)\mathrm{d}B_u\Big\|_{p,C,T} \\ &= \sup_{t\in[0,T]} e^{-Ct}\Big\|\sup_{s\in[0,t]}\Big|\int_0^s \big[b(u, X_{\cdot\wedge u}, \mu_{\cdot\wedge u}) - b(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u})\big]\mathrm{d}u\Big|\Big\|_p \qquad (3.21) \\ &\qquad + \sup_{t\in[0,T]} e^{-Ct}\Big\|\sup_{s\in[0,t]}\Big|\int_0^s \big[\sigma(u, X_{\cdot\wedge u}, \mu_{\cdot\wedge u}) - \sigma(u, Y_{\cdot\wedge u}, \nu_{\cdot\wedge u})\big]\mathrm{d}Bu\Big|\Big\|_p\,. \end{aligned}$$

We treat the two terms in (3.20) separately. Owing to Lemma 3.7, we first have

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

On the other hand, using Lemma 3.7

$$
\begin{aligned}
&\sup_{t\in[0,T]} e^{-Ct}\left\| \sup_{s\in[0,t]}\left|\int_0^s \left[\sigma(u,X_{\cdot\wedge u},\mu_{\cdot\wedge u})-\sigma(u,Y_{\cdot\wedge u},\nu_{\cdot\wedge u})\right]\mathrm{d}Bu\right|\right\|_p \\
&\quad\le \sup_{t\in[0,T]} e^{-Ct}C_{d,p,L}\Big\{\int_0^t\Big[\Big\|\sup_{s\in[0,u]}|X_s-Y_s|\Big\|_p^2+\sup_{s\in[0,u]}\mathcal{W}_p^2(\mu_s,\nu_s)\Big]\mathrm{d}u\Big\}^{\frac12} \\
&\quad\le \sup_{t\in[0,T]} e^{-Ct}C_{d,p,L}\Big\{\int_0^t\Big[\Big\|\sup_{s\in[0,u]}|X_s-Y_s|\Big\|_p^2+\sup_{s\in[0,u]}\|X_s-Y_s\|_p^2\Big]\mathrm{d}u\Big\}^{\frac12} \\
&\quad\le \sqrt{2}C_{d,p,L}\sup_{t\in[0,T]} e^{-Ct}\Big\{\int_0^t e^{2Cu}\Big(e^{-Cu}\Big\|\sup_{s\in[0,u]}|X_s-Y_s|\Big\|_p\Big)^2\mathrm{d}u\Big\}^{\frac12} \\
&\quad\le \sqrt{2}C_{d,p,L}\|X-Y\|_{p,C,T}\sup_{t\in[0,T]} e^{-Ct}\Big\{\int_0^t e^{2Cu}\mathrm{d}u\Big\}^{\frac12} \\
&\quad\le \sqrt{2}\,C_{d,p,L}\,\|X-Y\|_{p,C,T}\sup_{t\in[0,T]} e^{-Ct}\Big[\frac{e^{2Ct}-1}{2C}\Big]^{\frac12} \\
&\quad\le \frac{C_{d,p,L}}{\sqrt{C}}\,\|X-Y\|_{p,C,T}\,,
\end{aligned}
$$

since $\sup_{t\in[0,T]} e^{-Ct}\left[\frac{e^{2Ct}-1}{2C}\right]^{\frac12}\le\frac{1}{\sqrt{2C}}$. Injecting those two results in (3.20), we obtain (3.19).

Step 3. We combine the results of the previous two steps to conclude. We have, from the definition of $d_{\mathcal{H}\times\mathcal{P}}$ in (3.13),

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

and using Step 1, and (3.19),

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

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

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 [35, Lemma 7.3] for a proof (among many others).

Lemma 3.9 ("À la Gronwall" Lemma)**.** *Let $f:[0,T]\to\mathbb{R}_+$ be a Borel, locally bounded and non-decreasing function and let $\psi:[0,T]\to\mathbb{R}_+$ be a non-negative non-decreasing function satisfying*

$$
\forall t\in[0,T],\ f(t)\le A\int_0^t f(s)\mathrm{d}s+B\left(\int_0^t f^2(s)\mathrm{d}s\right)^{\frac12}+\psi(t),
$$

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

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

Proposition 3.8 directly implies the existence and uniqueness of a strong solution of the McKean-Vlasov equation (1.5) as shown below.

Proof of Theorem 1.1. Proposition 3.8 implies that Φ_C is a Lipschitz continuous function. Moreover, for a large enough constant C, we have $\left(\frac{K_1}{C}+\frac{K_2}{\sqrt{C}}\right)<1$, so that Φ_C is a contraction mapping. Therefore, Φ_C has a unique fixed point $\big(H,\iota(P_H)\big)$ in $\mathcal{H}_{p,C,T}\times\mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$ and this process H is the unique strong solution of the McKean-Vlasov equation (1.5).

We turn to the proof of (1.8). Let (X, P_X) be the unique strong solution of (1.5). Then, we have $\Phi_C^{(1)}(X, \iota(P_X)) = X$ since X is a fixed point of the application Φ_C. Therefore, (3.14) takes the following form

$$\Big\| \sup_{s\in[0,t]} |X_s| \Big\|_p \le \|X_0\|_p + C_{b,\sigma,L,T}(2T + C_{d,p}^{BDG}\sqrt{2T}) + 2C_{b,\sigma,L,T}\int_0^t \Big\| \sup_{s\in[0,u]} |X_s| \Big\|_p \mathrm{d}u$$
$$+ 2\sqrt{2}C_{d,p}^{BDG}C_{b,\sigma,L,T}\Big(\int_0^t \Big\| \sup_{s\in[0,u]} |X_s| \Big\|_p^2 \mathrm{d}u\Big)^{\frac{1}{2}}.$$

We let $f(t) := \Big\| \sup_{s\in[0,t]} |X_s| \Big\|_p$, and apply Lemma 3.9 to get

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

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

$$C_{p,d,b,\sigma,L,T} = \big(4C_{b,\sigma,L,T} + 8(C_{d,p}^{BDG}\, C_{b,\sigma,L,T})^2\big) \vee 2\big(1 \vee C_{b,\sigma,L,T}T + \sqrt{2T}\, C_{d,p}^{BDG}\, C_{b,\sigma,L,T}\big). \quad (3.22)$$

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

4. Propagation of chaos for the particle system

In this section, we study the particle system (1.9) and its convergence when $N \to \infty$. In Subsection 4.1, we prove the well-posedness of the particle system at fixed N. For this, we write the system (1.9) as an Nd-dimensional path-dependent diffusion to which we apply Theorem 1.1. In Subsection 4.2, we show the propagation of chaos. We use a synchronous coupling with i.i.d. particles sharing the marginal distributions of the solution to (1.5). This is of course reminiscent of the celebrated approach developed in dimension 1 by Sznitman [37], although our proof is more in the line of the recent exposition of Lacker [28].

4.1. Well-posedness of the particle system

We begin this subsection with a technical lemma. For the sake of clarity, for $\ell \ge 1$, we write $|\cdot|_\ell$ to denote the Euclidean norm in $\mathbb{R}^\ell$ in what follows.

Lemma 4.1. *Let $N \in \mathbb{N}^*$ be fixed, let $p \ge 2$ and consider the application*

$$L_N : \mathcal{C}\big([0,T], (\mathbb{R}^d)^N\big) \to \mathcal{C}\big([0,T], \mathcal{P}_p(\mathbb{R}^d)\big)$$
$$\bar\alpha = (\alpha^1, \dots, \alpha^N) \to L_N(\bar\alpha) = \Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}\Big)_{t\in[0,T]}.$$

Then,

(i) the application L_N is well-defined;

(ii) the application L_N is $N^{-\frac{1}{p}}$-Lipschitz continuous from $\big(\mathcal{C}\big([0,T], (\mathbb{R}^d)^N\big), \|\cdot\|_{\sup}\big)$ to the space $\big(\mathcal{C}\big([0,T], \mathcal{P}_p(\mathbb{R}^d)\big), d_p\big)$ where $\|\bar\alpha\|_{\sup} = \sup_{t\in[0,T]} |\bar\alpha_t|_{Nd}$ and d_p is defined by (1.7).

Proof. (i) We prove that for all $\bar\alpha = (\alpha^1, \dots, \alpha^N) \in \mathcal{C}\big([0,T], (\mathbb{R}^d)^N\big)$,

(a) for all $t \in [0,T]$, $N^{-1}\sum_{i=1}^N \delta_{\alpha_t^i} \in \mathcal{P}_p(\mathbb{R}^d)$;
(b) the map $t \to N^{-1}\sum_{i=1}^N \delta_{\alpha_t^i}$ is continuous with respect to $\mathcal{W}_p$.

For (a), the property follows from the fact that $N^{-1}\sum_{i=1}^N \delta_{\alpha_t^i}$ is obviously a probability measure on $\mathbb{R}^d$, which satisfies

$$\int_{\mathbb{R}^d} |\zeta|_d^p \Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}\Big)(\mathrm{d}\zeta) = \frac{1}{N}\sum_{i=1}^N |\alpha_t^i|_d^p \le \max_{1\le i\le N} \sup_{t\in[0,T]} |\alpha_t^i|_d^p < \infty$$

by assumptions on $(\alpha^1,\dots,\alpha^N)$. To prove (b), we consider $s,t\in[0,T]$. Noting that $\frac{1}{N}\sum_{i=1}^N \delta_{(\alpha_t^i,\alpha_s^i)}$ is a coupling of

$$\Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}, \frac{1}{N}\sum_{i=1}^N \delta_{\alpha_s^i}\Big),$$

the definition of the Wasserstein distance yields

$$\begin{aligned}\mathcal{W}_p^p\Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}, \frac{1}{N}\sum_{i=1}^N \delta_{\alpha_s^i}\Big) &\le \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|_d^p \Big(\frac{1}{N}\sum_{i=1}^N \delta_{(\alpha_t^i,\alpha_s^i)}\Big)(\mathrm{d}x,\mathrm{d}y)\\ &= \frac{1}{N}\sum_{i=1}^N \Big|\alpha_t^i-\alpha_s^i\Big|_d^p,\end{aligned}$$

and this last term converges to 0 as $|t-s|\to 0$ by continuity of each $(\alpha_u^i)_{u\in[0,T]}$.

(ii) Let $\bar\alpha=(\alpha^1,\dots,\alpha^N)$, $\bar\beta=(\beta^1,\dots,\beta^N)$ be elements of $\mathcal{C}([0,T],(\mathbb{R}^d)^N)$. Then, using that $\frac{1}{N}\sum_{i=1}^N \delta_{(\alpha_t^i,\beta_t^i)}$ is a coupling of

$$\Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}, \frac{1}{N}\sum_{i=1}^N \delta_{\beta_t^i}\Big),$$

and using the definition of the Wasserstein distance, we get

$$\begin{aligned} d_p^p\Big(L_N(\bar\alpha),L_N(\bar\beta)\Big) &= \sup_{t\in[0,T]} \mathcal{W}_p^p\Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}, \frac{1}{N}\sum_{i=1}^N \delta_{\beta_t^i}\Big)\\ &\le \sup_{t\in[0,T]}\Big[\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|_d^p \Big(\frac{1}{N}\sum_{i=1}^N \delta_{(\alpha_t^i,\beta_t^i)}\Big)(\mathrm{d}x,\mathrm{d}y)\Big]\\ &= \sup_{t\in[0,T]} \frac{1}{N}\sum_{i=1}^N \Big|\alpha_t^i-\beta_t^i\Big|_d^p\\ &\le \frac{1}{N}\sup_{t\in[0,T]} \Big|\bar\alpha_t-\bar\beta_t\Big|_{Nd}^p \end{aligned}$$

where we used that $p\ge 2$ and the convexity of $x\mapsto x^{\frac{p}{2}}$ in $\mathbb{R}_+$. We conclude by noticing that the last inequality implies

$$d_p\Big(L_N(\bar\alpha),L_N(\bar\beta)\Big) \le N^{-\frac{1}{p}} \sup_{t\in[0,T]} \Big|\bar\alpha_t-\bar\beta_t\Big|_{Nd} = N^{-\frac{1}{p}}\|\bar\alpha-\bar\beta\|_{\sup}.\square$$

Proposition 4.2. *Let $N\in\mathbb{N}^*$ be fixed and suppose that Assumption (I) holds for some $p\ge 2$. Then the particle system (1.9) is well-posed, and for all $t\in[0,T]$, $i\in\{1,\dots,N\}$, $X^{i,N}$ belongs to $L^p_{\mathcal{C}([0,T],\mathbb{R}^d)}(\mathbb{P})$.*

Proof of Proposition 4.2. For all $t\in[0,T]$, we write $\bar X_t=(X_t^{1,N},\dots,X_t^{N,N})$ in $(\mathbb{R}^d)^N$ and $\bar W_t=(B_t^1,\dots B_t^N)$. We rewrite the system (1.9) as follows

$$\begin{cases} \mathrm{d}\bar X_t &= B(t,\bar X_{\cdot\wedge t})\mathrm{d}t+\Sigma(t,\bar X_{\cdot\wedge t})\mathrm{d}\bar W_t, \qquad t\in[0,T],\\ \bar X_0 &= (X_0^{1,N},\dots,X_0^{N,N}). \end{cases} \tag{4.1}$$

Here, $B:[0,T]\times\mathcal{C}([0,T],(\mathbb{R}^d)^N)\to\mathbb{R}^{Nd}$ and $\Sigma:[0,T]\times\mathcal{C}([0,T],(\mathbb{R}^d)^N)\to\mathbb{M}_{Nd\times Nq}(\mathbb{R})$, are defined, for $t\in[0,T]$, for $\bar\alpha=(\alpha^1,\dots,\alpha^N)\in\mathcal{C}([0,T],(\mathbb{R}^d)^N)$, by

$$B(t,\bar\alpha):=\begin{pmatrix} b\big(t,\alpha^1,L_N(\bar\alpha)\big)\\ b\big(t,\alpha^2,L_N(\bar\alpha)\big)\\ \vdots\\ b\big(t,\alpha^N,L_N(\bar\alpha)\big)\end{pmatrix},$$

$$\Sigma(t,\bar{\alpha}) := \begin{pmatrix} \sigma\big(t,\alpha^1,L_N(\bar{\alpha})\big) & & & \\ & \sigma\big(t,\alpha^2,L_N(\bar{\alpha})\big) & & \\ & & \ddots & \\ & & & \sigma\big(t,\alpha^N,L_N(\bar{\alpha})\big) \end{pmatrix}.$$

Let $t \in [0,T]$ and $\bar{\alpha},\bar{\beta} \in \mathcal{C}([0,T],\mathbb{R}^d)^N$. By Assumption (I),

$$\begin{aligned} \big|B(t,\bar{\alpha}) - B(t,\bar{\beta})\big|_{Nd} &\le \sum_{i=1}^N \Big|b\big(t,\alpha^i,L_N(\bar{\alpha})\big) - b\big(t,\beta^i,L_N(\bar{\beta})\big)\Big|_d \\ &\le L\sum_{i=1}^N \Big(\|\alpha^i-\beta^i\|_{\sup} + d_p\big(L_N(\bar{\alpha}),L_N(\bar{\beta})\big)\Big) \\ &\le L\Big(N + N^{1-\frac{1}{p}}\Big)\|\bar{\alpha}-\bar{\beta}\|_{\sup} \end{aligned} \tag{4.2}$$

where we used Lemma 4.1. Similarly,

$$\begin{aligned} \|\!|\Sigma(t,\bar{\alpha}) - \Sigma(t,\bar{\beta})|\!\| &= \sup_{\substack{z=(z^1,\dots,z^N)\in(\mathbb{R}^q)^N \\ |z|_{Nq}\le 1}} \Big|\Big(\Sigma(t,\bar{\alpha}) - \Sigma(t,\bar{\beta})\Big)z\Big|_{Nd} \\ &\le \sup_{\substack{z=(z^1,\dots,z^N)\in(\mathbb{R}^q)^N \\ |z|_{Nq}\le 1}} \sum_{i=1}^N \Big|\Big(\sigma(t,\alpha^i,L_N(\bar{\alpha})) - \sigma(t,\beta^i,L_N(\bar{\beta}))\Big)z^i\Big|_d \\ &\le \sup_{\substack{z=(z^1,\dots,z^N)\in(\mathbb{R}^q)^N \\ \forall 1\le i\le N, |z^i|_q\le 1}} \sum_{i=1}^N \Big|\Big(\sigma(t,\alpha^i,L_N(\bar{\alpha})) - \sigma(t,\beta^i,L_N(\bar{\beta}))\Big)z^i\Big|_d \\ &\le \sum_{i=1}^N \|\!|\sigma(t,\alpha^i,L_N(\bar{\alpha})) - \sigma(t,\beta^i,L_N(\bar{\beta}))|\!\| \end{aligned}$$

and we apply again Assumption (I) and Lemma 4.1 to conclude. It follows that B and Σ satisfy Assumption (I) with ambient dimension Nd. We conclude by using Theorem 1.1 which gives the well-posedness of the system (4.1) (note that there is no dependency in the measure arguments for the coefficients). □

4.2. Propagation of chaos

In this subsection we prove Theorem 1.2.

Let $(Y_t^i)_{t\ge 0}$, $1 \le i \le N$ be N processes solving

$$Y_t^i = X_0^{i,N} + \int_0^t b\big(s,Y^i_{\cdot\wedge s},\mu_{\cdot\wedge s}\big)\,\mathrm{d}s + \int_0^t \sigma\big(s,Y^i_{\cdot\wedge s},\mu_{\cdot\wedge s}\big)\,\mathrm{d}B_s^i, \qquad t\in[0,T], \tag{4.3}$$

where $(\mu_s)_{s\in[0,T]}$ in the coefficient functions are the marginal distributions of the unique solution X of (1.5) given by Theorem 1.1, and where $B^i = (B_t^i)_{t\in[0,T]}$, $1 \le i \le N$ are the same i.i.d. standard Brownian motions considered in the particle system (1.9). Recall that $X_0^{i,N}$, $1 \le i \le N$ are i.i.d. copies of X_0. It follows from the uniqueness in Theorem 1.1 that $Y^i = (Y_t^i)_{t\in[0,T]}$ are i.i.d. copies of X.

With the help of the following lemma, we define, for all $\omega \in \Omega$,

$$\nu^N(\omega) := \frac{1}{N}\sum_{i=1}^N \delta_{Y^i(\omega)}, \tag{4.4}$$

which is a random measure valued in $\mathcal{P}_p(\mathcal{C}([0,T],\mathbb{R}^d))$ for all $p \ge 1$.

Lemma 4.3. *Let $\alpha^i = (\alpha_t^i)_{t\in[0,T]}$, $1 \le i \le N$ be elements of $\mathcal{C}([0,T],\mathbb{R}^d)$. Then*

(1) the empirical measure $\nu^{N,\alpha} := \frac{1}{N}\sum_{i=1}^N \delta_{\alpha^i} \in \mathcal{P}_p(\mathcal{C}[0,T],\mathbb{R}^d))$ *for all* $p \geq 1$*;*
(2) let $\iota : \mathcal{P}_p(\mathcal{C}([0,T],\mathbb{R}^d)) \to \mathcal{C}([0,T],\mathcal{P}_p(\mathbb{R}^d))$ *be the application defined in Lemma 3.2. Then*

$$\iota\big(\nu^{N,\alpha}\big) = \Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}\Big)_{t\in[0,T]}.$$

Proof. (1) For all $p \geq 1$,

$$\begin{aligned}\int_{\mathcal{C}([0,T],\mathbb{R}^d)} \|x\|_{\sup}^p \, \nu^{N,\alpha}(\mathrm{d}x) &= \int_{\mathcal{C}([0,T],\mathbb{R}^d)} \|x\|_{\sup}^p \Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha^i}\Big)(\mathrm{d}x)\\ &= \frac{1}{N}\sum_{i=1}^N \|\alpha^i\|_{\sup}^p < \infty.\end{aligned}$$

(2) Recall the definition of the coordinate map π_t from (3.5). We only need to prove that for a fixed $t \in [0,T]$,

$$\nu^{N,\alpha} \circ \pi_t^{-1} = \frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}. \tag{4.5}$$

Obviously both sides are probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Let $B \in \mathcal{B}(\mathbb{R}^d)$, we have

$$\begin{aligned}\nu^{N,\alpha} \circ \pi_t^{-1}(B) &= \nu^{N,\alpha}\big(\pi_t^{-1}(B)\big)\\ &= \Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha^i}\Big)\Big(\big\{\beta \in \mathcal{C}([0,T],\mathbb{R}^d) : \pi_t(\beta) \in B\big\}\Big)\\ &= \frac{1}{N}\sum_{i=1}^N \delta_{\alpha^i}\Big(\big\{\beta \in \mathcal{C}([0,T],\mathbb{R}^d) : \beta_t \in B\big\}\Big),\end{aligned}$$

where we used that $\pi_t(\beta) = \beta_t$. Notice in addition that

$$\delta_{\alpha^i}\Big(\big\{\beta \in \mathcal{C}([0,T],\mathbb{R}^d) : \beta_t \in B\big\}\Big) = \begin{cases} 1 & \text{if } \alpha_t^i \in B,\\ 0 & \text{otherwise.}\end{cases}$$

On the other hand,

$$\Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}\Big)(B) = \frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}(B)$$

where

$$\delta_{\alpha_t^i}(B) = \begin{cases} 1 & \text{if } \alpha_t^i \in B,\\ 0 & \text{otherwise.}\end{cases}$$

It follows that for all $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\Big(\nu^{N,\alpha} \circ \pi_t^{-1}\Big)(B) = \Big(\frac{1}{N}\sum_{i=1}^N \delta_{\alpha_t^i}\Big)(B),$$

and finally that (4.5) holds. □

With the processes Y^i, $1 \leq i \leq N$ from (4.3) at hand, we introduce a family of random distributions $(\nu_t^N)_{t\in[0,T]}$ defined by

$$\forall\, \omega \in \Omega, \quad t \in [0,T], \qquad \nu_t^N(\omega) := \frac{1}{N}\sum_{i=1}^N \delta_{Y_t^i(\omega)}. \tag{4.6}$$

Lemma 4.1 guarantees that for every ω, $(\nu_t^N(\omega))_{t\in[0,T]} \in \mathcal{C}([0,T],\mathcal{P}_p(\mathbb{R}^d))$ since

$$(\nu_t^N(\omega))_{t\in[0,T]} = L_N\Big(\big(Y^1(\omega),\dots,Y^N(\omega)\big)\Big).$$

Moreover, by Lemma 4.3, for every $\omega \in \Omega$,

$$\Big(\nu_t^N(\omega)\Big)_{t\in[0,T]} = \iota\Big(\nu^N(\omega)\Big)$$

so that $(\nu_t^N(\omega))_{t\in[0,T]}$ can be identified with the marginal distributions of $\nu^N(\omega)$.

Proof of Theorem 1.2. Fix $N > 1$ and $i \in \{1,\dots,N\}$. Let Y^i, $1 \le i \le N$ be the solutions to (4.3) and ν^N the associate empirical measure defined by (4.4), as detailed above. We have, for all $s \in [0,T]$,

$$\begin{aligned} X_s^{i,N} - Y_s^i &= \int_0^s \Big[b\big(u, X_{\cdot\wedge u}^{i,N}, \mu_{\cdot\wedge u}^N\big) - b\big(u, Y_{\cdot\wedge u}^i, \mu_{\cdot\wedge u}\big)\Big]\mathrm{d}u \\ &\quad + \int_0^s \Big[\sigma\big(u, X_{\cdot\wedge u}^{i,N}, \mu_{\cdot\wedge u}^N\big) - \sigma\big(u, Y_{\cdot\wedge u}^i, \mu_{\cdot\wedge u}\big)\Big]\mathrm{d}B_u^i. \end{aligned}$$

We set, for all $t \in [0,T]$, for all $i \in \{1,\dots,N\}$,

$$f_i(t) := \Big\| \sup_{s\in[0,t]} \big|X_s^{i,N} - Y_s^i\big| \Big\|_p .$$

Then

$$\begin{aligned} f_i(t) &= \Big\| \sup_{s\in[0,t]} \big|X_s^{i,N} - Y_s^i\big| \Big\|_p \\ &\le \Big\| \int_0^t \Big|b\big(s, X_{\cdot\wedge s}^{i,N}, \mu_{\cdot\wedge s}^N\big) - b\big(s, Y_{\cdot\wedge s}^i, \mu_{\cdot\wedge s}\big)\Big|\mathrm{d}s \\ &\qquad + \sup_{s\in[0,t]} \Big| \int_0^s \Big[\sigma\big(u, X_{\cdot\wedge u}^{i,N}, \mu_{\cdot\wedge u}^N\big) - \sigma\big(u, Y_{\cdot\wedge u}^i, \mu_{\cdot\wedge u}\big)\Big]\mathrm{d}B_u^i \Big| \Big\|_p \\ &\le \int_0^t \Big\|b\big(s, X_{\cdot\wedge s}^{i,N}, \mu_{\cdot\wedge s}^N\big) - b\big(s, Y_{\cdot\wedge s}^i, \mu_{\cdot\wedge s}\big)\Big\|_p \mathrm{d}s \\ &\qquad + C_{d,p}^{BDG}\Big[\int_0^t \Big\| \Big\|\sigma\big(s, X_{\cdot\wedge s}^{i,N}, \mu_{\cdot\wedge s}^N\big) - \sigma\big(s, Y_{\cdot\wedge s}^i, \mu_{\cdot\wedge s}\big)\Big\|\Big\|_p^2 \mathrm{d}s\Big]^{\frac{1}{2}} \end{aligned} \tag{4.7}$$

using Lemma 3.6. With a computation similar to (3.10), recalling that the marginal distributions $(\mu_t^N)_{t\in[0,T]}$ are themselves random, so that for all $u \in [0,T]$,

$$\Big\|d_p\big((\mu_{v\wedge u}^N)_{v\in[0,T]}, (\mu_{v\wedge u})_{v\in[0,T]}\big)\Big\|_p = \Big\| \sup_{v\in[0,u]} \mathcal{W}_p(\mu_v^N, \mu_v)\Big\|_p ,$$

we find

$$\begin{aligned} &\int_0^t \Big\|b\big(s, X_{\cdot\wedge s}^{i,N}, \mu_{\cdot\wedge s}^N\big) - b\big(s, Y_{\cdot\wedge s}^i, \mu_{\cdot\wedge s}\big)\Big\|_p \mathrm{d}s \\ &\qquad \le L\int_0^t \Big\| \sup_{v\in[0,s]} \big|X_v^{i,N} - Y_v^i\big|\Big\|_p \mathrm{d}s + L\int_0^t \Big\| \sup_{v\in[0,s]} \mathcal{W}_p\big(\mu_v^N, \mu_v\big)\Big\|_p \mathrm{d}s \\ &\qquad \le L\int_0^t f_i(s)\,\mathrm{d}s + L\int_0^t \Big\| \sup_{v\in[0,s]} \mathcal{W}_p\big(\mu_v^N, \mu_v\big)\Big\|_p \mathrm{d}s. \end{aligned} \tag{4.8}$$

Using a derivation similar to (3.11), one also finds

$$\begin{aligned} &C_{d,p}^{BDG}\Big[\int_0^t \Big\|\Big\|\sigma\big(s, X_{\cdot\wedge s}^{i,N}, \mu_{\cdot\wedge s}^N\big) - \sigma\big(s, Y_{\cdot\wedge s}^i, \mu_{\cdot\wedge s}\big)\Big\|\Big\|_p^2 \mathrm{d}s\Big]^{\frac{1}{2}} \\ &\qquad \le \sqrt{2}C_{d,p}^{BDG} L\Big\{\Big[\int_0^t \Big\| \sup_{v\in[0,s]} \big|X_v^{i,N} - Y_v^i\big|\Big\|_p^2 \mathrm{d}s\Big]^{\frac{1}{2}} \\ &\qquad\qquad + \Big[\int_0^t \Big\| \sup_{v\in[0,s]} \mathcal{W}_p\big(\mu_v^N, \mu_v\big)\Big\|_p^2 \mathrm{d}s\Big]^{\frac{1}{2}}\Big\} \\ &\qquad = \sqrt{2}C_{d,p}^{BDG} L\Big\{\Big[\int_0^t f_i(s)^2 \mathrm{d}s\Big]^{\frac{1}{2}} + \Big[\int_0^t \Big\| \sup_{v\in[0,s]} \mathcal{W}_p\big(\mu_v^N, \mu_v\big)\Big\|_p^2 \mathrm{d}s\Big]^{\frac{1}{2}}\Big\}. \end{aligned} \tag{4.9}$$

By using the triangle inequality, for all $s \in [0,T]$, we get first

$$\sup_{v\in[0,s]} \mathcal{W}_p^p(\mu_v^N,\mu_v) \le \Big(\sup_{v\in[0,s]} \mathcal{W}_p(\mu_v^N,\nu_v^N) + \sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N,\mu_v)\Big)^p \le 2^p\Big(\sup_{v\in[0,s]} \mathcal{W}_p^p(\mu_v^N,\nu_v^N) + \sup_{v\in[0,s]} \mathcal{W}_p^p(\nu_v^N,\mu_v)\Big). \tag{4.10}$$

In addition, the empirical measure defined for all $t \in [0,T]$ by $\frac{1}{N}\sum_{i=1}^N \delta_{(X_t^{i,N},Y_t^i)}$ is a random coupling of the empirical measures μ_t^N and ν_t^N. Thus, for all $v \in [0,T]$,

$$\mathcal{W}_p^p(\mu_v^N,\nu_v^N) \le \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p \frac{1}{N}\sum_{i=1}^N \delta_{(X_v^{i,N},Y_v^i)}(\mathrm{d}x,\mathrm{d}y) = \frac{1}{N}\sum_{i=1}^N \big|X_v^{i,N}-Y_v^i\big|^p.$$

Taking the supremum over $[0,s]$ and the expectation, noticing that

$$\sup_{v\in[0,s]} \sum_{i=1}^N |X_v^{i,N}-Y_v^i|^p \le \sum_{i=1}^N \sup_{v\in[0,s]} |X_v^{i,N}-Y_v^i|^p$$

almost surely, we get

$$\mathbb{E}\Big[\sup_{v\in[0,s]} \mathcal{W}_p^p(\mu_v^N,\nu_v^N)\Big] \le \frac{1}{N}\sum_{i=1}^N \mathbb{E}\Big[\sup_{v\in[0,s]} \big|X_v^{i,N}-Y_v^i\big|^p\Big] \le \frac{1}{N}\sum_{i=1}^N f_i(s)^p \le \sup_{1\le i\le N} f_i(s)^p.$$

Taking the expectation in (4.10) and using this last inequality, we find

$$\Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\mu_v^N,\mu_v)\Big\|_p^p \le 2^p\Big(\sup_{1\le i\le N} f_i(s)^p + \Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N,\mu_v)\Big\|_p^p\Big)$$

and since $p \ge 2$, a concave inequality gives,

$$\Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\mu_v^N,\mu_v)\Big\|_p \le 2\Big(\sup_{1\le i\le N} f_i(s) + \Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N,\mu_v)\Big\|_p\Big). \tag{4.11}$$

Bringing together (4.7), (4.8), (4.9) and (4.11), denoting $\bar f(t) := \sup_{1\le i\le N} f_i(t)$ for $t \in [0,T]$, we find, for all $i \in \{1,\dots,N\}$

$$f_i(t) \le L\int_0^t f_i(s)\,\mathrm{d}s + 2L\int_0^t \bar f(s)\,\mathrm{d}s + 2L\int_0^t \Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N,\mu_v)\Big\|_p \mathrm{d}s + \sqrt{2}C_{d,p}^{BDG}L\Big[\int_0^t f_i(s)^2\,\mathrm{d}s\Big]^{\frac12} + 2\sqrt{2}C_{d,p}^{BDG}L\Big[\int_0^t \Big(\bar f(s) + \Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N,\mu_v)\Big\|_p\Big)^2 \mathrm{d}s\Big]^{\frac12}.$$

Using the positivity of the Wasserstein distance and of the $(f_i)_{1\le i\le N}$, we obtain

$$f_i(t) \le 3L\int_0^t \bar f(s)\,\mathrm{d}s + 2L\int_0^t \Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N,\mu_v)\Big\|_p \mathrm{d}s + \big(\sqrt{2}+4\big)C_{d,p}^{BDG}L\Big[\int_0^t \bar f(s)^2\,\mathrm{d}s\Big]^{\frac12} + 4C_{d,p}^{BDG}L\Big[\int_0^t \Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N,\mu_v)\Big\|_p^2 \mathrm{d}s\Big]^{\frac12},$$

and taking the supremum over $\{1,\dots,N\}$ on the left-hand-side, we can apply Lemma 3.9 to get

$$\bar f(t) \le 4Le^{\kappa_0 t}\Big\{\int_0^t \Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N,\mu_v)\Big\|_p \mathrm{d}s + 2C_{d,p}^{BDG}\Big[\int_0^t \Big\| \sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N,\mu_v)\Big\|_p^2 \mathrm{d}s\Big]^{\frac12}\Big\} \tag{4.12}$$

with $\kappa_0 := 6L + ((\sqrt{2}+4)C_{d,p}^{BDG}L)^2 > 0$. Injecting this result into (4.11) and using that for all $s \in [0,T]$, $\sup_{v\in[0,s]} \mathcal{W}_p(\nu_v^N, \mu_v) \le \sup_{v\in[0,T]} \mathcal{W}_p(\nu_v^N, \mu_v)$ almost surely, we get

$$\Big\| \sup_{v\in[0,T]} \mathcal{W}_p\big(\mu_v^N, \mu_v\big)\Big\|_p \le 2\Big(1 + 4Le^{\kappa_0 T}(T + 2C_{d,p}^{BDG}\sqrt{T})\Big)\Big\| \sup_{v\in[0,T]} \mathcal{W}_p\big(\nu_v^N, \mu_v\big)\Big\|_p. \tag{4.13}$$

By Lemma 4.3, $(\nu_v^N)_{v\in[0,T]}$ can be identified with the marginal distributions of ν^N. Moreover, by Lemma 3.2, the map ι is 1-Lipschitz continuous. Thus,

$$\Big\| \sup_{v\in[0,T]} \mathcal{W}_p\big(\nu_v^N, \mu_v\big)\Big\|_p = \Big\| d_p\Big(\iota(\nu^N), \iota(\mu)\Big)\Big\|_p \le \Big\| \mathbb{W}_p(\nu^N, \mu)\Big\|_p.$$

Combining this inequality with (4.13) concludes the proof of (1.10). The limit is obtained by applying the convergence of $\|\mathcal{W}_p(\nu^N,\mu)\|_p$ with ν^N being an empirical measure of i.i.d. processes with distribution μ on the separable metric space $\mathcal{C}([0,T],\mathbb{R}^d)$, see for instance [28, Corollary 2.14].

To prove (1.11), we simply note that for all $k \in \{1,\dots,N\}$

$$\mathbb{E}\Big[\sup_{1\le i\le k}\sup_{t\in[0,T]} \big|X_t^{i,N} - Y_t^i\big|^p\Big] \le \sum_{i=1}^{k} \mathbb{E}\Big[\sup_{s\in[0,T]} \big|X_s^{i,N} - Y_s^i\big|^p\Big] \le k\bar{f}(t)^p \le C_{p,d,T,L}k\Big\|\mathbb{W}_p(\nu^N,\mu)\Big\|_p^p,$$

where we applied (4.12) to obtain the last inequality, with a constant $C_{p,d,T,L} > 0$. We conclude by using again [28, Corollary 2.14]. □

5. Interpolated Euler scheme and associated convergences

This section is devoted to the proof of Theorem 1.5. For this purpose, we first define the associated theoretical continuous Euler scheme $(\widetilde{X}_t^M)_{t\in[0,T]}$ of $(\widetilde{X}_{t_m}^M)_{0\le m\le M}$ from (1.15). We use the same temporal discretization as Definition 1.4: let $M \in \mathbb{N}^*$, $h = \frac{T}{M}$. For every $m = 0,...,M$, we set $t_m = m\,h$. As the size of the discretization parameter M will sometimes play a role, we write $(\widetilde{X}_t^M)_{t\ge0}$ when we wish to emphasize the dependency of the process in M and omit this superscript when it is clear from context.

Definition 5.1. *Given the discretized scheme $(\widetilde{X}_{t_0:t_M}^M)$ with the associated probability distributions $(\widetilde{\mu}_{t_0:t_m}^M)$ from Definition 1.4, using the same notations b_m, σ_m as in (1.16), we define the continuous Euler scheme, $(\widetilde{X}_t^M)_{t\in[0,T]}$ by setting, for all $t \in (t_m, t_{m+1}]$,*

$$\begin{cases} \widetilde{X}_0^M &= X_0, \\ \widetilde{X}_t^M &= \widetilde{X}_{t_m}^M + (t-t_m)\, b_m(t_m, \widetilde{X}_{t_0:t_m}^M, \widetilde{\mu}_{t_0:t_m}^M) + \sigma_m\,(t_m, \widetilde{X}_{t_0:t_m}^M, \widetilde{\mu}_{t_0:t_m}^M)(B_t - B_{t_m}). \end{cases} \tag{5.1}$$

According to the definition of b_m and σ_m in (1.16), the continuous Euler scheme (5.1) writes, for $t \in (t_m, t_{m+1}]$, $m \in \{0,\dots,M-1\}$,

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

In order to compare this with equation (1.5), 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

$$\underline{t} := t_m, \qquad [\underline{t}] := m \qquad \text{if} \quad t \in [t_m, t_{m+1}). \tag{5.2}$$

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

$$\widetilde{X}_t^M = \widetilde{X}_0^M + \int_0^t b\Big(\underline{s}, i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}^M\big), i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}^M\big)\Big)\,\mathrm{d}s + \int_0^t \sigma\Big(\underline{s}, i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}^M\big), i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}^M\big)\Big)\,\mathrm{d}B_s. \tag{5.3}$$

Theorem 1.5 is a direct result of the following proposition.

Proposition 5.2. *Under Assumptions (I) and (II), for $(X_t)_{t\in[0,T]}$ the unique strong solution to (1.5) given by Theorem 1.1, for $(\widetilde{X}_t^M)_{t\in[0,T]}$ the Euler scheme from Definition 5.1 with parameter M large enough, for $h = \frac{T}{M}$, one has,*

$$\left\| \sup_{t\in[0,T]} \left| X_t - \widetilde{X}_t^M \right| \right\|_p \leq \tilde{C}\Big(h^\gamma + \big(h|\ln(h)|\big)^{\frac{1}{2}} \Big), \tag{5.4}$$

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

This section is organized as follows. Subsection 5.1 shows several preliminary results for the interpolator i_m that will be used for the proof of Proposition 5.2. Subsection 5.2 gathers several properties of the process $(\widetilde{X}_t^M)_{t\geq 0}$ from Definition 5.1. Finally, in Subsection 5.3, we prove Proposition 5.2, Theorem 1.5 and Corollary 1.6.

Remark 5.3. 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 (1.5) can then be proved using similar arguments as those developed in this section.

5.1. Preliminary results

We gather several properties that will be used for the proof of Proposition 5.2. 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

$$\forall\, B \in \mathcal{B}\big(\mathbb{R}^d\big), \quad \big(\lambda\mu + (1-\lambda)\nu\big)(B) \coloneqq \lambda\mu(B) + (1-\lambda)\nu(B). \tag{5.5}$$

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

Lemma 5.4. *Let μ, $\nu \in \mathcal{P}_p(\mathbb{R}^d)$ 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

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

(b) Let δ_0 denote the Dirac measure on $0 \in \mathbb{R}^d$. Then

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

Remark that Lemma 5.4 implies that the interpolator i_m defined by (1.13) and (1.14) is well defined.

Proof of Lemma 5.4. 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 $\lambda \in [0,1]$,

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

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

$$\begin{aligned}
&\mathcal{W}_p^p\big(\tau(\lambda_1), \tau(\lambda_2)\big) \\
&\quad\leq \mathbb{E}\left[\left| \mathbb{1}_{\{U\leq\lambda_1\}} X + \mathbb{1}_{\{U>\lambda_1\}} Y - \mathbb{1}_{\{U\leq\lambda_2\}} X - \mathbb{1}_{\{U>\lambda_2\}} Y \right|^p \right] \\
&\quad= \mathbb{E}\left[\left| -\mathbb{1}_{\{\lambda_1<U\leq\lambda_2\}} X + \mathbb{1}_{\{\lambda_1<U\leq\lambda_2\}} Y \right|^p \right] \\
&\quad= \mathbb{E}\left[\mathbb{1}_{\{\lambda_1<U\leq\lambda_2\}} \left| X - Y \right|^p \right] = \mathbb{P}\left(\lambda_1 < U \leq \lambda_2\right) \mathbb{E}\left[\left| X - Y \right|^p \right] \quad (\text{as } U \perp\!\!\!\perp (X,Y)) \\
&\quad= (\lambda_2 - \lambda_1) \mathbb{E}\left[|X - Y|^p \right].
\end{aligned} \tag{5.6}$$

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

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

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 $\lambda\in[0,1]$,

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

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

Lemma 5.5 (Properties of the interpolator i_m)**.** *Let $m\in\{1,...,M\}$.*

(*a*) *For every $x_{0:m}\in(\mathbb{R}^d)^{m+1}$, $\|i_m(x_{0:m})\|_{\sup}=\sup_{0\le k\le m}|x_k|$.*
(*b*) *For every $\mu_{0:m}\in\big(\mathcal{P}_p(\mathbb{R}^d)\big)^{m+1}$, $\sup_{t\in[0,T]}\mathcal{W}_p\big(i_m(\mu_{0:m})_t,\delta_0\big)=\sup_{0\le k\le m}\mathcal{W}_p(\mu_k,\delta_0)$.*

Proof of Lemma 5.5. (*a*) First, it is obvious that $\sup_{0\le k\le m}|x_k|\le\big\|i_m(x_{0:m})\big\|_{\sup}$ by the definition of i_m. For every $k\in\{0,...,m-1\}$, for every $t\in[t_k,t_{k+1}]$, we have

$$\big|i_m(x_{0:m})_t\big|\le|x_k|\vee|x_{k+1}|\le\sup_{0\le k\le m}|x_k|\tag{5.7}$$

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

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

$$\begin{aligned}\sup_{t\in[t_k,t_{k+1}]}\mathcal{W}_p\Big(i_m(\mu_{0:m})_t,\delta_0\Big)&\le\mathcal{W}_p(\mu_k,\delta_0)\vee\mathcal{W}_p(\mu_{k+1},\delta_0)\quad\text{(by Lemma 5.4-(}b\text{))}\\&\le\sup_{0\le k\le m}\mathcal{W}_p(\mu_k,\delta_0)\end{aligned}$$

and

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

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

5.2. Properties of the discretization scheme

We gather here several properties of the process $(\widetilde{X}_t^M)_{t\ge0}$ from Definition 5.1.

Proposition 5.6. *For all $M\in\mathbb{N}^*$, write $(\widetilde{X}_t^M)_{t\in[0,T]}$ for the process from Definition 5.1 with parameter M. Then under Assumptions (I), we have*

(*a*) *For every $M\in\mathbb{N}^*$, $\big\|\sup_{t\in[0,T]}\big|\widetilde{X}_t^M\big|\big\|_p\le\Gamma\big(1+\|X_0\|_p\big)$ with the same constant Γ from Theorem 1.1.*
(*b*) *There exists a constant κ depending on $L,b,\sigma,\|X_0\|_p,p,d,T$ such that for $M\in\mathbb{N}^*$ large enough, there holds*

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

Proposition 5.6 directly implies the following result.

Corollary 5.7. *Under Assumptions (I), we have, for a large enough time discretization number* $M \in \mathbb{N}^*$,

$$\left\| \left\| \widetilde{X}^M - i_M\big(\widetilde{X}^M_{t_0:t_M}\big) \right\|_{\sup} \right\|_p \leq 2\kappa \left(h \left| \ln(h) \right|\right)^{\frac{1}{2}}$$
$$\text{and} \quad d_p\Big((\widetilde{\mu}_t)_{t\in[0,T]}, i_M\big(\widetilde{\mu}_{t_0:t_M}\big)\Big) \leq 3\kappa \left(h \left| \ln(h) \right|\right)^{\frac{1}{2}}. \tag{5.8}$$

Proof of Corollary 5.7. Let M be fixed and large enough. We drop the superscript in $\widetilde{X}^M$ for simplicity. It is obvious that

$$\begin{aligned}
\left\| \widetilde{X} - i_M\big(\widetilde{X}_{t_0:t_M}\big) \right\|_{\sup} &= \sup_{0\leq m\leq M-1} \sup_{t\in[t_m,t_{m+1}]} \left| \widetilde{X}_t - i_M\big(\widetilde{X}_{t_0:t_M}\big)_t \right| \\
&\leq \sup_{0\leq m\leq M-1} \sup_{t\in[t_m,t_{m+1}]} \left[\left| \widetilde{X}_t - \widetilde{X}_{t_m} \right| + \left| i_M\big(\widetilde{X}_{t_0:t_M}\big)_t - \widetilde{X}_{t_m} \right| \right] \\
&\leq \sup_{0\leq m\leq M-1} \sup_{t\in[t_m,t_{m+1}]} \left[\left| \widetilde{X}_t - \widetilde{X}_{t_m} \right| + \left| \widetilde{X}_{t_{m+1}} - \widetilde{X}_{t_m} \right| \right] \\
&\leq 2 \sup_{0\leq m\leq M-1} \sup_{t\in[t_m,t_{m+1}]} \left| \widetilde{X}_t - \widetilde{X}_{t_m} \right|.
\end{aligned} \tag{5.9}$$

Then we conclude by applying Proposition 5.6-(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 $(\widetilde{X}_t)_{t\in[0,T]}$. For every $m \in \{0, ..., M-1\}$ and for every $t \in [t_m, t_{m+1}]$,

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

Then

$$\begin{aligned}
d_p\Big((\widetilde{\mu}_t)_{t\in[0,T]}, i_M\big(\widetilde{\mu}_{t_0:t_M}\big)\Big) &= \sup_{t\in[0,T]} \mathcal{W}_p\Big(\widetilde{\mu}_t, i_M\big(\widetilde{\mu}_{t_0:t_M}\big)_t\Big) \\
&= \sup_{0\leq m\leq M-1} \sup_{t\in[t_m,t_{m+1}]} \mathcal{W}_p\Big(\widetilde{\mu}_t, i_M\big(\widetilde{\mu}_{t_0:t_M}\big)_t\Big) \\
&\leq \sup_{0\leq m\leq M-1} \sup_{t\in[t_m,t_{m+1}]} \left\| \widetilde{X}_t - \mathbb{1}_{\{U_m > \frac{t-t_m}{h}\}} \widetilde{X}_{t_m} - \mathbb{1}_{\{U_m \leq \frac{t-t_m}{h}\}} \widetilde{X}_{t_{m+1}} \right\|_p \\
&\leq \sup_{0\leq m\leq M-1} \sup_{t\in[t_m,t_{m+1}]} \Big(\left\| (\widetilde{X}_t - \widetilde{X}_{t_m}) \mathbb{1}_{\{U_m > \frac{t-t_m}{h}\}} \right\|_p \\
&\qquad + \left\| (\widetilde{X}_t - \widetilde{X}_{t_{m+1}}) \mathbb{1}_{\{U_m \leq \frac{t-t_m}{h}\}} \right\|_p \Big) \\
&\leq 3 \sup_{0\leq m\leq M-1} \sup_{t\in[t_m,t_{m+1}]} \left\| \widetilde{X}_t - \widetilde{X}_{t_m} \right\|_p \leq 3\kappa \left(h \left| \ln(h) \right|\right)^{\frac{1}{2}},
\end{aligned} \tag{5.10}$$

where the last inequality comes from Proposition 5.6-(b). □

Proof of Proposition 5.6. (*a*) **Step 1.** In this first step, we prove that for every fixed $M \in \mathbb{N}^*$

$$\left\| \sup_{0\leq k\leq M} |\widetilde{X}_{t_k}| \right\|_p < +\infty \tag{5.11}$$

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

$$\begin{aligned}
\left\| \sup_{0\leq k\leq l+1} |\widetilde{X}_{t_k}| \right\|_p &\leq \left\| \sup_{0\leq k\leq l} |\widetilde{X}_{t_k}| + \Big(|\widetilde{X}_{t_{l+1}}| - \sup_{0\leq k\leq l} |\widetilde{X}_{t_k}| \Big)_+ \right\|_p \\
&\leq \left\| \sup_{0\leq k\leq l} |\widetilde{X}_{t_k}| \right\|_p + \left\| \Big(|\widetilde{X}_{t_{l+1}}| - \sup_{0\leq k\leq l} |\widetilde{X}_{t_k}| \Big)_+ \right\|_p \\
&\leq \left\| \sup_{0\leq k\leq l} |\widetilde{X}_{t_k}| \right\|_p + \left\| \Big(|\widetilde{X}_{t_{l+1}}| - |\widetilde{X}_{t_l}| \Big)_+ \right\|_p \\
&\leq \left\| \sup_{0\leq k\leq l} |\widetilde{X}_{t_k}| \right\|_p + \left\| \left| |\widetilde{X}_{t_{l+1}}| - |\widetilde{X}_{t_l}| \right| \right\|_p \\
&\leq \left\| \sup_{0\leq k\leq l} |\widetilde{X}_{t_k}| \right\|_p + \left\| \widetilde{X}_{t_{l+1}} - \widetilde{X}_{t_l} \right\|_p.
\end{aligned} \tag{5.12}$$

Moreover,

$$\begin{aligned}\left\|\widetilde{X}_{t_{l+1}}-\widetilde{X}_{t_l}\right\|_p &= \left\|h\, b_l(t_l,\widetilde{X}_{t_0:t_l},\widetilde{\mu}_{t_0:t_l})+\sqrt{h}\,\sigma_l(t_l,\widetilde{X}_{t_0:t_l},\widetilde{\mu}_{t_0:t_l})Z_{l+1}\right\|_p\\ &\le h\left\|b\Big(t_l,i_l\big(\widetilde{X}_{t_0:t_l}\big),i_l\big(\widetilde{\mu}_{t_0:t_l}\big)\Big)\right\|_p+\sqrt{h}\left\|\left\|\sigma\Big(t_l,i_l\big(\widetilde{X}_{t_0:t_l}\big),i_l\big(\widetilde{\mu}_{t_0:t_l}\big)\Big)\right\|\right\|_p\left\|Z_{l+1}\right\|_p\\ &\quad\text{(by the definition of } b_l \text{ and } \sigma_l \text{ in Definition 1.3 and as } Z_{l+1}\perp\!\!\!\perp\sigma(\mathcal{F}_{t_l}))\\ &\le \Big(h+\sqrt{h}C_{p,q}\Big)\left\|C_{b,\sigma,L,T}\Big(1+\big\|i_l\big(\widetilde{X}_{t_0:t_l}\big)\big\|_{\sup}+\sup_{t\in[0,T]}\mathcal{W}_p\big(i_l\big(\widetilde{\mu}_{t_0:t_l}\big)_t,\delta_0\big)\Big)\right\|_p,\end{aligned}$$

where we used Lemma 3.3, where $C_{p,q}=\|Z_{l+1}\|_p<+\infty$ is a constant depending only on p and q, as $Z_{l+1}\sim\mathcal{N}(0,\mathbf{I}_q)$. We now invoke Lemma 5.5 to obtain

$$\begin{aligned}\left\|\widetilde{X}_{t_{l+1}}-\widetilde{X}_{t_l}\right\|_p &\le \Big(h+\sqrt{h}C_{p,q}\Big)\left\|C_{b,\sigma,L,T}\Big(1+\sup_{0\le k\le l}|\widetilde{X}_{t_k}|+\sup_{0\le k\le l}\mathcal{W}_p(\widetilde{\mu}_{t_k},\delta_0)\Big)\right\|_p\\ &\le C_{b,\sigma,L,T}\Big(h+\sqrt{h}C_{p,q}\Big)\Big(1+\Big\|\sup_{0\le k\le l}|\widetilde{X}_{t_k}|\Big\|_p+\sup_{0\le k\le l}\|\widetilde{X}_{t_k}\|_p\Big)\\ &\le C_{b,\sigma,L,T}\Big(h+\sqrt{h}C_{p,q}\Big)\Big(1+2\Big\|\sup_{0\le k\le l}|\widetilde{X}_{t_k}|\Big\|_p\Big)\\ &<+\infty\end{aligned}\tag{5.13}$$

where we used the induction hypothesis to obtain the last inequality. Thus

$$\Big\|\sup_{0\le k\le l+1}|\widetilde{X}_{t_k}|\Big\|_p<+\infty$$

which concludes the proof of (5.11) by induction.

Step 2. We prove that $\left\|\sup_{t\in[0,T]}|\widetilde{X}_t|\right\|_p<+\infty$. First, by (5.3) we get for every $t\in[0,T]$,

$$\begin{aligned}\Big\|\sup_{u\in[0,t]}|\widetilde{X}_u|\Big\|_p &= \Big\|\sup_{u\in[0,t]}\Big|X_0+\int_0^u b\Big(\underline{s},i_{\underline{[s]}}\big(\widetilde{X}_{t_0:t_{\underline{[s]}}}\big),i_{\underline{[s]}}\big(\widetilde{\mu}_{t_0:t_{\underline{[s]}}}\big)\Big)\mathrm{d}s\\ &\qquad+\int_0^u\sigma\Big(\underline{s},i_{\underline{[s]}}\big(\widetilde{X}_{t_0:t_{\underline{[s]}}}\big),i_{\underline{[s]}}\big(\widetilde{\mu}_{t_0:t_{\underline{[s]}}}\big)\Big)\mathrm{d}B_s\Big|\Big\|_p\\ &\le \|X_0\|_p+\Big\|\int_0^t\Big|b\Big(\underline{s},i_{\underline{[s]}}\big(\widetilde{X}_{t_0:t_{\underline{[s]}}}\big),i_{\underline{[s]}}\big(\widetilde{\mu}_{t_0:t_{\underline{[s]}}}\big)\Big)\Big|\mathrm{d}s\Big\|_p\\ &\qquad+\Big\|\sup_{u\in[0,t]}\Big|\int_0^u\sigma\Big(\underline{s},i_{\underline{[s]}}\big(\widetilde{X}_{t_0:t_{\underline{[s]}}}\big),i_{\underline{[s]}}\big(\widetilde{\mu}_{t_0:t_{\underline{[s]}}}\big)\Big)\mathrm{d}B_s\Big|\Big\|_p,\end{aligned}\tag{5.14}$$

where we used Minkowski's inequality to obtain the second inequality. The second term in (5.14) can be upper bounded as follows: using Lemma 3.4,

$$\begin{aligned}&\Big\|\int_0^t\Big|b\Big(\underline{s},i_{\underline{[s]}}\big(\widetilde{X}_{t_0:t_{\underline{[s]}}}\big),i_{\underline{[s]}}\big(\widetilde{\mu}_{t_0:t_{\underline{[s]}}}\big)\Big)\Big|\mathrm{d}s\Big\|_p\\ &\quad\le\int_0^t\Big\|b\Big(\underline{s},i_{\underline{[s]}}\big(\widetilde{X}_{t_0:t_{\underline{[s]}}}\big),i_{\underline{[s]}}\big(\widetilde{\mu}_{t_0:t_{\underline{[s]}}}\big)\Big)\Big\|_p\mathrm{d}s\\ &\quad\le\int_0^t\Big\|C_{b,\sigma,L,T}\Big(1+\big\|i_{\underline{[s]}}\big(\widetilde{X}_{t_0:t_{\underline{[s]}}}\big)\big\|_{\sup}+\sup_{u\in[0,T]}\mathcal{W}_p\big(i_{\underline{[s]}}\big(\widetilde{\mu}_{t_0:t_{\underline{[s]}}}\big)_u,\delta_0\big)\Big)\Big\|_p\mathrm{d}s\\ &\quad=\int_0^t\Big\|C_{b,\sigma,L,T}\Big(1+\sup_{0\le k\le\underline{[s]}}|\widetilde{X}_{t_k}|+\sup_{0\le k\le\underline{[s]}}\mathcal{W}_p(\widetilde{\mu}_{t_k},\delta_0)\Big)\Big\|_p\mathrm{d}s\\ &\quad=\int_0^t C_{b,\sigma,L,T}\Big(1+2\Big\|\sup_{0\le k\le\underline{[s]}}|\widetilde{X}_{t_k}|\Big\|_p\Big)\mathrm{d}s\\ &\quad\le T\,C_{b,\sigma,L,T}+2\,C_{b,\sigma,L,T}\int_0^t\Big\|\sup_{0\le k\le\underline{[s]}}|\widetilde{X}_{t_k}|\Big\|_p\mathrm{d}s\end{aligned}\tag{5.15}$$

which is finite by (5.11). We used Lemma 3.3 to obtain the second inequality, and Lemma 5.5 to deduce the next equality.

Moreover, combining Lemmas 3.6 and 3.3, the third term in (5.14) can be upper bounded as follows

$$
\begin{aligned}
&\Big\| \sup_{u\in[0,t]} \Big| \int_0^u \sigma\Big(\underline{s}, i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big), i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)\Big) \mathrm{d}B_s \Big| \Big\|_p \\
&\quad \le C_{d,p}^{BDG} \Big\{ \int_0^t \Big\| C_{b,\sigma,L,T}\Big(1 + \big\| i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big)\big\|_{\sup} \\
&\qquad\qquad + \sup_{u\in[0,T]} \mathcal{W}_p\big(i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)_u, \delta_0\big)\Big)\Big\|_p^2 \mathrm{d}s \Big\}^{\frac{1}{2}} \\
&\quad = C_{d,p}^{BDG} \Big\{ \int_0^t \Big\| C_{b,\sigma,L,T}\Big(1 + \sup_{0\le k\le[\underline{s}]} |\widetilde{X}_{t_k}| + \sup_{0\le k\le[\underline{s}]} \mathcal{W}_p\big(\widetilde{\mu}_{t_k}, \delta_0\big)\Big)\Big\|_p^2 \mathrm{d}s \Big\}^{\frac{1}{2}} \\
&\quad \le \sqrt{2T}\, C_{d,p}^{BDG}\, C_{b,\sigma,L,T} + 2C_{d,p}^{BDG}\, C_{b,\sigma,L,T} \Big\{ \int_0^t \Big\| \sup_{0\le k\le[\underline{s}]} |\widetilde{X}_{t_k}| \Big\|_p^2 \mathrm{d}s \Big\}^{\frac{1}{2}}
\end{aligned} \tag{5.16}
$$

which is again finite by (5.11). We used again Lemma 5.5 to get the third line. We conclude that $\big\| \sup_{t\in[0,T]} |\widetilde{X}_t| \big\|_p < +\infty$.

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

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

by the definition of $[\underline{s}]$, see (5.2), the inequalities (5.14), (5.15) and (5.16) in the previous step imply that for every $t \in [0,T]$

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

Hence, by applying Lemma 3.9 with $f(t) := \Big\| \sup_{u\in[0,t]} |\widetilde{X}_u| \Big\|_p$, we obtain

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

with the constant $C_{p,d,b,\sigma,L,T} > 0$ defined by (3.22). Then

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

and we conclude by recognizing $\Gamma = C_{p,d,b,\sigma,L,T}\, e^{C_{p,d,b,\sigma,L,T} T}$ from Theorem 1.1.

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

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

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. [13, Theorem 1.1.1] to handle the Brownian component in the last inequality: there exists $M_0 \geq 0$ such that for $M \geq M_0$,

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

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

$$\begin{aligned}
\Big\| \sup_{0\leq m\leq M-1} \big|b_m(t_m, \widetilde{X}_{t_0:t_m}, \widetilde{\mu}_{t_0:t_m})\big| \Big\|_p &= \Big\| \sup_{0\leq m\leq M-1} \big|b\Big(t_m, i_m\big(\widetilde{X}_{t_0:t_m}\big), i_m\big(\widetilde{\mu}_{t_0:t_m}\big)\Big)\big| \Big\|_p \\
&\leq \Big\| \sup_{0\leq m\leq M-1} C_{b,\sigma,L,T}\Big(1 + \big\|i_m\big(\widetilde{X}_{t_0:t_m}\big)\big\|_{\sup} \\
&\qquad + \sup_{u\in[0,T]} \mathcal{W}_p\Big(i_m\big(\widetilde{\mu}_{t_0:t_m}\big)_u, \delta_0\Big)\Big) \Big\|_p \\
&\leq \Big\| \sup_{0\leq m\leq M-1} C_{b,\sigma,L,T}\Big(1 + \sup_{0\leq k\leq m} |\widetilde{X}_k| + \sup_{0\leq k\leq m} \mathcal{W}_p\big(\widetilde{\mu}_{t_k}, \delta_0\big)\Big) \Big\|_p \\
&\leq \Big\| C_{b,\sigma,L,T}\Big(1 + \sup_{0\leq k\leq M} |\widetilde{X}_k| + \sup_{0\leq k\leq M} \mathcal{W}_p\big(\widetilde{\mu}_{t_k}, \delta_0\big)\Big) \Big\|_p \\
&\leq C_{b,\sigma,L,T}\Big(1 + 2\Big\| \sup_{0\leq k\leq M} |\widetilde{X}_k|\Big\|_p\Big) \leq C_{b,\sigma,L,T}\Big(1 + 2\,\Gamma(1 + \|X_0\|_p)\Big) < +\infty,
\end{aligned} \tag{5.17}$$

where we used Lemma 3.3 to obtain the first inequality, and Lemma 5.5 to get the second one. Let $C_\star := C_{b,\sigma,L,T}\Big(1 + 2\,\Gamma(1 + \|X_0\|_p)\Big)$, where we recall that Γ is given in Theorem 1.1. By a similar computation, we obtain

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

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

$$\Big\| \sup_{0\leq m\leq M-1} \sup_{v\in[t_m,t_{m+1}]} \big|\widetilde{X}_v - \widetilde{X}_{t_m}\big| \Big\|_p \leq 3C_\star\big(h|\ln(h)|\big)^{\frac{1}{2}} \tag{5.18}$$

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

5.3. *Proof of Proposition 5.2, Theorem 1.5 and Corollary 1.6*

Before turning to the proof of Proposition 5.2, we briefly prove Theorem 1.5 and Corollary 1.6, which are two easy consequences of Proposition 5.2.

Proof of Theorem 1.5. The proof is straightforward since

$$\Big\| \sup_{0\leq m\leq M} \big|X_{t_m} - \widetilde{X}^M_{t_m}\big| \Big\|_p \leq \Big\| \sup_{t\in[0,T]} \big|X_t - \widetilde{X}^M_t\big| \Big\|_p$$

by the definitions of $\big(\widetilde{X}^M_{t_m}\big)_{0\leq m\leq M}$ and $\big(\widetilde{X}^M_t\big)_{t\in[0,T]}$ in Definition 1.4 and Definition 5.1. □

Proof of Corollary 1.6. Corollary 5.7 implies that

$$\Big\| \big\|\widehat{X} - \widetilde{X}\big\|_{\sup} \Big\|_p \leq 2\kappa\big(h|\ln(h)|\big)^{\frac{1}{2}}.$$

Then the result is a direct application of Theorem 1.5. □

Proof of Proposition 5.2. For every $s \in [0,T]$, we have

$$X_s - \widetilde{X}_s = \int_0^s \Big[b(u, X_{\cdot\wedge u}, \mu_{\cdot\wedge u}) - b\Big(\underline{u}, i_{[\underline{u}]}\big(\widetilde{X}_{t_0:t_{[\underline{u}]}}\big), i_{[\underline{u}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{u}]}}\big)\Big)\Big]\,\mathrm{d}u + \int_0^s \Big[\sigma(u, X_{\cdot\wedge u}, \mu_{\cdot\wedge u}) - \sigma\Big(\underline{u}, i_{[\underline{u}]}\big(\widetilde{X}_{t_0:t_{[\underline{u}]}}\big), i_{[\underline{u}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{u}]}}\big)\Big)\Big]\,\mathrm{d}B_u,$$

and we set

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

It follows from Proposition 5.6-(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, $\widetilde{\mu} \in \mathcal{P}_p\big(\mathcal{C}([0,T],\mathbb{R}^d)\big)$ and $\iota(\widetilde{\mu}) = (\widetilde{\mu}_t)_{t\in[0,T]} \in \mathcal{C}\big([0,T],\mathcal{P}_p(\mathbb{R}^d)\big)$ by applying Lemma 3.2. Hence,

$$\begin{aligned} f(t) &= \Big\| \sup_{s\in[0,t]} \big|X_s - \widetilde{X}_s\big| \Big\|_p \\ &\le \Big\| \int_0^t \Big|b(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - b\Big(\underline{s}, i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big), i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)\Big)\Big|\,\mathrm{d}s \\ &\quad + \sup_{s\in[0,t]} \Big| \int_0^s \Big[\sigma(u, X_{\cdot\wedge u}, \mu_{\cdot\wedge u}) - \sigma\Big(\underline{u}, i_{[\underline{u}]}\big(\widetilde{X}_{t_0:t_{[\underline{u}]}}\big), i_{[\underline{u}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{u}]}}\big)\Big)\Big]\,\mathrm{d}B_u \Big| \Big\|_p \\ &\le \int_0^t \Big\| b(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - b\Big(\underline{s}, i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big), i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)\Big)\Big\|_p \,\mathrm{d}s \\ &\quad + C^{BDG}_{d,p}\Big[\int_0^t \Big\| \big|\sigma(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - \sigma\Big(\underline{u}, i_{[\underline{u}]}\big(\widetilde{X}_{t_0:t_{[\underline{u}]}}\big), i_{[\underline{u}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{u}]}}\big)\Big)\big|^2 \Big\|_p \mathrm{d}s\Big]^{\frac{1}{2}} \end{aligned} \tag{5.19}$$

using Lemma 3.6. The first term in (5.19) can be upper bounded by

$$\begin{aligned} &\int_0^t \Big\| b(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - b\Big(\underline{s}, i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big), i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)\Big)\Big\|_p \,\mathrm{d}s \\ &\le \int_0^t \big\| b(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - b(\underline{s}, X_{\cdot\wedge s}, \mu_{\cdot\wedge s})\big\|_p \,\mathrm{d}s \\ &\quad + \int_0^t \Big\| b(\underline{s}, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - b\Big(\underline{s}, i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big), i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)\Big)\Big\|_p \,\mathrm{d}s. \end{aligned} \tag{5.20}$$

For the first term in (5.20), we use Assumption (II) to obtain

$$\begin{aligned} &\int_0^t \big\| b(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - b(\underline{s}, X_{\cdot\wedge s}, \mu_{\cdot\wedge s})\big\|_p \,\mathrm{d}s \\ &\quad \le \int_0^t L\Big\| 1 + \|X_{\cdot\wedge s}\|_{\sup} + \sup_{u\in[0,T]} W_p(\mu_{u\wedge s}, \delta_0)\Big\|_p |s - \underline{s}|^\gamma \mathrm{d}s \\ &\quad \le \Big(LT + 2LT\Big\| \sup_{t\in[0,T]} |X_t|\Big\|_p\Big) h^\gamma \\ &\quad \le h^\gamma 2LT\Gamma(1 + \|X_0\|_p), \end{aligned} \tag{5.21}$$

where we used (1.8) to obtain the last inequality. For the second term of (5.20), we have

$$\begin{aligned} &\int_0^t \Big\| b(\underline{s}, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - b\Big(\underline{s}, i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big), i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)\Big)\Big\|_p \,\mathrm{d}s \\ &\le \int_0^t \Big\| L\Big[\big\|X_{\cdot\wedge s} - i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big)\big\|_{\sup} + d_p\Big((\mu_{v\wedge s})_{v\in[0,T]}, i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)\Big)\Big]\Big\|_p \mathrm{d}s \\ &\le L\int_0^t \Big\| \big\|X_{\cdot\wedge s} - i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big)\big\|_{\sup} \Big\|_p \,\mathrm{d}s + L\int_0^t d_p\Big((\mu_{v\wedge s})_{v\in[0,T]}, i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)\Big)\mathrm{d}s \\ &\le L\int_0^t \Big\| \big\|X_{\cdot\wedge s} - \widetilde{X}_{\cdot\wedge s}\big\|_{\sup} \Big\|_p \,\mathrm{d}s + L\int_0^t \Big\| \big\|\widetilde{X}_{\cdot\wedge s} - i_{[\underline{s}]}\big(\widetilde{X}_{t_0:t_{[\underline{s}]}}\big)\big\|_{\sup} \Big\|_p \,\mathrm{d}s \\ &\quad + L\int_0^t d_p\Big((\mu_{v\wedge s})_{v\in[0,T]}, (\widetilde{\mu}_{v\wedge s})_{v\in[0,T]}\Big)\mathrm{d}s + L\int_0^t d_p\Big((\widetilde{\mu}_{v\wedge s})_{v\in[0,T]}, i_{[\underline{s}]}\big(\widetilde{\mu}_{t_0:t_{[\underline{s}]}}\big)\Big)\mathrm{d}s \end{aligned}$$

and, using Corollary 5.7,

$$\begin{aligned}
&\int_0^t \Big\| b(\underline{s}, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - b\Big(\underline{s}, i_{\lfloor \underline{s}\rfloor}\big(\widetilde{X}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big), i_{\lfloor \underline{s}\rfloor}\big(\widetilde{\mu}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big)\Big)\Big\|_p \mathrm{d}s \\
&\le L\int_0^t \Big\| \sup_{v\in[0,s]} |X_v - \widetilde{X}_v|\Big\|_p \mathrm{d}s + 2LT\kappa\big(h\big|\ln(h)\big|\big)^{\frac12} \\
&\quad + L\int_0^t \sup_{v\in[0,s]} \mathcal{W}_p(\mu_v, \widetilde{\mu}_v)\mathrm{d}s + 3LT\kappa\big(h\big|\ln(h)\big|\big)^{\frac12} \\
&\le L\int_0^t f(s)\mathrm{d}s + LT\,5\kappa\left(h\big|\ln(h)\big|\right)^{\frac12} + L\int_0^t \sup_{v\in[0,s]} \big\|X_v - \widetilde{X}_v\big\|_p \mathrm{d}s \\
&\le 2L\int_0^t f(s)\mathrm{d}s + 5LT\kappa\big(h\big|\ln(h)\big|\big)^{\frac12}, \qquad (5.22)
\end{aligned}$$

Now we consider the second term of (5.19). It follows by applying Lemma 3.6 and norm inequalities that

$$\begin{aligned}
&C_{d,p}^{BDG}\Big[\int_0^t \Big|\Big|\Big|\sigma(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - \sigma\Big(\underline{s}, i_{\lfloor \underline{s}\rfloor}\big(\widetilde{X}_{t_0:t_{\lfloor u\rfloor}}\big), i_{\lfloor \underline{s}\rfloor}\big(\widetilde{\mu}_{t_0:t_{\lfloor u\rfloor}}\big)\Big)\Big|\Big|\Big|_p^2 \mathrm{d}s\Big]^{\frac12} \\
&\le \sqrt{2}C_{d,p}^{BDG}\Big[\int_0^t \Big|\Big|\Big|\sigma(\underline{s}, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - \sigma\Big(\underline{s}, i_{\lfloor \underline{s}\rfloor}\big(\widetilde{X}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big), i_{\lfloor \underline{s}\rfloor}\big(\widetilde{\mu}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big)\Big)\Big|\Big|\Big|_p^2 \mathrm{d}s\Big]^{\frac12} \\
&\quad + \sqrt{2}C_{d,p}^{BDG}\Big[\int_0^t \Big\|\|\sigma(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - \sigma(\underline{s}, X_{\cdot\wedge s}, \mu_{\cdot\wedge s})\|\Big\|_p^2 \mathrm{d}s\Big]^{\frac12} \qquad (5.23)
\end{aligned}$$

For the first term in (5.23), we use the same argument as the one giving (5.21) to get

$$\Big[\int_0^t \Big\|\|\sigma(s, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - \sigma(\underline{s}, X_{\cdot\wedge s}, \mu_{\cdot\wedge s})\|\Big\|_p^2 \mathrm{d}s\Big]^{\frac12} \le h^\gamma\Big(\sqrt{2T} + 2\sqrt{T}\Gamma_2(1 + \|X_0\|_p)\Big) \qquad (5.24)$$

for some constant $\Gamma_2 > 0$ depending explicitely on κ from (1.8) and the constants of Lemma 3.4 and Assumptions (I) and (II). The second term of (5.23) can be upper bounded as follows

$$\begin{aligned}
&\sqrt{2}C_{d,p}^{BDG}\Big[\int_0^t \Big|\Big|\Big|\sigma(\underline{s}, X_{\cdot\wedge s}, \mu_{\cdot\wedge s}) - \sigma\Big(\underline{s}, i_{\lfloor \underline{s}\rfloor}\big(\widetilde{X}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big), i_{\lfloor \underline{s}\rfloor}\big(\widetilde{\mu}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big)\Big)\Big|\Big|\Big|_p^2 \mathrm{d}s\Big]^{\frac12} \\
&\le \sqrt{2}C_{d,p}^{BDG}\Big[\int_0^t \Big\|L\left[\|X_{\cdot\wedge s} - i_{\lfloor \underline{s}\rfloor}\big(\widetilde{X}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big)\|_{\sup} + d_p\Big((\mu_{v\wedge s})_{v\in[0,T]}, i_{\lfloor \underline{s}\rfloor}\big(\widetilde{\mu}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big)\Big)\right]\Big\|_p^2 \mathrm{d}s\Big]^{\frac12} \\
&\le 2LC_{d,p}^{BDG}\Big[\int_0^t \Big\|\|X_{\cdot\wedge s} - i_{\lfloor \underline{s}\rfloor}\big(\widetilde{X}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big)\|_{\sup}\Big\|_p^2 \mathrm{d}s\Big]^{\frac12} \\
&\qquad + 2LC_{d,p}^{BDG}\Big[\int_0^t d_p\Big((\mu_{v\wedge s})_{v\in[0,T]}, i_{\lfloor \underline{s}\rfloor}\big(\widetilde{\mu}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big)\Big)^2 \mathrm{d}s\Big]^{\frac12} \\
&\le 2\sqrt{2}LC_{d,p}^{BDG}\Big[\int_0^t \Big\| \|X_{\cdot\wedge s} - \widetilde{X}_{\cdot\wedge s}\|_{\sup}\Big\|_p^2 \mathrm{d}s\Big]^{\frac12} \\
&\quad + 2\sqrt{2}LC_{d,p}^{BDG}\Big[\int_0^t \Big\| \|\widetilde{X}_{\cdot\wedge s} - i_{\lfloor \underline{s}\rfloor}\big(\widetilde{X}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big)\|_{\sup}\Big\|_p^2 \mathrm{d}s\Big]^{\frac12} \\
&\quad + 2\sqrt{2}LC_{d,p}^{BDG}\Big[\int_0^t d_p\Big((\mu_{v\wedge s})_{v\in[0,T]}, (\widetilde{\mu}_{v\wedge s})_{v\in[0,T]}\Big)^2 \mathrm{d}s\Big]^{\frac12} \\
&\quad + 2\sqrt{2}LC_{d,p}^{BDG}\Big[\int_0^t d_p\Big((\widetilde{\mu}_{v\wedge s})_{v\in[0,T]}, i_{\lfloor \underline{s}\rfloor}\big(\widetilde{\mu}_{t_0:t_{\lfloor \underline{s}\rfloor}}\big)\Big)^2 \mathrm{d}s\Big]^{\frac12} \\
&\le 4\sqrt{2}LC_{d,p}^{BDG}\Big[\int_0^t f(s)^2\mathrm{d}s\Big]^{\frac12} + 2\sqrt{2}LC_{d,p}^{BDG}\sqrt{T}\,5\kappa\big(h\big|\ln(h)\big|\big)^{\frac12} \qquad (5.25)
\end{aligned}$$

by a similar reasoning as the one leading to (5.22). Bringing those inequalities together, we find

$$\begin{aligned} f(t) \leq & L\, 2h^\gamma\, T\Gamma\left(1+\|X_0\|_p\right) + 2L\int_0^t f(s)\mathrm{d}s + 5LT\kappa\big(h\big|\ln(h)\big|\big)^{\frac{1}{2}} \\ & + h^\gamma\Big(\sqrt{2T} + 2\sqrt{T}\,\Gamma_2(1+\|X_0\|_p)\Big) + 4\sqrt{2}\,L\,C_{d,p}^{BDG}\Big[\int_0^t f(s)^2\mathrm{d}s\Big]^{\frac{1}{2}} \\ & + 10\sqrt{2}\,LC_{d,p}^{BDG}\,\sqrt{T}\,\kappa\big(h\big|\ln(h)\big|\big)^{\frac{1}{2}}. \end{aligned} \tag{5.26}$$

The conclusion follows by applying Lemma 3.9. □

Appendix A: Proof of Section 2

A.1. Proof for Subsection 2.1

We provide in this appendix the proofs of the results from Subsection 2.1.

Proof of Proposition 2.2. We write (2.1) in the form of (1.9) and apply Theorem 1.2. Letting, for all $i \in \{1,\dots,N\}$, $\tilde{V}_i = (V_{1,i}, V_{2,i}, V_{3,i})$, the system (2.1) writes,

$$\mathrm{d}\tilde{V}_i(t) = b\big(t, \tilde{V}_i(\cdot\wedge t), \tilde{\mu}_{\cdot\wedge t}\big)\mathrm{d}t + \sigma(t)d\tilde{W}_t, \qquad t\in[0,T],$$

where, for all $t\in[0,T]$, $\sigma(t) = \mathrm{diag}\big(f_1(t), f_2(t), f_3(t)\big)$, $\tilde{W}_t^i = (W^{1,i}(t), W^{2,i}(t), W^{3,i}(t))$, and, for s in $[0,T]$, $x = (x_t)_{t\in[0,T]} \in \mathcal{C}([0,T],\mathbb{R}^d)$, $(\mu_t)_{t\in[0,T]} \in \mathcal{C}([0,T],\mathcal{P}_p(\mathbb{R}^3))$, $\mu = (\mu^1,\mu^2,\mu^3)$, $b = (b_1,b_2,b_3)$ where for all $j\in\{1,2,3\}$,

$$\begin{aligned} b_j(t,x,\mu) := & -\frac{(x_T)_j}{\tau_j} + \sum_{k=1}^3 D_{j,k}\Big(1+\varepsilon\int_0^t \varphi(x_s)\,\mathrm{d}s\Big)\int_{\mathbb{R}^3} S(x_k)\mu_{t-\triangle}(\mathrm{d}x_1,\mathrm{d}x_2,\mathrm{d}x_3) \\ & + I_j(t). \end{aligned}$$

This definition should be understood in the sense that $\mu_{t-\triangle}$ is replaced by $\delta_{0_{\mathbb{R}^3}}$ when $t\leq\triangle$. Note that, in the sense of Assumption (I), the first term on the right-hand side of the definition of b is clearly Lipschitz. Moreover, writing, for $(\mu_t)_{t\in[0,T]} \in \mathcal{C}([0,T],\mathcal{P}_p(\mathbb{R}^3))$, $(x_t)_{t\in[0,T]}$ in $\mathcal{C}([0,T],\mathbb{R}^d)$,

$$H_k(t,x) := D_{j,k}\Big(1+\varepsilon\int_0^t\varphi(x_s)\mathrm{d}s\Big), \quad L_k(t,\mu) = \int_{\mathbb{R}^3} S(x_k)\mu_{t-\delta}(\mathrm{d}x_1,\mathrm{d}x_2,\mathrm{d}x_3), \tag{A.1}$$

it follows from our assumptions that both H_k and L_k are bounded, and Lipschitz in the sense of Assumption (I). Since the product of bounded Lipschitz functions is Lipschitz, it follows that b satisfies Assumption (I). The fact that σ also checks this hypothesis follows from the assumptions on $(f_j)_{1\leq j\leq 3}$. □

A.2. Proofs for Subsection 2.2

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

Proof of Proposition 2.3. Since the diffusion matrix is the identity, we only need to focus on the drift coefficient. We recall from [39, Proof of Proposition 3.9] that there exists $C_\epsilon > 0$ such that for all $t\in[0,T]$ and $x,y\in\mathbb{R}^2$,

$$|b_0^\epsilon(t,x) - b_0^\epsilon(t,y)| + |K_t^\epsilon(x) - K_t^\epsilon(y)| \leq C_\epsilon|x-y|, \tag{A.2}$$

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

$$b_0^\epsilon(0,x) = 0 \quad \text{and} \quad K_0^\epsilon(x) = 0. \tag{A.3}$$

Using this, one sees easily that b_0^ϵ 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

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

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,

$$\begin{aligned}
&\Big|A\big(t, x_1, (\mu_s)_{s\in[0,T]}\big) - A\big(t, x_2, (\nu_s)_{s\in[0,T]}\big)\Big| \\
&\quad \le \chi \Big| \int_0^t e^{-\lambda(t-s)} \Big[\int_{\mathbb{R}^d} \Big(K^\epsilon_{t-s}(x_1 - y) - K^\epsilon_{t-s}(x_2 - y)\Big)\mu_s(\mathrm{d}y)\Big]\Big|\mathrm{d}s \\
&\qquad + \chi \Big| \int_0^t e^{-\lambda(t-s)} \int_{\mathbb{R}^d} K^\epsilon_{t-s}(x_2 - y)\big(\mu_s - \nu_s\big)(\mathrm{d}y)\Big|\mathrm{d}s \\
&\quad := B_t^1 + B_t^2,
\end{aligned}$$

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

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

For the second term, we use the dual representation of the Wasserstein distance $\mathcal{W}_1$ (see e.g. [19], [40, Remark 6.5]), namely, for every $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$,

$$\mathcal{W}_1(\mu,\nu) = \sup\Big\{ \int_{\mathbb{R}^d} \varphi \,\mathrm{d}\mu - \int_{\mathbb{R}^d} \varphi \,\mathrm{d}\nu \;\Big|\; \varphi : \mathbb{R}^d \to \mathbb{R} \text{ Lipschitz continuous} \\ \text{with Lipschitz constant } [\varphi]_{\mathrm{Lip}} \le 1 \Big\} \tag{A.4}$$

and the fact that for every $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$, $p \ge 1$, $\mathcal{W}_1(\mu,\nu) \le \mathcal{W}_p(\mu,\nu)$. This implies

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

which concludes the proof that the drift coefficient satisfies Assumption (I), and the proposition follows by applying Theorem 1.1. □

Proof of Proposition 2.6. We treat the two terms of the drift separately. Again from [39, Proof of Proposition 3.9], there exists $C_\epsilon > 0$ such that for all $t \in [0,T]$, $x \in \mathbb{R}^d$,

$$\big|K_t^\epsilon(x)\big| \le C_\epsilon, \tag{A.5}$$

where for $t = 0$, we consider the same extension of K_t^ϵ as in (A.3).

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

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

Hence

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

From (A.5), 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 $y \to (x-y)/\sqrt{t}$ 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 $t \ge 0$, $x \in \mathbb{R}^d$,

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

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

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

$$\begin{aligned}
A\big(t, x, (\mu_s)_{0\le s\le T}\big) - A\big(u, x, (\mu_s)_{0\le s\le T}\big) = \chi\Big\{ &\int_u^t e^{-\lambda(t-s)} \int_{\mathbb{R}^d} K^\epsilon_{t-s}(x-y)\mu_s(\mathrm{d}y)\mathrm{d}s \\
&+ \int_0^u \Big[e^{-\lambda(t-s)} \int_{\mathbb{R}^d} K^\epsilon_{t-s}(x-y)\mu_s(\mathrm{d}y) \\
&- e^{-\lambda(u-s)} \int_{\mathbb{R}^d} K^\epsilon_{u-s}(x-y)\mu_s(\mathrm{d}y)\Big]\mathrm{d}s.\Big\}
\end{aligned} \tag{A.6}$$

Using (A.5) leads to

$$\Big| \int_u^t e^{-\lambda(t-s)} \int_{\mathbb{R}^d} K^\epsilon_{t-s}(x-y)\mu_s(\mathrm{d}y)\mathrm{d}s \Big| \leq C_2|t-u|$$

for some constant $C_2 > 0$. For the second term on the right-hand-side of (A.6), we notice that the function defined for $t \geq 0$, $x \in \mathbb{R}^d$

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

is such that

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

so that

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

hence g is Lipschitz in time. We conclude that Assumption (II) holds with $\gamma = 1$, and applying Theorem 1.5, the result follows. □

Acknowledgments

Acknowledgment

A.B. thanks Corentin Bernou for fruitful discussions regarding the neuroscience behind the model in Section 2.1.

Funding

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). Y. L. acknowledges partial funding from Institut National des Sciences Mathématiques et de leurs Interactions (INSMI) through the PEPS JCJC 2023 program and funding from Université Paris-Dauphine.

References

[1] Antonelli, F. and Kohatsu-Higa, A. (2002). Rate of convergence of a particle method to the solution of the McKean-Vlasov equation. *The Annals of Applied Probability* **12** 423–476. MR1910635

[2] Bolley, F. (2008). Separability and completeness for the Wasserstein distance. In *Séminaire de probabilités XLI* 371–377. Springer.

[3] Bossy, M. and Talay, D. (1997). A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.* **66** 157–192. MR1370849

[4] Bouleau, N. (1988). *Processus stochastiques et applications*. Hermann Paris.

[5] Cáceres, M. J., Carrillo, J. A. and Perthame, B. (2011). Analysis of Nonlinear Noisy Integrate& Fire Neuron Models: blow-up and steady states. *The Journal of Mathematical Neuroscience* **1** 7.

[6] Caines, P. E., Huang, M. and Malhamé, R. P. (2006). Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information and Systems* **6** 221–252.

[7] Carmona, R. and Delarue, F. (2018a). *Probabilistic Theory of Mean Field Games with Applications I*. Springer International Publishing.

[8] Carmona, R. and Delarue, F. (2018b). *Probabilistic Theory of Mean Field Games with Applications II*. Springer International Publishing.

[9] Carrillo, J. A., Perthame, B., Salort, D. and Smets, D. (2015). Qualitative properties of solutions for the noisy integrate and fire model in computational neuroscience. *Nonlinearity* **28** 3365–3388.

[10] Chaintron, L. P. and Diez, A. (2021). Propagation of Chaos: A Review of Models, Methods and Applications. Ii. Applications. *arXiv : 2106.14812*.

[11] Chaintron, L. P. and Diez, A. (2022). Propagation of Chaos: A Review of Models, Methods and Applications. I. Models and Methods. *arXiv: 2203.00446*.
[12] Cosso, A., Gozzi, F., Kharroubi, I., Pham, H. and Rosestolato, M. (2020). Optimal control of path-dependent McKean-Vlasov SDEs in infinite dimension. *arXiv: 2012.14772*.
[13] Csörgo, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics*. Elsevier Science & Techn.
[14] Cudmore, R. H. and Turrigiano, G. G. (2004). Long-Term Potentiation of Intrinsic Excitability in LV Visual Cortical Neurons. *Journal of Neurophysiology* **92** 341–348.
[15] Delarue, F., Inglis, J., Rubenthaler, S. and Tanré, E. (2015). Global solvability of a networked integrate-and-fire model of McKean-Vlasov type. *The Annals of Applied Probability* **25**.
[16] Deschle, N., Gossn, J. I., Tewarie, P., Schelter, B. and Daffertshofer, A. (2021). On the Validity of Neural Mass Models. *Frontiers in Computational Neuroscience* **14**.
[17] Djete, M. F., Possamaï, D. and Tan, X. (2022). McKean-Vlasov optimal control: The dynamic programming principle. *The Annals of Probability* **50**.
[18] Dobrushin, R. L. (1970). Prescribing a System of Random Variables by Conditional Distributions. *Theory of Probability & Its Applications* **15** 458–486.
[19] Edwards, D. A. (2011). On the Kantorovich-Rubinstein theorem. *Expo. Math.* **29** 387–398. MR2861765
[20] Faugeras, O., Touboul, J. and Cessac, B. (2009). A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in Computational Neuroscience* **3**.
[21] Fournier, N. and Tomasevic, M. (2022). Particle approximation of the doubly parabolic Keller-Segel equation in the plane. *arXiv preprint : 2211.09627*.
[22] Hegselmann, R. and Krause, U. (2002). Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation. *Journal of Artifical Societies and Social Simulation* **5**.
[23] Hoffmann, M. and Liu, Y. (2023). A statistical approach for simulating the density solution of a McKean-Vlasov equation. *In preparation*.
[24] Horstmann, D. (2003). From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. I. *Jahresber, Deutsch. Math.-Verein* **105** 103-165.
[25] Horstmann, D. (2004). From 1970 until present: The Keller–Segel model in chemotaxis and its conse-quences. II. *Jahresber, Deutsch. Math.-Verein* **106** 51-69.
[26] Jansen, B. H. and Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics* **73** 357–366.
[27] Karatzas, I. and Fernholz, R. (2009). Stochastic Portfolio Theory: an Overview. In *Special Volume: Mathematical Modeling and Numerical Methods in Finance* 89–167. Elsevier.
[28] Lacker, D. (2018a). Mean field games and interacting particle systems. *Preprint*. Available at http://www.columbia.edu/ dl3133/MFGSpring2018.pdf.
[29] Lacker, D. (2018b). On a Strong Form of Propagation of Chaos for McKean-Vlasov Equations. *Electronic Communications in Probability* **23**.
[30] Liu, Y. (2019). Optimal Quantization: Limit Theorems, Clustering and Simulation of the McKean-Vlasov Equation, PhD thesis, Sorbonne Université.
[31] Liu, Y. (2022). Particle Method and Quantization-Based Schemes for the Simulation of the McKean-Vlasov Equation. *arXiv : 2212.14853*.
[32] Liu, Y. and Pagès, G. (2020). Functional convex order for the scaled McKean-Vlasov processes. *arXiv:2005.03154*.
[33] Liu, Y. and Pagès, G. (2022). Monotone convex order for the McKean–Vlasov processes. *Stochastic Processes and their Applications* **152** 312-338.
[34] McKean, H. P. (1967). Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)* 41–57.
[35] Pagès, G. (2018). *Numerical Probability: An Introduction with Applications to Finance*. Springer.
[36] Roth, A. and van Rossum, M. C. W. (2009). Modeling Synapses. In *Computational Modeling Methods for Neuroscientists* 139–160. The MIT Press.
[37] Sznitman, A. S. (1991). Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989. Lecture Notes in Math.* **1464** 165–251. Springer, Berlin. MR1108185

[38] Talay, D. and Tomašević, M. (2020). A new McKean-Vlasov stochastic interpretation of the parabolic–parabolic Keller-Segel model: The one-dimensional case. *Bernoulli* **26**.

[39] Tomašević, M. (2021). A new McKean-Vlasov stochastic interpretation of the parabolic-parabolic Keller-Segel model: The two-dimensional case. *The Annals of Applied Probability* **31**.

[40] Villani, C. (2009). *Optimal transport, Old and new. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Springer-Verlag, Berlin. MR2459454