INTERACTING PARTICLES SYSTEMS WITH DELAY AND RANDOM DELAY DIFFERENTIAL EQUATIONS

JUAN PABLO PINASCO, MAURO RODRIGUEZ CARTABIA, NICOLAS SAINTIER

ABSTRACT. In this work we study a kinetic model of active particles with delayed dynamics, and its limit when the number of particles goes to infinity. This limit turns out to be related to delayed differential equations with random initial conditions. We analyze two different dynamics, one based on the full knowledge of the individual trajectories of each particle, and another one based only on the trace of the particle cloud, loosing track of the individual trajectories. Notice that in the first dynamic the state of a particles is its path, whereas it is simply a point in $\mathbb{R}^d$ in the second case. We analyse in both cases the corresponding mean-field dynamic obtaining an equation for the time evolution of the distribution of the particles states. Well-posedness of the equation is proved by a fixed-point argument. We conclude the paper with some possible future research directions and modelling applications.

1. INTRODUCTION

The study of large systems of interacting particles attracted a great deal of attention in the last years. Among many other topics, we can find a growing literature on flocking, crowds, traffic, opinion dynamics, wealth distribution, and many other natural and social phenomena studied using kinetic equations (see e.g. [2], [21].). Here, we will consider delayed dynamics, where the agent reaction depends on the recent history of other agents behavior, in some sense that will be clarified later.

Let us observe that in many cases a delay makes a priori no sense, since nobody try to convince us of some political position that she/he held on the past, or our reactions in a middle of a flock (a crowd, a traffic jam) depend only on the present state of the world. However, the presence of some communication delay, or of a reaction time between the signal and the response, are reasonable assumption which immediately lead to introducing a delay in the model (see for instance [8], [20]). There are many other contexts in which the recent history is clearly relevant. For instance suppose that many agents are randomly matched in a two player game, each one is using a mixed strategy, and they update their strategies after each pairwise encounter depending only on the last game played, see [24] for an example. On the other hand, using no-regret algorithms [16], the players try to improve their strategies by considering the full history of games played.

1

So, let us suppose we have an interacting population of N agents labeled $i = 1, \ldots, N$ evolving in $\mathbb{R}^d$. We assume that interactions among agents are the only factor of movement. In absence of delay we can model the evolution of each agent by the following coupled system:

$$(1.1) \qquad \frac{d}{dt} x^{(i)}(t) = \frac{1}{N} \sum_{j=1}^{N} K\left(x^{(i)}(t), x^{(j)}(t)\right), \qquad i = 1, \ldots, N,$$

where the function $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ models the interaction between any pair of agents. This is a mean-field framework in the sense that each agent feels the mean influence of the rest of the population and that interactions are only binary and modelled by the same function K regardless of the particular agents involved in the interaction. It is a classical setting relevant in various applications ranging from physics to biology and sociology. In biology for instance we can think of $x^{(i)}$ as the pair (position, velocity) characterizing animal i. The kernel K can model a velocity- alignment mechanism as e.g. in the classical Cucker-Smale flocking model, or synchronization of oscillators as in the Kuramoto model. In sociology, $x^{(i)}$ could be the opinion of individual i, and K then models the variation of opinions due to interactions among individuals (see e.g. [27] or [32] and the references therein). We refer e.g. to the survey [13] for a detailed mathematical introduction to this subject in the absence of delay.

The main purpose of this work is to extend this theory by incorporating a general delay. To take delay into account in the interaction between any two agents, say i and j, we suppose that agent i reacts to the history $\{x_t^{(j)}(s) := x^{(j)}(t - s), s \in [-\tau, 0]\}$ of agent j in the time window $[t - \tau, t]$. Notice that $x_t^{(j)} \in C([-\tau, 0], \mathbb{R}^d)$. In the mean-field framework this leads to replace (1.1) by

$$(1.2) \qquad \frac{d}{dt} x^{(i)}(t) = \frac{1}{N} \sum_{j=1}^{N} K\left(x^{(i)}(t), x_t^{(j)}\right), \qquad i = 1, \ldots, N,$$

where $K : \mathbb{R}^d \times C([-\tau, 0], \mathbb{R}^d) \to \mathbb{R}^d$. Notice that in the delayed setting the natural variable to characterize the state of agent i is not the value $x^{(i)}(t) \in \mathbb{R}^d$ at time t but his history $x_t^{(i)}$ in the time window $[t - \tau, t]$. Thus the natural state space is not $\mathbb{R}^d$ but $E := C([-\tau, 0], \mathbb{R}^d)$. From a modelling point of view this means that

(a) each agents knows exactly the trajectories $x_t^{(j)}$, $t \geq 0$, of any other agents in the population.

It may happen however that some particular choice of kernel K leads to a loss of information in the sense that agent i does not react to the precise trajectories $x_t^{(j)}$ of others agents j in the population but only to the global distribution of agents at each time $t + s$, $s \in [-\tau, 0]$ without caring for the individual trajectories. Consider e.g. the kernel K given by

$$(1.3) \qquad K(x^{(i)}(t), x_t^{(j)}) := \int_{-\tau}^{0} \tilde{K}\left(x^{(i)}(t), x^{(j)}(t + s)\right) \rho(ds),$$
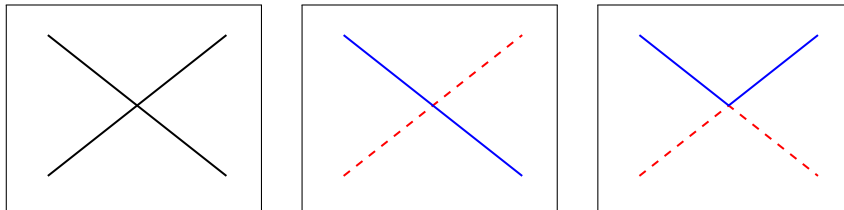
FIGURE 1. In the first model, a player who receives the information of the image at the left panel can not distinguish between two players who crossed (central panel) or two who touched and bounced (right panel).

where $\tilde{K} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, and $\rho(ds)$ is a given probability measure on $[-\tau, 0]$. System (1.2) becomes

$$(1.4) \qquad \frac{d}{dt} x^{(i)}(t) = \int_{-\tau}^{0} \frac{1}{N} \sum_{j=1}^{N} \tilde{K}\left(x^{(i)}(t), x^{(j)}(t+s)\right) \rho(ds), \qquad i = 1, \dots, N.$$

Thus this particular kernel K leads to a situation where information of the individual trajectories is lost so that

    (b) each agent knows if one agent was located at position x at some time $t + s$, without knowing exactly which specific agent was there.

This means that agents have imperfect memory: they know the trace of the paths followed by others agents, but they are not able to identify which agent travel each path. A typical situation depicted in Figure 1 corresponds to considering the path in the time window $[t - \tau, t]$ at time t of two agents $x^{(1)}$ and $x^{(2)}$ namely $x_t^{(1)}, x_t^{(2)} \in E$. For ease of notation we let $\sigma^{(i)} := x_t^{(i)}$, $i = 1, 2$. We suppose that these paths intersect ar time $t - h$ for some $h \in [-\tau, 0]$ i.e. $\sigma^{(1)}(h) = \sigma^{(2)}(h)$, and consider the paths

$$\tilde{\sigma}^{(1)} = \sigma^{(1)} \mathbb{1}_{[-\tau, h]} + \sigma^{(2)} \mathbb{1}_{[h, 0]}, \qquad \tilde{\sigma}^{(2)} = \sigma^{(2)} \mathbb{1}_{[-\tau, h]} + \sigma^{(1)} \mathbb{1}_{[h, 0]}.$$

Then it is easily seen that for any $x \in \mathbb{R}^d$,

$$\int_{-\tau}^{0} K(x, \sigma^{(1)}(s)) + K(x, \sigma^{(2)}(s)) \, d\rho(s) = \int_{-\tau}^{0} K(x, \tilde{\sigma}^{(1)}(s)) + K(x, \tilde{\sigma}^{(2)}(s)) \, d\rho(s).$$

Thus an agent located at x cannot distinguish the precise individual trajectories but only the trails they left without knowing who is where.

    This setting is well known in other problems in partial differential equations. We mention for instance materials with memory as studied by Dafermos [10]. As observed in [11], these kind of problems does not have the property of minimality, since different initial data generate different solutions.

    One of the main concern in dealing with interacting particles system like (1.1) or (1.2) is the determination of its limit as $N \to +\infty$. The treatment of this question in absence of delay, namely for (1.1), is classic (see e.g. [13]) and relies on rewriting system (1.1) as a single equation for the empirical measure $\mu^N(t) := \frac{1}{N} \sum_{i=1}^{N} \delta\{x^{(i)}(t)\}$, where $\delta\{x^{(i)}(t)\}$ is the Dirac mass at $x^{(i)}(t)$. Notice that $\mu^N(t)$

belongs to $\mathbb{P}(\mathbb{R}^d)$, the set of probability measure on $\mathbb{R}^d$. It is indeed easy to see that $\mu^N(t)$ is the unique solution to

(1.5) $$\frac{d}{dt}\mu + div(\mathcal{F}[\mu]\,\mu) = 0,$$

with initial condition $\mu^N(0)$. Here, for a given probability measure μ on $\mathbb{R}^d$, $\mathcal{F}[\mu]$ is the vector field given by $\mathcal{F}[\mu](x) = \int_{\mathbb{R}^d} K(x,y)\mu(dy)$. Well-posedness in $C([0,+\infty), \mathbb{P}(\mathbb{R}^d))$ of (1.5) for a wide class of initial condition is well-known (see e.g. [13] and references therein) assuming that K is bounded Lipschitz. In particular it is known that the solution depends continuously on the initial condition. Thus if $\mu^N(0) \to \mu(0)$ as $N \to +\infty$, we have convergence of the corresponding solutions namely $\mu^N(t) \to \mu(t)$. In that sense we can consider (1.5) as the correct equation to treat the interacting population whether the number of agents is finite or not. We mention that measure-valued solutions are proving increasingly useful to model various phenomena ranging from biology (see e.g. flocking e.g. [7], population dynamic e.g. [1],[6],[15]), game theory (e.g. [3],[25]), sociology (opinion formation process [22],[23], political polarizaition [26],[29]) to traffic (see e.g. [9]).

In this paper we extend this approach to the delayed system (1.2). To do so we notice that the solution $\mu(t)$ of (1.5) can also be thought of as the law of the solution to the ODE $x'(t) = \mathcal{F}[\mu(t)](x(t))$ when the initial condition is chosen randomly according to the probability distribution $\mu(0) \in \mathbb{P}(\mathbb{R}^d)$. In the delayed setting we are thus led to consider delay differential equation with random initial condition. We prove the well-posedness of such equations and argue that they are the correct limit of the system (1.2) as $N \to +\infty$. Notice that now the state of the population at time t is given by a probability measure $\mu(t)$ on E in case (a) or on $\mathbb{R}^d$ in case (b). In any case

(c) at any time $t \geq 0$, there exists a family of probability distributions $\mu(t+s)$, $s \in [-\tau, 0]$, giving the density of agents in the time-window $[t - \tau, t]$ which is common knowledge for all agents.

Remark 1.1. *Let us mention briefly that in* (c) *we are using the notion of common knowledge introduced by Aumann, see for instance* [17], *which implies that everybody knows* $\mu(x, t + s)$, *and also everybody knows that everybody knows that, and so on. So, everybody will use this family of measures in order to predict the evolution of the system. Let us note that if a particular agent knows exactly the position of finitely many agents, this information does not enter in her/his analysis, since it is a set of zero measure respect to* μ.

As an application of our model, we can consider a population of ants moving on some planar domain. Usually, they leave traces of pheromones that serve as a guide to other ants, and the concentration decay with time. This chemical concentration in some ball around an ant can be easily modeled with a kernel supported in this ball, and weighting different the paths at different times, for instance, using

$$\int_{-\tau}^{0} \sum_{j=1}^{N} K\left(x^{(i)}(t), x^{(j)}(t+s)\right) \rho(s)ds$$

with $\rho \to 0$ when $t \to -\tau$, being τ some characteristic time for the duration of the pheromones. Let us mention the pioneer work of Fontelos and Friedman [12], where the behavior of ants was modeled using a system of two partial differential

equations, one governing the distribution of ants, and the other one devoted to the generation and decay of the pheromones, see also [5]. Since new pheromones appear in the path each ant travelled, it is enough to know their paths and weighting them with a function of t in order to model the decay of the chemical trail. We will study this problem in a separate work.

The paper is organized as follow. We first recall some preliminary results about probability measures and differential equations, with and without delay. We then study interacting particles system like (1.2) and its limit as $N \to +\infty$ in the form of delayed differential equations with random initial condition. The main result of this paper is the well-posedness of such equation. We go on examining in details the case of a kernel like (1.3). We conclude by examining some possible direction for future research. The proofs are given at the very end of the paper for an easy flow.

2. Preliminaries

In this section we establish some notations and recall known results that we use throughout this article.

Let us fix some constants $\tau > 0$, the maximum value of the delay. We denote $x_t^{(i)}$ the path in the time-window $[t - \tau, t]$ of a an agent i. It is an element of the space

$$(2.1) \qquad E := C([-\tau, 0], \mathbb{R}^d).$$

with $x_t^{(i)}(s) := x^{(i)}(t - s)$, $s \in [-\tau, 0]$. We will usually denote σ a generic element of E. We endow E with the usual sup-norm $\|\sigma\|_\infty = \max_{s \in [-\tau, 0]} |\sigma(s)|$, $\sigma \in E$.

2.1. Preliminaries on probability measures.
Let (X, d) be a metric space. We denote $\mathbb{P}(X)$ the set of Borel probability measures on (X, d). For instance $\delta\{x\}$ is the Dirac measure at the point $x \in X$ which is defined for any $A \subset X$ Borel by $\delta\{x\}(A) = 1$ if $x \in A$, and $\delta\{x\}(A) = 0$ if $x \notin A$. Notice in particular that

$$(2.2) \qquad \int_X \varphi(y) \, d\delta\{x\}(y) = \varphi(x)$$

for any continuous function $\varphi : X \to \mathbb{R}$.

A measure $\mu \in \mathbb{P}(X)$ has finite first moment if $\int_X d(x, x_0) \, d\mu(x) < \infty$ for some (hence for any) $x_0 \in X$. We denote $\mathbb{P}_1(X)$ the set of such probability measures.

As a matter of notation we will indifferently denote $\int_X f(x) \, d\mu(x)$ or $\int f(x)\mu(dx)$.

Given another measure space Y with some σ-algebra Σ_Y, the push forward of a measure μ on X by a measurable function $f : X \to Y$ is the measure $f\#\mu$ on Y defined by

$$(2.3) \qquad f\#\mu(A) := \mu(f^{-1}(A))$$

for each $A \in \Sigma_Y$. This is equivalent to saying that

$$\int_Y \phi \, d(f\#\mu) = \int_X \phi \circ f \, d\mu$$

for any bounded measurable $\phi : Y \to \mathbb{R}$.

On $\mathbb{P}_1(X)$ we consider the Monge-Kantorovich distance W_1 defined by

$$(2.4) \qquad W_1(\mu, \tilde{\mu}) := \sup \left\{ \int_X \varphi \, d(\mu - \tilde{\mu}) : \varphi : X \to \mathbb{R}, \text{and } Lip(\varphi) \leq 1 \right\}$$

for $\mu, \tilde{\mu} \in \mathbb{P}_1(X)$. Here $Lip(\varphi)$ denotes the Lipschitz constant of φ i. e. the least constant $C > 0$ such that $|\phi(x) - \phi(y)| \leq Cd(x, y)$ for any $x, y \in X$. It is known that if (X, d) is complete (resp. Polish, compact) then so is $(\mathbb{P}_1(X), W_1)$. Moreover given $\mu_k, \mu \in \mathbb{P}_1(X)$, the convergence $W_1(\mu_k, \mu) \overset{k \to +\infty}{\longrightarrow} 0$ is equivalent to the convergence $\int \phi \, d\mu_k \to \int \phi \, d\mu$ for any continuous function $\phi : X \to \mathbb{R}$ with at most linear growth (i. e. $|\phi(x)| \leq C(1 + |x|)$ for some $C > 0$). We refer e. g. to [33] for the proof of these statements and more details on Monge-Kantorovich distance.

When $X = \mathbb{R}^d$, we will need to consider space

$$(2.5) \qquad \mathcal{E} := C\left([-\tau, 0], \mathbb{P}_1(\mathbb{R}^d)\right)$$

that we endow with the sup distance $\mathcal{W}_1$ defined by

$$(2.6) \qquad \mathcal{W}_1(\mu, \nu) := \max_{s \in [-\tau, 0]} W_1\left(\mu(s), \nu(s)\right) \qquad \mu, \nu \in \mathcal{E}.$$

Notice that if (X, d) is complete, so that $(\mathbb{P}_1(\mathbb{R}^d), W_1)$ is complete, then $(\mathcal{E}, \mathcal{W}_1)$ is complete. Eventually given some $\mu \in C\left([-\tau, T], \mathbb{P}_1(\mathbb{R}^d)\right)$ and $t \in [0, T]$ we let $\mu_t \in \mathcal{E}$ be defined as $\mu_t(s) := \mu(t + s)$, $s \in [-\tau, 0]$.

2.2. **Preliminaries on Ordinary Differential Equations.**

Consider a vector field $\mathcal{F} : (x, t) \in \mathbb{R}^d \times [0, T] \to \mathcal{F}(x, t) \in \mathbb{R}^d$ that we suppose continuous in (x, t) and also globally Lipschitz in x uniformly in t. Then it is well-known that for any $x \in \mathbb{R}^d$ and $s \in \mathbb{R}$, the differential equation

$$(2.7) \qquad y'(t) = \mathcal{F}(y(t), t) \qquad t \in \mathbb{R}$$

with initial condition $y(s) = x$, has a unique global solution that we denote $\mathcal{T}(s, t, x)$. We then have the flow property $\mathcal{T}(s, t + t', x) = \mathcal{T}(t, t + t', \mathcal{T}(s, t, x))$. Eventually for any $s, t \in \mathbb{R}$, $\mathcal{T}(s, t, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$ is a C^1-diffeomorphism with inverse $\mathcal{T}(t, s, \cdot)$. For simplicity we will denote $\mathcal{T}(t, \cdot) := \mathcal{T}(0, t, \cdot)$, $t \in \mathbb{R}$.

Take $s = 0$ and suppose now that the initial condition x in (2.7) is chosen at random following some probability measure $\mu(0) \in \mathbb{P}(\mathbb{R}^d)$. Then the solution y to (2.7) is a random vector whose distribution $\mu(t)$ at time t is the probability measure $\mu(t) := \mathcal{T}(t, \cdot) \sharp \mu(0)$. It is well-known that for any $T > 0$, μ is the unique solution in $C([0, T], \mathbb{P}(\mathbb{R}^d))$ to the first order equation

$$(2.8) \qquad \partial_t \mu + \mathrm{div}(\mathcal{F}\mu) = 0$$

in the sense that

$$(2.9) \qquad \int_{\mathbb{R}^d} \phi \, d\mu(t) = \int_{\mathbb{R}^d} \phi \, d\mu(0) + \int_0^t \int_{\mathbb{R}^d} \mathcal{F}(x, t) \nabla \phi(x) \, \mu(s)(dx) ds$$

for any $\phi \in C_c^1(\mathbb{R}^d)$. We refer e.g. to [33]. Moreover, given two initial condition $\mu(0), \tilde{\mu}(0) \in \mathbb{P}_1(\mathbb{R}^d)$, the corresponding solutions $\mu(t), \tilde{\mu}(t)$ to (2.8) satisfy

$$(2.10) \qquad W_1(\mu(t), \tilde{\mu}(t)) \leq e^{Lt} W_1(\mu(0), \tilde{\mu}(0)) \qquad t \in \mathbb{R}.$$

where $L > 0$ is such that $|\mathcal{F}(x, t) - \mathcal{F}(y, t)| \leq L|x - y|$, $x, y \in \mathbb{R}^d$, $t \geq 0$. This follows easily from (2.4) and $\mu(t) := \mathcal{T}(t, \cdot) \sharp \mu(0)$, $\tilde{\mu}(t) := \mathcal{T}(t, \cdot) \sharp \tilde{\mu}(0)$.

To end this section we recall the classical Grönwall's inequality:

Lemma 2.1 (Grönwall's inequality). *Let be $u, a, b : [a, b] \to [0, \infty)$ continuous and such that*

$$u(t) \leq a(t) + b(t) \int_0^t u(h) \, dh \qquad \text{for any } t \geq 0.$$

Then

$$u(t) \leq a(t) + b(t) \int_0^t a(h) e^{\int_h^t b} \, dh \qquad \text{for any } t \geq 0.$$

2.3. **Preliminaries on Delay Differential Equations.** A Delay Differential Equation (DDE) is an equation of the form

$$(2.11) \qquad x'(t) = F(t, x_t)$$

where x takes in values in $\mathbb{R}^p$ for a given p, $x_t \in C([-\tau, 0], \mathbb{R}^p)$ is defined as before by $x_t(s) := x(t+s)$, and $F : (t, \sigma) \in \mathbb{R} \times C([-\tau, 0], \mathbb{R}^p) \to F(t, \sigma) \in \mathbb{R}^p$. Thus the evolution of x at time t depends on the past history of x in $[t - \tau, t]$. For instance if $F(t, \sigma) = \tilde{F}(t, \sigma(0))$ for some $\tilde{F} : \mathbb{R} \times \mathbb{R}^p \to \mathbb{R}$, then the DDE (2.11) becomes the ODE $x'(t) = \tilde{F}(t, x(t))$. In general given a probability measure $\rho \in \mathbb{P}([-\tau, 0])$, we can consider $F(\sigma) = \int_{-\tau}^0 \tilde{F}(s, \sigma(s)) \, d\rho(s)$. In that case the DDE (2.11) becomes

$$x'(t) = \int_{-\tau}^0 \tilde{F}(s, x(t+s)) \, d\rho(s)$$

so that the evolution of x at time t depends on the previous states $x(t+s)$ weighted by ρ. We recover the ODE $x'(t) = \tilde{F}(x(t), t)$ taking $\rho = \delta\{0\}$.

To start the evolution from $t = 0$, we must prescribe the values of $x(s)$, $s \in [-\tau, 0]$. Thus the DDE must be complemented with an initial condition of the form

$$x(s) = \phi(s), \qquad s \in [-\tau, 0]$$

i. e. $x_0 = \phi$ in $[-\tau, 0]$, for some function $\phi \in C([-\tau, 0], \mathbb{R}^N)$.

We refer to the classical book [14] and also to the more recent book [30] for a thorough treatment of DDE. We quote in particular the following existence and uniqueness result from [30], see also section 2.3 in [14]:

Theorem 2.1. *[Thm. 3.7, p.32] Suppose that $F :\in C([-\tau, 0], \mathbb{R}^p) \times \mathbb{R} \to \mathbb{R}^p$ is continuous and satisfies the following local Lipschitz condition: for all a, $b \in \mathbb{R}$ and $M > 0$, there exists $L > 0$ such that:*

$$|F(t, \sigma) - F(t, \tilde{\sigma})| \leq L \|\sigma - \tilde{\sigma}\|$$

for any $t \in [a, b]$ and $\sigma, \tilde{\sigma} \in C([-\tau, 0], \mathbb{R}^p)$ such that $\|\sigma\|, \|\tilde{\sigma}\| \leq M$.

Then for any $M > 0$, there exists $\delta > 0$ depending only on M, such that for any initial condition $\phi \in C([-\tau, 0], \mathbb{R}^p)$ such that $\|\phi\| \leq M$, there exists a unique solution x to the DDE defined up to time δ:

$$\begin{cases} x'(t) & = F(x_t, t) \qquad 0 \leq t < \delta, \\ x_0 & = \phi. \end{cases}$$

Observation 2.1. *It follows from Remark 3.8 in [30] that in the setting of Theorem 2.1, if F satisfies a global Lipschitz condition instead of a local one, that is, if L can be chosen independent of a, b, M, then the solution x exists for all $t \geq 0$.*

We will see in the next section how the study of interacting particles systems with delay naturally leads to consider DDE with random initial condition in the same spirit that (2.8) stems from considering the ODE (2.7) where the initial condition is chosen at random following some probability distribution $\mu(0)$.

3. Interacting particles system with delay.

As explained in the introduction this paper is devoted to the study of a interacting particles system with delay of the form (1.2), namely

$$(3.1) \qquad \frac{d}{dt}x^{(i)}(t) = \frac{1}{N} \sum_{j=1}^{N} K\left(x^{(i)}(t), x_t^{(j)}\right), \qquad i = 1, \dots, N,$$

where the interaction kernel is $K : \mathbb{R}^d \times E \to \mathbb{R}^d$ with $E = C([-\tau, 0], \mathbb{R}^d)$.

We assume that

(H) $K : (x, \sigma) \in \mathbb{R}^d \times E \to K(x, \sigma) \in \mathbb{R}^d$ is globally Lipschitz: there exists $L > 0$ such that for any $x, \tilde{x} \in \mathbb{R}^d$ and any $\sigma, \tilde{\sigma} \in E$,

$$|K(x, \sigma) - K(\tilde{x}, \tilde{\sigma})| \leq L(|x - \tilde{x}| + |\sigma - \tilde{\sigma}|).$$

A direct application of Theorem 2.1 and Remark 2.1 shows that

Theorem 3.1. *If K satisfies asssumption (H) then for any $\sigma_{in}^1, \dots, \sigma_{in}^N \in E$, there exists a unique solution to* (3.1) *with initial condition $x_0^i = \sigma_{in}^i$, $i = 1, \dots, N$.*

Proof. For each $i = 1, \dots, N$, define $F_i : C([-\tau, 0], (\mathbb{R}^d)^N) \to \mathbb{R}^d$ by

$$F_i(\sigma) = \frac{1}{N} \sum_{j=1}^{N} K(\sigma^i(0), \sigma^j), \qquad \sigma = (\sigma^1, \dots, \sigma^N).$$

Then (3.1) is $\frac{d}{dt}\sigma^i(t) = F_i(\sigma_t^1, \dots, \sigma_t^N)$, $i = 1, \dots, N$. It is easily seen using assumption (H) that $F_1, \dots, F_N$ are globally Lipschitz. The result then follows from Theorem 2.1 and Remark 2.1. □

Denote $\mu^N(t)$ the empirical measure associated to $x_t^1, \dots, x_t^N$ namely

$$(3.2) \qquad \mu^N(t) := \frac{1}{N} \sum_{i=1}^{N} \delta\left\{x_t^{(i)}\right\},$$

where $\delta\left\{x_t^{(i)}\right\}$ is the Dirac mass at the point x_t^i of E. Notice that $\mu^N \in C([0, +\infty), \mathbb{P}(E))$. Independently for any $\mu \in C([0, +\infty), \mathbb{P}(E))$, consider the vector fields $\mathcal{F}[\mu] : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ defined by

$$(3.3) \qquad \mathcal{F}[\mu](t, x) = \int_E K(x, \sigma)\, \mu(t)(d\sigma).$$

Then (3.1) can be written as

$$(3.4) \qquad \frac{d}{dt}x^{(i)}(t) = \mathcal{F}[\mu^N](t, x^{(i)}(t)) \qquad i = 1, \dots, N.$$

We can interpret this system of equations thinking that we are solving the equation

$$\frac{d}{dt}x(t) = \mathcal{F}[\mu^N](t, x(t))$$

choosing the inital condition in the set $\{x_0^{(1)}, \dots, x_0^{(N)}\}$ with equiprobability i. e. following $\mu^N(0)$. Then $\mu^N(t)$ is the distribution of x_t. Thus as $N \to +\infty$ the limit $\mu(t) := \lim_{N \to +\infty} \mu^N(t)$, if it exists, should satisfy the equation

$$(3.5) \qquad \frac{d}{dt}X(t) = \mathcal{F}[\mu](t, X(t))$$

where X_t is a random variable with values in E whose distribution is $\mu(t)$, and the initial condition is a given random variable $X_0 : \Omega \to E$ (where Ω is a underlying probability space) with distribution $\mu(0) := \lim_{N \to +\infty} \mu^N(0)$ (if the limit exists). This equation is understood in a path-wise sense: for almost any $\omega \in \Omega$,

$$(3.6) \qquad \begin{aligned} X(t)(\omega) &= X_0(0)(\omega) + \int_0^t \mathcal{F}[\mu](s, X(s)(\omega)) \, ds \qquad t \geq 0, \\ X(t)(\omega) &= X_0(t)(\omega) \qquad t \in [-\tau, 0]. \end{aligned}$$

We are thus led to consider the random delay differential equation

$$(3.7) \qquad \begin{cases} X'(t) &= \mathcal{F}[\mu](t, X(t)) = \int_{\mathcal{C}} K(X(t), \sigma) \, d\mu(t, \sigma), \ \mu(t) = \mathcal{L}(X_t), \\ \mu(0) &= \mathcal{L}(X_0), \end{cases}$$

where $\mathcal{L}(X_t) \in Prob(E)$ denotes the distribution of X_t. We have the following well-posedness result:

Theorem 3.2. *Suppose that $K : \mathbb{R}^d \times E \to \mathbb{R}^d$ satisfies assumption (H) above. For any $R_0 > 0$ and any initial distribution $\mu(0) \in Prob(E)$ supported in the ball $B_0(R_0) \subset E$, there exists a unique solution to (3.7) in a path-wise sense.*

Moreover, given two initial condition $\mu_{in}, \nu_{in} \in \mathbb{P}(E)$ supported in some ball $B_0(R_0)$, the law $\mu(t)$ and $\nu(t)$ of the corresponding solution satisfy

$$(3.8) \qquad W_1(\mu(t), \nu(t)) \leq r(t) W_1(\mu_{in}, \nu_{in})$$

for some explicit continuous function $r : [0, +\infty) \to [0, +\infty)$ such that $r(0) = 1$.

The proof of this result is given in section 6 below. It consists first in rewriting the problem as fixed-point equation for curves in $\mathbb{P}(E)$, and then applying the standard Banach fixed-point Theorem. In particular the law $\mu(t)$ of X_t is shown to be characterized by the fixed-point equation

$$(3.9) \qquad (res[t] \circ ext[\mu]) \sharp \mu_{in} = \mu(t) \qquad t \geq 0,$$

where the restriction operator $res[t]$, $t \geq 0$, is defined by

$$res[t] : \quad \begin{aligned} C([-\tau, +\infty), \mathbb{R}^d) &\to E \\ \sigma &\mapsto \sigma_t, \end{aligned}$$

and the extension operator $ext[\mu]$ is

$$(3.10) \qquad ext[\mu] : \quad \begin{aligned} E &\to C([-\tau, +\infty), \mathbb{R}^d) \\ \sigma &\mapsto \begin{cases} \sigma(t) & \text{if } t \leq 0 \\ \mathcal{T}(t, \sigma(0); \mu) & \text{if } t > 0, \end{cases} \end{aligned}$$

being $\mathcal{T}(s, t, x; \mu)$ the flow of te vector field $(t, x) \to \mathcal{F}[\mu](t, x)$.

Coming back to the system (3.1), suppose that the empirical measure associated to the initial conditions $x_0^{(i)}$, namey $\mu^N(0) = \frac{1}{N} \sum_{i=1}^N \delta\{x_0^{(i)}\} \in Prob(E)$ converges as $N \to +\infty$ to some measure $\mu(0) \in \mathbb{P}(E)$. Denote $\mu(t)$ the solution of the fixed-point equation (3.9) i. e. $\mu(t)$ is the law of the unique solution $X(t)$ to (3.7). Then according to (3.8),

$$W_1(\mu^N(t), \mu(t)) \leq r(t) W_1(\mu^N(0), \mu(0)) \qquad t \geq 0,$$

so that for any $t \geq 0$,

$$\lim_{N \to +\infty} W_1(\mu^N(t), \mu(t)) = 0.$$

We can thus consider that (3.9), or equivalently (3.7), is the correct equation to describe an arbitrary population of agents interacting through binary interaction via the kernel K.

4. Imperfect vs perfect memory

In our general model an agent at some time t updates his trajectory reacting to the past history in $[t - \tau, t]$ of all the others individuals. We can thus consider he/she has a perfect memory of all individual history in the recent past $[t - \tau, t]$. However, it may happen that for some specific interaction kernel K this complete knowledge is not fully used. Consider for instance the kernel

$$(4.1) \qquad K(x, \sigma) = \int_{-\tau}^{0} \tilde{K}(x, \sigma(s)) \, d\rho(s) \qquad x \in \mathbb{R}^d, \, \sigma \in E,$$

for some probability measure $\rho \in \mathbb{P}([-\tau, 0])$ and kernel $\tilde{K} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$. Then the system (3.1) of interacting agents becomes

$$(4.2) \qquad \frac{d}{dt} x^{(i)}(t) = \frac{1}{N} \sum_j \int_{-\tau}^{0} \tilde{K}(x^{(i)}(t), x^{(j)}(t + s)) \, d\rho(s).$$

We assume that

(H') $\tilde{K} : \mathbb{R}^d \times \mathbb{R}^d \to \in \mathbb{R}^d$ is globally Lipschitz: there exists $L > 0$ such that for any $x, \bar{x}, y, \bar{y} \in \mathbb{R}^d$,

$$|K(x, y) - K(\bar{x}, \bar{y})| \leq L(|x - \bar{x}| + |y - \bar{y}|).$$

It follows that K satisfies assumption (H) so that (4.2) has a unique solution for any given initial condition $x_0^{(1)}, \ldots, x_0^{(N)}$.

Introducing

$$\tilde{\mu}^N(t) = \frac{1}{N} \sum_{j=1}^{N} \delta\{x^{(j)}(t)\} \in \mathbb{P}(\mathbb{R}^d), \qquad s \in \mathbb{R},$$

equation (4.2) can be rewritten

$$(4.3) \quad \frac{d}{dt} x^{(i)}(t) = \int_{-\tau}^{0} \left(\int_{\mathbb{R}^d} \tilde{K}(x^{(i)}(t), y) \, \tilde{\mu}^N(t + s)(dy) \right) d\rho(s) \qquad i = 1, \ldots, N.$$

The measure $\tilde{\mu}^N(s)$ gives the distribution at time s of the agents in $\mathbb{R}^d$. Thus the time evolution of each $x^{(i)}(t)$ depends only on this distribution at time $t + s$, $s \in [-\tau, 0]$, and not on the full knowledge of individual trajectories. In that sense we can say that agents have imperfect memory of the past: they only need to know the past distribution of the whole population and not the precise trajectory of each individual. So it seems that in this case the natural state space is $\mathbb{P}(\mathbb{R}^d)$.

For any $\tilde{\mu} \in C([-\tau, +\infty], \mathbb{P}(\mathbb{R}^d))$, let us introduce the vector field

$$(4.4) \qquad \tilde{\mathcal{F}}[\tilde{\mu}](t, x) := \int_{-\tau}^{0} \left(\int_{\mathbb{R}^d} \tilde{K}(x, y) \, \tilde{\mu}(t + s)(dy) \right) d\rho(s) \qquad t \geq 0, \, x \in \mathbb{R}^d.$$

Thus,

$$\frac{d}{dt} x^{(i)}(t) = \tilde{\mathcal{F}}[\tilde{\mu}^N](x^{(i)}(t)) \qquad i = 1, \ldots, N, \quad t \geq 0.$$

Given some test function $\varphi \in C_c^1(\mathbb{R}^d)$, taking the time derivative of $\int \phi \, d\mu^N(t) = \frac{1}{N} \sum_{i=1}^N \varphi(x^{(i)}(t))$, we obtain

$$\frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) \, d\mu^N(x,t) = \frac{1}{N} \sum_{i=1}^N \nabla\varphi(x^{(i)}(t)) \cdot \tilde{\mathcal{F}} \left[\mu^N \right] (x^{(i)}(t))$$

$$= \int_{\mathbb{R}^d} \nabla\varphi(x) \cdot \tilde{\mathcal{F}} \left[\mu^N \right] (x) \, \mu^N(t)(dx).$$

Thus μ^N is a solution, in the sense of definition (2.9), of the equation

$$(4.5) \qquad \frac{d}{dt}\mu + \mathrm{div}\left(\tilde{\mathcal{F}}[\mu] \, \mu \right) = 0,$$

with initial condition $\mu_0 = \tilde{\mu}_0^N$ i. e. $\mu(s) = \tilde{\mu}^N(s)$ for $s \in [-\tau, 0]$.

The next result states this equation is well-posed for arbitrary initial condition:

Theorem 4.1. *For any $R_0 > 0$ and any initial condition $\mu_{in} \in \mathcal{E}$ such that $\mu_{in}(t)$ is suported in the ball $B_0(R_0) \subset \mathbb{R}^d$ for any $t \in [-\tau, 0]$, there exists a unique solution $\mu \in C\left([-\tau, \infty), \mathbb{P}_1\left(\mathbb{R}^d\right) \right)$ to (4.5) with intial condition $\mu_0 = \mu_{in}$. Moreover $\mu(t)$ has compact support for any $t \geq 0$.*

Eventually, the solution depends continuously on the initial data: there exists a continuous increasing function $r : [0, \infty) \to [1, \infty)$ with $r(0) = 1$ such that for any compactly supported initial conditions μ_{in} and ν_{in}, the corresponding solutions $\mu(t)$ and $\nu(t)$ of (4.5) satisfy

$$(4.6) \qquad \mathcal{W}_1\left(\mu_t, \nu_t \right) \leq r(t) \mathcal{W}_1(\mu_{in}, \nu_{in}),$$

where the distance $\mathcal{W}$ is defined in (2.6).

The proof of this result is given in section 7 below.

As before, the estimate (4.6) allows us to consider equation (4.5) as the correct limit as $N \to +\infty$ of the interacting agents system (4.3).

However, system (4.3) is an example of the general system (3.1), we must study how both equations (4.5) and (3.9) are related. For any $s \in [-\tau, 0]$, denote $ev(s)$ the evaluation operator defined by

$$ev[s]: \quad \begin{array}{ccc} E & \to & \mathbb{R}^d \\ \sigma & \mapsto & \sigma(s). \end{array}$$

Notice in particular that for any $\mu(0) \in \mathbb{P}(E)$ and any $s \in [-\tau, 0]$, we have $ev(s)\sharp\mu(0) \in \mathbb{P}(\mathbb{R}^d)$. Moreover if we assume that $\mu(0)$ is supported in some ball $B_0(R_0) \subset E$ then the curve $s \in [-\tau, 0] \to ev(s)\sharp\mu(0) \in \mathbb{P}(\mathbb{R}^d)$ is continuous.

So, we have

Theorem 4.2. *Fix some $R_0 > 0$ and an initial distribution $\mu(0) \in \mathbb{P}(E)$ supported in the ball $B_0(R_0) \subset E$, and denote $\mu(t)$ be the unique solution of the fixed-point equation (3.9) as given by Theorem 3.2. Then $\tilde{\mu}(t) := ev(0)\sharp\mu(t) \in \mathbb{P}(\mathbb{R}^d)$ is the unique solution of (4.5) with initial condition $\tilde{\mu}(s) := ev(s)\sharp\mu(0)$, $s \in [-\tau, 0]$.*

The proof of this result is deferred to section 8 below.

5. Conclusions and future work

In this work we considered an interacting population of agents incorporating a delay in the dynamic. We argued that a useful way to study this dynamic for an arbitrary number of agents is through a delayed differential equation with random initial condition where the vector field depends on the distribution of the solution itself. We showed well-posedness of this equation using a fixed-point argument. Eventually we studied a particular case where information is lost and how it is related to the general case.

To conclude we want to comment briefly on some possible extensions of this work.

First we could add a noise to the random delayed differential equation (3.7), and consider for instance the stochastic delayed differential equation

$$(5.1) \quad \begin{cases} dX(t) & = \mathcal{F}[\mu(t)](t, X(t))dt + \sigma(\mu(t))dW(t), \qquad \mu(t) = \mathcal{L}(X_t) \\ \mu(0) & = \mathcal{L}(X_0) \end{cases}$$

This is a delayed version of the classic McKean-Vlasov process [19]. In the absence of delay, well-posedness results are known under various regularity and growh assumptions of the coefficients (see e.g. the classical [31], the references in [4], and also Chapter 5 in [28] and Chapter 6 in [18] for a very general theory). We also mention the recent preprint [4] where the authors study a McKean stochastic equation, with a *fixed* delay τ, of the form

$$\begin{aligned} dX(t) =& b(X(t), X(t-\tau), \mathcal{L}(X(t), \mathcal{L}(X(t-\tau))))dt \\ &+ \sigma(X(t), X(t-\tau), \mathcal{L}(X(t), \mathcal{L}(X(t-\tau))))dW(t). \end{aligned}$$

Delayed dynamic for interacting population could also be interesting from a modelling point of view. Indeed, an agent updates its position reacting to the past trajectories of other agents. In particular he could react to some geometric features of these trajectories thus allowing to anticipate to some extent the future movements of others agents. This should be interesting in the modelling of crowds or interactions predator-preys.

The remaining sections of the paper are devoted to the proof of Theorem 3.2, Theorem 4.1 and Theorem 4.2.

6. Proof of Theorem 3.2

This section is devoted to the proof of Theorem 3.2 concerning the well-posedness of

$$(6.1) \quad \begin{cases} X'(t) = \mathcal{F}[\mu](t, X(t)) = \displaystyle\int_{\mathcal{C}} K(X(t), \sigma) \, d\mu(t, \sigma), \, \mu(t) = \mathcal{L}(X_t) \qquad t \geq 0, \\ \mu(0) = \mathcal{L}(X_0) \end{cases}$$

where $K : \mathbb{R}^d \times E \to \mathbb{R}^d$, with $E = C([-\tau, 0], \mathbb{R}^d)$, satisfies assumption (H) namely there exists $L > 0$ such that for any $x, \tilde{x} \in \mathbb{R}^d$ and any $\sigma, \tilde{\sigma} \in E$,

$$|K(x, \sigma) - K(\tilde{x}, \tilde{\sigma})| \leq L(|x - \tilde{x}| + |\sigma - \tilde{\sigma}|).$$

The proof consists in two steps. First we want to rewrite (6.1) as a fixed-point problem. We fix some $T > 0$. Notice first that, as a consequence of assumption

(H), for any $\nu \in C([0,T], \mathbb{P}_1(E))$, the vector field $(t,x) \to \mathcal{F}[\nu](t,x)$ is continuous and Lipschitz in $x \in \mathbb{R}^d$ uniformly in $t \in [0,T]$. Indeed

$$|\mathcal{F}[\nu](t,x) - \mathcal{F}[\nu](\tilde{t},\tilde{x})|$$

$$\leq \left| \int K(x,\sigma)\Big(\nu(t)(d\sigma) - \nu(\tilde{t})(d\sigma)\Big) \right| + \int |K(x,\sigma) - K(\tilde{x},\sigma)| \, \nu(\tilde{t})(d\sigma)$$

$$\leq L\Big(W_1(\nu(t), \nu(\tilde{t})) + |x - \tilde{x}| \Big)$$

where we used that the function $\sigma \to K(x,\sigma)$ is L-Lipschitz for any $x \in \mathbb{R}^d$. Let $\mathcal{T}(s,t,x;\nu)$ be the flow of $\mathcal{F}[\nu](t,x)$, i. e.,

$$\begin{cases} \dfrac{d}{dt}\mathcal{T}(s,t,x;\nu) = \mathcal{F}[\nu](\mathcal{T}(s,t,x;\nu),t) & t \in \mathbb{R}, \\ \mathcal{T}(s,s,x;\nu) = x. \end{cases}$$

Then the solution of the DDE

$$\frac{d}{dt}x(t) = \mathcal{F}[\nu](t, x(t))$$

starting from $x_0 \in E$ is

(6.2) $$x(t) = \begin{cases} x_0(t) & t \in [-\tau, 0], \\ \mathcal{T}(0, t, x(0); \nu) & t \geq 0. \end{cases}$$

If the initial condition is a random variable $X_0 : \Omega \to E$, then the unique pathwise solution to

(6.3) $$X'(t) = \mathcal{F}[\nu](t, X(t))$$

is the random variable $X(t)$ defined by

(6.4) $$X(t)(\omega) = \begin{cases} X_0(t)(\omega) & t \in [-\tau, 0], \\ \mathcal{T}(0, t, X_0(0); \nu)(\omega)) & t \geq 0 \end{cases}$$

for any $\omega \in \Omega$.

To simplify notation we introduce the following extension operator

(6.5) $$\begin{aligned} ext[\nu]: \quad E \quad &\to \quad C([-\tau, +\infty), \mathbb{R}^d) \\ \sigma \quad &\mapsto \quad \begin{cases} \sigma(t) & \text{if } t \leq 0 \\ \mathcal{T}(0, t, \sigma(0); \mu) & \text{if } t > 0. \end{cases} \end{aligned}$$

Thus $X(t) = ext[\nu](X_0)$. Snce E is the natural state space for DDE, we also need the restriction operator $res[t] : C([-\tau, T], \mathbb{R}^d) \to E$ defined by

$$\begin{aligned} res[t]: \quad C([-\tau, +\infty), \mathbb{R}^d) \quad &\to \quad E \\ \sigma \quad &\mapsto \quad \sigma_t. \end{aligned}$$

Notice that

$$res[t] \circ ext[\nu] : E \to E$$

and that

$$X_t = (res[t] \circ ext[\nu])(X_0).$$

Denoting $\tilde{\nu}(t)$ the distribution of X_t and $\mu(0)$ that of X_0, it follows that

(6.6) $$\tilde{\nu}_t = (res[t] \circ ext[\nu])\sharp\mu_0.$$

Indeed, for any $\phi : E \to \mathbb{R}$ measurable and bounded, we have

$$
\begin{aligned}
\int_E \phi \, d\tilde{\nu}(t) &= \mathbb{E}[\phi(X_t)] \\
&= \mathbb{E}[\phi((res[t] \circ ext[\nu])(X_0))] \\
&= \int \phi \circ (res[t] \circ ext[\nu]) \, d\mu(0) \\
&= \int \phi \, d((res[t] \circ ext[\nu]) \sharp \mu(0)).
\end{aligned}
$$

In conclusion, for any $\nu \in C([0, +\infty), \mathbb{P}_1(E))$ the solution X_t to (6.3) with an initial condition distributed as μ_0 has distribution $\tilde{\nu}_t$ given by (6.6). We thus deduce that

Step 1. *A stochastic process X_t with distribution μ_t is solution of* (6.1) *path-wise with initial condition μ_0 if and only if μ_t satisfies the fixed-point equation*

$$(6.7) \qquad \mu_t = (res[t] \circ ext[\mu]) \sharp \mu_0 \qquad t \geq 0.$$

We must therefore prove that the fixed-point equation (6.7) has a unique solution. From now on we assume that the distribution μ_0 of the initial condition is supported in the $B_0(R_0) \subset E$. For a time $T > 0$ to be specified later, we consider the map Γ defined by

$$
(6.8) \qquad
\begin{array}{rcl}
\Gamma : C([0, T], \mathbb{P}_1(E)) & \to & C([0, T], \mathbb{P}_1(E)) \\
\mu(\cdot) & \mapsto & (res[\cdot] \circ ext[\mu]) \# \mu_0.
\end{array}
$$

We also consider the set

$$\mathcal{A}_T := \Big\{ \mu \in C([0, T], \mathbb{P}_1(E)) : \mu(0) = \mu_{in}, \text{ and } \operatorname{supp} \mu(t) \subset B_0(2R_0) \, \forall \, t \in [0, T] \Big\}.$$

We endow $\mathbb{P}_1(E)$ with the W_1 distance and $C([0, T], \mathbb{P}_1(E))$ with the associated sup distance $\max_{t \in [0,T]} W_1(\mu(t), \nu(t))$ for $\mu, \nu \in C([0, T], \mathbb{P}_1(E))$. Since $(\mathbb{P}_1(E), W_1)$ is complete, $C([0, T], \mathbb{P}_1(E))$ is also complete. Endowed with the same distance, $\mathcal{A}_T$ is therefore complete as a closed subspace of $C([0, T], \mathbb{P}_1(E))$.

We will apply the classical Banach fixed point to Γ and $\mathcal{A}_T$ for T small enough.

Step 2. *There exists $T > 0$ depending only on L (as defined in assumption (H)) such that $\Gamma(\mathcal{A}_T) \subset \mathcal{A}_T$ and Γ is a contraction in $\mathcal{A}_T$ in the sense that there is a $C \in [0, 1)$ such that*

$$\max_{t \in [0,T]} W_1(\Gamma(\mu)(t), \Gamma(\nu)(t)) \leq C \max_{t \in [0,T]} W_1(\mu(t), \nu(t)) \,.$$

It follows that Γ has a unique fixed point in $\mathcal{A}_T$. We can then repeat the argument in $[T, 2T], [2T, 3T], \ldots$ to obtain the unique μ solving (6.7).

Before beginning the proof of Step 2, we state an easy lemma concerning the flow $\mathcal{T}(s, t, x; \mu)$ of the vector field $(t, x) \to \mathcal{F}[\mu](t, x)$ for $\mu \in \mathcal{A}_T$.

Lemma 6.1. *There exists $C > 0$ depending only on R_0 and K such that for any $x, y \in \mathbb{R}^d$, any $\mu, \nu \in \mathcal{A}_T$, and any $t \in [0, T]$, there holds*

$$(6.9) \qquad |\mathcal{T}(t, x; \mu)| \leq |x| e^{Ct} + (e^{Ct} - 1)(1 + 2R_0),$$

$$(6.10) \qquad |\mathcal{T}(t, x; \mu) - \mathcal{T}(t, y; \mu)| \leq e^{Lt}|x - y|,$$

$$|\mathcal{T}(t, x; \mu) - \mathcal{T}(t, x; \nu)|$$

$$(6.11) \qquad \leq L \int_0^t \left(W_1(\mu(h), \nu(h)) + L \int_0^h W_1(\mu(r), \nu(r)) \, dr \, e^{L(t-h)} \right) dh.$$

Proof. In view of assumption (H) there exists $C > 0$ depending only on K such that $|K(x, \sigma)| \leq C(1 + |x| + |\sigma|)$ for any $x \in \mathbb{R}^d$ and $\sigma \in E$. Since $\mu(t)$ is supported in $B_0(R_0) \subset E$ for any $t \in [0, T]$ we obtain

$$\mathcal{F}[\mu](t, x) \leq C \int_E (1 + |x| + |\sigma|) \, \mu(t)(d\sigma) \leq C(1 + |x| + 2R_0).$$

From this we obtain

$$\begin{aligned}|\mathcal{T}(t, x; \mu)| &\leq |x| + \int_0^t |\mathcal{F}[\mu](\mathcal{T}(h, x; \mu), h)| \, dh \\ &\leq |x| + C(1 + 2R_0)t + C \int_0^t |\mathcal{T}(h, x; \mu)| \, dh.\end{aligned}$$

We deduce (6.9) applying Grönwall's lemma 2.1.

To prove (6.10) we first notice that

$$(6.12) \qquad |\mathcal{F}[\mu](t, x) - \mathcal{F}[\mu](t, y)| \leq \int_E |K(x, \sigma) - K(y, \sigma)| \, \mu(t)(d\sigma) \leq L|x - y|.$$

Thus

$$\begin{aligned}|\mathcal{T}(t, x; \mu) - \mathcal{T}(t, y; \mu)| &\leq |x - y| + \int_0^t |\mathcal{F}[\mu](h, \mathcal{T}(h, x; \mu)) - \mathcal{F}[\mu](h, \mathcal{T}(h, y; \mu))| \, dh \\ &\leq |x - y| + L \int_0^t |\mathcal{T}(h, x; \mu) - \mathcal{T}(h, y; \mu)| \, dh.\end{aligned}$$

We deduce (6.10) applying Grönwall's lemma 2.1.

We eventually prove (6.11). First for any $h \in [0, T]$ and $x \in \mathbb{R}^d$,

$$|\mathcal{F}[\mu](h, x) - \mathcal{F}[\nu](h, x)| \leq \Big| \int_E K(x, \sigma) \, (\mu(h) - \nu(h))(d\sigma) \Big| \leq LW_1(\mu(h), \nu(h))$$

where we used that $K(x, \sigma)$ is L-Lipschitz in σ for any $x \in \mathbb{R}^d$. Thus using (6.12) we obtain

$$\begin{aligned}|\mathcal{F}[\mu](h, x) - \mathcal{F}[\nu](h, y)| &\leq |\mathcal{F}[\mu](h, x) - \mathcal{F}[\mu](h, y)| + |\mathcal{F}[\mu](h, y) - \mathcal{F}[\nu](h, y)| \\ &\leq L|x - y| + LW_1(\mu(h), \nu(h)).\end{aligned}$$

Thus

$$\begin{aligned}|\mathcal{T}(t, x; \mu) - \mathcal{T}(t, x; \nu)| &\leq \int_0^t |\mathcal{F}[\mu](h, \mathcal{T}(h, x; \mu)) - \mathcal{F}[\nu](h, \mathcal{T}(h, x; \nu))| \, dh \\ &\leq L \int_0^t W_1(\mu(h), \nu(h)) \, dh + L \int_0^t |\mathcal{T}(h, x; \mu) - \mathcal{T}(h, x; \nu)| \, dh.\end{aligned}$$

Applying Grönwall's lemma gives (6.11). $\qquad\square$

We are now in position to prove Step 2:

Proof of Step 2. We first prove that $\Gamma(\mathcal{A}_T) \subset \mathcal{A}_T$. Fix some $\mu \in \mathcal{A}_T$. Since μ_{in} is a probability measure it follows from the definition (2.3) of the push-forward that $\Gamma(\mu)(t)$ is a probability measure for all $t \geq 0$. Also because $res[0] \circ ext[\mu]$ is the identity in E we have $\Gamma(\mu)(0) = \mu_{in}$.

For any $\sigma \in E$ such that $\|\sigma\| \leq R_0$ we have by definition of $ext[\mu]$ that for any $t \in [0, T]$,

$$\|res[t] \circ ext[\mu](\sigma)\| \leq \max_{h \in [-\tau, T]} |ext[\mu](\sigma)(h)| \leq \max_{h \in [0, T]} \{|\mathcal{T}_h(\sigma(0); \mu)|, \|\sigma\|\}.$$

Using (6.9) we deduce that

$$\|res[t] \circ ext[\mu](\sigma)\| \le \max \left\{ R_0 e^{CT} + (e^{CT} - 1)(1 + 2R_0), R_0 \right\}$$

which is $\le 2R_0$ for T small enough depending only on C i. e. on K. As a consequence for any $\varphi : E \to \mathbb{R}$ supported outside the ball $B_0(2R_0)$, we have $\varphi(res[t] \circ ext[\mu](\sigma)) = 0$ for any $\|\sigma\| \le R_0$. Since μ_{in} is supported in $B_0(R_0)$ we obtain

$$\int_E \varphi(\sigma) \, d\Gamma(\mu)(\sigma, t) = \int_E \varphi(res[t] \circ ext[\mu](\sigma)) \, d\mu_{in}(\sigma) = 0.$$

Thus, $\operatorname{supp} \Gamma(\mu)(t) \subset B_0(2R_0)$ for any $t \in [0, T]$.

We eventually prove that $\Gamma(\mu)(t)$ is continuous in t for the distance W_1. Since $\operatorname{supp} \Gamma(\mu)(t) \subset B_0(2R_0)$, it is enough to prove that

$$\lim_{s \to t} \int \phi \, d(\Gamma(\mu)(s) - \Gamma(\mu)(t)) = 0$$

for any $\phi : E \to \mathcal{E}$ bounded continuous. Fix such a ϕ. Then

$$\int_E \phi \, d(\Gamma(\mu)(s) - \Gamma(\mu)(t)) = \int_E \phi(res[s] \circ ext[\mu](\sigma)) - \phi(res[t] \circ ext[\mu](\sigma)) \, d\mu(0)(\sigma).$$

We can pass to the limit $s \to$ by applying the Dominated Convergence Theorem using that $res[s] \circ ext[\mu](\sigma)$ is continuous in t.

We now verify that Γ is a strict contraction for T small enough. Let $\tau_0 = \min\{\tau, t\}$. For any $\mu, \nu \in \mathcal{A}_T$ we have using (6.11) that

$$
\begin{aligned}
&\|res[t] \circ ext[\mu](\sigma) - res[t] \circ ext[\nu](\sigma)\| \\
&= \max_{s \in [-\tau_0, 0]} |\mathcal{T}_{t+s}(\sigma(0); \mu) - \mathcal{T}_{t+s}(\sigma(0); \nu)| \\
&\le \max_{s \in [-\tau_0, 0]} \left| L \int_0^{t+s} W_1(\mu(h), \nu(h)) \, dh \right. \\
&\qquad\qquad \left. + L^2 \int_0^{t+s} \int_0^h W_1(\mu(r), \nu(r)) \, dr \, e^{L(t+s-h)} \, dh \right| \\
&\le L \int_0^t W_1(\mu(h), \nu(h)) \, dh + L^2 \int_0^t \int_0^h W_1(\mu(r), \nu(r)) \, dr \, e^{L(t-h)} \, dh.
\end{aligned}
$$
(6.13)

Then, for every $\varphi : E \to \mathbb{R}^d$ 1-Lipschitz and any $t \in [0, T]$, we have using (6.13) that

$$
\begin{aligned}
\left| \int_E \varphi(\sigma) \, d(\Gamma(\mu)(t) - \Gamma(\nu)(t))(\sigma) \right| \\
= \int_E |\varphi(res[t] \circ ext[\mu](\sigma)) - \varphi(res[t] \circ ext[\nu](\sigma))| \, d\mu_{in}(\sigma) \\
= \int_E \|(res[t] \circ ext[\mu](\sigma)) - (res[t] \circ ext[\nu](\sigma))\| \, d\mu_{in}(\sigma) \\
\le \max\{L, L^2\} C_T \max_{h \in [0, T]} W_1(\mu(h), \nu(h)),
\end{aligned}
$$

where $C_T \to 0$ as $T \to 0$. Taking the supremum over all such φ we deduce

$$\max_{t \in [0, T]} W_1(\Gamma(\mu)(t), \Gamma(\nu)(t)) \le \max\{L, L^2\} C_T \max_{h \in [0, T]} W_1(\mu(h), \nu(h)).$$

We thus take T small enough so that $\max\{L, L^2\} C_T < 1$. $\square$

We eventually prove that the unique fixed point of Γ depends continuously on the initial condition.

Step 3. *Given two initial condition* $\mu_{ino}, \nu_{inp} \in \mathbb{P}(E)$ *supported in some ball* $B_0(R_0)$, *the corresponding fixed-point* μ *and* ν *solution for* $t \geq 0$ *of*

$$\mu_t = (res[t] \circ ext[\mu])\sharp\mu_{inp} \qquad and \qquad \nu_t = (res[t] \circ ext[\nu])\sharp\nu_{inp}$$

satisfy

(6.14) $$W_1(\mu(t), \nu(t)) \leq r(t)W_1(\mu_{in}, \nu_{in}).$$

for some continuous increasing function $r : [0, +\infty) \to [1, \infty)$ *such that* $r(0) = 1$.

Proof. Observe first that for any $\mu \in \mathcal{A}_T$ and any $\sigma, \tilde{\sigma} \in E$, we have using (6.10) that

$$\|res[t] \circ ext[\mu](\sigma) - res[t] \circ ext[\mu](\tilde{\sigma})\|$$

$$\leq \max\left\{ \max_{h \in [0,t]} |\mathcal{T}_h(\sigma(0); \mu) - T_h(\tilde{\sigma}(0); \mu)|, \|\sigma - \tilde{\sigma}\| \right\}$$

$$\leq \max\left\{ \max_{h \in [0,t]} e^{Lh} |\sigma(0) - \tilde{\sigma}(0)|, \|\sigma - \tilde{\sigma}\| \right\}$$

$$\leq e^{Lt} \|\sigma - \tilde{\sigma}\|$$

Therefore, $(res[t] \circ ext[\mu])$ is e^{Lt}-Lipschitz.

Then, for every $\varphi : \mathcal{C} \to \mathbb{R}^d$ 1-Lipschitz we have using (6.13) that

$$\left| \int_E \varphi(\sigma) \, d(\mu(t) - \nu(t))(\sigma) \right|$$

$$= \left| \int_E \varphi(res[t] \circ ext[\mu](\sigma)) \, d\mu_{in}(\sigma) - \int_E \varphi(res[t] \circ ext[\nu](\sigma)) \, d\nu_{in}(\sigma) \right|$$

$$\leq \left| \int_E \varphi(res[t] \circ ext[\mu](\sigma)) \, d(\mu_{in} - \nu_{in})(\sigma) \right|$$

$$\qquad + \int_E |\varphi(res[t] \circ ext[\mu](\sigma)) - \varphi(res[t] \circ ext[\nu](\sigma))| \, d\nu_{in}(\sigma).$$

Since $(res[t] \circ ext[\mu])$ is e^{Lt}-Lipschitz, the first term is bounded by $e^{Lt} W_1(\mu_{in}, \nu_{in})$. We bound the 2nd term using (6.13) by

$$\int_E |(res[t] \circ ext[\mu](\sigma)) - (res[t] \circ ext[\nu](\sigma))| \, d\nu_{in}(\sigma)$$

$$\leq L \int_0^t W_1(\mu(h), \nu(h)) \, dh + L^2 \int_0^t \int_0^h W_1(\mu(r), \nu(r)) \, dr \, e^{L(t-h)} \, dh$$

$$\leq L \int_0^t W_1(\mu(r), \nu(r)) \, dr \left(1 + L \int_0^t e^{L(t-h)} \, dh \right)$$

We thus obtain

$$\left| \int_E \varphi(\sigma) \, d(\mu(t) - \nu(t))(\sigma) \right| \leq e^{Lt} W_1(\mu_{in}, \nu_{in}) + Le^{Lt} \int_0^t W_1(\mu(r), \nu(r)) \, dr.$$

Taking the supremum over all such ϕ gives

$$W_1(\mu(t), \nu(t)) \leq e^{Lt} W_1(\mu_{in}, \nu_{in}) + Le^{Lt} \int_0^t W_1(\mu(r), \nu(r)) \, dr.$$

We deduce (6.14) applying Gronwall's Lemma 2.1. $\qquad\square$

7. Proof of Theorem 4.1.

The proof of Theorem 4.1 is very similar to the proof of Theorem 3.2 so we will only sketch it.

We fix $R_0 > 0$ and an initial condition $\mu_{in} \in C([-\tau, 0], \mathbb{P}(\mathbb{R}^d))$ such that $\mu_{in}(s)(B_0(R_0)) = 1$ for any $s \in [-\tau, 0]$.

Recall that for any $\mu \in C([-\tau, T], \mathbb{P}(\mathbb{R}^d))$, the unique solution in $C([-\tau, T], \mathbb{P}(\mathbb{R}^d))$ of

$$\partial_t \nu + \mathrm{div}(\tilde{\mathcal{F}}[\mu]\nu) = 0, \qquad t \geq 0$$

with initial condition $\nu_0 = \mu_{in}$ and where $\tilde{F}[\mu]$ is defined in (4.4), is given by $\nu = \Gamma\mu$ where $\Gamma : C([-\tau, T], \mathbb{P}_1(\mathbb{R}^d)) \to C([-\tau, T], \mathbb{P}_1(\mathbb{R}^d))$ is defined by

$$\Gamma[\mu](t) = \begin{cases} \mu_{in}(t) & \text{if } t \leq 0, \\ \tilde{\mathcal{T}}(t, \cdot; \mu) \# \mu_{in}(0) & \text{if } t \geq 0. \end{cases}$$

Here $\tilde{\mathcal{T}}(t, x; \mu)$ is the flow of $\tilde{F}[\mu]$. We must therefore prove that Γ has a unique fixed point.

We look for such a fixed point in the set $\mathcal{A}_T$ defined by

$$\mathcal{A}_T = \{\mu \in C([-\tau, T], \mathbb{P}(\mathbb{R}^d)) : \mu_0 = \mu_{in}, \mathrm{supp}\,\mu(t) \subset B_0(2R_0), \forall\, t \in [0, T]\}.$$

We endow $C([-\tau, T], \mathbb{P}(\mathbb{R}^d))$ with the distance $\max_{s \in [-\tau, T]} W_1(\mu(s), \nu(s))$ that makes it complete. With the induced distance, $\mathcal{A}_T$ is closed and thus complete.

Because we assumed that $\tilde{K}$ satisfies assumption (H'), properties (6.9) and (6.10) stated in Lemma 6.1 hold with $\tilde{T}(t, x; \mu)$ in place of $T(t, x; \mu)$. Concerning (6.11), it must be replaced by

$$(7.1) \qquad |\tilde{\mathcal{T}}(t, x; \mu) - \tilde{\mathcal{T}}(t, x; \nu)| \leq Le^{Lt} \int_0^t \mathcal{W}_1(\mu_s, \nu_s)\, ds$$

for any $\mu, \nu \in \mathcal{A}_T$, $t \geq 0$, and $x \in \mathbb{R}^d$, and where $\mathcal{W}_1(\mu_s, \nu_s) = \max_{h \in [-\tau, 0]} W_1(\mu(s + h), \nu(s + h))$. Indeed,

$$|\tilde{\mathcal{F}}[\mu](t, x) - \tilde{\mathcal{F}}[\nu](t, \bar{x})| \leq |\tilde{\mathcal{F}}[\mu](t, x) - \tilde{\mathcal{F}}[\mu](t, \bar{x})| + |\tilde{\mathcal{F}}[\mu](t, \bar{x}) - \tilde{\mathcal{F}}[\nu](t, \bar{x})|$$

$$\leq L|x - \bar{x}|$$

$$+ \int_{-\tau}^0 \int_{\mathbb{R}^d} \tilde{K}(\bar{x}, y)(\mu(t + s)(dy) - \nu(t + s)(dy))\, d\rho(s).$$

Since $K(\bar{x}, .)$ is L-Lipschitz, the second term can be bounded by

$$L \int_{-\tau}^0 W_1(\mu(t + s)(dy), \nu(t + s)(dy))\, d\rho(s) \leq L\mathcal{W}_1(\mu_t, \nu_t).$$

Thus

$$|\tilde{\mathcal{F}}[\mu](t, x) - \tilde{\mathcal{F}}[\nu](t, \bar{x})| \leq L|x - \bar{x}| + L\mathcal{W}_1(\mu_t, \nu_t).$$

We then obtain

$$|\tilde{\mathcal{T}}(t, x; \mu) - \tilde{\mathcal{T}}(t, x; \nu)| \leq \int_0^t |\tilde{\mathcal{F}}[\mu](s, \tilde{\mathcal{T}}(s, x; \mu)) - \tilde{\mathcal{F}}[\nu](s, \tilde{\mathcal{T}}(t, x; \nu))|\, ds$$

$$\leq L \int_0^t \mathcal{W}_1(\mu_s, \nu_s)\, ds + L \int_0^t |\tilde{\mathcal{T}}(s, x; \mu) - \tilde{\mathcal{T}}(s, x; \nu)|\, ds.$$

Gronwall's inequality then gives (7.1).

We can then prove, as in the proof of Theorem 3.2, that for a small enough T depending only on L (defined in assumption (H')), we have $\Gamma(\mathcal{A}_T) \subset \mathcal{A}_T$ and Γ a strict contraction, i. e., there is a constant $C \in [0,1)$ such that

$$\max_{t\in[0,T]} W_1\left(\Gamma(\mu)(t), \Gamma(\nu)(t)\right) \leq C \max_{t\in[0,T]} W_1(\mu(t), \nu(t)) \qquad \mu, \nu \in \mathcal{A}_T.$$

We deduce the existence of a unique fixed point of Γ in $\mathcal{A}_T$. Iterating this argument gives the existence and uniqueness statement of Theorem 4.1.

To finish the proof we have to prove the continuous dependence of the solution with respect to the initial data. Fix another initial condition ν_{in} and denote ν the corresponding solution. For every 1-Lipschitz $\varphi : \mathbb{R}^d \to \mathbb{R}$ we have for $t \geq 0$ that

$$\int_{\mathbb{R}^d} \varphi(x) \left(\mu(t)(dx) - \nu(t)(dx)\right)$$

$$= \int_{\mathbb{R}^d} \varphi(\tilde{\mathcal{T}}(t,x;\mu)) \, \mu_{in}(0)(dx) - \int_{\mathbb{R}^d} \varphi(\tilde{\mathcal{T}}(t,x;\nu)) \, \nu_{in}(0)(dx)$$

$$= \int_{\mathbb{R}^d} \varphi(\tilde{\mathcal{T}}(t,x;\mu)) \left(\mu_{in}(0)(dx) - \nu_{in}(0)(dx)\right)$$

$$\quad + \int_{\mathbb{R}^d} \varphi(\tilde{\mathcal{T}}(t,x;\mu)) - \varphi(\tilde{\mathcal{T}}(t,x;\nu)) \, \nu_{in}(0)(dx)$$

$$=: A + B.$$

We bound A by

$$A \leq Lip(\varphi(\tilde{\mathcal{T}}(t,\cdot;\mu)))W_1(\mu_{in}(0), \nu_{in}(0)) \leq e^{Lt} \, W_1(\mu_{in}(0), \nu_{in}(0)).$$

Concerning B we use that ϕ is 1-Lipschitz and (7.1) to write

$$B \leq \int_{\mathbb{R}^d} |\tilde{\mathcal{T}}(t,x;\mu)) - \tilde{\mathcal{T}}(t,x;\nu)| \, \nu_{in}(0)(dx) \leq Le^{Lt} \int_0^t \mathcal{W}_1(\mu_s, \nu_s) \, ds.$$

Thus after taking the supremum over all 1-Lipschitz functions φ we obtain

$$W_1(\mu(t), \nu(t)) \leq e^{Lt} \left(W_1(\mu_{in}(0), \nu_{in}(0)) + L \int_0^t \mathcal{W}_1(\mu_s, \nu_s) \, dh\right) \qquad t \geq 0.$$

Incorporating times $t \in [-\tau, 0]$, we obtain

$$W_1(\mu(t), \nu(t)) \leq e^{Lt} \left(\mathcal{W}_1(\mu_{in}, \nu_{in}) + L \int_0^t \mathcal{W}_1(\mu_s, \nu_s) \, dh\right) \qquad t \geq -\tau.$$

We have in particular that for any $t \geq 0$ and $h \in [-\tau, 0]$,

$$W_1(\mu(t+h), \nu(t+h)) \leq e^{Lt} \left(\mathcal{W}_1(\mu_{in}, \nu_{in}) + L \int_0^t \mathcal{W}_1(\mu_s, \nu_s) \, dh\right)$$

so that

$$\mathcal{W}_1(\mu_t, \nu_t) \leq e^{Lt} \left(\mathcal{W}_1(\mu_{in}, \nu_{in}) + L \int_0^t \mathcal{W}_1(\mu_s, \nu_s) \, dh\right) \qquad t \geq 0.$$

Gronwall's lemma yields the result.

8. Proof of Theorem 4.2.

Fix some $R_0 > 0$ and an initial distribution $\mu(0) \in \mathbb{P}(E)$ supported in the ball $B_0(R_0) \subset E$, and denote $\mu(t)$ be the unique solution of the fixed-point equation (3.9) as given by Theorem 3.2. We must prove that $\tilde{\mu}(t) := ev(0)\sharp\mu(t) \in \mathbb{P}(\mathbb{R}^d)$ is the unique solution of (4.5) with initial condition $\tilde{\mu}(s) := ev(s)\sharp\mu(0)$, $s \in [-\tau, 0]$.

Let us first observe that

(8.1) $ev(s)\sharp\mu(t) = \tilde{\mu}(t + s)$ for any $t \geq 0$ and $s \in [-\tau, 0]$.

Indeed for any $\sigma \in C([-\tau, +\infty), \mathbb{R}^d)$, we have

$$(ev(s) \circ res(t))(\sigma) = \sigma_t(s) = \sigma(t + s) = \sigma_{t+s}(0) = (ev(0) \circ res(t + s))(\sigma)$$

so that $ev(s) \circ res(t) = ev(0) \circ res(t+s)$. Recalling that $\mu(t) = (res(t) \circ ext(\mu))\sharp\mu(0)$ we obtain

$$\begin{aligned}
ev(s)\sharp\mu(t) &= (ev(s) \circ res(t) \circ ext(\mu))\sharp\mu(0) \\
&= (ev(0) \circ res(t + s) \circ ext(\mu))\sharp\mu(0) \\
&= ev(0)\sharp(res(t + s) \circ ext(\mu))\sharp\mu(0) \\
&= ev(0)\sharp\mu(t + s) \\
&= \tilde{\mu}(t + s),
\end{aligned}$$

which is (8.1).

It follows that the vector field $\mathcal{F}[\mu]$ and $\tilde{\mathcal{F}}[\tilde{\mu}]$ defined in (3.3) and (4.4) coincide:

$$\mathcal{F}[\mu](t, x) = \tilde{\mathcal{F}}[\tilde{\mu}](t, x)   \text{ for any } t \geq 0 \text{ and } x \in \mathbb{R}^d.$$

Indeed, in view of the definition (4.1) of K, we have

$$\begin{aligned}
\mathcal{F}[\mu](t, x) &= \int_E K(x, \sigma)\,\mu(t)(d\sigma) = \int_{-\tau}^0 \Big(\int_E \tilde{K}(x, \sigma(s))\,\mu(t)(d\sigma) \Big)\, d\rho(s) \\
&= \int_{-\tau}^0 \Big(\int_{\mathbb{R}^d} \tilde{K}(x, y)\,(ev(s)\sharp\mu(t))(dy) \Big)\, d\rho(s).
\end{aligned}$$

Using (8.1) this is

$$\begin{aligned}
\mathcal{F}[\mu](t, x) &= \int_{-\tau}^0 \Big(\int_{\mathbb{R}^d} \tilde{K}(x, y)\,\tilde{\mu}(t + s)(dy) \Big)\, d\rho(s) \\
&= \tilde{\mathcal{F}}[\tilde{\mu}](t, x).
\end{aligned}$$

We can now prove that $\tilde{\mu}(t) := ev(0)\sharp\mu(t)$ is the unique solution of (4.5) with initial condition $ev(t)\sharp\mu(0)$, $t \in [-\tau, 0]$. According to the proof of Theorem 4.1 (see the definition of Γ) this is equivalent to proving that

$$\tilde{\mu}(t) = \begin{cases} ev(t)\sharp\mu(0), & t \in [-\tau, 0], \\ \tilde{T}(t, \cdot; \tilde{\mu})\sharp\tilde{\mu}(0) & t \geq 0, \end{cases}$$

where $\tilde{T}(t, x; \tilde{\mu})$ is the flow of the vector field $\tilde{F}[\tilde{\mu}]$. The equality for $t \in [-\tau, 0]$ follows from (8.1). For $t \geq 0$, using the definition (3.10) of $ext(\mu)$, we have for any $\sigma \in E$,

$$(ev(0) \circ res(t) \circ ext(\mu))(\sigma) = ext(\mu)(\sigma)(t) = \mathcal{T}(t, \sigma(0); \mu)$$

where $\mathcal{T}(t, x; \mu)$ is the flow of the vector field $\mathcal{F}[\mu]$. Thus,

$$ev(0) \circ res(t) \circ ext(\mu) = \mathcal{T}(t, \cdot; \mu) \circ ev(0).$$

It follows that

$$\begin{aligned}
\tilde{\mu}(t) &= ev(0)\sharp\mu(t) \\
&= (ev(0) \circ res(t) \circ ext(\mu))\sharp\mu(0) \\
&= \mathcal{T}(t,\cdot;\mu)\sharp(ev(0)\sharp\mu(0)) \\
&= \mathcal{T}(t,\cdot;\mu)\sharp\tilde{\mu}(0).
\end{aligned}$$

Since $\mathcal{F}[\mu] = \tilde{\mathcal{F}}[\tilde{\mu}]$, we have $T(t,x;\mu) = \tilde{T}(t,x;\tilde{\mu})$. Thus, $\tilde{\mu}(t) = \tilde{\mathcal{T}}(t,\cdot;\tilde{\mu})\sharp\tilde{\mu}(0)$ for any $t \geq 0$, which completes the proof.

REFERENCES

[1] Azmy S. Ackleh and Nicolas Saintier. Diffusive limit to a selection-mutation equation with small mutation formulated on the space of measures. *Discrete and Continuous Dynamical Systems Serie B.* in press. 4

[2] Giacomo Albi, Nicola Bellomo, Luisa Fermo, Seung Yeal Ha, J. Kim, Lorenzo Pareschi, David Poyato, and Juan Soler. Vehicular traffic, crowds, and swarms: From kinetic theory and multiscale methods to applications and research perspectives. *Mathematical Models and Methods in Applied Sciences*, 29(10):1901–2005, 2019. 1

[3] Luigo Ambrosio, M Fornassier, Marco Morandotti, and Giuseppe Savare. Spatially inhomogeneous evolutionary games. https://arxiv.org/pdf/1805.04027.pdf. 4

[4] Jianhai Bao, Christoph Reisinger, Panpan Ren, and Wolfgang Stockinger. Milstein schemes for delay McKean equations and interacting particle systems. *arXiv preprint arXiv:2005.01165*, 2020. 12

[5] Emmanuel Boissard, Pierre Degond, and Sebastien Motsch. Trail formation based on directed pheromone deposition. *Journal of mathematical biology*, 66(6):1267–1301, 2013. 5

[6] Jose A. Canizo, Jose A. Carrillo, and Silvia Cuadrado. Measure solutions for some models in population dynamics. *Acta Appl Math*, 123:141–156, 2013. 4

[7] Jose A. Canizo, Jose A. Carrillo, and Jesus Rosado. A well-posedness theory in measures for some kinetic models of collective motion. *Mathematical Models and Methods in Applied Sciences*, 21(3):515–539, 2011. 4

[8] Young-Pil Choi and Jan Haskovec. Cucker-Smale model with normalized communication weights and time delay. *Kinet. Relat. Models*, 10(4):1011–1033, 2017. 1

[9] Emiliano Cristiani, Benedetto Piccoli, and Andrea Tosin. *Multiscale Modeling of Pedestrian Dynamics*, volume 12. Springer International Publishing, 2014. 4

[10] Constantine M Dafermos. Asymptotic stability in viscoelasticity. *Archive for rational mechanics and analysis*, 37(4):297–308, 1970. 3

[11] Mauro Fabrizio, Claudio Giorgi, and Vittorino Pata. A new approach to equations with memory. *Archive for rational mechanics and analysis*, 198(1):189–232, 2010. 3

[12] Marco A Fontelos and Avner Friedman. A pde model for the dynamics of trail formation by ants. *Journal of Mathematical Analysis and Applications*, 425(1):1–19, 2015. 4

[13] François Golse. On the dynamics of large particle systems in the mean field limit. In *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*, volume 3 of *Lect. Notes Appl. Math. Mech.*, pages 1–144. Springer, [Cham], 2016. 2, 3, 4

[14] Jack K. Hale. Theory of functional differential equations. *Applied Mathematical Sciences*, 3, 1977. 7

[15] Carrillo J.A., R.M. Colombo, P Gwiazda, and A Ulikowska. Structured populations, cell growth and measure valued balance laws. *Journal of differential equations*, 252:3245–3277, 2012. 4

[16] Amir Jafari, Amy Greenwald, David Gondek, and Gunes Ercal. On no-regret learning, fictitious play, and nash equilibrium. In *ICML*, volume 1, pages 226–233, 2001. 1

[17] Frédéric Koessler. Common knowledge and interactive behaviors: A survey. *European Journal of Economic and Social Systems*, 14(3):271–308, 2000. 4

[18] Vassili N Kolokoltsov. *Nonlinear Markov processes and kinetic equations*, volume 182. Cambridge University Press, 2010. 12

[19] Henry McKean. A class of Markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences of the USA*, 56:1907–1911, 1966. 12

[20] Ulrich Munz, Antonis Papachristodoulou, and Frank Allgower. Delay-dependent rendezvous and flocking of large scale multi-agent systems with communication delays. In *2008 47th IEEE Conference on Decision and Control*, pages 2038–2043. IEEE, 2008. 1

[21] Lorenzo Pareschi and Giuseppe Toscani. *Interacting Multiagent Systems*. Oxford University Press, 2014. 1

[22] Lucia Pedraza, Juan Pablo Pinasco, and N Saintier. Measure-valued opinion dynamics. *Mathematical Models and Methods in Applied Sciences*, 30(2):225–260, 2020. 4

[23] Mayte Perez-Llanos, Juan Pablo Pinasco, N. Saintier, and Analia Silva. Opinion formation models with heterogeneous persuasion and zealotry. *SIAM Journal on Mathematical Analysis*, 50(5):4812–4837, 2018. 4

[24] Juan Pablo Pinasco, Mauro Rodriguez Cartabia, and Nicolas Saintier. A game theoretic model of wealth distribution. *Dynamic Games and Applications*, 8(4):874–890, 2018. 1

[25] Juan Pablo Pinasco, Mauro Rodriguez Cartabia, and Nicolas Saintier. Evolutionary game theory in mixed strategies: from microscopic interactions to kinetic equations. `https://mate.dm.uba.ar/~nsaintie/juegos_Cartabia_Pinasco_Saintier.pdf`. 4

[26] Juan Pablo Pinasco, Nicolas Saintier, and Federico Vazquez. Role of voting intention in public opinion polarization. *Physical Review E*, 101, 012101, 2020. 4

[27] Mayte Pérez-Llanos, Juan P. Pinasco, Nicolas Saintier, and Analía Silva. Opinion formation models with heterogeneous persuasion and zealotry. *SIAM Journal on Mathematical Analysis*, 50(5):4812–4837, 2018. 2

[28] Dialid Santiago Ramírez. *On the Existence of a Certain Class of Nonlinear Stochastic Processes*. PhD thesis, University of Warwick, 2015. 12

[29] Nicolas Saintier, Juan Pablo Pinasco, and Federico Vazquez. A model for a phase transition between political mono-polarization and bi-polarization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, in press. 4

[30] Hal L Smith. *An introduction to delay differential equations with applications to the life sciences*, volume 57. Springer New York, 2011. 7

[31] Alain-Sol Sznitman. *Topics in Propagation of Chaos - Ecole d'été de probabilités de Saint-Flour XIX - 1989*, volume 1464. Springer-Verlag, 1991. 12

[32] Giuseppe Toscani et al. Kinetic models of opinion formation. *Communications in mathematical sciences*, 4(3):481–496, 2006. 2

[33] Cédric Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003. 6

Juan Pablo Pinasco,
IMAS UBA-CONICET and Departamento de Matemática,
Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires,
Av Cantilo s/n, Ciudad Universitaria
(1428) Buenos Aires, Argentina.
Email address: `jpinasco@dm.uba.ar`

Mauro Rodriguez Cartabia,
IMAS UBA-CONICET and Departamento de Matemática,
Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires,
Av Cantilo s/n, Ciudad Universitaria
(1428) Buenos Aires, Argentina.
Email address: `mrodriguezcartabia@gmail.com`

Nicolas Saintier,
IMAS UBA-CONICET and Departamento de Matemática,
Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires,
Av Cantilo s/n, Ciudad Universitaria
(1428) Buenos Aires, Argentina.
Email address: `nsaintie@dm.uba.ar`