

## EXISTENCE AND APPROXIMATION OF PROBABILITY MEASURE SOLUTIONS TO MODELS OF COLLECTIVE BEHAVIORS

ANDREA TOSIN AND PAOLO FRASCA

Department of Mathematics, Politecnico di Torino  
Corso Duca degli Abruzzi 24  
10129 Torino, Italy

**ABSTRACT.** In this paper we consider first order differential models of collective behaviors of groups of agents, based on the mass conservation equation. Models are formulated taking the spatial distribution of the agents as the main unknown, expressed in terms of a probability measure evolving in time. We develop an existence and approximation theory of the solutions to such models and we show that some recently proposed models of crowd and swarm dynamics fit our theoretic paradigm.

**1. Introduction.** This paper deals with mathematical models of collective behaviors of groups of interacting agents, such as human crowds and swarms. The reference framework is that of first order differential models ruled by the principle of conservation of mass and supplemented by a kinematic description of the behavioral strategy developed by the agents. The number of individuals is assumed to be finite, though arbitrarily large, their state being represented by their position evolving in time in the Euclidean space  $\mathbb{R}^d$ ,  $d \geq 1$ . Rather than looking at the agents singularly, we abstract their spatiotemporal evolution into that of a suitable probability measure, representing the law of their positions understood as random variables. This allows us to provide a unified theory for both discrete and continuous models.

There are three main contributions of this paper. (i) We outline a basic set of modeling assumptions, which allow us to prove the existence of probability measure solutions to a broad class of models of collective behaviors of the kind described above. (ii) Under the very same assumptions, complemented with a suitable condition on the time and space discretization, we analyze an *ad hoc* numerical scheme, originally proposed in [9], for approximating the solutions to such models. (iii) We reinterpret the *rendez-vous*, swarm, and crowd dynamics models developed in [4, 5, 8, 9] in the light of our probabilistic description, showing that they comply with the above modeling assumptions and are therefore in the scope of our existence and approximation theory.

The reader should be aware that our aforementioned modeling assumptions, whence we let all the theory follow, are not meant to be sharp from the strict analytic point of view. They are primarily some reasonable ones in view of the

---

2000 *Mathematics Subject Classification.* Primary: 35L65, 35Q70; Secondary: 35Q91.

*Key words and phrases.* Systems of interacting agents, probability distribution, continuity equation, nonlocal flux.

The first author is supported by a post-doctoral research scholarship “Compagnia di San Paolo” awarded by the National Institute for Advanced Mathematics “F. Severi” (INdAM, Italy).

application of the models to real world problems; in addition, they allow for proofs which do not require a sophisticated technical background in optimal transport.

It is also worth pointing out that our existence result is actually a standard one in optimal transport. However, we state it with proof because, in our opinion, it contributes to the general presentation of the qualitative properties of our models of collective behaviors. In addition, it introduces techniques useful for approaching later the analysis of the approximation scheme. Finally, though classical, to our knowledge there is no precise reference for it in the specialized literature.

In more detail, the paper is organized as follows. After this Introduction, Section 1.1 briefly introduces and explains the main notations and notions used throughout the other sections. Section 2 proposes a probabilistic interpretation of the dynamics of systems of interacting agents, discussing both the indefinite mass conservation equation and the related Cauchy problems. Then it presents the modeling assumptions and offers an overview of the results proved in the subsequent sections. Section 3 deals with the existence of solutions and Section 4 with the analysis of the approximation scheme. These two sections may be skipped by readers not interested in the technical details of the proofs. Section 5 addresses the above-cited crowd and swarm dynamics models, showing that they fit the theory in all cases interesting for the applications. Finally, Section 6 gives an example of how individual-based models can be explicitly recovered from our measure-theoretic framework. In addition, it shows by means of numerical tests the convergence of the computational scheme discussed in Section 4 to the ODE solutions of such models.

**1.1. Notations and background.** In this section we quickly review the main notations and notions that we will extensively use in the paper.

*Functions and function spaces.* We denote by  $C(A; B)$  the space of continuous functions  $f : A \rightarrow B$ . The set  $B$  is usually omitted if it is  $\mathbb{R}$ . Coherently, we denote by  $C_c^\infty(\mathbb{R}^d)$  the space of real-valued infinitely differentiable functions with compact support in  $\mathbb{R}^d$ .

The indicator function of a set  $A$  is  $\mathbb{1}_A$ , namely  $\mathbb{1}_A(x) = 1$  if  $x \in A$  and  $\mathbb{1}_A(x) = 0$  if  $x \in A^c$ .

*Norms.* We use  $|\cdot|$  for the Euclidean norm in  $\mathbb{R}^d$  and  $\cdot$  for the corresponding inner product. The open Euclidean ball with center  $x \in \mathbb{R}^d$  and radius  $R \geq 0$  is denoted by  $B_R(x)$ . If  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is bounded, its  $\infty$ -norm is  $\|f\|_\infty := \sup_{x \in \mathbb{R}^d} |f(x)|$ .

*Measures.* Let  $\mathcal{B}(\mathbb{R}^d)$  be the Borel  $\sigma$ -algebra on  $\mathbb{R}^d$ . If  $\mu$  is a measure on the measurable space  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  and  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is Borel, the integral of  $f$  w.r.t.  $\mu$  over a measurable set  $A$  is denoted by  $\int_A f d\mu$ , or by  $\int_A f(x) d\mu(x)$  when it is necessary to emphasize the variable of integration. The Lebesgue measure in  $\mathbb{R}^d$  is denoted  $\mathcal{L}^d$ . However, for integrals with respect to  $\mathcal{L}^d$  we will prefer the usual symbol  $dx$  to  $d\mathcal{L}^d(x)$ . If  $f$  is Borel, we denote by  $f\#\mu$  the push forward of  $\mu$  through  $f$ . Specifically,  $f\#\mu$  is the measure defined by the relation

$$\int_{\mathbb{R}^d} \eta d(f\#\mu) = \int_{\mathbb{R}^d} \eta \circ f d\mu \quad (1)$$

for all bounded (or  $f\#\mu$ -integrable) and Borel  $\eta : \mathbb{R}^d \rightarrow \mathbb{R}$ . Taking  $\eta = \mathbb{1}_A$ ,  $A \in \mathcal{B}(\mathbb{R}^d)$ , gives in particular  $(f\#\mu)(A) = \mu(f^{-1}(A))$ .

*Probability spaces.* We denote by  $\mathcal{P}_1(\mathbb{R}^d)$  the space of probability measures on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  whose first moment is finite, i.e.,  $\int_{\mathbb{R}^d} |x| d\mu(x) < +\infty$ . The space  $\mathcal{P}_2(\mathbb{R}^d)$ , that we will also occasionally mention, is defined analogously using the second moment. Given two probability measures  $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$ , their Wasserstein distance is defined to be

$$W_1(\mu, \nu) = \sup_{\varphi \in \text{Lip}_1(\mathbb{R}^d)} \int_{\mathbb{R}^d} \varphi d(\nu - \mu),$$

where  $\text{Lip}_1(\mathbb{R}^d)$  is the space of Lipschitz continuous functions  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  with Lipschitz constant  $\text{Lip}(\varphi) \leq 1$ . It can be shown that  $W_1$  is a metric on  $\mathcal{P}_1(\mathbb{R}^d)$  and that  $(\mathcal{P}_1(\mathbb{R}^d), W_1)$  is complete (see e.g., [1, Proposition 7.1.5]).

Finally, to deal with curves in  $\mathcal{P}_1(\mathbb{R}^d)$  parameterized by time,  $t \mapsto \mu_t \in \mathcal{P}_1(\mathbb{R}^d)$ , we introduce the space  $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ , which is complete with the metric

$$\text{dist}(\mu_\bullet, \nu_\bullet) := \sup_{t \in [0, T]} W_1(\mu_t, \nu_t).$$

**2. Problem statement and main results.** In this section we present our approach to the modeling of systems of interacting agents by means of probability measures, and we give an overview of our results.

**2.1. Probabilistic description of systems of interacting agents.** Crowds and swarms can be thought of, in abstract, as systems of  $N$  interacting agents in the physical space  $\mathbb{R}^d$ . The evolution of such systems in time is described by tracing the agent positions at successive instants. Assume that a stochastic basis  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, P)$  is given,  $(\Omega, \mathcal{F}, P)$  being an abstract probability space and  $(\mathcal{F}_t)_t$  a filtration of  $\mathcal{F}$ . We model the position of the  $i$ -th agent at time  $t$  as an  $\mathcal{F}_t$ -measurable random variable  $X_t^i$  with values in the measurable space  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ . The probability  $P$  is naturally transported by  $X_t^i$  onto the new probability  $\mu_t := X_t^i \# P$  on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ , called the law of  $X_t^i$ . We explicitly assume that  $\mu_t$  does not depend on the agent label  $i$ . This means that agents are *indistinguishable*, i.e., the probability of finding a certain agent somewhere in  $\mathbb{R}^d$  at time  $t$  is the same for all agents: given  $A \in \mathcal{B}(\mathbb{R}^d)$ ,  $P(\{X_t^i \in A\}) = \mu_t(A)$  for all  $i = 1, \dots, N$ .

Let us now fix a measurable set  $A \subseteq \mathbb{R}^d$  and count the average number of agents contained in  $A$  at time  $t$ . This amounts to introducing the new random variable  $Y_{t,A} : \Omega \rightarrow \mathbb{N}$  defined as

$$Y_{t,A} = \sum_{i=1}^N \mathbb{1}_{\{X_t^i \in A\}}$$

and taking its expectation, that we can compute as follows:

$$\mathbb{E}[Y_{t,A}] = \sum_{i=1}^N P(\{X_t^i \in A\}) = N\mu_t(A).$$

Notice that  $\mathbb{E}[Y_{t,\bullet}]$ , thought of as a mapping on  $\mathcal{B}(\mathbb{R}^d)$ , is a positive measure, say  $m_t$ , over the measurable space  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ . It is straightforward to identify  $m_t$  with the *mass* of the system at time  $t$ . From the above calculation we see that  $m_t$  is proportional to the probability measure of the distribution of the agents:  $m_t(A) = N\mu_t(A)$ , where  $m_t(A)$  is the mass of the set  $A \in \mathcal{B}(\mathbb{R}^d)$  at time  $t$ . In addition,  $m_t(\mathbb{R}^d) = N$  for all  $t$ , i.e., the total mass of the system is constant in time.

The latter observation suggests that we can assume the principle of conservation of the mass, stating that the mass of any measurable set  $A$  may change in time only because of inflow or outflow of mass from the boundary  $\partial A$ . In other words, the mass is neither created nor destroyed but only moved across the domain. This is expressed by postulating the *continuity equation* (or *mass conservation equation*) for the evolution of the measure  $m_t$ :

$$\frac{\partial m_t}{\partial t} + \nabla \cdot (m_t v_t) = 0, \quad (2)$$

where  $v_t(x)$  is the velocity at time  $t$  in the point  $x \in \mathbb{R}^d$ . In systems of interacting agents the velocity is likely to be affected by the distribution of the agents themselves. Due to the proportionality between  $m_t$  and  $\mu_t$ , this implies that  $v_t$  may ultimately depend on the probability  $\mu_t$ . We assume in particular that  $v_t$  depends on  $t$  only through  $\mu_t$  itself, i.e., that the system is autonomous. Finally, we write  $v_t = v[\mu_t]$  to emphasize such a structure of the velocity and notice that 2 can be converted into an evolution equation for the probability  $\mu_t$ :

$$\frac{\partial \mu_t}{\partial t} + \nabla \cdot (\mu_t v[\mu_t]) = 0 \quad (3)$$

for  $x \in \mathbb{R}^d$  and  $t \in (0, T]$ , where  $T > 0$  is the final time. If  $\mu_t$  solves 3, then  $m_t$  formally solves 2 with  $v_t = v[m_t/N]$ .

**2.2. Cauchy problems.** By supplementing 3 with an initial condition  $\bar{\mu}$ , the following Cauchy problem is obtained:

$$\begin{cases} \frac{\partial \mu_t}{\partial t} + \nabla \cdot (\mu_t v[\mu_t]) = 0 & \text{in } \mathbb{R}^d \times (0, T] \\ \mu_0 = \bar{\mu} & \text{in } \mathbb{R}^d, \end{cases} \quad (4)$$

which models the spatiotemporal evolution of the agent distribution starting from the initial configuration described by  $\bar{\mu}$ . Derivatives in problem 4 are meant in the sense of distributions, which leads us to consider the following notion of weak solution:

**Definition 2.1** (Weak solutions). Given  $\bar{\mu} \in \mathcal{P}_1(\mathbb{R}^d)$ ,  $\mu_\bullet \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  is a weak solution to problem 4 if

$$\int_{\mathbb{R}^d} \phi d\mu_t = \int_{\mathbb{R}^d} \phi d\bar{\mu} + \int_0^t \int_{\mathbb{R}^d} v[\mu_\tau] \cdot \nabla \phi d\mu_\tau d\tau, \quad \forall \phi \in C_c^\infty(\mathbb{R}^d), \forall t \in [0, T]. \quad (5)$$

As far as the probabilistic interpretation is concerned, we take the initial condition into account by understanding  $\mu_t$  in 4 as the law of the random variable  $\hat{X}_t^i := \mathbb{E}[X_t^i | X_0^i]$ , the expectation of  $X_t^i$  conditioned to its initial value  $X_0^i$ . In practice,  $\mu_t$  is reinterpreted as the distribution of the position of the  $i$ -th agent subject to the distribution of the corresponding initial position. From the theory of conditional expectation, we know that  $\hat{X}_t^i$  is a function of  $X_0^i$ , i.e., there exists a Borel mapping  $\gamma_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that  $\hat{X}_t^i = \gamma_t(X_0^i)$ . This implies, in particular,

$$\mu_t = \gamma_t \# \bar{\mu}. \quad (6)$$

In order for 6 to result in a representation formula for the solutions of problem 4, a more precise characterization of the function  $\gamma_t$  is needed. Formally, we plug 6

into 5 and compute

$$\int_{\mathbb{R}^d} [\phi(\gamma_t(x)) - \phi(x)] d\bar{\mu}(x) = \int_0^t \int_{\mathbb{R}^d} v[\gamma_\tau \# \bar{\mu}](\gamma_\tau(x)) \cdot \nabla \phi(\gamma_\tau(x)) d\bar{\mu}(x) d\tau$$

for an arbitrarily fixed test function  $\phi$ . Next we notice that the integrand at the right-hand side can be read as the derivative w.r.t.  $\tau$  of the function  $\phi(\gamma_\tau(x))$ , provided we identify  $\frac{\partial}{\partial t} \gamma_t(x)$  with  $v[\gamma_t \# \bar{\mu}](\gamma_t(x))$ . Under this assumption we get

$$\int_{\mathbb{R}^d} [\phi(\gamma_t(x)) - \phi(x)] d\bar{\mu}(x) = \int_0^t \int_{\mathbb{R}^d} \frac{\partial}{\partial \tau} \phi(\gamma_\tau(x)) d\bar{\mu}(x) d\tau$$

and further, interchanging the order of integration at the right-hand side,

$$= \int_{\mathbb{R}^d} [\phi(\gamma_t(x)) - \phi(\gamma_0(x))] d\bar{\mu}(x).$$

With the additional condition  $\gamma_0(x) = x$  (i.e.,  $\gamma_0$  is the identity function in  $\mathbb{R}^d$ ), this shows that  $\mu_t$  represented by 6 is formally a weak solution to the Cauchy problem 4.

To sum up,  $\gamma_t$  has been characterized as a function such that

$$\begin{cases} \frac{\partial \gamma_t(x)}{\partial t} = v[\gamma_t \# \bar{\mu}](\gamma_t(x)), & t \in (0, T] \\ \gamma_0(x) = x \end{cases} \tag{7}$$

for every  $x \in \mathbb{R}^d$ . In transport theory, such a function is called a *flow map*. The physical interpretation is that  $\gamma_t(A)$  is the configuration assumed by the set  $A$  at time  $t > 0$  when transported by the velocity field  $v$ . Alternatively, for  $x \in \mathbb{R}^d$  the mapping  $t \mapsto \gamma_t(x)$  is the trajectory of system 7 issuing from  $x$ .

The method of representing solutions to the Cauchy problem 4 via flow maps is called *method of the characteristics*. To develop our theory we will mostly prefer a different approach, more suited to treat, by common ideas, existence and approximation of solutions to the models we are interested in. The reader interested in the method of the characteristics is referred to [3] and references therein for further details.

**2.3. Basic assumptions and results overview.** Models based on 3 require to specify the velocity, namely its dependence on the probability  $\mu_t$  and on the space variable  $x$ . Rather than considering a specific model, we outline here a small set of assumptions, which can be possibly regarded as modeling guidelines, whence the whole theory will follow.

**Assumption 1** (Properties of  $v$ ). We assume that the velocity field  $v = v[\mu](x)$  satisfies the following properties.

1. *Uniform boundedness:* there exists  $V > 0$  such that

$$|v[\mu](x)| \leq V, \quad \forall x \in \mathbb{R}^d, \forall \mu \in \mathcal{P}_1(\mathbb{R}^d).$$

2. *Lipschitz continuity:* there exists a constant  $\text{Lip}(v) > 0$ , independent of both the space variable and the probability, such that

$$|v[\nu](y) - v[\mu](x)| \leq \text{Lip}(v)(|y - x| + W_1(\mu, \nu)), \quad \forall x, y \in \mathbb{R}^d, \forall \mu, \nu \in \mathcal{P}_1(\mathbb{R}^d).$$

3. *Linearity w.r.t. the measure for convex combinations:*

$$v[\alpha\mu + (1 - \alpha)\nu] = \alpha v[\mu] + (1 - \alpha)v[\nu], \quad \forall \mu, \nu \in \mathcal{P}_1(\mathbb{R}^d), \forall \alpha \in [0, 1].$$

It is worth stressing that these assumptions are not meant to be sharp from the strict analytic point of view. In particular, although we are mainly concerned with existence of solutions, they might be sufficient to obtain uniqueness as well. Actually, they are primarily intended to cope with real world applications, and in this respect we believe that they are suitable for getting reasonable models. In addition, they have the merit of making possible proofs of our results not requiring a sophisticated technical background in optimal transport. Finally, they allow us to approach both existence and approximation of solutions by common working hypotheses and techniques.

From the basic set of hypotheses delivered by Assumption 1 we will be able to prove, in Section 3, existence of solutions to the Cauchy problem 4 in the appropriate weak sense of Definition 2.1. More precisely, to this end we need the further technical assumption that the initial condition have finite first and second order moments:

**Assumption 2** (Initial condition). We assume that  $\bar{\mu} \in \mathcal{P}_1(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$ .

Then, our main result in Section 3 reads:

**Main Theorem** (cf. Theorem 3.4). *Under Assumptions 1, 2 there exists a weak solution to problem 4.*

Notice that Assumption 2 is readily satisfied if, for instance,  $\bar{\mu}$  has compact support. Indeed, in such a case  $\text{supp } \bar{\mu}$  is bounded, i.e., there exists a ball  $B_R(0)$  of sufficiently large radius  $R > 0$  such that, for every  $p \geq 0$ ,

$$\int_{\mathbb{R}^d} |x|^p d\bar{\mu}(x) = \int_{\text{supp } \bar{\mu}} |x|^p d\bar{\mu}(x) \leq \int_{B_R(0)} |x|^p d\bar{\mu} \leq R^p.$$

The compactness of the support of  $\bar{\mu}$  makes perfectly sense from the modeling point of view, in fact a crowd or a swarm spread on the whole space would sound quite unrealistic. Assumption 2 is therefore not restrictive for our purposes.

In Section 4 we turn our attention to the approximation of solutions to problem 4. We introduce a sequence of grids in  $\mathbb{R}^d \times [0, T]$  with mesh parameters  $h_k$  in space and  $\Delta t_k$  in time. The index  $k$  relates to the grid refinement, in such a way that  $h_k, \Delta t_k \rightarrow 0$  when  $k \rightarrow \infty$ . Specifically, we consider the numerical scheme proposed in [9], which at each time step seeks an approximation of  $\mu_t$  via a probability measure absolutely continuous w.r.t. to Lebesgue and piecewise constant in space. By introducing a linear-in-time interpolation of such approximate solutions and passing to the limit  $k \rightarrow \infty$ , we obtain the following convergence result:

**Main Theorem** (cf. Theorem 4.7). *Under Assumptions 1, 2, suppose that  $h_k = o(\Delta t_k)$  for  $k \rightarrow \infty$ . If the sequence of approximate solutions converges to some  $\mu_\bullet \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  when the grid is refined then  $\mu_\bullet$  is a weak solution to problem 4.*

Note that convergence when the grid is refined is an assumption of this theorem. In this respect, this result resembles the Lax-Wendroff's Theorem for the numerical approximation of hyperbolic conservation laws (see e.g., [7]). However, we anticipate that if there exists a bounded subset of  $\mathbb{R}^d$ , that at each time step and for all level

of refinement of the grid contains the supports of all the approximate solutions, then the sequence does converge in  $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  to some limit, which is then a weak solution to problem 4 (cf. Corollary 1).

In Section 5 we apply the above theory to the models of swarm and crowd dynamics presented in [5, 8]. As shown in [6], these models can be derived from the common framework provided by 3, with a velocity field of the form

$$v[\mu_t](x) = v_d(x) + N \int_{\mathbb{R}^d} f(|y - x|)r(y - x)\chi_{U_x}(y) d\mu_t(y),$$

where the integral expresses the interactions among the agents. Assuming some minimal regularity of the functions  $v_d, f, r, \chi_{U_x}$ , we prove that this velocity complies with Assumption 1 in case of both isotropic and anisotropic interactions, cf. Sections 5.1 and 5.2, respectively. As a consequence, for the above-mentioned models we deduce existence of probability measure solutions, that can be duly approximated via the numerical scheme discussed in Section 4.

Our results are exemplified in Section 6, where we show that, given a purely atomic initial measure  $\bar{\mu}$ , i.e.,

$$\bar{\mu} = \frac{1}{N} \sum_{l=1}^N \delta_{x_0^l} \quad (x_0^l \in \mathbb{R}^d),$$

a solution to problem 4 can be found by solving a system of ODEs whose unknowns are the trajectories of the agents. In addition, using a numerical solution of these ODEs as a benchmark, we are able to visualize the convergence of the numerical scheme presented in Section 4.

**3. Existence of solutions.** This section is devoted to give a constructive proof of the existence of solutions to the Cauchy problem 4. Under Assumptions 1, 2, the solution is constructed as the limit of a suitable sequence of curves in  $\mathcal{P}_1(\mathbb{R}^d)$  parameterized by time  $t$ .

As anticipated in the Introduction, this result is classical in optimal transport, yet to our knowledge there is no precise reference for it in the specialized literature. In the present context, it has a twofold merit: on the one hand, it contributes to a general presentation of basic qualitative properties of the models we are interested in; on the other hand, it introduces some techniques that will be useful later for the study of the approximation scheme.

Let  $(\Delta t_k)_{k \geq 0}$  be a sequence of time steps such that  $\Delta t_k \rightarrow 0$  when  $k \rightarrow \infty$ . We consider the measures  $(\mu_n^k)_{n \geq 0}$  generated recursively as

$$\begin{cases} \mu_{n+1}^k = \gamma_n^k \# \mu_n^k, & n = 0, 1, \dots, N_k - 1, \\ \mu_0^k = \bar{\mu}, \end{cases} \tag{8}$$

where  $N_k \in \mathbb{N}$  is such that  $N_k \Delta t_k = T$  and  $\gamma_n^k$  is the (one-step) flow map

$$\gamma_n^k(x) = x + v[\mu_n^k](x) \Delta t_k. \tag{9}$$

It can be shown (see e.g., [9]) that 8 is the explicit time discretization of 4 at the time instants  $t_n^k = n \Delta t_k$ . In particular, the mapping  $\gamma_n^k$  results from the explicit time discretization of problem 7. Since  $\bar{\mu}$  is a probability measure, by induction it is immediate to check that so are all of the  $\mu_n^k$ 's.

By linear interpolation in time, we define the following curves:

$$M_t^k = \sum_{n=0}^{N_k-1} \left[ \left(1 - \frac{t - t_n^k}{\Delta t_k}\right) \mu_n^k + \frac{t - t_n^k}{\Delta t_k} \mu_{n+1}^k \right] \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t).$$

Obviously,  $M_t^k$  is a probability measure for each  $t \in [0, T]$  and each  $k \geq 0$ .

We will use the curves  $M_\bullet^k$  to construct, in the limit  $k \rightarrow \infty$ , a weak solution to problem 4. The proof is divided in two parts, each of which proceeds through a series of intermediate steps developed in the next two sections. First, we show that the measures  $M_\bullet^k$  converge to a limit; later, we show that such a limit satisfies 5.

**3.1. Convergence of the measures  $M_t^k$ .** In this section we prove that, when  $k \rightarrow \infty$  and up to subsequences,  $M_\bullet^k$  converges to a limit  $\mu_\bullet$  in  $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ .

We begin by establishing the necessary regularity properties of the iterates  $\mu_n^k$  and the curves  $M_\bullet^k$ .

**Lemma 3.1.** *We have  $\mu_n^k \in \mathcal{P}_1(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$  for all  $n = 1, \dots, N_k$  and all  $k \geq 0$ . In addition:*

$$\sup_{k \geq 0} \sup_{1 \leq n \leq N_k} \int_{\mathbb{R}^d} |x|^p d\mu_n^k(x) < +\infty, \quad p = 1, 2,$$

*i.e., first and second moments of the  $\mu_n^k$ 's are uniformly bounded.*

*Proof.* 1. Let us begin by considering the case  $p = 1$ . Note that

$$\begin{aligned} \int_{\mathbb{R}^d} |x| d\mu_{n+1}^k(x) &= \int_{\mathbb{R}^d} |\gamma_n^k(x)| d\mu_n^k(x) \\ &\leq \int_{\mathbb{R}^d} |x| d\mu_n^k(x) + \Delta t_k \int_{\mathbb{R}^d} |v[\mu_n^k](x)| d\mu_n^k(x) \\ &\leq \int_{\mathbb{R}^d} |x| d\mu_n^k(x) + V\Delta t_k, \end{aligned}$$

then

$$\int_{\mathbb{R}^d} |x| d\mu_{n+1}^k(x) - \int_{\mathbb{R}^d} |x| d\mu_n^k(x) \leq V\Delta t_k,$$

and summing telescopically over  $n$  we get

$$\int_{\mathbb{R}^d} |x| d\mu_n^k(x) \leq \int_{\mathbb{R}^d} |x| d\bar{\mu}(x) + VT. \tag{10}$$

Since  $\bar{\mu} \in \mathcal{P}_1(\mathbb{R}^d)$ , this implies  $\mu_n^k \in \mathcal{P}_1(\mathbb{R}^d)$  for all  $n = 1, \dots, N_k$  and all  $k \geq 0$ . But the right-hand side of 10 is independent of both  $n$  and  $k$ , hence this also provides the uniform bound for  $p = 1$ .

2. We argue analogously for the case  $p = 2$ . We note that

$$\int_{\mathbb{R}^d} |x|^2 d\mu_{n+1}^k(x) = \int_{\mathbb{R}^d} |\gamma_n^k(x)|^2 d\mu_n^k(x) = \int_{\mathbb{R}^d} |x + v[\mu_n^k](x)\Delta t_k|^2 d\mu_n^k(x)$$

and that

$$\begin{aligned} |x + v[\mu_n^k](x)\Delta t_k|^2 &= |x|^2 + 2\Delta t_k x \cdot v[\mu_n^k](x) + \Delta t_k^2 |v[\mu_n^k](x)|^2 \\ &\leq |x|^2 + 2V\Delta t_k |x| + V^2\Delta t_k^2, \end{aligned}$$



therefore

$$\int_{\mathbb{R}^d} |x|^2 d\mu_{n+1}^k(x) \leq \int_{\mathbb{R}^d} |x|^2 d\mu_n^k(x) + 2V\Delta t_k \int_{\mathbb{R}^d} |x| d\mu_n^k(x) + V^2\Delta t_k^2.$$

Moreover, using the bound 10 we deduce

$$\int_{\mathbb{R}^d} |x|^2 d\mu_{n+1}^k(x) - \int_{\mathbb{R}^d} |x|^2 d\mu_n^k(x) \leq 2V\Delta t_k \int_{\mathbb{R}^d} |x| d\bar{\mu}(x) + 2V^2T\Delta t_k + V^2\Delta t_k^2$$

and, summing telescopically over  $n$ ,

$$\int_{\mathbb{R}^d} |x|^2 d\mu_n^k(x) \leq \int_{\mathbb{R}^d} |x|^2 d\bar{\mu}(x) + 2VT \int_{\mathbb{R}^d} |x| d\bar{\mu}(x) + 3V^2T^2$$

whence the claims of the lemma follow also for  $p = 2$ . □

**Lemma 3.2.** *For all  $k \geq 0$  we have*

$$W_1(M_s^k, M_t^k) \leq V |t - s|, \quad \forall s, t \in [0, T], \tag{11}$$

*i.e., the curves  $M_\bullet^k \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  are Lipschitz continuous uniformly in  $k$ . Moreover,*

$$\sup_{k \geq 0} \sup_{t \in [0, T]} \int_{\mathbb{R}^d} |x|^p dM_t^k(x) < +\infty, \quad p = 1, 2.$$

*Proof.* 1. We claim  $M_t^k \in \mathcal{P}_1(\mathbb{R}^d)$  for all  $t, k$ . To show this, we fix  $k \geq 0$  and  $t \in [0, T]$  and observe that there exists  $0 \leq n \leq N_k$  such that  $t \in [t_n^k, t_{n+1}^k]$ . Hence, using 10, we obtain

$$\begin{aligned} \int_{\mathbb{R}^d} |x| dM_t^k(x) &= \left(1 - \frac{t - t_n^k}{\Delta t_k}\right) \int_{\mathbb{R}^d} |x| d\mu_n^k(x) + \frac{t - t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} |x| d\mu_{n+1}^k(x) \\ &\leq \int_{\mathbb{R}^d} |x| d\bar{\mu}(x) + VT. \end{aligned}$$

From the arbitrariness of  $t \in [0, T]$ ,  $k \geq 0$  our claim follows, along with the uniform boundedness of the first moment of  $M_t^k$  in both  $t$  and  $k$ .

2. We prove now the estimate 11. Let  $s, t \in [0, T]$  and assume, without loss of generality, that  $s \leq t$ . Then there exist two integers  $m$  and  $n$ , such that  $0 \leq m \leq n \leq N_k$ , with the property that  $t_m^k \leq s \leq t_{m+1}^k$  and  $t_n^k \leq t \leq t_{n+1}^k$ . Therefore we can write

$$\begin{aligned} M_s^k &= \left(1 - \frac{s - t_m^k}{\Delta t_k}\right) \mu_m^k + \frac{s - t_m^k}{\Delta t_k} \mu_{m+1}^k, \\ M_t^k &= \left(1 - \frac{t - t_n^k}{\Delta t_k}\right) \mu_n^k + \frac{t - t_n^k}{\Delta t_k} \mu_{n+1}^k \end{aligned}$$

and further, owing to the triangle inequality,

$$\begin{aligned} W_1(M_s^k, M_t^k) &\leq W_1(M_s^k, \mu_{m+1}^k) \\ &\quad + \sum_{j=m+1}^{n-1} W_1(\mu_j^k, \mu_{j+1}^k) + W_1(\mu_n^k, M_t^k). \end{aligned} \tag{12}$$

Notice that

$$\begin{aligned} W_1(M_s^k, \mu_{m+1}^k) &= \sup_{\varphi \in \text{Lip}_1(\mathbb{R}^d)} \int_{\mathbb{R}^d} \varphi d(\mu_{m+1}^k - M_s^k) \\ &= \left(1 - \frac{s - t_m^k}{\Delta t_k}\right) \sup_{\varphi \in \text{Lip}_1(\mathbb{R}^d)} \int_{\mathbb{R}^d} \varphi d(\mu_{m+1}^k - \mu_m^k) \\ &= \frac{t_{m+1}^k - s}{\Delta t_k} W_1(\mu_m^k, \mu_{m+1}^k) \end{aligned}$$

and analogously  $W_1(\mu_n^k, M_t^k) = (t - t_n^k)W_1(\mu_n^k, \mu_{n+1}^k)/\Delta t_k$ , thus, according to [12](#), estimating  $W_1(M_s^k, M_t^k)$  amounts to estimating  $W_1(\mu_i^k, \mu_{i+1}^k)$  for arbitrary  $0 \leq i \leq N_k - 1$ .

For any  $\varphi \in \text{Lip}_1(\mathbb{R}^d)$ , using  $\mu_{i+1}^k = \gamma_i^k \# \mu_i^k$  yields

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x) d(\mu_{i+1}^k - \mu_i^k)(x) &= \int_{\mathbb{R}^d} (\varphi(\gamma_i^k(x)) - \varphi(x)) d\mu_i^k(x) \\ &\leq \int_{\mathbb{R}^d} |\gamma_i^k(x) - x| d\mu_i^k(x) \\ &= \Delta t_k \int_{\mathbb{R}^d} |v[\mu_i^k](x)| d\mu_i(x) \leq V \Delta t_k, \end{aligned}$$

consequently  $W_1(\mu_i^k, \mu_{i+1}^k) \leq V \Delta t_k$  all  $i$ . From [12](#) we deduce

$$\begin{aligned} W_1(M_s^k, M_t^k) &\leq V[(t_{m+1}^k - s) + (n - m - 1)\Delta t_k + (t - t_n^k)] \\ &= V(t - s + t_{m+1}^k - t_m^k - \Delta t_k) = V(t - s), \end{aligned}$$

which proves our claim.

3. Finally, we claim  $M_t^k \in \mathcal{P}_2(\mathbb{R}^d)$  for all  $t, k$ . Arguing like in [1](#) we have

$$\begin{aligned} \int_{\mathbb{R}^d} |x|^2 dM_t^k(x) &= \left(1 - \frac{t - t_n^k}{\Delta t_k}\right) \int_{\mathbb{R}^d} |x|^2 d\mu_n^k(x) + \frac{t - t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} |x|^2 d\mu_{n+1}^k(x) \\ &\leq \int_{\mathbb{R}^d} |x|^2 d\bar{\mu}(x) + 2VT \int_{\mathbb{R}^d} |x| d\bar{\mu}(x) + 3V^2T^2, \end{aligned}$$

where [Lemma 3.1](#) has been used. This proves the claim and also the uniform boundedness of the second moment of  $M_t^k$  in both  $t$  and  $k$ .  $\square$

We are now ready to prove the main convergence result of this section.

**Proposition 1.** *There exists  $\mu_\bullet \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  and there exists a subsequence  $(M_\bullet^{k_j})_{j \geq 0}$  such that*

$$\lim_{j \rightarrow \infty} \sup_{t \in [0, T]} W_1(M_t^{k_j}, \mu_t) = 0.$$

*Proof.* It suffices to prove that  $\{M_\bullet^k\}_{k \geq 0}$  is relatively compact in  $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ . Owing to Ascoli-Arzelà's Theorem, this happens if  $\{M_\bullet^k\}_{k \geq 0}$  is equicontinuous and  $\{M_t^k\}_{k \geq 0}$  is relatively compact in  $\mathcal{P}_1(\mathbb{R}^d)$  for all  $t \in [0, T]$ .

1. Equicontinuity follows from the estimate 11. Indeed, let  $\varepsilon > 0$  then for  $\delta = \varepsilon/(2V)$ , which does not depend on  $k$ , we have  $|t - s| < \delta \Rightarrow W_1(M_s^k, M_t^k) \leq \varepsilon/2 < \varepsilon$ .
2. According to [1, Proposition 7.1.5], the relative compactness of  $\{M_t^k\}_{k \geq 0}$  in  $\mathcal{P}_1(\mathbb{R}^d)$  is equivalent to the fact that  $\{M_t^k\}_{k \geq 0}$  be tight and have uniformly integrable first moments.
  - (a) Using [1, Remark 5.1.5], a sufficient condition for tightness is that there exists a function  $\varphi : \mathbb{R}^d \rightarrow [0, +\infty]$ , whose sublevels  $\{x \in \mathbb{R}^d : \varphi(x) \leq c\}$  are compact in  $\mathbb{R}^d$ , such that

$$\sup_{k \geq 0} \int_{\mathbb{R}^d} \varphi dM_t^k < +\infty.$$

Taking  $\varphi(x) = |x|$  and invoking Lemma 3.2 we see that this condition is fulfilled, hence  $\{M_t^k\}_{k \geq 0}$  is tight.

- (b) Using [1, (5.1.20)], a sufficient condition for the uniform integrability of the first moments of  $\{M_t^k\}_{k \geq 0}$  is that there exists  $p > 1$  such that

$$\sup_{k \geq 0} \int_{\mathbb{R}^d} |x|^p dM_t^k(x) < +\infty.$$

From Lemma 3.2 we know that this actually holds for  $p = 2$ .

Since  $\{M_{\bullet}^k\}_{k \geq 0}$  is relatively compact in  $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ , up to subsequences we obtain that the sequence  $(M_{\bullet}^k)_{k \geq 0}$  converges in  $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  and we are done.  $\square$

**3.2. The limit  $\mu_{\bullet}$  solves problem 4.** In the previous section we have constructed a map  $\mu_{\bullet}$  as the limit of the sequence  $(M_{\bullet}^k)_{k \geq 0}$ . In this section we prove that such a  $\mu_{\bullet}$  is a weak solution to problem 4. To this end, we first derive an equation solved by the  $M_t^k$ 's, then we pass to the limit  $k \rightarrow \infty$  thanks to Proposition 1.

In the following,  $\phi \in C_c^\infty(\mathbb{R}^d)$  is a fixed test function. Using the definition of  $M_t^k$ , let us compute

$$\begin{aligned} \int_{\mathbb{R}^d} \phi dM_t^k &= \sum_{n=0}^{N_k-1} \left\{ \left(1 - \frac{t - t_n^k}{\Delta t_k}\right) \int_{\mathbb{R}^d} \phi d\mu_n^k + \frac{t - t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} \phi d\mu_{n+1}^k \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) \\ &= \sum_{n=0}^{N_k-1} \left\{ \int_{\mathbb{R}^d} \phi d\mu_n^k + \frac{t - t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} (\phi \circ \gamma_n^k - \phi) d\mu_n^k \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t). \end{aligned} \tag{13}$$

A Taylor expansion of  $\phi \circ \gamma_n^k$  with Lagrange's reminder gives

$$\begin{aligned} \phi(\gamma_n^k(x)) &= \phi(x + v[\mu_n^k](x)\Delta t_k) \\ &= \phi(x) + \nabla \phi(x) \cdot v[\mu_n^k](x)\Delta t_k + \frac{1}{2}(D^2 \phi(\bar{x})v[\mu_n^k](x)) \cdot v[\mu_n^k](x)\Delta t_k^2, \end{aligned}$$

where  $D^2\phi$  is the Hessian of  $\phi$  and  $\bar{x}$  is a point of the segment connecting  $x$  and  $x + v[\mu_n^k](x)\Delta t_k$ . Hence the previous computation specializes as

$$\begin{aligned} \int_{\mathbb{R}^d} \phi dM_t^k &= \sum_{n=0}^{N_k-1} \left\{ \int_{\mathbb{R}^d} \phi d\mu_n^k + (t - t_n^k) \int_{\mathbb{R}^d} v[\mu_n^k] \cdot \nabla \phi d\mu_n^k \right. \\ &\quad \left. + \frac{1}{2} \Delta t_k (t - t_n^k) \int_{\mathbb{R}^d} (D^2\phi(\bar{x})v[\mu_n^k]) \cdot v[\mu_n^k] d\mu_n^k \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t). \end{aligned}$$

We claim now that the mapping  $t \mapsto \int_{\mathbb{R}^d} \phi dM_t^k$  is Lipschitz continuous, hence a.e. differentiable by Rademacher's Theorem. To see this, observe that  $x \mapsto \phi(x)/\text{Lip}(\phi)$  is Lipschitz continuous with Lipschitz constant at most 1, so that

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \phi dM_t^k - \int_{\mathbb{R}^d} \phi dM_s^k \right| &= \text{Lip}(\phi) \left| \int_{\mathbb{R}^d} \frac{\phi}{\text{Lip}(\phi)} d(M_t^k - M_s^k) \right| \\ &\leq \text{Lip}(\phi) W_1(M_s^k, M_t^k) \leq \text{Lip}(\phi) V |t - s|. \end{aligned}$$

Thus, using [13](#), we compute the derivative

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} \phi dM_t^k &= \sum_{n=0}^{N_k-1} \left\{ \int_{\mathbb{R}^d} v[\mu_n^k] \cdot \nabla \phi d\mu_n^k \right. \\ &\quad \left. + \frac{1}{2} \Delta t_k \int_{\mathbb{R}^d} (D^2\phi(\bar{x})v[\mu_n^k]) \cdot v[\mu_n^k] d\mu_n^k \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t). \end{aligned} \quad (14)$$

Let us consider now:

$$\begin{aligned} \int_{\mathbb{R}^d} v[M_t^k] \cdot \nabla \phi dM_t^k &= \sum_{n=0}^{N_k-1} \left\{ \left( 1 - \frac{t - t_n^k}{\Delta t_k} \right) \int_{\mathbb{R}^d} v[M_t^k] \cdot \nabla \phi d\mu_n^k \right. \\ &\quad \left. + \frac{t - t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} v[M_t^k] \cdot \nabla \phi d\mu_{n+1}^k \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) \end{aligned}$$

and invoke [Assumption 1\(3\)](#) to get

$$\begin{aligned} &= \sum_{n=0}^{N_k-1} \int_{\mathbb{R}^d} v[\mu_n^k] \cdot \nabla \phi d\mu_n^k \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) \\ &\quad + \sum_{n=0}^{N_k-1} \left\{ \frac{t - t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} v[\mu_n^k] \cdot \nabla \phi d(\mu_{n+1}^k - \mu_n^k) \right. \\ &\quad \left. - \left( \frac{t - t_n^k}{\Delta t_k} \right)^2 \int_{\mathbb{R}^d} (v[\mu_{n+1}^k] - v[\mu_n^k]) \cdot \nabla \phi d\mu_n^k \right. \\ &\quad \left. + \left( \frac{t - t_n^k}{\Delta t_k} \right)^2 \int_{\mathbb{R}^d} (v[\mu_{n+1}^k] - v[\mu_n^k]) \cdot \nabla \phi d\mu_{n+1}^k \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t). \end{aligned}$$

Using this in 14 gives

$$\begin{aligned}
 \frac{d}{dt} \int_{\mathbb{R}^d} \phi dM_t^k - \int_{\mathbb{R}^d} v[M_t^k] \cdot \nabla \phi dM_t^k &= \sum_{n=0}^{N_k-1} \left\{ \frac{1}{2} \Delta t_k \int_{\mathbb{R}^d} (D^2 \phi(\bar{x}) v[\mu_n^k]) \cdot v[\mu_n^k] d\mu_n^k \right. \\
 &\quad - \frac{t-t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} v[\mu_n^k] \cdot \nabla \phi d(\mu_{n+1}^k - \mu_n^k) \\
 &\quad + \left( \frac{t-t_n^k}{\Delta t_k} \right)^2 \int_{\mathbb{R}^d} (v[\mu_{n+1}^k] - v[\mu_n^k]) \cdot \nabla \phi d\mu_n^k \\
 &\quad \left. - \left( \frac{t-t_n^k}{\Delta t_k} \right)^2 \int_{\mathbb{R}^d} (v[\mu_{n+1}^k] - v[\mu_n^k]) \cdot \nabla \phi d\mu_{n+1}^k \right\} \\
 &\quad \times \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t). \tag{15}
 \end{aligned}$$

Formally we can regard this expression as an equation satisfied by  $M_t^k$ . The remaining of this section is devoted to relate this equation to 5.

Set  $H := \|D^2 \phi\|_\infty = \text{Lip}(\nabla \phi)$  and notice that, owing to Assumption 1(1, 2), the function  $x \mapsto v[\mu_n^k](x) \cdot \nabla \phi(x)$  is Lipschitz continuous with

$$\text{Lip}(v \cdot \nabla \phi) \leq HV + \text{Lip}(v) \|\nabla \phi\|_\infty =: L. \tag{16}$$

Consequently:

$$\begin{aligned}
 \left| \frac{d}{dt} \int_{\mathbb{R}^d} \phi dM_t^k - \int_{\mathbb{R}^d} v[M_t^k] \cdot \nabla \phi dM_t^k \right| &\leq \sum_{n=0}^{N_k-1} \left\{ \frac{1}{2} HV^2 \Delta t_k \right. \\
 &\quad + \frac{t-t_n^k}{\Delta t_k} L \left| \int_{\mathbb{R}^d} \frac{v[\mu_n^k] \cdot \nabla \phi}{L} d(\mu_{n+1}^k - \mu_n^k) \right| \\
 &\quad \left. + 2 \left( \frac{t-t_n^k}{\Delta t_k} \right)^2 \|\nabla \phi\|_\infty W_1(\mu_n^k, \mu_{n+1}^k) \right\} \\
 &\quad \times \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) \\
 &\leq C \sum_{n=0}^{N_k-1} \left\{ \Delta t_k + \frac{t-t_n^k}{\Delta t_k} W_1(\mu_n^k, \mu_{n+1}^k) \right. \\
 &\quad \left. + \left( \frac{t-t_n^k}{\Delta t_k} \right)^2 W_1(\mu_n^k, \mu_{n+1}^k) \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t)
 \end{aligned}$$

where  $C := \max\{\frac{1}{2}HV^2, L, 2\|\nabla \phi\|_\infty\}$ . But  $W_1(\mu_n^k, \mu_{n+1}^k) \leq V\Delta t_k$ , therefore

$$\begin{aligned}
 &\leq C \sum_{n=0}^{N_k-1} \left\{ \Delta t_k + V(t-t_n^k) \right. \\
 &\quad \left. + V \frac{(t-t_n^k)^2}{\Delta t_k} \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t).
 \end{aligned}$$

Now, for  $t \in [t_n^k, t_{n+1}^k]$  it results  $t - t_n^k \leq \Delta t_k$ . In addition,  $\sum_{n=0}^{N_k-1} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) = \mathbb{1}_{[0, T]}(t)$ , hence finally

$$\leq C(1 + 2V)\Delta t_k \mathbb{1}_{[0, T]}(t).$$

Integrating 15 between 0 and  $t \leq T$  and using the last inequality yields

$$\begin{aligned} & \left| \int_0^t \left( \frac{d}{d\tau} \int_{\mathbb{R}^d} \phi dM_\tau^k - \int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi dM_\tau^k \right) d\tau \right| \\ & \leq \int_0^t \left| \frac{d}{d\tau} \int_{\mathbb{R}^d} \phi dM_\tau^k - \int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi dM_\tau^k \right| d\tau \leq C(1 + 2V)T\Delta t_k \end{aligned}$$

so that, by further manipulating the left-hand side and taking the limit for  $k \rightarrow \infty$ , we obtain

$$\lim_{k \rightarrow \infty} \left| \int_{\mathbb{R}^d} \phi dM_t^k - \int_{\mathbb{R}^d} \phi d\mu_t - \int_0^t \int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi dM_\tau^k d\tau \right| = 0. \tag{17}$$

In order to infer from 17 that  $\mu_t$  solves 5, we need the following convergence result.

**Lemma 3.3.** *When  $k \rightarrow \infty$  we have, up to subsequences,*

$$\int_{\mathbb{R}^d} \phi dM_t^k \rightarrow \int_{\mathbb{R}^d} \phi d\mu_t \quad \text{and} \quad \int_0^t \int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi dM_\tau^k d\tau \rightarrow \int_0^t \int_{\mathbb{R}^d} v[\mu_\tau] \cdot \nabla \phi d\mu_\tau d\tau$$

for all  $t \in [0, T]$ .

*Proof.* 1. For the first limit we write

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} \phi d\mu_t - \int_{\mathbb{R}^d} \phi dM_t^k \right| = \left| \int_{\mathbb{R}^d} \phi d(\mu_t - M_t^k) \right| \\ & = \text{Lip}(\phi) \left| \int_{\mathbb{R}^d} \frac{\phi}{\text{Lip}(\phi)} d(\mu_t - M_t^k) \right| \\ & \leq \text{Lip}(\phi) W_1(M_t^k, \mu_t), \end{aligned}$$

and, up to passing to a suitable subsequence of  $(M_\bullet^k)_{k \geq 0}$ , we conclude by applying Proposition 1.

2. For the second limit we preliminarily observe that  $|\int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi dM_\tau^k| \leq V \|\nabla \phi\|_\infty$ , thus, by dominated convergence,

$$\lim_{k \rightarrow \infty} \int_0^t \int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi dM_\tau^k d\tau = \int_0^t \left( \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi dM_\tau^k \right) d\tau. \tag{18}$$

Now

$$\begin{aligned} \left| \int_{\mathbb{R}^d} v[\mu_\tau] \cdot \nabla \phi \, d\mu_\tau - \int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi \, dM_\tau^k \right| &\leq \int_{\mathbb{R}^d} |v[\mu_\tau] - v[M_\tau^k]| \cdot |\nabla \phi| \, d\mu_\tau \\ &\quad + \left| \int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi \, d(\mu_\tau - M_\tau^k) \right| \\ &\leq (\|\nabla \phi\|_\infty \operatorname{Lip}(v) + L)W_1(M_\tau^k, \mu_\tau), \end{aligned}$$

$L$  being the constant defined in 16. Proposition 1 implies then  $\int_{\mathbb{R}^d} v[M_\tau^k] \cdot \nabla \phi \, dM_\tau^k \rightarrow \int_{\mathbb{R}^d} v[\mu_\tau] \cdot \nabla \phi \, d\mu_\tau$  for all  $\tau \in [0, T]$ , and the thesis follows from 18.  $\square$

Combining 17 and Lemma 3.3, and thanks to the arbitrariness of  $\phi$ , we obtain that  $\mu_\bullet$  solves problem 4 in the sense of Definition 2.1. In conclusion, we have proved:

**Theorem 3.4** (Existence). *Let Assumptions 1, 2 hold. Then there exists a weak solution  $\mu_\bullet \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  to the Cauchy problem 4.*

**4. Approximation of the solutions.** This section is devoted to a convergence analysis of the numerical scheme proposed in [9] for the approximation of the solutions to problem 4. We begin by sketching the main ideas which underlie the construction of the scheme, referring the reader to the above-cited paper for a detailed derivation. The scheme is obtained from a twofold approximation of problem 4, in time and in space. To this goal, we introduce a discretization of the time interval  $[0, T]$  by means of discrete instants  $t_n^k = n\Delta t_k$ , where the index  $n$  ranges from 0 to a value  $N_k$  such that  $N_k\Delta t_k = T$  and the time step  $\Delta t_k > 0$  tends to 0 when  $k \rightarrow \infty$ . By this discretization, we obtain the discrete-time dynamical system 8. Then, we introduce a space discretization in the following way. We define a pairwise disjoint partition of  $\mathbb{R}^d$  made of measurable elements  $E_i^k \in \mathcal{B}(\mathbb{R}^d)$ ,  $i = (i_1, \dots, i_d) \in \mathbb{Z}^d$ ,  $k \geq 0$ , such that

$$\bigcup_{i \in \mathbb{Z}^d} E_i^k = \mathbb{R}^d, \quad E_i^k \cap E_j^k = \emptyset \quad \forall i \neq j$$

for all  $k \geq 0$ . For the sake of simplicity, we assume that the  $E_i^k$ 's are hypercubes of edge length  $h_k > 0$ , with  $h_k \rightarrow 0$  when  $k \rightarrow \infty$ . Specifically,

$$E_i^k = \prod_{l=1}^d \left[ i_l - \frac{1}{2}, i_l + \frac{1}{2} \right) h_k.$$

Note that the index  $k$  identifies the level of refinement of the numerical grid. Using this space discretization, we approximate both the initial condition  $\bar{\mu}$  and the measures  $\mu_n^k$  by means of piecewise constant measures  $\lambda_n^k \ll \mathcal{L}^d$ ,  $n = 0, \dots, N_k$ ,  $k \geq 0$ . More precisely,  $d\lambda_n^k = \rho_n^k dx$  with

$$\rho_n^k(x) = \sum_{i \in \mathbb{Z}^d} \rho_i^n \mathbb{1}_{E_i^k}(x).$$

The space discretization makes it necessary to approximate also the flow map  $\gamma_n^k$  defined in 9. This is accomplished by the mapping

$$\tilde{\gamma}_n^k(x) = x + \tilde{v}_n^k(x)\Delta t_k, \quad \tilde{v}_n^k(x) = \sum_{i \in \mathbb{Z}^d} v[\lambda_n^k](x_i^k) \mathbb{1}_{E_i^k}(x),$$

$x_i^k$  being a point of the grid cell  $E_i^k$  (e.g., its center  $x_i^k = ih_k$ ). In practice, the velocity  $v[\mu_n^k](x)$  is approximated by the piecewise constant field  $\tilde{v}_n^k(x)$  taking in  $E_i^k$  the value that  $v$ , computed w.r.t. the measure  $\lambda_n^k$ , takes in  $x_i^k$ . Notice that  $|\tilde{v}_n^k(x)| \leq V$  for all  $x \in \mathbb{R}^d$ , all  $n = 0, \dots, N_k$ , and all  $k \geq 0$  because of Assumption 1(1).

By imposing  $\lambda_{n+1}^k(E_i^k) = (\tilde{\gamma}_n^k \# \lambda_n^k)(E_i^k)$  for each  $i \in \mathbb{Z}^d$ , we deduce the following explicit-in-time scheme relating recursively the coefficients  $\{\rho_i^n\}$  at two successive time steps:

$$\rho_i^{n+1} = \frac{1}{h_k^d} \sum_{j \in \mathbb{Z}^d} \rho_j^n \mathcal{L}^d(E_j^k \cap (\tilde{\gamma}_n^k)^{-1}(E_i^k)), \quad (19)$$

where  $h_k^d$  is  $\mathcal{L}^d(E_i^k)$ . To start up the scheme one has to provide the coefficients  $\rho_i^0$ , that we obtain as

$$\rho_i^0 = \frac{1}{h_k^d} \bar{\mu}(E_i^k), \quad i \in \mathbb{Z}^d. \quad (20)$$

We are now ready to start the analysis of the proposed scheme. To this end, we begin with a simple property of the approximation measures.

**Lemma 4.1.** *For every  $n = 0, \dots, N_k$ ,  $k \geq 0$ , the  $\lambda_n^k$ 's are probability measures.*

*Proof.* 1. The claim is certainly true for  $\lambda_0^k$ , since 20 shows that  $\rho_i^0 \geq 0$  all  $i$  and furthermore

$$\lambda_0^k(\mathbb{R}^d) = \int_{\mathbb{R}^d} \rho_0^k(x) dx = h_k^d \sum_{i \in \mathbb{Z}^d} \rho_i^0 = \sum_{i \in \mathbb{Z}^d} \bar{\mu}(E_i^k) = \bar{\mu}(\mathbb{R}^d) = 1,$$

where we have used the  $\sigma$ -additivity of  $\bar{\mu}$ .

2. If we assume now that  $\lambda_n^k$  is a probability measure for a certain  $n$ , using 19 we get  $\rho_i^{n+1} \geq 0$  all  $i$  and moreover

$$\begin{aligned} \lambda_{n+1}^k(\mathbb{R}^d) &= h_k^d \sum_{i \in \mathbb{Z}^d} \rho_i^{n+1} = \sum_{i \in \mathbb{Z}^d} \sum_{j \in \mathbb{Z}^d} \rho_j^n \mathcal{L}^d(E_j^k \cap (\tilde{\gamma}_n^k)^{-1}(E_i^k)) \\ &= \sum_{j \in \mathbb{Z}^d} \rho_j^n \sum_{i \in \mathbb{Z}^d} \mathcal{L}^d(E_j^k \cap (\tilde{\gamma}_n^k)^{-1}(E_i^k)) \\ &= \sum_{j \in \mathbb{Z}^d} \rho_j^n \mathcal{L}^d(E_j^k \cap (\tilde{\gamma}_n^k)^{-1}(\mathbb{R}^d)) \\ &= h_k^d \sum_{j \in \mathbb{Z}^d} \rho_j^n = \lambda_n^k(\mathbb{R}^d) = 1, \end{aligned}$$

where we have used the fact that  $(\tilde{\gamma}_n^k)^{-1}(E_i^k)$  are pairwise disjoint and that set operations, e.g., union, commute with the inverse image of a function. By induction on  $n$ , and by the arbitrariness of  $k \geq 0$ , the claim follows.  $\square$

The next result shows that the link between two successive measures  $\lambda_n^k, \lambda_{n+1}^k$  is a push forward as defined by 1, provided we restrict test functions to *simple*



functions adapted to the spatial grid  $\{E_i^k\}_{i \in \mathbb{Z}^d}$ , i.e., piecewise constant functions  $s : \mathbb{R}^d \rightarrow \mathbb{R}$  of the form

$$s(x) = \sum_{i \in \mathbb{Z}^d} \alpha_i \mathbb{1}_{E_i^k}(x) \quad (\alpha_i \in \mathbb{R}).$$

**Lemma 4.2.** *Let  $s : \mathbb{R}^d \rightarrow \mathbb{R}$  be simple over the grid  $\{E_i^k\}_{i \in \mathbb{Z}^d}$ . Then*

$$\int_{\mathbb{R}^d} s \, d\lambda_{n+1}^k = \int_{\mathbb{R}^d} s \circ \tilde{\gamma}_n^k \, d\lambda_n^k.$$

*Proof.* We have

$$\begin{aligned} \int_{\mathbb{R}^d} s \, d\lambda_{n+1}^k &= h_k^d \sum_{i \in \mathbb{Z}^d} \alpha_i \rho_i^{n+1} = \sum_{i \in \mathbb{Z}^d} \alpha_i \sum_{j \in \mathbb{Z}^d} \rho_j^n \mathcal{L}^d(E_j^k \cap (\tilde{\gamma}_n^k)^{-1}(E_i^k)) \\ &= \sum_{j \in \mathbb{Z}^d} \rho_j^n \sum_{i \in \mathbb{Z}^d} \alpha_i \mathcal{L}^d(E_j^k \cap (\tilde{\gamma}_n^k)^{-1}(E_i^k)). \end{aligned}$$

Notice that  $s(\tilde{\gamma}_n^k(x)) = \sum_{i \in \mathbb{Z}^d} \alpha_i \mathbb{1}_{(\tilde{\gamma}_n^k)^{-1}(E_i^k)}(x)$ , i.e.,  $s \circ \tilde{\gamma}_n^k$  is piecewise constant over the sets  $(\tilde{\gamma}_n^k)^{-1}(E_i^k)$ , which form a pairwise disjoint partition of  $\mathbb{R}^d$ . Therefore, for any given  $A \in \mathcal{B}(\mathbb{R}^d)$ ,

$$\int_A s(\tilde{\gamma}_n^k(x)) \, dx = \sum_{i \in \mathbb{Z}^d} \int_{A \cap (\tilde{\gamma}_n^k)^{-1}(E_i^k)} s(\tilde{\gamma}_n^k(x)) \, dx = \sum_{i \in \mathbb{Z}^d} \alpha_i \mathcal{L}^d(A \cap (\tilde{\gamma}_n^k)^{-1}(E_i^k)).$$

For  $A = E_j^k$  this enables us to continue the previous computation as

$$\int_{\mathbb{R}^d} s \, d\lambda_{n+1}^k = \sum_{j \in \mathbb{Z}^d} \rho_j^n \int_{E_j^k} s(\tilde{\gamma}_n^k(x)) \, dx = \int_{\mathbb{R}^d} s \circ \tilde{\gamma}_n^k \, d\lambda_n^k$$

and to obtain the thesis. □

An immediate consequence of this lemma is the following result, that will be fundamental for the sequel.

**Lemma 4.3.** *For all Lipschitz continuous  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  we have*

$$\left| \int_{\mathbb{R}^d} \varphi \, d\lambda_{n+1}^k - \int_{\mathbb{R}^d} \varphi \circ \tilde{\gamma}_n^k \, d\lambda_n^k \right| \leq 2 \operatorname{Lip}(\varphi) \sqrt{d} h_k, \quad \forall n = 0, \dots, N_k.$$

*Proof.* We consider the simple function

$$s(x) = \sum_{i \in \mathbb{Z}^d} \varphi(x_i^k) \mathbb{1}_{E_i^k}(x)$$

and, using Lemma 4.2, we compute:

$$\begin{aligned}
\int_{\mathbb{R}^d} \varphi(x) d\lambda_{n+1}^k(x) &= \int_{\mathbb{R}^d} (\varphi(x) - s(x)) d\lambda_{n+1}^k + \int_{\mathbb{R}^d} s(x) d\lambda_{n+1}^k(x) \\
&= \int_{\mathbb{R}^d} (\varphi(x) - s(x)) d\lambda_{n+1}^k(x) + \int_{\mathbb{R}^d} [s(\tilde{\gamma}_n^k(x)) - \varphi(\tilde{\gamma}_n^k(x))] d\lambda_n^k(x) \\
&\quad + \int_{\mathbb{R}^d} \varphi(\tilde{\gamma}_n^k(x)) d\lambda_n^k(x) \\
&= \sum_{i \in \mathbb{Z}^d} \int_{E_i^k} (\varphi(x) - \varphi(x_i^k)) d\lambda_{n+1}^k(x) \\
&\quad + \sum_{i \in \mathbb{Z}^d} \int_{E_i^k} (\varphi(x_i^k) - \varphi(x)) d(\tilde{\gamma}_n^k \# \lambda_n^k)(x) \\
&\quad + \int_{\mathbb{R}^d} \varphi(\tilde{\gamma}_n^k(x)) d\lambda_n^k(x).
\end{aligned}$$

Notice that  $\tilde{\gamma}_n^k \# \lambda_n^k$  is a probability measure. The Lipschitz continuity of  $\varphi$  entails

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} \varphi(x) d\lambda_{n+1}^k(x) - \int_{\mathbb{R}^d} \varphi(\tilde{\gamma}_n^k(x)) d\lambda_n^k(x) \right| \\
&\leq \text{Lip}(\varphi) \sum_{i \in \mathbb{Z}^d} \left\{ \int_{E_i^k} |x - x_i^k| d\lambda_{n+1}^k(x) + \int_{E_i^k} |x_i^k - x| d(\tilde{\gamma}_n^k \# \lambda_n^k)(x) \right\}
\end{aligned}$$

whence, since  $|x - x_i^k| \leq \text{diam } E_i^k = \sqrt{d}h_k$  for all  $x \in E_i^k$ ,

$$\leq \text{Lip}(\varphi) \sqrt{d}h_k \sum_{i \in \mathbb{Z}^d} \{ \lambda_{n+1}^k(E_i^k) + (\tilde{\gamma}_n^k \# \lambda_n^k)(E_i^k) \} = 2 \text{Lip}(\varphi) \sqrt{d}h_k$$

and, due to the arbitrariness of  $\varphi$  and  $n$ , the thesis follows.  $\square$

To address the convergence of the measures  $\lambda_n^k$ , it is convenient to introduce the following linear-in-time interpolation:

$$\Lambda_t^k = \sum_{n=0}^{N_k-1} \left[ \left( 1 - \frac{t - t_n^k}{\Delta t_k} \right) \lambda_n^k + \frac{t - t_n^k}{\Delta t_k} \lambda_{n+1}^k \right] \mathbb{1}_{[t_n^k, t_{n+1}^k)}(t), \quad (21)$$

with  $\Lambda_t^k$  a probability measure for all  $t \in [0, T]$  and all  $k \geq 0$ . This definition is reminiscent of the definition of  $M_t^k$  (cf. Section 3).

In order to prove a convergence result about  $\Lambda_t^k$ , from now on we make the following assumption:

**Assumption 3.** The mesh parameters  $h_k, \Delta t_k$  satisfy

$$h_k = o(\Delta t_k) \quad \text{for } k \rightarrow \infty.$$

As a consequence, there exists a sequence  $(\beta_k)_{k \geq 0} \subset \mathbb{R}$ , with  $\beta_k > 0$  all  $k$  and  $\lim_{k \rightarrow \infty} \beta_k = 0$ , such that  $h_k = \beta_k \Delta t_k$ . By convergence, it further results  $\sup_{k \geq 0} h_k, \sup_{k \geq 0} \Delta t_k, \sup_{k \geq 0} \beta_k < +\infty$ .

4.1. **Regularity of  $\Lambda_{\bullet}^k$ .** This section is devoted to establish some regularity properties of the curves  $\Lambda_{\bullet}^k$ .

**Lemma 4.4.** *We have  $\lambda_n^k \in \mathcal{P}_1(\mathbb{R}^d)$  for all  $n = 0, \dots, N_k$  and all  $k \geq 0$ . In particular,*

$$\sup_{k \geq 0} \sup_{0 \leq n \leq N_k} \int_{\mathbb{R}^d} |x| d\lambda_n^k(x) < +\infty,$$

*i.e., first order moments of the  $\lambda_n^k$ 's are uniformly bounded.*

*Proof.* 1. We begin by proving an inequality for  $n = 0$ . Recalling 20, we have

$$\int_{\mathbb{R}^d} |x| d\lambda_0^k(x) = \sum_{i \in \mathbb{Z}^d} \rho_i^0 \int_{E_i^k} |x| dx = \sum_{i \in \mathbb{Z}^d} \frac{1}{h_k^d} \int_{E_i^k} |x| dx \bar{\mu}(E_i^k) = \int_{\mathbb{R}^d} s(x) d\bar{\mu}(x),$$

where  $s$  is the simple function  $s = \sum_{i \in \mathbb{Z}^d} \alpha_i \mathbb{1}_{E_i^k}$  with  $\alpha_i = \frac{1}{h_k^d} \int_{E_i^k} |x| dx$ . Thus

$$\begin{aligned} \left| \int_{\mathbb{R}^d} |x| d\lambda_0^k(x) - \int_{\mathbb{R}^d} |x| d\bar{\mu}(x) \right| &\leq \int_{\mathbb{R}^d} |s(x) - |x|| d\bar{\mu}(x) \\ &= \sum_{i \in \mathbb{Z}^d} \int_{E_i^k} |\alpha_i - |x|| d\bar{\mu}(x). \end{aligned}$$

Considering that  $|\alpha_i - |x|| = \left| \frac{1}{h_k^d} \int_{E_i^k} (|y| - |x|) dy \right| \leq \frac{1}{h_k^d} \int_{E_i^k} |x - y| dy$ , we further deduce

$$\leq \sum_{i \in \mathbb{Z}^d} \frac{1}{h_k^d} \int_{E_i^k} \int_{E_i^k} |x - y| dy d\bar{\mu}(x) \leq \sqrt{d} h_k,$$

where we have used that, in the double integral, both  $x$  and  $y$  are points of the same grid cell  $E_i^k$ , hence  $|x - y| \leq \text{diam } E_i^k = \sqrt{d} h_k$ . It follows

$$\int_{\mathbb{R}^d} |x| d\lambda_0^k(x) \leq \int_{\mathbb{R}^d} |x| d\bar{\mu}(x) + \sqrt{d} h_k,$$

whence the thesis for  $n = 0$ .

2. We obtain the general case  $0 < n \leq N_k$  by induction from this and Lemma 4.3. Choosing  $\varphi(x) = |x|$  in the latter yields

$$\begin{aligned} \int_{\mathbb{R}^d} |x| d\lambda_{n+1}^k(x) &\leq \int_{\mathbb{R}^d} |\tilde{\gamma}_n^k(x)| d\lambda_n^k(x) + 2\sqrt{d} h_k \\ &\leq \int_{\mathbb{R}^d} |x| d\lambda_n^k(x) + \Delta t_k \int_{\mathbb{R}^d} |\tilde{v}_n^k(x)| d\lambda_n^k(x) + 2\sqrt{d} h_k \\ &\leq \int_{\mathbb{R}^d} |x| d\lambda_n^k(x) + V \Delta t_k + 2\sqrt{d} h_k \end{aligned}$$

whence, summing telescopically and using  $h_k = \beta_k \Delta t_k$ ,

$$\begin{aligned} \int_{\mathbb{R}^d} |x| d\lambda_n^k(x) &\leq \int_{\mathbb{R}^d} |x| d\lambda_0^k(x) + n(V\Delta t_k + 2\sqrt{d}h_k) \\ &\leq \int_{\mathbb{R}^d} |x| d\bar{\mu}(x) + (V + 2\sqrt{d}\beta_k)t_n^k + \sqrt{d}h_k. \end{aligned}$$

The thesis now follows taking the supremum of both sides in  $n$  and  $k$  while considering that  $t_n^k \leq T$ .  $\square$

**Lemma 4.5.** *For all  $k \geq 0$ ,*

$$W_1(\Lambda_s^k, \Lambda_t^k) \leq (V + 2\sqrt{d}\bar{\beta}) |t - s|, \tag{22}$$

where  $\bar{\beta} := \sup_{k \geq 0} \beta_k$ , i.e., the curves  $\Lambda_\bullet^k \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  are Lipschitz continuous uniformly in  $k$ . Moreover,

$$\sup_{k \geq 0} \sup_{t \in [0, T]} \int_{\mathbb{R}^d} |x| d\Lambda_t^k(x) < +\infty.$$

*Proof.* 1. Fix  $k \geq 0$  and  $t \in [0, T]$ . There exists  $0 \leq n \leq N_k$  such that  $t \in [t_n^k, t_{n+1}^k]$ , hence

$$\int_{\mathbb{R}^d} |x| d\Lambda_t^k(x) = \left(1 - \frac{t - t_n^k}{\Delta t_k}\right) \int_{\mathbb{R}^d} |x| d\lambda_n^k(x) + \frac{t - t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} |x| d\lambda_{n+1}^k(x).$$

Owing to Lemma 4.4, we can find a uniform upper bound on the first moments of the  $\lambda_n^k$ 's: there exists a constant  $C > 0$ , independent of  $n$  and  $k$ , such that

$$\leq \left(1 - \frac{t - t_n^k}{\Delta t_k}\right) C + \frac{t - t_n^k}{\Delta t_k} C = C,$$

which says that  $\Lambda_t^k \in \mathcal{P}_1(\mathbb{R}^d)$  for all  $k \geq 0$  and all  $t \in [0, T]$  with uniformly bounded first moment.

2. We argue the Lipschitz continuity of the mapping  $t \mapsto \Lambda_t^k$  as in Lemma 3.2. In particular, for  $s, t \in [0, T]$  with  $s \leq t$  we obtain

$$\begin{aligned} W_1(\Lambda_s^k, \Lambda_t^k) &\leq \frac{t_{m+1}^k - s}{\Delta t_k} W_1(\lambda_m^k, \lambda_{m+1}^k) \\ &\quad + \sum_{j=m+1}^{n-1} W_1(\lambda_j^k, \lambda_{j+1}^k) + \frac{t - t_n^k}{\Delta t_k} W_1(\lambda_n^k, \lambda_{n+1}^k), \end{aligned} \tag{23}$$

where  $0 \leq m \leq n \leq N_k$  are such that  $s \in [t_m^k, t_{m+1}^k]$  and  $t \in [t_n^k, t_{n+1}^k]$ . In order to estimate  $W_1(\lambda_i^k, \lambda_{i+1}^k)$  for a generic  $0 \leq i \leq N_k$  we fix  $\varphi \in \text{Lip}_1(\mathbb{R}^d)$

and use Lemma 4.3:

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi d(\lambda_{i+1}^k - \lambda_i^k) &\leq \int_{\mathbb{R}^d} [\varphi(\tilde{\gamma}_n^k(x)) - \varphi(x)] d\lambda_n^k(x) + 2\sqrt{d}h_k \\ &\leq \int_{\mathbb{R}^d} |\tilde{\gamma}_n^k(x) - x| d\lambda_n^k(x) + 2\sqrt{d}h_k \\ &= \Delta t_k \int_{\mathbb{R}^d} |\tilde{v}_n^k(x)| d\lambda_n^k(x) + 2\sqrt{d}h_k \\ &\leq V\Delta t_k + 2\sqrt{d}h_k = (V + 2\sqrt{d}\beta_k)\Delta t_k. \end{aligned}$$

Consequently  $W_1(\lambda_i^k, \lambda_{i+1}^k) \leq C\Delta t_k$  all  $i$ , where  $C := V + 2\sqrt{d}\sup_{k \geq 0} \beta_k$ . Computing as in Lemma 3.2, in view of 23 this yields  $W_1(\Lambda_s^k, \Lambda_t^k) \leq C(t - s)$  and we have the thesis.  $\square$

**4.2. Limit equation for  $\Lambda_\bullet^k$ .** The next step is to find, similarly to what we did in Section 3.2, an equation satisfied by  $\Lambda_\bullet^k$  in which to pass to the limit  $k \rightarrow \infty$ . Fix a test function  $\phi \in C_c^\infty(\mathbb{R}^d)$  and notice that the mapping  $t \mapsto \int_{\mathbb{R}^d} \phi d\Lambda_t^k$  is Lipschitz continuous because so is the mapping  $t \mapsto \Lambda_t^k$  in view of Lemma 4.5. Then, owing to Rademacher’s Theorem, it is a.e. differentiable. Using expression 21, we find that its derivative is

$$\frac{d}{dt} \int_{\mathbb{R}^d} \phi d\Lambda_t^k = \frac{1}{\Delta t_k} \sum_{n=0}^{N_k-1} \int_{\mathbb{R}^d} \phi d(\lambda_{n+1}^k - \lambda_n^k) \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t). \tag{24}$$

Let us introduce now the function  $g_n^k : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,

$$g_n^k(x) = x + v[\lambda_n^k](x)\Delta t_k,$$

i.e., the flow map  $\gamma_n^k$  (cf. 9) computed w.r.t. the measure  $\lambda_n^k$  instead of  $\mu_n^k$ . Then

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} \phi d\Lambda_t^k &= \frac{1}{\Delta t_k} \sum_{n=0}^{N_k-1} \left\{ \int_{\mathbb{R}^d} \phi d\lambda_{n+1}^k - \int_{\mathbb{R}^d} \phi \circ g_n^k d\lambda_n^k \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) \\ &\quad + \frac{1}{\Delta t_k} \sum_{n=0}^{N_k-1} \int_{\mathbb{R}^d} (\phi \circ g_n^k - \phi) d\lambda_n^k \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t), \end{aligned}$$

whence, expanding  $\phi \circ g_n^k$  in the second term at the right-hand side with Lagrange’s reminder (cf. the analogous calculation performed in Section 3.2 for the function  $\phi \circ \gamma_n^k$ ),

$$\begin{aligned} &= \frac{1}{\Delta t_k} \sum_{n=0}^{N_k-1} \left\{ \int_{\mathbb{R}^d} \phi d\lambda_{n+1}^k - \int_{\mathbb{R}^d} \phi \circ g_n^k d\lambda_n^k \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) \\ &\quad + \sum_{n=0}^{N_k-1} \left\{ \int_{\mathbb{R}^d} v[\lambda_n^k] \cdot \nabla \phi d\lambda_n^k + \frac{1}{2} \Delta t_k \int_{\mathbb{R}^d} (D^2 \phi(\bar{x})v[\lambda_n^k]) \cdot v[\lambda_n^k] d\lambda_n^k \right\} \\ &\quad \times \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t), \end{aligned}$$

where  $\bar{x}$  is a point of the segment connecting  $x$  and  $x + v[\lambda_n^k]\Delta t_k$ .

On the other hand, computing as in Section 3.2 we find

$$\begin{aligned}
\int_{\mathbb{R}^d} v[\Lambda_t^k] \cdot \nabla \phi \, d\Lambda_t^k &= \sum_{n=0}^{N_k-1} \int_{\mathbb{R}^d} v[\lambda_n^k] \cdot \nabla \phi \, d\lambda_n^k \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) \\
&+ \sum_{n=0}^{N_k-1} \left\{ \frac{t-t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} v[\lambda_n^k] \cdot \nabla \phi \, d(\lambda_{n+1}^k - \lambda_n^k) \right. \\
&- \left( \frac{t-t_n^k}{\Delta t_k} \right)^2 \int_{\mathbb{R}^d} (v[\lambda_{n+1}^k] - v[\lambda_n^k]) \cdot \nabla \phi \, d\lambda_n^k \\
&\left. + \left( \frac{t-t_n^k}{\Delta t_k} \right)^2 \int_{\mathbb{R}^d} (v[\lambda_{n+1}^k] - v[\lambda_n^k]) \cdot \nabla \phi \, d\lambda_{n+1}^k \right\} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t),
\end{aligned}$$

hence finally

$$\begin{aligned}
\frac{d}{dt} \int_{\mathbb{R}^d} \phi \, d\Lambda_t^k - \int_{\mathbb{R}^d} v[\Lambda_t^k] \cdot \nabla \phi \, d\Lambda_t^k &= \frac{1}{\Delta t_k} \sum_{n=0}^{N_k-1} \left\{ \int_{\mathbb{R}^d} \phi \, d\lambda_{n+1}^k - \int_{\mathbb{R}^d} \phi \circ g_n^k \, d\lambda_n^k \right\} \\
&\times \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) \\
&+ \sum_{n=0}^{N_k-1} \left\{ \frac{1}{2} \Delta t_k \int_{\mathbb{R}^d} (D^2 \phi(\bar{x}) v[\lambda_n^k]) \cdot v[\lambda_n^k] \, d\lambda_n^k \right. \\
&- \frac{t-t_n^k}{\Delta t_k} \int_{\mathbb{R}^d} v[\lambda_n^k] \cdot \nabla \phi \, d(\lambda_{n+1}^k - \lambda_n^k) \\
&+ \left( \frac{t-t_n^k}{\Delta t_k} \right)^2 \int_{\mathbb{R}^d} (v[\lambda_{n+1}^k] - v[\lambda_n^k]) \cdot \nabla \phi \, d\lambda_n^k \\
&\left. - \left( \frac{t-t_n^k}{\Delta t_k} \right)^2 \int_{\mathbb{R}^d} (v[\lambda_{n+1}^k] - v[\lambda_n^k]) \cdot \nabla \phi \, d\lambda_{n+1}^k \right\} \\
&\times \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t).
\end{aligned}$$

This is formally an equation satisfied by  $\Lambda_\bullet^k$ . A derivation analogous to that performed in Section 3.2 gives

$$\begin{aligned}
\left| \frac{d}{dt} \int_{\mathbb{R}^d} \phi \, d\Lambda_t^k - \int_{\mathbb{R}^d} v[\Lambda_t^k] \cdot \nabla \phi \, d\Lambda_t^k \right| &\leq \frac{1}{\Delta t_k} \sum_{n=0}^{N_k-1} \left| \int_{\mathbb{R}^d} \phi \, d\lambda_{n+1}^k - \int_{\mathbb{R}^d} \phi \circ g_n^k \, d\lambda_n^k \right| \\
&\times \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) \\
&+ C \sum_{n=0}^{N_k-1} \{ \Delta t_k + W_1(\lambda_n^k, \lambda_{n+1}^k) \} \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t),
\end{aligned}$$

where  $C > 0$  is a constant depending only on  $V$ ,  $\text{Lip}(v)$ ,  $\text{Lip}(\nabla\phi) = \|D^2\phi\|_\infty$ . In addition, from 22 we know  $W_1(\lambda_n^k, \lambda_{n+1}^k) \leq (V + 2\sqrt{d\bar{\beta}})\Delta t_k$ , thus

$$\begin{aligned} &\leq \frac{1}{\Delta t_k} \sum_{n=0}^{N_k-1} \left| \int_{\mathbb{R}^d} \phi d\lambda_{n+1}^k - \int_{\mathbb{R}^d} \phi \circ g_n^k d\lambda_n^k \right| \\ &\quad \times \mathbb{1}_{[t_n^k, t_{n+1}^k]}(t) + C\Delta t_k, \end{aligned}$$

$C$  being a new constant including the previous one and  $V + 2\sqrt{d\bar{\beta}}$ .

To estimate the remaining term at the right-hand side, we need to adapt Lemma 4.3 to  $g_n^k$ .

**Lemma 4.6.** *For all Lipschitz continuous  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  we have*

$$\left| \int_{\mathbb{R}^d} \varphi d\lambda_{n+1}^k - \int_{\mathbb{R}^d} \varphi \circ g_n^k d\lambda_n^k \right| \leq C \text{Lip}(\varphi) \sqrt{dh_k}$$

where  $C := 2 + \text{Lip}(v) \sup_{k \geq 0} \Delta t_k$ .

*Proof.* It suffices to observe that

$$\begin{aligned} &\left| \int_{\mathbb{R}^d} \varphi(\tilde{\gamma}_n^k(x)) d\lambda_n^k(x) - \int_{\mathbb{R}^d} \varphi(g_n^k(x)) d\lambda_n^k(x) \right| \leq \text{Lip}(\varphi) \int_{\mathbb{R}^d} |\tilde{\gamma}_n^k(x) - g_n^k(x)| d\lambda_n^k(x) \\ &= \Delta t_k \text{Lip}(\varphi) \sum_{i \in \mathbb{Z}^d} \int_{E_i^k} |v[\lambda_n^k](x_i^k) - v[\lambda_n^k](x)| d\lambda_n^k(x) \\ &= \Delta t_k \text{Lip}(\varphi) \text{Lip}(v) \sum_{i \in \mathbb{Z}^d} \int_{E_i^k} |x_i^k - x| d\lambda_n^k(x) \\ &\leq \sqrt{dh_k} \Delta t_k \text{Lip}(\varphi) \text{Lip}(v). \end{aligned}$$

The thesis follows from this inequality, combining the triangle inequality and Lemma 4.3.  $\square$

With this result, we can further manipulate the previous inequality and obtain

$$\left| \frac{d}{dt} \int_{\mathbb{R}^d} \phi d\Lambda_t^k - \int_{\mathbb{R}^d} v[\Lambda_t^k] \cdot \nabla \phi d\Lambda_t^k \right| \leq C(\beta_k + \Delta t_k)$$

which, integrating both sides in time from 0 to  $t \leq T$ , implies

$$\begin{aligned} &\left| \int_0^t \left( \frac{d}{d\tau} \int_{\mathbb{R}^d} \phi d\Lambda_\tau^k - \int_{\mathbb{R}^d} v[\Lambda_\tau^k] \cdot \nabla \phi d\Lambda_\tau^k \right) d\tau \right| \\ &\leq \int_0^t \left| \frac{d}{d\tau} \int_{\mathbb{R}^d} \phi d\Lambda_\tau^k - \int_{\mathbb{R}^d} v[\Lambda_\tau^k] \cdot \nabla \phi d\Lambda_\tau^k \right| d\tau \leq CT(\beta_k + \Delta t_k) \end{aligned}$$

and finally, taking the limit  $k \rightarrow \infty$ ,

$$\lim_{k \rightarrow \infty} \left| \int_{\mathbb{R}^d} \phi d\Lambda_t^k - \int_{\mathbb{R}^d} \phi d\lambda_0^k - \int_0^t \int_{\mathbb{R}^d} v[\Lambda_\tau^k] \cdot \nabla \phi d\Lambda_\tau^k d\tau \right| = 0. \quad (25)$$

Thanks to 25, we are in a position to prove our convergence result of the numerical scheme.

**Theorem 4.7.** *Let Assumptions 1–3 hold and assume that the sequence  $(\Lambda_{\bullet}^k)_{k \geq 0}$  converges to some  $\mu_{\bullet}$  in  $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  when  $k \rightarrow \infty$ . Then  $\mu_{\bullet}$  is a weak solution to problem 4.*

*Proof.* Arguing as in Lemma 3.3, if  $\lim_{k \rightarrow \infty} \sup_{t \in [0, T]} W_1(\Lambda_t^k, \mu_t) = 0$  we know that, up to subsequences,

$$\int_{\mathbb{R}^d} \phi \, d\Lambda_t^k \rightarrow \int_{\mathbb{R}^d} \phi \, d\mu_t, \quad \int_0^t \int_{\mathbb{R}^d} v[\Lambda_{\tau}^k] \cdot \nabla \phi \, d\Lambda_{\tau}^k \, d\tau \rightarrow \int_0^t \int_{\mathbb{R}^d} v[\mu_{\tau}] \cdot \nabla \phi \, d\mu_{\tau} \, d\tau \quad (k \rightarrow \infty).$$

The claim of the theorem then follows from 25 and the arbitrariness of  $\phi$ , provided we have also  $\int_{\mathbb{R}^d} \phi \, d\lambda_0^k \rightarrow \int_{\mathbb{R}^d} \phi \, d\bar{\mu}$  when  $k \rightarrow \infty$ . To show the latter fact, it is enough to prove that  $W_1(\lambda_0^k, \bar{\mu}) \rightarrow 0$  when  $k \rightarrow \infty$ .

Recall that  $d\lambda_0^k = \rho_0^k \, dx$ , with  $\rho_0^k$  given by 20. Fixing  $\varphi \in \text{Lip}_1(\mathbb{R}^d)$  and reasoning like in the proof of Lemma 4.4(1), we find that

$$\int_{\mathbb{R}^d} \varphi \, d\lambda_0^k = \int_{\mathbb{R}^d} s \, d\bar{\mu},$$

where  $s(x) = \sum_{i \in \mathbb{Z}^d} \alpha_i \mathbb{1}_{E_i^k}(x)$  and  $\alpha_i = \frac{1}{h_k^d} \int_{E_i^k} \varphi(x) \, dx$ . Therefore:

$$\int_{\mathbb{R}^d} \varphi \, d(\bar{\mu} - \lambda_0^k) = \int_{\mathbb{R}^d} (\varphi - s) \, d\bar{\mu} = \sum_{i \in \mathbb{Z}^d} \int_{E_i^k} (\varphi(x) - \alpha_i) \, d\bar{\mu}(x).$$

But  $|\varphi(x) - \alpha_i| \leq \frac{1}{h_k^d} \int_{E_i^k} |\varphi(x) - \varphi(y)| \, dy \leq \frac{1}{h_k^d} \int_{E_i^k} |x - y| \, dy$ , so that from the previous calculations we deduce

$$\int_{\mathbb{R}^d} \varphi \, d(\bar{\mu} - \lambda_0^k) \leq \frac{1}{h_k^d} \sum_{i \in \mathbb{Z}^d} \int_{E_i^k} \int_{E_i^k} |x - y| \, dy \, d\bar{\mu}(x) \leq \sqrt{d} h_k,$$

whence  $W_1(\lambda_0^k, \bar{\mu}) \leq \sqrt{d} h_k \rightarrow 0$  for  $k \rightarrow \infty$  as desired. □

We conclude this section with a simple criterion which implies the convergence property assumed in Theorem 4.7.

**Corollary 1.** *Assume there exists a bounded set  $K \subset \mathbb{R}^d$  such that  $\text{supp } \lambda_n^k \subseteq K$  for all  $0 \leq n \leq N_k$  and all  $k \geq 0$ . Then  $(\Lambda_{\bullet}^k)_{k \geq 0}$  converges in  $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$  to a weak solution of problem 4 when  $k \rightarrow \infty$ .*

*Proof.* Fix  $k \geq 0$ ,  $t \in [0, T]$  and let  $n$  be such that  $t_n^k \leq t \leq t_{n+1}^k$ . Let moreover  $A \in \mathcal{B}(\mathbb{R}^d)$  be contained in  $K^c$ , then  $\lambda_i^k(A) = 0$  for all  $i, k$  and therefore

$$\Lambda_t^k(A) = \left(1 - \frac{t - t_n^k}{\Delta t_k}\right) \lambda_n^k(A) + \frac{t - t_n^k}{\Delta t_k} \lambda_{n+1}^k(A) = 0.$$

From the arbitrariness of  $k, t$  it follows  $\text{supp } \Lambda_t^k \subseteq K$  for all  $k \geq 0$  and all  $t \in [0, T]$ . Since  $K$  is bounded, we can find a ball  $B_R(0)$  with radius  $R > 0$  so large that  $K \subseteq B_R(0)$ . Consequently

$$\int_{\mathbb{R}^d} |x|^p \, d\Lambda_t^k(x) = \int_K |x|^p \, d\Lambda_t^k(x) \leq R^p < +\infty,$$



i.e., the  $\Lambda_t^k$ 's have uniformly bounded moments of any order  $p \geq 0$ . This implies that  $\{\Lambda_t^k\}_{k \geq 0}$  is relatively compact in  $\mathcal{P}_1(\mathbb{R}^d)$  for all  $t \in [0, T]$ , which, together with the equicontinuity of the family  $\{\Lambda_{\bullet}^k\}_{k \geq 0}$  entailed by the estimate 22, allows us to apply Ascoli-Arzelà's Theorem and conclude that the sequence  $(\Lambda_{\bullet}^k)_{k \geq 0}$  is relatively compact in  $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ . Thus, up to subsequences, it converges to some  $\mu_{\bullet} \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ , which is a weak solution to problem 4 because of Theorem 4.7.  $\square$

**Remark 1** (CFL condition). Let us define  $\alpha_k := V/\beta_k$  for all  $k \geq 0$ . Then the condition expressed by Assumption 3 is equivalent to

$$V\Delta t_k = \alpha_k h_k, \tag{26}$$

which implies  $|\tilde{\gamma}_n^k(x) - x| \leq \alpha_k h_k$  all  $x \in \mathbb{R}^d$ , i.e., the displacements produced by the mapping  $\tilde{\gamma}_n^k$  are bounded by the quantity  $\alpha_k h_k$ . As a consequence, the number of nonempty intersections  $\tilde{\gamma}_n^k(E_j^k) \cap E_i^k$ ,  $j \in \mathbb{Z}^d$ , is finite and bounded from above uniformly in  $i$ , which in particular makes the series in 19 actually a finite sum for all  $i$ .

We observe that condition 26 allows for  $\alpha_k \rightarrow +\infty$  when  $k \rightarrow \infty$ , in order to ensure convergence to continuous-in-time-and-space solutions as stated by Theorem 4.7. With the time step  $\Delta t$  frozen, 26 is the classical Courant-Friedrichs-Lewy (CFL) condition, used in [9] to prove the stability of the numerical scheme 19 in approximating the solutions to the *discrete-in-time* model 8.

**5. Models of crowd and swarm dynamics.** In [5, 8] a class of models describing the collective dynamics of swarms and crowds has been introduced. These models are based on the idea that each individual of the group is an *intelligent* (or *active*) agent, able to develop a behavioral strategy for pursuing specific goals. Since agents are not passively dragged by external forces, the Newtonian approach is abandoned in favor of a kinematic one, in which the velocity of the agents stems from few basic behavioral rules. These concepts are formalized in [6], where it is shown that the above-cited models can be given a common formulation within the framework of 3. In particular, the corresponding velocity of the agents is

$$v[\mu_t](x) = v_d(x) + N \int_{\mathbb{R}^d} f(|y - x|)r(y - x)\chi_{U_x}(y) d\mu_t(y). \tag{27}$$

Adopting the terminology of the kinetic theory for active particles, see [2], we call *test agent* an agent potentially concerned with interactions, whose position is described in the above formula by the variable  $x$ , and *field agents* the agents distributed in space (variable  $y$ ), which the test agent may interact with. In more detail:

1.  $v_d : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the test agent's *desired velocity*, i.e., the velocity of an isolated test agent in the absence of interactions;
2.  $U_x$  is the *interaction neighborhood* of the test agent. It conveys the idea that the test agent experiences *nonlocal* interactions with some selected field agents, namely those inside  $U_x$ .
3.  $\chi_A : \mathbb{R}^d \rightarrow \mathbb{R}$  is a cut-off function for the set  $A \subseteq \mathbb{R}^d$ , such that

$$\chi_A(x) = 0 \quad \forall x \in A^c, \quad 0 \leq \chi_A(x) \leq 1 \quad \forall x \in \mathbb{R}^d.$$

In particular, the effect of the term  $\chi_{U_x}$  in 27 is to rule out interactions of the test agent with field agents outside  $U_x$ ;

4.  $r : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the *direction of the interaction*, depending on the relative position of the interacting agents and oriented in such a way that  $r(y-x) \cdot (y-x) \geq 0$ , with in addition  $|r| \leq 1$ ;
5.  $f : [0, +\infty) \rightarrow \mathbb{R}$  is the *interaction strength*, which depends on the distance between the interacting agents. If  $f < 0$  the interaction is *repulsive*, i.e., the test agent tries to avoid local aggregation with the field agents (as it is common in crowds under normal – i.e., non-panic – conditions). If  $f > 0$  the interaction is *attractive*, i.e., the test agent aims at staying close to the surrounding field agents (as in swarms, where group cohesion is advantageous e.g., for food search or predator avoidance).

The shape of  $U_x$  depends partly on the criteria used for selecting the field agents to interact with. In general, when  $U_x$  is symmetric w.r.t.  $x$ , interactions are said to be *isotropic*, as opposed to *anisotropic* when  $U_x$  is not symmetric. Notice also that the interaction integral in 27 is actually computed w.r.t. the mass measure  $N\mu_t$ , because interactions depend on the number of field agents in  $U_x$ .

Prototypes of the above objects are, among others:  $U_x = B_R(x)$  for isotropic interactions,  $U_x = S_R^\alpha(x)$  for anisotropic interactions ( $S_R^\alpha(x)$  being the sector of  $B_R(x)$  with angular width  $\alpha \in (0, 2\pi]$ ),  $\chi_{U_x} = \mathbb{1}_{U_x}$ ,  $f(|y-x|) = -1/|y-x|$  for repulsive interactions,  $f(|y-x|) = |y-x|$  for attractive interactions,  $r(y-x) = (y-x)/|y-x|$  (i.e., the unit vector in the direction of  $y-x$ ). Although meaningful from the modeling point of view, some of these choices need to be adapted in order for the velocity field 27 to comply with Assumption 1.

Concerning this, let us introduce the vector-valued function  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,

$$F(x) := f(|x|)r(x),$$

so that 27 becomes

$$v[\mu_t](x) = v_d(x) + N \int_{\mathbb{R}^d} F(y-x)\chi_{U_x}(y) d\mu_t(y). \quad (28)$$

Such a form encompasses a broader class of models of coordinated behavior based on 3, for instance those in [4]. From now on, we will take 28 as our velocity model.

**Assumption 4** (Properties of  $v$  as in 28). We assume that the velocity field in 28 satisfies the following properties.

1.  $v_d$  is Lipschitz continuous and bounded in  $\mathbb{R}^d$ .
2. There exists  $R > 0$  such that, for each  $x \in \mathbb{R}^d$ , the set  $U_x \subseteq B_R(x)$  is measurable and isometric to a reference set  $U_0 \subseteq B_R(0)$ .
3.  $F$  is Lipschitz continuous in  $B_R(0)$ .
4.  $\chi_A$  is Lipschitz continuous in  $\mathbb{R}^d$  for each measurable  $A \subseteq \mathbb{R}^d$ .

**Remark 2.** Some comments on Assumption 4 are in order.

1. In the sequel we denote

$$V_d := \sup_{x \in \mathbb{R}^d} |v_d(x)| < +\infty.$$

2. Saying that  $U_x$  is isometric to  $U_0$  means that there is a rigid transformation  $\xi_x : \mathbb{R}^d \rightarrow \mathbb{R}^d$  mapping  $U_0$  onto  $U_x$ :  $\xi_x(U_0) = U_x$ . Specifically,  $\xi_x$  has the form

$$\xi_x(z) = \mathcal{R}_x z + x, \quad (29)$$

where  $\mathcal{R}_x \in \mathbb{R}^{d \times d}$  is a rotation matrix possibly depending on the point  $x$ .

3. Lipschitz continuity of  $F$  is required *only* in the ball  $B_R(0)$ , not in the whole space, i.e., the condition  $|F(z_2) - F(z_1)| \leq \text{Lip}(F) |z_2 - z_1|$  must hold only for  $|z_1|, |z_2| \leq R$ . This implies that  $F$  is bounded in  $B_R(0)$ , that is,  $|F(z)| \leq C_F$  for  $|z| \leq R$ , with  $C_F := \max\{|F(0)|, \text{Lip}(F)\}(1 + R)$ .
4. The function  $\chi_A$  can be thought of as a mollification of  $\mathbb{1}_A$ . In particular, by continuity  $\chi_A = 0$  on  $\partial A$ . Because of the isometry between  $U_0$  and  $U_x$ , one can check that the following relation holds:

$$\chi_{U_x}(y) = \chi_{U_0}(\xi_x^{-1}(y)), \quad \forall y \in \mathbb{R}^d. \tag{30}$$

In the next two sections we show that, with both isotropic and anisotropic interactions, Assumption 4 is sufficient to apply our existence and approximation results to the above-mentioned crowd and swarm models. First, we consider the case of a spherical neighborhood, which is the most important prototype for isotropic interactions. Later, we extend the analysis to any bounded neighborhood, including the significant cases of anisotropic interactions.

**5.1. Spherical interaction neighborhood.** In case of isotropic interactions we set  $U_x = B_R(x)$ , and consequently  $U_0 = B_R(0)$ . Since  $B_R(x)$  is invariant under rotations, the mapping  $\xi_x$  is simply  $\xi_x(z) = z + x$ .

By the change of variables  $z = \xi_x^{-1}(y)$  in 28, and recalling furthermore 30, we can rewrite the velocity as

$$v[\mu_t](x) = v_d(x) + N \int_{\mathbb{R}^d} F(z) \chi_{B_R(0)}(z) d(\xi_x^{-1} \# \mu_t)(z),$$

where  $\xi_x^{-1} \# \mu_t$  is in  $\mathcal{P}_1(\mathbb{R}^d)$  for all  $\mu_t \in \mathcal{P}_1(\mathbb{R}^d)$  and all fixed  $x$ .

We state a preliminary result, which is useful to verify that such a velocity field complies with Assumption 1.

**Lemma 5.1.** *The function  $z \mapsto F(z) \chi_{B_R(0)}(z)$  is Lipschitz continuous in  $\mathbb{R}^d$ .*

*Proof.* To study the expression

$$e(z_1, z_2) := |F(z_2) \chi_{B_R(0)}(z_2) - F(z_1) \chi_{B_R(0)}(z_1)|, \quad z_1, z_2 \in \mathbb{R}^d$$

it is convenient to distinguish three cases.

1. If  $z_1, z_2 \in B_R(0)$  we have

$$\begin{aligned} e(z_1, z_2) &\leq |F(z_2) \chi_{B_R(0)}(z_2) - F(z_2) \chi_{B_R(0)}(z_1)| \\ &\quad + |F(z_2) \chi_{B_R(0)}(z_1) - F(z_1) \chi_{B_R(0)}(z_1)| \\ &\leq (C_F \text{Lip}(\chi_{B_R(0)}) + \text{Lip}(F)) |z_2 - z_1|. \end{aligned}$$

2. If  $z_1 \notin B_R(0)$  and  $z_2 \in B_R(0)$  (or vice versa),  $e(z_1, z_2) = |F(z_2) \chi_{B_R(0)}(z_2)|$ . Let  $z_\theta := \theta z_1 + (1 - \theta) z_2$ ,  $\theta \in [0, 1]$ , be a point of the segment connecting  $z_1$  to  $z_2$  and pick  $\bar{\theta}$  such that  $z_{\bar{\theta}} \in \partial B_R(0)$ . Since  $\chi_{B_R(0)}(z_{\bar{\theta}}) = 0$ , we have

$$\begin{aligned} e(z_1, z_2) &= |F(z_2) \chi_{B_R(0)}(z_2) - F(z_{\bar{\theta}}) \chi_{B_R(0)}(z_{\bar{\theta}})| \\ &\leq (C_F \text{Lip}(\chi_{B_R(0)}) + \text{Lip}(F)) |z_2 - z_{\bar{\theta}}|. \end{aligned}$$

On the other hand,  $|z_2 - z_{\bar{\theta}}| = \bar{\theta} |z_2 - z_1| \leq |z_2 - z_1|$ .

3. If  $z_1, z_2 \notin B_R(0)$  then  $e(z_1, z_2) = 0 \leq |z_2 - z_1|$ . □

**Proposition 2** (Velocity with isotropic interactions). *Let Assumption 4 hold with  $U_0 = B_R(0)$ . Then the velocity field 28 complies with Assumption 1.*

*Proof.* In the sequel,  $x \in \mathbb{R}^d$  and  $\mu \in \mathcal{P}_1(\mathbb{R}^d)$  are fixed but arbitrary. We verify the items of Assumption 1 in order.

1. The field  $v$  is uniformly bounded thanks to

$$\begin{aligned} |v[\mu](x)| &\leq V_d + N \int_{\mathbb{R}^d} |F(z)\chi_{B_R(0)}(z)| d(\xi_x^{-1}\#\mu)(z) \\ &\leq V_d + NC_F\mu(\xi_x(\mathbb{R}^d)) = V_d + NC_F. \end{aligned}$$

2. As for Lipschitz continuity, we check it separately w.r.t. to  $x$  and to  $\mu$ .

(a) Let  $x_1, x_2 \in \mathbb{R}^d$ , then, recalling Lemma 5.1,

$$\begin{aligned} |v[\mu](x_2) - v[\mu](x_1)| &\leq |v_d(x_2) - v_d(x_1)| \\ &\quad + N \left| \int_{\mathbb{R}^d} F(z)\chi_{B_R(0)}(z) d(\xi_{x_2}^{-1}\#\mu - \xi_{x_1}^{-1}\#\mu)(z) \right| \\ &\leq \text{Lip}(v_d) |x_2 - x_1| \\ &\quad + N \text{Lip}(F\chi_{B_R(0)})W_1(\xi_{x_1}^{-1}\#\mu, \xi_{x_2}^{-1}\#\mu). \end{aligned}$$

In addition,

$$\begin{aligned} W_1(\xi_{x_1}^{-1}\#\mu, \xi_{x_2}^{-1}\#\mu) &= \sup_{\varphi \in \text{Lip}_1(\mathbb{R}^d)} \int_{\mathbb{R}^d} (\varphi \circ \xi_{x_2}^{-1} - \varphi \circ \xi_{x_1}^{-1}) d\mu \\ &\leq \int_{\mathbb{R}^d} |\xi_{x_2}^{-1}(y) - \xi_{x_1}^{-1}(y)| d\mu(y) = |x_2 - x_1|, \end{aligned}$$

whence the Lipschitz continuity of  $x \mapsto v[\mu](x)$ .

(b) Let now  $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$ , then, invoking again Lemma 5.1,

$$\begin{aligned} |v[\nu](x) - v[\mu](x)| &= N \left| \int_{\mathbb{R}^d} F(z)\chi_{B_R(0)}(z) d(\xi_x^{-1}\#\nu - \xi_x^{-1}\#\mu)(z) \right| \\ &\leq N \text{Lip}(F\chi_{B_R(0)})W_1(\xi_x^{-1}\#\mu, \xi_x^{-1}\#\nu) \\ &= N \text{Lip}(F\chi_{B_R(0)}) \sup_{\varphi \in \text{Lip}_1(\mathbb{R}^d)} \int_{\mathbb{R}^d} \varphi \circ \xi_x^{-1} d(\nu - \mu) \\ &= N \text{Lip}(F\chi_{B_R(0)})W_1(\mu, \nu), \end{aligned}$$

since  $\xi_x^{-1} \in \text{Lip}_1(\mathbb{R}^d)$  implies that  $\varphi \circ \xi_x^{-1}$  spans the whole space  $\text{Lip}_1(\mathbb{R}^d)$  when  $\varphi$  varies in  $\text{Lip}_1(\mathbb{R}^d)$ . Hence also  $\mu \mapsto v[\mu](x)$  is Lipschitz continuous.

3. Finally we examine the linearity w.r.t. to the probability for convex combinations. Let  $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$  and  $\alpha \in [0, 1]$ , then

$$\begin{aligned} v[\alpha\mu + (1 - \alpha)\nu](x) &= v_d(x) + \alpha N \int_{\mathbb{R}^d} F(z)\chi_{B_R(0)}(z) d(\xi_x^{-1}\#\mu)(z) \\ &\quad + (1 - \alpha)N \int_{\mathbb{R}^d} F(z)\chi_{B_R(0)}(z) d(\xi_x^{-1}\#\nu)(z). \end{aligned}$$

Writing  $v_d(x) = \alpha v_d(x) + (1 - \alpha)v_d(x)$  and collecting the coefficients  $\alpha$  and  $1 - \alpha$  gives the result.  $\square$

Owing to Proposition 2, we can state that models based on 3, with a velocity field featuring isotropic interactions as in 28, see e.g., [8], admit probability measure solutions for any initial distribution of the agents with finite first and second order moments (e.g., an initial distribution with compact support). Moreover, such solutions can be duly approximated using the numerical scheme 19.

**Remark 3** (Unbounded interaction neighborhood). As a modification to Assumption 4, we can allow  $U_0 = U_x = \mathbb{R}^d$  along with the boundedness assumption  $F(x) \leq C_F$  for all  $x \in \mathbb{R}^d$ . The above arguments can be promptly adapted to show that also in this case the velocity field 28 complies with Assumption 1. However, it should be noted that, for the applications we have in mind, the physically relevant cases are those in which the interaction neighborhood is bounded.

**5.2. Bounded interaction neighborhood.** In this section we drop the specific hypothesis that  $U_x$  be a ball. We allow it to have a generic shape, with the only requirement of being bounded, i.e., contained in a ball. This encompasses the important case of anisotropic interactions. In such cases, the neighborhood  $U_x$  may not be invariant under rotations, therefore we have to consider the full form 29 of the transformation  $\xi_x$ . Performing again the change of variables  $z = \xi_x^{-1}(y)$  in the integral 28, the velocity takes now the form

$$v[\mu_t](x) = v_d(x) + N \int_{\mathbb{R}^d} F(\mathcal{R}_x z) \chi_{U_0}(z) d(\xi_x^{-1} \# \mu_t)(z)$$

with the rotation matrix explicitly appearing in the argument of the function  $F$ . To deal with it, for the sake of simplicity we confine ourselves to the two-dimensional case ( $d = 2$ ), for then a simple representation of  $\mathcal{R}_x$  is available:

$$\mathcal{R}_x = \begin{pmatrix} \cos \vartheta_x & -\sin \vartheta_x \\ \sin \vartheta_x & \cos \vartheta_x \end{pmatrix},$$

$\vartheta_x \in [0, 2\pi)$  being the angle of rotation which determines the local orientation in the point  $x$  of the neighborhood of interaction.

The choice of  $\vartheta_x$  has to do with the way in which the anisotropy of the interactions is modeled. In the models of crowd and swarm dynamics we are considering,  $\vartheta_x$  is the angle formed by the desired velocity  $v_d(x)$  w.r.t. a fixed reference direction, for instance the horizontal one identified by the unit vector  $\mathbf{i}$ . Assuming for simplicity that  $v_d$  has constant unit modulus, this implies

$$\cos \vartheta_x = v_d(x) \cdot \mathbf{i}, \quad \sin \vartheta_x = (v_d(x) \times \mathbf{i}) \cdot \mathbf{k}. \tag{31}$$

In the second formula,  $v_d(x)$  and  $\mathbf{i}$  are thought of as embedded into  $\mathbb{R}^3$ , with  $\times$  denoting vector product and  $\mathbf{k}$  the unit vector orthogonal to the plane of  $v_d(x)$  and  $\mathbf{i}$ .

**Remark 4.** More in general, 31 holds with  $v_d(x)$  replaced by  $v_d(x)/|v_d(x)|$ , which is a Lipschitz continuous field provided  $|v_d|$  is uniformly bounded away from zero.

The formalism introduced above allows us to prove the following technical fact.

**Lemma 5.2.** *For all  $x_1, x_2, z \in \mathbb{R}^2$  it results*

$$|(\mathcal{R}_{x_2} - \mathcal{R}_{x_1})z| \leq \sqrt{2} \text{Lip}(v_d) |x_2 - x_1| |z|.$$

*The same holds with  $\mathcal{R}_{x_1}, \mathcal{R}_{x_2}$  replaced by  $\mathcal{R}_{x_1}^{-1}, \mathcal{R}_{x_2}^{-1}$ .*

*Proof.* A straightforward computation shows that

$$|(\mathcal{R}_{x_2} - \mathcal{R}_{x_1})z| = \sqrt{(\cos \vartheta_{x_2} - \cos \vartheta_{x_1})^2 + (\sin \vartheta_{x_2} - \sin \vartheta_{x_1})^2} |z|,$$

and the same is true also using inverse matrices. In addition,

$$\begin{aligned} |\cos \vartheta_{x_2} - \cos \vartheta_{x_1}| &= |(v_d(x_2) - v_d(x_1)) \cdot \mathbf{i}| \leq |v_d(x_2) - v_d(x_1)|, \\ |\sin \vartheta_{x_2} - \sin \vartheta_{x_1}| &= |[(v_d(x_2) - v_d(x_1)) \times \mathbf{i}] \cdot \mathbf{k}| \leq |v_d(x_2) - v_d(x_1)|, \end{aligned}$$

hence the thesis follows from the Lipschitz continuity of  $v_d$ .  $\square$

With Lemma 5.2 we are in a position to prove that the velocity 28 complies with Assumption 1 also in case of anisotropic interactions.

**Proposition 3** (Velocity with anisotropic interactions). *Let Assumption 4 hold and assume  $d = 2$ . Then the velocity field 28 complies with Assumption 1.*

*Proof.* 1. Uniform boundedness, Lipschitz continuity w.r.t. the probability, and linearity w.r.t. to the probability for convex combinations follow straightforwardly from calculations entirely analogous to those performed in the proof of Proposition 2. In fact, it is sufficient to observe that,  $\mathcal{R}_x$  being an isometry, the function  $F(\mathcal{R}_x \cdot)$  is Lipschitz continuous and bounded in  $B_R(0)$  with  $\text{Lip}(F(\mathcal{R}_x \cdot)) = \text{Lip}(F)$  for all  $x \in \mathbb{R}^d$ . Moreover, by the same argument as in the proof of Lemma 5.1, the function  $F(\mathcal{R}_x \cdot) \chi_{U_0}(\cdot)$  is Lipschitz continuous in  $\mathbb{R}^d$  with the same Lipschitz constant as  $F \chi_{U_0}$ , thus in particular independent of  $x$ .

2. Lipschitz continuity w.r.t. to  $x$  is instead more delicate, because it involves directly the rotation matrix  $\mathcal{R}_x$ . Let  $x_1, x_2 \in \mathbb{R}^2$  and fix  $\mu \in \mathcal{P}_1(\mathbb{R}^2)$ , then

$$\begin{aligned} |v[\mu](x_2) - v[\mu](x_1)| &\leq |v_d(x_2) - v_d(x_1)| \\ &\quad + N \left| \int_{\mathbb{R}^d} F(\mathcal{R}_{x_2} z) \chi_{U_0}(z) d(\xi_{x_2}^{-1} \# \mu)(z) \right. \\ &\quad \left. - \int_{\mathbb{R}^d} F(\mathcal{R}_{x_1} z) \chi_{U_0}(z) d(\xi_{x_1}^{-1} \# \mu)(z) \right| \\ &\leq \text{Lip}(v_d) |x_2 - x_1| \\ &\quad + N \int_{\mathbb{R}^d} |F(\mathcal{R}_{x_2} z) - F(\mathcal{R}_{x_1} z)| \chi_{U_0}(z) d(\xi_{x_2}^{-1} \# \mu)(z) \\ &\quad + N \left| \int_{\mathbb{R}^d} F(\mathcal{R}_{x_1} z) \chi_{U_0}(z) d(\xi_{x_2}^{-1} \# \mu - \xi_{x_1}^{-1} \# \mu)(z) \right|. \quad (32) \end{aligned}$$

In the first integral at the right-hand side of 32 we can assume  $z \in U_0$ , for otherwise  $\chi_{U_0}(z) = 0$ . Hence  $|z| \leq R$  and moreover  $\mathcal{R}_x z \in B_R(0)$  for all  $x \in \mathbb{R}^2$ . Lipschitz continuity of  $F$  in that ball, along with Lemma 5.2, implies

$$\begin{aligned} |F(\mathcal{R}_{x_2} z) - F(\mathcal{R}_{x_1} z)| &\leq \text{Lip}(F) |(\mathcal{R}_{x_1} - \mathcal{R}_{x_2})z| \\ &\leq \sqrt{2} \text{Lip}(F) \text{Lip}(v_d) R |x_2 - x_1|, \end{aligned}$$

so that finally

$$\begin{aligned} N \int_{\mathbb{R}^d} |F(\mathcal{R}_{x_2} z) - F(\mathcal{R}_{x_1} z)| \chi_{U_0}(z) d(\xi_{x_2}^{-1} \# \mu)(z) \\ \leq N \int_{U_0} |F(\mathcal{R}_{x_2} z) - F(\mathcal{R}_{x_1} z)| d(\xi_{x_2}^{-1} \# \mu)(z) \\ \leq N \sqrt{2} \text{Lip}(F) \text{Lip}(v_d) R |x_2 - x_1|, \end{aligned} \tag{33}$$

having observed that  $(\xi_{x_2}^{-1} \# \mu)(U_0) = \mu(U_{x_2}) \leq 1$ .

As far as the second integral at the right-hand side of 32 is concerned, we have

$$\begin{aligned} N \left| \int_{\mathbb{R}^d} F(\mathcal{R}_{x_1} z) \chi_{U_0}(z) d(\xi_{x_2}^{-1} \# \mu - \xi_{x_1}^{-1} \# \mu)(z) \right| \\ \leq N \int_{U_{x_1} \cup U_{x_2}} |F(\mathcal{R}_{x_1} \xi_{x_2}^{-1}(y)) \chi_{U_0}(\xi_{x_2}^{-1}(y)) \\ - F(\mathcal{R}_{x_1} \xi_{x_1}^{-1}(y)) \chi_{U_0}(\xi_{x_1}^{-1}(y))| d\mu(y). \end{aligned} \tag{34}$$

Notice that we can confine ourselves to  $y \in U_{x_1} \cup U_{x_2}$ , for otherwise it results  $\chi_{U_0}(\xi_{x_j}^{-1}(y)) = \chi_{U_{x_j}}(y) = 0$  for both  $j = 1, 2$ . We distinguish two cases.

(a)  $|x_2 - x_1| > 2R$ .

In this case  $U_{x_1} \cap U_{x_2} = \emptyset$  because the balls  $B_R(x_1), B_R(x_2)$  are disjoint. Thus:

$$\begin{aligned} (34) &= N \int_{U_{x_1}} |F(\mathcal{R}_{x_1} \xi_{x_1}^{-1}(y))| \chi_{U_{x_1}}(y) d\mu(y) \\ &\quad + N \int_{U_{x_2}} |F(\mathcal{R}_{x_1} \xi_{x_2}^{-1}(y))| \chi_{U_{x_2}}(y) d\mu(y). \end{aligned}$$

For all  $y \in U_{x_j}, j = 1, 2$ , it results  $\xi_{x_j}^{-1}(y) \in U_0 \subset B_R(0)$ , hence  $\mathcal{R}_{x_1} \xi_{x_j}^{-1}(y) \in B_R(0)$  and we can use the boundedness of  $F$  in that ball to get

$$\begin{aligned} &\leq NC_F \left( \int_{U_{x_1}} \chi_{U_{x_1}}(y) d\mu(y) + \int_{U_{x_2}} \chi_{U_{x_2}}(y) d\mu(y) \right) \\ &\leq NC_F \mu(U_{x_1} \cup U_{x_2}) \leq NC_F. \end{aligned}$$

But  $1 < \frac{|x_2 - x_1|}{2R}$ , therefore we conclude

$$(34) \leq \frac{NC_F}{2R} |x_2 - x_1|. \tag{35}$$

(b)  $|x_2 - x_1| \leq 2R$ .

In this case the neighborhoods  $U_{x_1}, U_{x_2}$  need not be disjoint, however we

can resort to the Lipschitz continuity of the function  $F(\mathcal{R}_x \cdot) \chi_{U_0}(\cdot)$ :

$$\begin{aligned}
 (34) &\leq \text{Lip}(F\chi_{U_0})N \int_{U_{x_1} \cup U_{x_2}} |\xi_{x_2}^{-1}(y) - \xi_{x_1}^{-1}(y)| d\mu(y) \\
 &= \text{Lip}(F\chi_{U_0})N \int_{U_{x_1} \cup U_{x_2}} |\mathcal{R}_{x_2}^{-1}(y - x_2) - \mathcal{R}_{x_1}^{-1}(y - x_1)| d\mu(y) \\
 &= \text{Lip}(F\chi_{U_0})N \int_{U_{x_1} \cup U_{x_2}} |(\mathcal{R}_{x_2}^{-1} - \mathcal{R}_{x_1}^{-1})(y - x_1) \\
 &\quad + \mathcal{R}_{x_2}^{-1}(x_2 - x_1)| d\mu(y)
 \end{aligned}$$

and further, thanks to Lemma 5.2 and to the fact that  $\mathcal{R}_{x_2}^{-1}$  is an isometry,

$$\begin{aligned}
 &\leq \text{Lip}(F\chi_{U_0})N \left( \sqrt{2} \text{Lip}(v_d) \int_{U_{x_1} \cup U_{x_2}} |y - x_1| d\mu(y) \right. \\
 &\quad \left. + \mu(U_{x_1} \cup U_{x_2}) \right) |x_2 - x_1|.
 \end{aligned}$$

Let us examine the term with the integral. If  $y \in U_{x_1}$  then  $|y - x_1| \leq R$  whereas if  $y \in U_{x_2}$  then  $|y - x_1| \leq |y - x_2| + |x_2 - x_1| \leq 3R$ . Finally,  $|y - x_1| \leq 3R$  for all  $y \in U_{x_1} \cup U_{x_2}$ , which says

$$(34) \leq \text{Lip}(F\chi_{U_0})N(3\sqrt{2} \text{Lip}(v_d)R + 1) |x_2 - x_1|. \quad (36)$$

In conclusion, from 35 and 36 we deduce that there exists a constant  $C > 0$  such that (34)  $\leq C|x_2 - x_1|$  for all  $x_1, x_2 \in \mathbb{R}^2$ . This, together with the estimate 33, completes the proof of Lipschitz continuity of the mapping  $x \mapsto v[\mu](x)$ .  $\square$

In view of Proposition 3 we conclude that two-dimensional models based on the velocity 28 with anisotropic interactions have probability measure solutions, which can be approximated arbitrarily well using the scheme 19 on finer and finer numerical grids. Notice that, as far as e.g., crowd dynamics is concerned, two-dimensional problems are enough for applications.

**Remark 5** (Higher dimension). For  $d > 2$  additional technicalities arise, due to a more complex structure of the rotation matrix. Nevertheless, in the special case that the desired velocity is constant in  $x$ , it is straightforward to extend the results to any spatial dimension. In fact, the rotation matrix being independent of  $x$ , Proposition 3 can be proved without using Lemma 5.2, which is the only point where we use the explicit representation of the matrix. Models with constant desired velocity have been recently proposed for swarm dynamics problems [5] and for *rendez-vous* algorithms [4].

**Remark 6** (Zero desired velocity). When interactions are anisotropic and the desired velocity is zero [5], the orientation of the neighborhood of interaction cannot be defined by 31. However, this issue can be solved by replacing  $v_d$  in 31 with any other Lipschitz continuous unit vector field, e.g., a nonzero constant one, with the only purpose of defining a rotation angle. Clearly, this problem does not arise if the desired velocity is zero but interactions are isotropic, as in [4].



**6. Case study: Discrete models.** In this last section we put the theory into operation by giving an example of explicit solution to the Cauchy problem 4. Furthermore, we visualize the convergence to such solution of the approximations produced by the numerical scheme discussed in Section 4.

For illustrative purposes, we consider the simple case of a system of agents featuring isotropic interactions with zero desired velocity:

$$v[\mu_t](x) = N \int_{\mathbb{R}^d} F(y-x) \chi_{B_R(x)}(y) d\mu_t(y). \quad (37)$$

Moreover, we prescribe as initial condition the discrete probability measure

$$\bar{\mu} = \frac{1}{N} \sum_{l=1}^N \delta_{x_0^l}, \quad (38)$$

where  $\delta_{x_0^l}$  is the Dirac mass concentrated in  $x_0^l$  (i.e., for any  $A \in \mathcal{B}(\mathbb{R}^d)$  it results  $\delta_{x_0^l}(A) = 1$  if  $x_0^l \in A$ ,  $\delta_{x_0^l}(A) = 0$  otherwise) and  $x_0^1, \dots, x_0^N$  are  $N$  selected points in  $\mathbb{R}^d$ .

We recall that  $\bar{\mu}$  is the common law of the random variables  $X_0^i$ ,  $i = 1, \dots, N$ , expressing the initial positions of the agents. The structure 38 of  $\bar{\mu}$  implies that each  $X_0^i$  is a random variable taking almost surely the values  $x_0^1, \dots, x_0^N$ , each with probability  $1/N$ . Therefore we are considering a discrete model of a group of indistinguishable agents initially concentrated in  $x_0^1, \dots, x_0^N$ . The indistinguishability is reflected by the fact that any of the points  $x_0^l$  can be, with equal probability, the initial position of the generic  $i$ -th agent.

We find a solution to the Cauchy problem 4 with initial condition 38 by the method of the characteristics (cf. Section 2.2). In particular, we know from 6 that  $\mu_t = \gamma_t \# \bar{\mu}$ , where  $\gamma_t$  is the flow map. From the linearity of the push forward we first deduce  $\mu_t = \frac{1}{N} \sum_{l=1}^N \gamma_t \# \delta_{x_0^l}$ , then we observe that for any measurable set  $A$  it results

$$(\gamma_t \# \delta_{x_0^l})(A) = \delta_{x_0^l}(\gamma_t^{-1}(A)) = \begin{cases} 1 & \text{if } x_0^l \in \gamma_t^{-1}(A) \Leftrightarrow \gamma_t(x_0^l) \in A \\ 0 & \text{otherwise.} \end{cases}$$

Hence  $\gamma_t \# \delta_{x_0^l} = \delta_{\gamma_t(x_0^l)}$  and we can write the solution as

$$\mu_t = \frac{1}{N} \sum_{l=1}^N \delta_{\gamma_t(x_0^l)}. \quad (39)$$

Recalling that  $\mu_t$  is the law of the random variables  $\hat{X}_t^i = \mathbb{E}[X_t^i | X_0^i]$ ,  $i = 1, \dots, N$ , from 39 we infer that each  $\hat{X}_t^i$  takes almost surely the values  $\gamma_t(x_0^1), \dots, \gamma_t(x_0^N)$ , each with probability  $1/N$ . Therefore, at every time  $t > 0$  the distribution of the group of agents is concentrated on the discrete set of points  $\gamma_t(x_0^1), \dots, \gamma_t(x_0^N)$ . Notice that again we cannot associate deterministically a given agent with its position because of the indistinguishability of the agents. However, we can describe the trajectories of the agents by means of the mappings  $t \mapsto \gamma_t(x_0^l)$ ,  $l = 1, \dots, N$ .

The flow map is defined by 7, which with the velocity 37 and the initial condition 38 yields

$$\begin{cases} \frac{\partial \gamma_t(x)}{\partial t} = \sum_{j=1}^N F(\gamma_t(x_0^j) - \gamma_t(x)) \chi_{B_R(\gamma_t(x))}(\gamma_t(x_0^j)) \\ \gamma_0(x) = x \end{cases} \quad (40)$$

for all  $x \in \mathbb{R}^d$ . The discrete structure 39 of  $\mu_t$  makes it actually sufficient to solve problem 40 for  $x = x_0^l$ . Setting  $x_l(t) := \gamma_t(x_0^l)$  and computing 40 for  $x = x_0^l$  we find the initial-value problem

$$\begin{cases} \dot{x}_l = \sum_{j=1}^N F(x_j - x_l) \chi_{B_R(x_l)}(x_j) \\ x_l(0) = x_0^l \end{cases} \quad (l = 1, \dots, N), \quad (41)$$

thus we conclude that constructing the measure 39 amounts to solving the system of ODEs 41, whose solutions are the positions of the agents at  $t > 0$ .

From the numerical point of view, we can either approximate the measure 39 by using the scheme 19 or integrate problem 41 via a standard numerical method for ODEs. The remaining part of this section is devoted to show the convergence of the approximations obtained from 19 to the numerical solutions of 41 in a toy model. Let us consider a one-dimensional problem ( $d = 1$ ) with  $N = 10$  agents, whose initial positions are sampled from a uniform distribution on the interval  $[0, 1]$ . Agents repel each other according to the following repulsion function:

$$F(z) = -\frac{az}{\max\{|z|, \varepsilon\}^2} \quad (a, \varepsilon > 0)$$

which is from the product of

$$f(z) = -\frac{a}{\max\{|z|, \varepsilon\}}, \quad r(z) = \frac{z}{\max\{|z|, \varepsilon\}}.$$

The repulsion strength  $f$  is inversely proportional to the distance between the interacting agents (up to a minimal threshold  $\varepsilon$ ) and the direction of the interaction  $r$  is a Lipschitz mollification of the unit vector  $z/|z|$ . The reference interaction neighborhood is  $B_R(0) = (-R, R)$  with cut-off function

$$\chi_{B_R(0)}(z) = e^{-\frac{b|z|^2}{R^2 - |z|^2}} \mathbb{1}_{B_R(0)}(z) \quad (b > 0),$$

a  $C^\infty$ -mollification of the indicator function of  $B_R(0)$ . Parameters are set to  $R = 0.1$ ,  $a = 0.01$ ,  $b = 0.02$ ,  $\varepsilon = R/4$ . The computational time is  $T = 0.1$ .

Simulations of the ODE system 41 and of problem 4 with initial condition 38 were run independently and their results visualized on the same graphs of Fig. 1 for duly comparison. In particular, the ODE system was numerically integrated using an explicit-in-time Euler scheme, whereas the conservation law for the probability  $\mu_t$  was solved through the scheme 19 on different meshes, choosing

$$\Delta t_k = \left(\frac{h_k}{V}\right)^\delta \quad (0 < \delta < 1).$$

Note that this entails  $\beta_k \sim h_k^{1-\delta}$  in Assumption 3 and  $\alpha_k \sim h_k^{\delta-1}$  in 26. Figure 1 displays the numerical solution computed with  $\delta = 0.9$  and  $h_k = 1/k$  in the three cases  $k = 10^2, 10^3, 10^4$ , at two different time instants. Convergence toward the

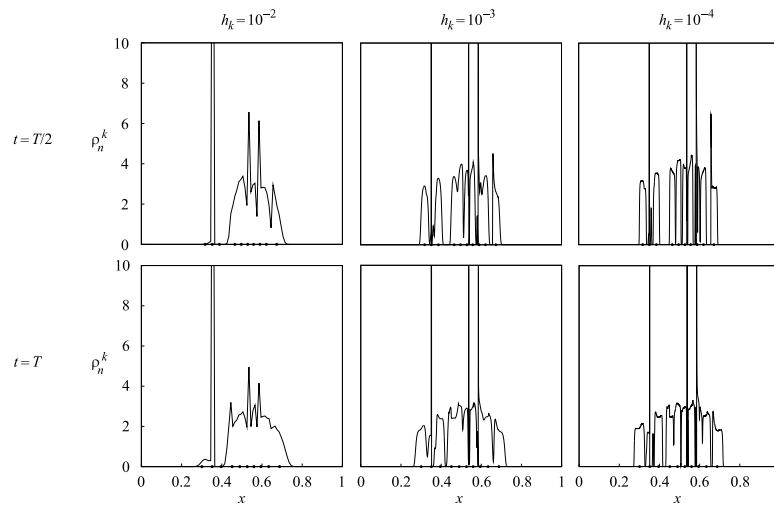


FIGURE 1. Numerical solution to 3 with discrete initial datum 38 at two successive times. Bullets are the solution of the ODE system 41 computed by the explicit Euler scheme. The continuous curve is the solution computed by the numerical scheme 19 on meshes with different level of refinement.

exact solution 39 as the mesh refines can be visually appreciated, although approximating singular measures with densities is a difficult task, which requires very fine and computationally expensive meshes to get accurate results. Therefore, when such a structure of the solution is numerically sought, it is more efficient to exploit the stated equivalence of the original problem with the discrete system of ODEs.

**Acknowledgments.** The authors wish to thank an anonymous referee for his instructive comments and suggestions, which helped them improve significantly the presentation of the work.

## REFERENCES

- [1] L. Ambrosio, N. Gigli and G. Savaré, “Gradient Flows in Metric Spaces and in the Space of Probability Measures,” 2<sup>nd</sup> edition, Birkhäuser Verlag, Basel, 2008.
- [2] N. Bellomo, “Modeling Complex Living Systems. A Kinetic Theory and Stochastic Game Approach,” Birkhäuser, Boston, 2008.
- [3] J. A. Cañizo, J. A. Carrillo and J. Rosado, *A well-posedness theory in measures for some kinetic models of collective motion*, Math. Models Methods Appl. Sci., **21** (2011), 515–539.
- [4] C. Canuto, F. Fagnani and P. Tilli, *A Eulerian approach to the analysis of rendez-vous algorithms*, in “Proceedings of the 17th IFAC World Congress”, (2008), 9039–9044.
- [5] E. Cristiani, P. Frasca and B. Piccoli, *Effects of anisotropic interactions on the structure of animal groups*, J. Math. Biol., **62** (2011), 569–588.
- [6] E. Cristiani, B. Piccoli and A. Tosin, *Modeling self-organization in pedestrian and animal groups from macroscopic and microscopic viewpoints*, in “Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences” (eds. G. Naldi, L. Pareschi and G. Toscani), Birkhäuser, Boston, (2010), 337–364.
- [7] R. J. LeVeque, “Numerical Methods for Conservation Laws,” 2<sup>nd</sup> edition, Birkhäuser Verlag, Basel, 1992.
- [8] B. Piccoli and A. Tosin, *Pedestrian flows in bounded domains with obstacles*, Contin. Mech. Thermodyn., **21** (2009), 85–107.

- [9] B. Piccoli and A. Tosin, *Time-evolving measures and macroscopic modeling of pedestrian flow*, Arch. Ration. Mech. Anal., **199** (2011), 707–738.

Received December 2010; revised February 2011.

*E-mail address:* [andrea.tosin@polito.it](mailto:andrea.tosin@polito.it)

*E-mail address:* [paolo.frasca@polito.it](mailto:paolo.frasca@polito.it)