

MEAN-FIELD ANALYSIS OF MULTI-POPULATION DYNAMICS WITH LABEL SWITCHING

MARCO MORANDOTTI AND FRANCESCO SOLOMBRINO

ABSTRACT. The mean-field analysis of a multi-population agent-based model is performed. The model couples a particle dynamics driven by a nonlocal velocity with a Markov-type jump process on the probability that each agent has of belonging to a given population. A general functional analytic framework for the well-posedness of the problem is established, and some concrete applications are presented, both in the case of discrete and continuous set of labels. In the particular case of a leader-follower dynamics, the existence and approximation results recently obtained in [2] are recovered and generalized as a byproduct of the abstract approach proposed.

Keywords: kinetic equations, mean-field limits, continuity equations, jump processes, superposition principle.

CONTENTS

1. Introduction	1
2. Preliminaries	4
2.1. Basic notation	4
2.2. Functional setting	6
2.3. Well-posedness of ODE's in Banach spaces	6
3. The abstract model	7
3.1. The discrete problem and statement of the main result	9
3.2. Lagrangian solutions	10
3.3. Proof of Theorem 3.5	12
4. A leader-follower dynamics	14
5. Further applications	20
5.1. Discrete and finite spaces of labels U	20
5.2. A continuum of labels	23
References	25

1. INTRODUCTION

The concept of mean-field interaction, originally used in statistical physics with Kac [30] and then McKean [33] to describe the collisions between particles in a gas, has later proved to be a powerful tool to analyze the asymptotic behavior of systems of interacting agents. Recent applications range from biological, social and economical phenomena [1, 11, 18] to automatic learning [13, 29] and optimization heuristics [24, 31]. The underlying idea is that the collective behavior of large systems of particles (agent-based models, usually consisting of a set of ODE's) can be efficiently treated by replacing the influence of all the other individuals in the dynamics on a given agent by a single averaged effect. From a mathematical point of view, this amounts to passing from a particle description to a kinetic description, consisting of a limit PDE whose unknown is the particle density distribution in the state space. The well-posedness of such models has therefore to be proven in spaces of measures (see, for instance, the results in [17]).

In some of the applications, the interacting agents are assumed to belong to a number of different species, or populations [3, 4, 20, 22, 26]. This is a useful modeling assumption, *e.g.*, in the theory of mean-field games, or in control theory, where it can be used to

Date: July 8, 2019.

2010 *Mathematics Subject Classification.* 35Q91, (60J75, 37C10, 47J35, 58D25) .

distinguish informed agents steering pedestrians to leave unknown environments, or to highlight the influence of few key investors in the stock market (see the discussion in [12, Introduction]). In a multi-population setting, source (or sink) terms can be added to the model, to account for the "birth" and "death" rate of a single population (see for instance [38, Sections 4-5]). Furthermore, even in the case where the number of agents remains constant along the evolution, the possibility of having some individuals switching from one population to another has been considered [39].

As a matter of fact, exchange rates among populations are a common feature in several applications. For instance, in models of chemical reaction networks a particle may change its type as a result of the interaction with the others, at a stochastic rate which may also depend on its position [32, 34, 36]. Another relevant example comes from social dynamics, where transition rates appear in some kinetic models of opinion formation in the presence of strong leaders, as the one proposed in [25], inspired by the earlier contribution [40]. In this model, the opinion of the agents, described by their position in the state space, evolves because of the exchange of opinion with the others. This is encoded by a nonlocal transport term taking into account the presence, among the overall population, of a restricted number of "leaders" promoting their opinion with a strong influence on the "followers". It is natural to postulate, as the authors do (see [25, Section 3.b] for a detailed discussion), that opinion leadership is not constant over time: someone who is an opinion leader today may lose this role tomorrow, or a follower may become a leader in the future.

Summarizing, the multi-population dynamics we are interested in attaches to an agent sitting at a position x a probability measure $\lambda \in \mathcal{P}(U)$, where U is a compact space of labels accounting for the population to which the agent belongs. In fact, we find it natural not to assume any deterministic knowledge of the label of a single agent, since transitions are usually modeled as the outcome of a stochastic process. Hence, λ expresses the probability that the agent has, at a certain moment in time, of belonging to a subset of U , and may itself evolve as a consequence of the interaction among the agents. The minimal model for evolution that we propose¹ consists of the coupling of a non-local transport dynamics with a Markov-type jump process. We namely assume that

- a particle at position x_i with probability distribution of labels λ_i experiences a velocity field $\dot{x}_i = v_{\Psi^N}(x_i, \lambda_i)$ influenced by the global state of the system Ψ^N , the empirical measure $\frac{1}{N} \sum_{i=1}^N \delta_{(x_i, \lambda_i)}$;
- the probability distribution of labels λ_i evolves according to $\dot{\lambda}_i = \mathcal{T}^*(x_i, \Psi^N)\lambda_i$, where the operator \mathcal{T}^* , accounting for the transitions among the labels, is the adjoint of an operator $\mathcal{T}(x, \Psi)$ defined on Lipschitz functions on U and may depend on the position of the particle and on the state of the system.

In the case of two populations $U = \{F, L\} = \{\text{followers, leaders}\}$, a model of this type has been recently analyzed in [2], as a (partial) discrete counterpart of the PDE model in [25]. Some simplifying assumptions had however to be added in order to perform a mean-field analysis. In particular, the velocity field

$$v_{\Psi^N}(x_i) = \frac{1}{N} \sum_{j=1}^N K^F(x_i - x_j) \lambda_j(\{F\}) + \frac{1}{N} \sum_{j=1}^N K^L(x_i - x_j) \lambda_j(\{L\}) \quad (1.1a)$$

did not depend on the probability on the labels λ_i and the transition operators

$$\begin{aligned} (\mathcal{T}^*(\Psi^N)\lambda_i)(\{F\}) &= -\alpha_F(\Psi^N)\lambda_i(\{F\}) + \alpha_L(\Psi^N)(1 - \lambda_i(\{F\})), \\ (\mathcal{T}^*(\Psi^N)\lambda_i)(\{L\}) &= \alpha_F(\Psi^N)\lambda_i(\{F\}) - \alpha_L(\Psi^N)(1 - \lambda_i(\{F\})) \end{aligned} \quad (1.1b)$$

did not depend on the position x_i . The goal of the present paper is to get rid of these simplifying assumptions, providing an *appropriate functional setting* as well as a *general set of assumptions* on the velocity field v_{Ψ} and on the operator \mathcal{T}^* (see Section 3) which allow one to perform a rigorous mean-field analysis of agent-based models of the kind discussed above.

¹We neglect in the present contribution the possible presence of diffusion terms.

In order to develop our analysis, we will borrow some functional-analytic and some measure-theoretic tools that have been recently specified for the different, although related, context of *spatially inhomogeneous evolutionary games* in [5]. The analogue of (1.1) is described in [5, Section 3.2], where the dynamics considered is

$$\begin{cases} \dot{x}_i = a(x_i, \sigma_i) = \int_U e(x_i, u) d\sigma_i(u), \\ \dot{\sigma}_i = \left(\frac{1}{N} \sum_{j=1}^N \left(\int_U J(x_i, \cdot, x_j, u') d\sigma_j(u') - \int_U \int_U J(x_i, w, x_j, u') d\sigma_j(u') d\sigma_i(w) \right) \right) \sigma_i, \end{cases} \quad (1.2)$$

for $x \in \mathbb{R}^d$ tracking the position of a player and $\sigma \in \mathcal{P}(U)$ denoting their mixed strategy. In (1.2), the velocity \dot{x}_i is determined by the x -dependent vector field $e(x, u)$ through an averaging process with respect to the strategies available to the player x_i ; the evolution of the mixed strategy σ_i is dictated by the *replicator equation* (see [28]): the term in the parenthesis in the right-hand side of the second equation above determines the performance of the strategy played by player x_i with respect to all of the available strategies, averaging out all of the possible distributions of the opponents x_j with their mixed strategies σ_j . In particular, $J(x_i, u, x_j, u')$ is the payoff of the game between player x_j with strategy u against player x_i with strategy u' . For the heuristic interpretation and the modeling issues related to (1.2) we refer the interested reader to [14, 16].

The paper [5] introduced suitable notions (namely, Lagrangian and Eulerian) of solutions to the mean-field limit of (1.2) showing convergence of the particle model to the limit description by means of stability estimates in the Wasserstein metric arising from the well-posedness theory of ODE's in Banach spaces [15, 19, 21].

While sharing some common features, the dynamics described by (1.1) and (1.2) show a relevant difference in the structures of the velocity fields in the right-hand sides. Indeed, in our setting, it is natural to postulate that the velocity of the agents is also depending on the behavior of the ones around them, at least in a small neighborhood. Such a feature is not present in the first equation in (1.2), whereas it is encoded in our model by considering a velocity field v_Ψ depending on the global state of the system. Beside this main distinction, also the right-hand sides of the equations for σ_i and λ_i present some differences; we refer the reader to Remark 5.9 for a more detailed comparison of the two models.

The novel contribution of this work is in specifying the tools from [5] to the case of our particle system (1.1). In doing this, we will pursue a more abstract point of view than the one considered in [5], by focusing on the structural assumptions on the velocity field v_Ψ and on the operator \mathcal{T} that guarantee the well-posedness of the mean-field analysis. However, we will show in Sections 4 and 5 that velocities and transition rates modeled by interaction kernels as those considered in (1.1) and (1.2) (in this last case, with the due difference that we pointed out before) also fit in the setting that we propose in Section 3. Thus, we can retrieve an even more general version of the results of [2] as a by-product of our analysis (see Theorem 4.8, where also an explicit dependence on the space variable x of the transition rates is allowed).

The general assumptions on v_Ψ and \mathcal{T} are presented at the beginning of Section 3, see (v1)-(v3) and (T0)-(T3) below. The dependence of v_Ψ on the global status of the system call for a locally Lipschitz dependence both on the status variable (x, λ) (in a suitable topology) and on Ψ with respect to the Wasserstein distance between probability measures, which is customary for the mean-field analysis of transport equation with non-local vector fields, see e.g. [6, 37, 23, 27, 41]. Furthermore, the sublinearity condition (v3) guarantees long-time existence of the solution. Concerning the operator \mathcal{T} , it must comply with analogous local Lipschitz continuity conditions and sublinearity conditions, see (T1)-(T2); in addition, we require that

- constants belong to the kernel of $\mathcal{T}(x, \Psi)$, namely $\mathcal{T}(x, \Psi)1 = 0$;
- there exists $\delta > 0$ such that $\mathcal{T}(x, \Psi) + \delta I \geq 0$.

The two conditions above (see (T0) and (T3) below for a precise statement) ensure that the velocities $\dot{\lambda}_i$'s lie on the tangent plane to the probabilities on $\{F, L\}$ and that the positivity of the measures are preserved, respectively. Altogether, conditions (v1)-(v3) and (T0)-(T3)

allow us to show that the particle model, which we can rewrite in the form

$$\dot{y}_i = \begin{pmatrix} \dot{x}_i \\ \dot{\lambda}_i \end{pmatrix} = \begin{pmatrix} v_{\Psi^N}(x_i, \lambda_i) \\ \mathcal{T}^*(x_i, \Psi^N)\lambda_i \end{pmatrix} =: b_{\Psi^N}(y_i), \quad (1.3)$$

is well-posed (see Proposition 3.3).

We subsequently show that in the limit as $N \rightarrow \infty$ the empirical measures Ψ^N converge to a continuous path of probability measures Ψ_t that solves the continuity-type equation

$$\partial_t \Psi_t + \operatorname{div}(b_{\Psi_t} \Psi_t) = 0 \quad (1.4)$$

in the space of probability measures on the product space $Y = \mathbb{R}^d \times \mathcal{P}(U)$.

A key point of the proof is that the solutions to (1.4), which are defined via duality with test functions (Eulerian solutions, see Definition 3.8), can be equivalently characterized as Lagrangian solution satisfying the fixed-point equation

$$\Psi_t = \mathbf{Y}_{\Psi}(t, 0, \cdot)_{\#} \bar{\Psi} \quad \text{for every } t \geq 0, \quad (1.5)$$

where $\mathbf{Y}(t, 0, \cdot)$ is the transition map (see Definition 3.7) associated with the ODE (1.3) (with Ψ in place of Ψ^N), $\bar{\Psi}$ is a given initial distribution of agents and probability on the labels, and the symbol $\#$ denotes the push-forward measure (see Definition 2.1).

In Section 4 we also retrieve and extend the main result of [2] as a special case of our analysis. We associate with a solution to (1.4) the followers and leaders distributions

$$\mu_{\Psi}^F(B) := \int_{B \times \mathcal{P}(\{F, L\})} \lambda(\{F\}) d\Psi(x, \lambda), \quad \mu_{\Psi}^L(B) := \int_{B \times \mathcal{P}(\{F, L\})} \lambda(\{L\}) d\Psi(x, \lambda), \quad (1.6)$$

for each Borel set $B \subset \mathbb{R}^d$. If the vector field v_{Ψ} has the special structure (1.1a) and under suitable structural assumptions on the transition rates α_F and α_L (see (4.13) below), using the definition of Eulerian solution, we are able to show that μ_{Ψ}^F and μ_{Ψ}^L are indeed the unique solutions to the system

$$\begin{cases} \partial_t \mu_t^F = -\operatorname{div}((K^F \star \mu_t^F + K^L \star \mu_t^L) \mu_t^F) - \alpha_F(x, \mu_t^F, \mu_t^L) \mu_t^F + \alpha_L(x, \mu_t^F, \mu_t^L) \mu_t^L, \\ \partial_t \mu_t^L = -\operatorname{div}((K^F \star \mu_t^F + K^L \star \mu_t^L) \mu_t^L) + \alpha_F(x, \mu_t^F, \mu_t^L) \mu_t^F - \alpha_L(x, \mu_t^F, \mu_t^L) \mu_t^L \end{cases} \quad (1.7)$$

considered in [2]. Notice that, as discussed in Remark 4.6, rewriting (3.27) as a system of equation in μ^F and μ^L is not possible for a velocity field explicitly depending on λ such as

$$\begin{aligned} v_{\Psi^N}(x_i, \lambda_i) = & \lambda_i(\{F\}) \left(\frac{1}{N} \sum_{j=1}^N K^{FF}(x_i - x_j) \lambda_j(\{F\}) + \frac{1}{N} \sum_{j=1}^N K^{LF}(x_i - x_j) \lambda_j(\{L\}) \right) + \\ & \lambda_i(\{L\}) \left(\frac{1}{N} \sum_{j=1}^N K^{FL}(x_i - x_j) \lambda_j(\{F\}) + \frac{1}{N} \sum_{j=1}^N K^{LL}(x_i - x_j) \lambda_j(\{L\}) \right) \end{aligned}$$

considered for instance in [22]. In the presence of exchange rates, the correct mean-field description of the above particle system is given by (3.27) in the product space $\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})$.

The paper is organized as follows. In Section 2, we recall the basic notions of measure theory that will be needed in the sequel, and prove a corollary of a theorem by Brezis [15, Sect. I.3, Thm. 1.4, Cor. 1.1] on the well-posedness of ODE's in Banach spaces. In Section 3 we introduce our abstract model and we prove our main result, Theorem 3.5, on the mean-field limit of the dynamics. Section 4 is devoted to the special case of $U = \{F, L\}$, modeling the leader-follower dynamics. We apply the abstract results to this case and recover the results of [2], extending them to the case of x -dependent transition rates. Finally, Section 5.1 we extend the results of Section 4 to any finite numbers of labels, whereas in Section 5.2 we discuss some explicit examples of velocity fields v_{Ψ} and transition operators $\mathcal{T}(x, \Psi)$ which are encompassed by our setting in the case of a continuum of labels, and we compare them with those considered in [5].

2. PRELIMINARIES

2.1. Basic notation. If (X, d_X) is a metric space, we denote by $\mathcal{M}(X)$ the space of signed Borel measures in X with finite total variation, by $\mathcal{M}_+(X)$ and $\mathcal{P}(X)$ the convex subsets of nonnegative measures and probability measures, respectively. The notation $\mathcal{P}_c(X)$ will

be used for measures having compact support in X . For $\mu \in \mathcal{M}(X)$, $|\mu|$ denotes the total variation of μ . If we denote by $C_0(X)$ the space of continuous functions vanishing at the boundary of X , and by $C_b(X)$ the space of bounded continuous functions, the weak* and narrow convergence in $\mathcal{M}(X)$ are defined by the convergence of the duality products

$$\int_X \phi d\mu_h \rightarrow \int_X \phi d\mu, \quad h \rightarrow \infty$$

for each $\phi \in C_0(X)$ and $\phi \in C_b(X)$, respectively.

Whenever $X = \mathbb{R}^d$, $d \geq 1$, it remains understood that it is endowed with the Euclidean norm (and induced distance), which shall be simply denoted by $|\cdot|$.

For a Lipschitz function $f: X \rightarrow \mathbb{R}$ we denote by

$$\text{Lip}(f) := \sup_{\substack{x, y \in X \\ x \neq y}} \frac{|f(x) - f(y)|}{d_X(x, y)}$$

the Lipschitz constant. The notations $\text{Lip}(X)$ and $\text{Lip}_b(X)$ will be used for the spaces of Lipschitz and bounded Lipschitz functions on X , respectively. Both are normed spaces with the norm $\|f\| := \|f\|_\infty + \text{Lip}(f)$.

In a complete and separable metric space (X, d_X) , we shall use the Kantorovich-Rubinstein distance \mathcal{W}_1 in the class $\mathcal{P}(X)$, defined as

$$\mathcal{W}_1(\mu, \nu) := \sup \left\{ \int_X \varphi d\mu - \int_X \varphi d\nu : \varphi \in \text{Lip}_b(X), \text{Lip}(\varphi) \leq 1 \right\}$$

or equivalently (thanks to the Kantorovich duality) as

$$\mathcal{W}_1(\mu, \nu) := \inf \left\{ \int_{X \times X} d_X(x, y) d\Pi(x, y) : \Pi(A \times X) = \mu(A), \Pi(X \times B) = \nu(B) \right\},$$

involving couplings Π of μ and ν . Notice that $\mathcal{W}_1(\mu, \nu)$ is finite if μ and ν belong to the space

$$\mathcal{P}_1(X) := \left\{ \mu \in \mathcal{P}(X) : \int_X d_X(x, \bar{x}) d\mu(x) < +\infty \text{ for some } \bar{x} \in X \right\}$$

and that $(\mathcal{P}_1(X), \mathcal{W}_1)$ is complete if (X, d_X) is complete. For a positive measure $\mu \in \mathcal{M}_+(E)$, for E being a Banach space, we define the first moment $m_1(\mu)$ as

$$m_1(\mu) := \int_E \|x\|_E d\mu.$$

Notice that, for a probability measure μ , finiteness of the integral above is equivalent to $\mu \in \mathcal{P}_1(E)$, whenever E is endowed with the distance induced by the norm $\|\cdot\|$.

We now recall the definition of push-forward measure.

Definition 2.1. *Let $\mu \in \mathcal{M}_+(X)$ and $f: X \rightarrow Z$ a μ -measurable function be given. The push-forward measure $f_{\#}\mu \in \mathcal{M}_+(Z)$ is defined by $f_{\#}\mu(B) = \mu(f^{-1}(B))$ for any Borel set $B \subset Z$. The push-forward measures has the same total mass as μ , namely $\mu(X) = f_{\#}\mu(Z)$. It also holds the change of variables formula*

$$\int_Z g df_{\#}\mu = \int_X g \circ f d\mu$$

whenever either one of the integrals makes sense.

For E being a Banach space, the notation $C_b^1(E)$ will be used to denote the subspace of $C_b(E)$ of functions having bounded continuous Fréchet differential at each point. The notation $D\phi(\cdot)$ will be used to denote the Fréchet differential. In the case of a function $\phi: [0, T] \times E \rightarrow \mathbb{R}$, the simbol ∂_t will be used to denote partial differentiation with respect to t , while D will only stand for differentiation with respect to the variables in E .

2.2. Functional setting. The space of labels U will be assumed to be a compact metric space. The state of the system is described by $y := (x, \lambda) \in \mathbb{R}^d \times \mathcal{P}(U) =: Y$. The component $x \in \mathbb{R}^d$ describes the location of an agent in space, whereas the component $\lambda \in \mathcal{P}(U)$ describes the distribution of labels of the agent. A probability distribution $\Psi \in \mathcal{P}(Y)$ denotes a distribution of agents with labels.

To define the functional setting for the dynamics, we need to consider the *free space*

$$\mathcal{F}(U) := \overline{\text{span}(\mathcal{P}(U))}^{\|\cdot\|_{\text{BL}}} \subset (\text{Lip}(U))', \quad (2.1)$$

which is also called in the literature Arens-Eells space, see [8, 10] and [42, Chapter 3]. The closure in (2.1) is taken with respect to the *bounded Lipschitz* norm $\|\cdot\|_{\text{BL}}$, which is defined, for $\mu \in (\text{Lip}(U))'$, by

$$\|\mu\|_{\text{BL}} := \sup \{ \langle \mu, \varphi \rangle : \varphi \in \text{Lip}(U), \|\varphi\|_{\text{Lip}} \leq 1 \}.$$

With the free space $\mathcal{F}(U)$ at hand, we define $\bar{Y} := \mathbb{R}^d \times \mathcal{F}(U)$ and the norm $\|\cdot\|_{\bar{Y}}$ by

$$\|y\|_{\bar{Y}} = \|(x, \lambda)\|_{\bar{Y}} := |x| + \|\lambda\|_{\text{BL}}; \quad (2.2)$$

and we pose $\|y\|_Y = \|y\|_{\bar{Y}}$.

For a given $R > 0$, we denote by B_R the closed ball of radius R in \mathbb{R}^d and by B_R^Y the ball of radius R in Y , namely $B_R^Y = \{y \in Y : \|y\|_{\bar{Y}} \leq R\}$, and observe that it is a compact set, since Y is locally compact by our assumptions on U . The Banach space structure of $\bar{Y} \supset Y$ allows us to define the first moment $m_1(\Psi)$ for a probability measure $\Psi \in \mathcal{P}(Y)$ as

$$m_1(\Psi) := \int_Y \|y\|_{\bar{Y}} d\Psi$$

so that the space $\mathcal{P}_1(Y)$ can be equivalently characterized as

$$\mathcal{P}_1(Y) = \{\Psi \in \mathcal{P}(Y) : m_1(\Psi) < +\infty\}.$$

We will sometimes use the notation $\mathcal{P}(K)$ to denote probability measures with support contained on a given compact subset $K \subset Y$. Notice that trivially we have $\mathcal{P}(K) \subset \mathcal{P}_1(Y)$.

2.3. Well-posedness of ODE's in Banach spaces. We recall here a theorem by Brezis [15, Sect. I.3, Thm. 1.4, Cor. 1.1] on the well-posedness of ODE's in Banach spaces.

Theorem 2.2. *Let $(E, \|\cdot\|_E)$ be a Banach space, C a closed convex subset of E , and let $A(t, \cdot) : C \rightarrow E$, $t \in [0, T]$, be a family of operators satisfying the following properties:*

(i) *there exists a constant $L \geq 0$ such that for every $c_1, c_2 \in C$ and $t \in [0, T]$*

$$\|A(t, c_1) - A(t, c_2)\|_E \leq L\|c_1 - c_2\|_E; \quad (2.3)$$

(ii) *for every $c \in C$ the map $t \mapsto A(t, c)$ is continuous in $[0, T]$;*
 (iii) *for every $R > 0$ there exists $\theta > 0$ such that*

$$c \in C, \|c\|_E \leq R \quad \Rightarrow \quad c + \theta A(t, c) \in C. \quad (2.4)$$

Then for every $\bar{c} \in C$ there exists a unique curve $c : [0, T] \rightarrow C$ of class C^1 satisfying $c_t \in C$ for all $t \in [0, T]$ and

$$\frac{d}{dt}c_t = A(t, c_t) \quad \text{in } [0, T], \quad c_0 = \bar{c}. \quad (2.5)$$

Moreover, if c^1, c^2 are the solutions starting from the initial data $\bar{c}^1, \bar{c}^2 \in C$ respectively, we have

$$\|c_t^1 - c_t^2\|_E \leq e^{Lt}\|\bar{c}^1 - \bar{c}^2\|_E, \quad \text{for every } t \in [0, T]. \quad (2.6)$$

For our purposes, we need the following generalization.

Corollary 2.3. *Let hypotheses (ii) and (iii) of Theorem 2.2 hold for a family of operators $A(t, \cdot) : C \rightarrow E$, $t \in [0, T]$. Assume, in addition that*

(i') *for every $R > 0$ there exists a constant $L_R \geq 0$ such that for every $c_1, c_2 \in C \cap B_R$ and $t \in [0, T]$*

$$\|A(t, c_1) - A(t, c_2)\|_E \leq L_R\|c_1 - c_2\|_E; \quad (2.7)$$

(i'') *there exists $M > 0$ such that for every $c \in C$, there holds*

$$\|A(t, c)\|_E \leq M(1 + \|c\|_E). \quad (2.8)$$

Then for every $\bar{c} \in C$ there exists a unique curve $c: [0, T] \rightarrow C$ of class C^1 satisfying $c_t \in C$ for all $t \in [0, T]$ and

$$\frac{d}{dt}c_t = A(t, c_t) \quad \text{in } [0, T], \quad c_0 = \bar{c}. \quad (2.9)$$

Moreover, if c^1, c^2 are the solutions starting from the initial data $\bar{c}^1, \bar{c}^2 \in C \cap B_R$ respectively, there exists a constant $L = L(M, R, T) > 0$ such that

$$\|c_t^1 - c_t^2\|_E \leq e^{Lt} \|\bar{c}^1 - \bar{c}^2\|_E, \quad \text{for every } t \in [0, T]. \quad (2.10)$$

Proof. Let us fix the initial datum $\bar{c} \in C$, and let us choose $\bar{R} := (\|\bar{c}\|_E + MT)e^{MT}$. Consider a smooth function with compact support $\chi: \mathbb{R}^+ \rightarrow [0, 1]$ such that $\chi(r) = 1$ for every $r \leq \bar{R}$ and set $B(t, c) := \chi(\|c\|_E)A(t, c)$. Then one can see that B satisfies hypotheses (i) and (ii) of Theorem 2.2. To see that hypothesis (iii) is also satisfied, it suffices to notice that, by convexity and since $0 \leq \chi \leq 1$, $c + \theta\chi(\|c\|_E)A(t, c) \in C$ whenever $c + \theta A(t, c) \in C$. Therefore there exists a unique solution $t \mapsto c(t)$ of class C^1 of

$$\frac{d}{dt}c_t = B(t, c_t) \quad \text{in } [0, T], \quad c_0 = \bar{c}. \quad (2.11)$$

Using again that $0 \leq \chi \leq 1$ and (2.8), one can see that

$$\|c_t\|_E \leq \|\bar{c}\|_E + MT + M \int_0^T \|c_s\|_E ds, \quad (2.12)$$

hence Gronwall's Lemma implies that $\|c_t\|_E \leq \bar{R}$ for every $t \in [0, T]$. With this, c_t solves (2.9). A similar argument shows that any other solution $t \mapsto \hat{c}_t$ to (2.9) must satisfy $\|\hat{c}_t\|_E \leq \bar{R}$ for every $t \in [0, T]$. Thus, uniqueness of solutions for (2.9) follows from the uniqueness of solutions to (2.11). A similar argument also yields (2.10). \square

3. THE ABSTRACT MODEL

The state of our system is described pairs $y := (x, \lambda)$. The element $x \in \mathbb{R}^d$ denotes the position of the agents, whereas the element $\lambda \in \mathcal{P}(U)$ denotes a (probability) distribution over the space U , which we assume to be a compact metric space, which can be interpreted as a space of strategies (as in [5], in which case an element of $\mathcal{P}(U)$ is a mixed strategy) or as a space of labels (as in [2], where the case $U = \{\text{leader}, \text{follower}\}$ was considered). A distribution of states will be described by an element $\Lambda \in \mathcal{P}(Y)$, where $Y := \mathbb{R}^d \times \mathcal{P}(U)$. We will be concerned with the evolution of Λ , given an initial Λ^0 , determined by the laws of evolution of x and λ , which are going to be discussed below.

For $y = (x, \lambda) \in Y$ and $\Psi \in \mathcal{P}_1(Y)$, we define a vector field $b_\Psi: Y \rightarrow \bar{Y}$ through

$$b_\Psi(y) := \begin{pmatrix} v_\Psi(y) \\ \mathcal{T}^*(x, \Psi)\lambda \end{pmatrix} \quad (3.1)$$

The first component of b_Ψ is a velocity field in \mathbb{R}^d determined by the global state of the system Ψ ; the second component is expressed in terms of the adjoint $\mathcal{T}^*(x, \Psi)$ of an operator $\mathcal{T}(x, \Psi)$ which sees the location of the agents and the global state of the system around them. In order to state the regularity assumptions that we make on b_Ψ , we will discuss separately the assumptions on v_Ψ and on \mathcal{T} .

We assume that the velocity field $v_\Psi: Y \rightarrow \mathbb{R}^d$ satisfies the following conditions:

- (v1) for every $R > 0$, for every $\Psi \in \mathcal{P}(B_R^Y)$, $v_\Psi \in \text{Lip}(B_R^Y; \mathbb{R}^d)$ uniformly with respect to Ψ , namely there exists a constant $L_{v,R} > 0$ such that

$$|v_\Psi(y^1) - v_\Psi(y^2)| \leq L_{v,R} \|y^1 - y^2\|_{\bar{Y}}; \quad (3.2)$$

- (v2) for every $R > 0$, for every $\Psi \in \mathcal{P}(B_R^Y)$, there exists a constant $L_{v,R} > 0$ such that for every $y \in B_R^Y$, and for every $\Psi^1, \Psi^2 \in \mathcal{P}(B_R^Y)$

$$|v_{\Psi^1}(y) - v_{\Psi^2}(y)| \leq L_{v,R} \mathcal{W}_1(\Psi^1, \Psi^2); \quad (3.3)$$

- (v3) there exists $M_v > 0$ such that for every $y \in Y$ and for every $\Psi \in \mathcal{P}_1(Y)$ there holds

$$|v_\Psi(y)| \leq M_v (1 + \|y\|_{\bar{Y}} + m_1(\Psi)). \quad (3.4)$$

We now describe the assumptions on \mathcal{T} . For $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$, let $\mathcal{T}(x, \Psi): \text{Lip}(U) \rightarrow \text{Lip}(U)$ be an operator such that

(T0) for every $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$, constants are in the kernel of $\mathcal{T}(x, \Psi)$, that is

$$\mathcal{T}(x, \Psi)1 = 0; \quad (3.5)$$

(T1) for every $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$, there exists a constant $M_{\mathcal{T}} > 0$ such that the operator norm satisfies

$$\|\mathcal{T}(x, \Psi)\|_{\mathcal{L}(\text{Lip}(U); \text{Lip}(U))} \leq M_{\mathcal{T}}(1 + |x| + m_1(\Psi)); \quad (3.6)$$

(T2) for every $R > 0$ there exists $L_{\mathcal{T}, R} > 0$ such that, for every $(x^1, \Psi^1), (x^2, \Psi^2) \in B_R \times \mathcal{P}(B_R^Y)$,

$$\|\mathcal{T}(x^1, \Psi^1) - \mathcal{T}(x^2, \Psi^2)\|_{\mathcal{L}(\text{Lip}(U); \text{Lip}(U))} \leq L_{\mathcal{T}, R}(|x^1 - x^2| + \mathcal{W}_1(\Psi^1, \Psi^2)); \quad (3.7)$$

(T3) for every $R > 0$ there exists $\delta_R > 0$ such that for every $(x, \Psi) \in B_R \times \mathcal{P}_1(Y)$ we have

$$\mathcal{T}(x, \Psi) + \delta_R I \geq 0. \quad (3.8)$$

The following lemma is easily proved.

Lemma 3.1. *Let $\mathcal{T}(x, \Psi): \text{Lip}(U) \rightarrow \text{Lip}(U)$ satisfy (T0)-(T3) above and let $\mathcal{T}^*(x, \Psi): \mathcal{F}(U) \rightarrow \mathcal{F}(U)$ be its adjoint. Then*

(T*1) for every $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$, there exists $M_{\mathcal{T}} > 0$ such that the operator norm satisfies

$$\|\mathcal{T}^*(x, \Psi)\|_{\mathcal{L}(\mathcal{F}(U); \mathcal{F}(U))} \leq M_{\mathcal{T}}(1 + |x| + m_1(\Psi)); \quad (3.9)$$

(T*2) for every $R > 0$ there exists $L_{\mathcal{T}, R} > 0$ such that, for every $(x^1, \Psi^1), (x^2, \Psi^2) \in B_R \times \mathcal{P}(B_R^Y)$,

$$\|\mathcal{T}^*(x^1, \Psi^1) - \mathcal{T}^*(x^2, \Psi^2)\|_{\mathcal{L}(\mathcal{F}(U); \mathcal{F}(U))} \leq L_{\mathcal{T}, R}(|x^1 - x^2| + \mathcal{W}_1(\Psi^1, \Psi^2)), \quad (3.10)$$

(T*3) for every $R > 0$ there exists $\delta_R > 0$ such that

$$\mathcal{T}^*(x, \Psi) + \delta_R I \geq 0, \quad (3.11)$$

Proof. To see that (T*2) holds, we apply the definition of operator norm

$$\|\mathcal{T}^*(x, \Psi)\|_{\mathcal{L}(\mathcal{F}(U); \mathcal{F}(U))} := \sup \{ \|\mathcal{T}^*(x, \Psi)\mu\|_{\text{BL}} : \mu \in \mathcal{F}(U), \|\mu\|_{\text{BL}} \leq 1 \}. \quad (3.12)$$

For $(x^1, \Psi^1), (x^2, \Psi^2) \in B_R \times \mathcal{P}(B_R^Y)$, we have,

$$\begin{aligned} & \|\mathcal{T}^*(x^1, \Psi^1) - \mathcal{T}^*(x^2, \Psi^2)\|_{\mathcal{L}(\mathcal{F}(U); \mathcal{F}(U))} \\ &= \sup \{ \|(\mathcal{T}^*(x^1, \Psi^1) - \mathcal{T}^*(x^2, \Psi^2))\mu\|_{\text{BL}} : \mu \in \mathcal{F}(U), \|\mu\|_{\text{BL}} \leq 1 \} \\ &= \sup \left\{ \sup \{ \langle (\mathcal{T}^*(x^1, \Psi^1) - \mathcal{T}^*(x^2, \Psi^2))\mu, \varphi \rangle : \|\varphi\|_{\text{Lip}} \leq 1 \} : \mu \in \mathcal{F}(U), \|\mu\|_{\text{BL}} \leq 1 \right\} \\ &= \sup \left\{ \sup \{ \langle \mu, (\mathcal{T}(x^1, \Psi^1) - \mathcal{T}(x^2, \Psi^2))\varphi \rangle : \|\varphi\|_{\text{Lip}} \leq 1 \} : \mu \in \mathcal{F}(U), \|\mu\|_{\text{BL}} \leq 1 \right\} \\ &\leq \sup \left\{ \sup \{ \|\mu\|_{\text{BL}} \|(\mathcal{T}(x^1, \Psi^1) - \mathcal{T}(x^2, \Psi^2))\varphi\|_{\text{Lip}} : \|\varphi\|_{\text{Lip}} \leq 1 \} : \mu \in \mathcal{F}(U), \|\mu\|_{\text{BL}} \leq 1 \right\} \\ &\leq \sup \{ \|(\mathcal{T}(x^1, \Psi^1) - \mathcal{T}(x^2, \Psi^2))\varphi\|_{\text{Lip}} : \|\varphi\|_{\text{Lip}} \leq 1 \} \\ &\leq \|\mathcal{T}(x^1, \Psi^1) - \mathcal{T}(x^2, \Psi^2)\|_{\mathcal{L}(\text{Lip}(U); \text{Lip}(U))} \leq L_{\mathcal{T}, R}(|x^1 - x^2| + \mathcal{W}_1(\Psi^1, \Psi^2)), \end{aligned}$$

and (T*2) follows from (T2). A similar argument using (3.12) also gives (T*1). To prove (T*3), let $R > 0$ be fixed and let $\delta_R > 0$ be such that (3.8) holds. Then

$$\mathcal{T}^*(x, \Psi) + \delta_R I = (\mathcal{T}(x, \Psi) + \delta_R I)^* \geq 0,$$

since the adjoint operator preserves positivity. \square

Proposition 3.2. *For $y \in Y$ and $\Psi \in \mathcal{P}_1(Y)$, define $b_{\Psi}(y)$ as in (3.1). Assume that $v_{\Psi}: Y \rightarrow \mathbb{R}^d$ satisfies (v1)-(v3) and $\mathcal{T}(x, \Psi): \text{Lip}(U) \rightarrow \text{Lip}(U)$ satisfies (T0)-(T3). Then*

(i) for every $R > 0$, for every $\Psi \in \mathcal{P}(B_R^Y)$, and for every $y^1, y^2 \in B_R^Y$, there exists $L_R > 0$ such that

$$\|b_{\Psi}(y^1) - b_{\Psi}(y^2)\|_{\overline{Y}} \leq L_R \|y^1 - y^2\|_{\overline{Y}}; \quad (3.13)$$

(ii) for every $R > 0$, for every $\Psi^1, \Psi^2 \in \mathcal{P}(B_R^Y)$, and for every $y \in B_R^Y$, there exists $L_R > 0$ such that

$$\|b_{\Psi^1}(y) - b_{\Psi^2}(y)\|_{\overline{Y}} \leq L_R \mathcal{W}_1(\Psi^1, \Psi^2); \quad (3.14)$$

(iii) for every $R > 0$, there exists $\theta > 0$ such that for every $y \in B_R^Y$ and for every $\Psi \in \mathcal{P}(B_R^Y)$

$$y + \theta b_\Psi(y) \in Y; \quad (3.15)$$

(iv) there exists $M > 0$ such that for every $y \in Y$ and for every $\Psi \in \mathcal{P}_1(Y)$ there holds

$$\|b_\Psi(y)\|_{\bar{Y}} \leq M(1 + \|y\|_{\bar{Y}} + m_1(\Psi)). \quad (3.16)$$

Proof. Property (i) is a consequence of (3.2) and (3.10); property (ii) follows from (3.3) and (3.10). Observing that for all $y = (x, \lambda) \in Y$ we have that $\|\lambda\|_{\text{BL}} \leq 1$, we obtain (3.16) from (3.4) and (3.9). Finally, to prove (iii), since \mathbb{R}^d is convex, we simply have to show that for every $R > 0$, there exists $\theta > 0$ such that for every $y = (x, \lambda) \in B_R^Y$ and for every $\Psi \in \mathcal{P}(B_R^Y)$

$$\lambda + \theta \mathcal{T}^*(x, \Psi)\lambda \in \mathcal{P}(U). \quad (3.17)$$

From (3.11), we have that, for $\theta = 1/\delta_R$, $\lambda + \theta \mathcal{T}^*(x, \Psi)\lambda$ is a positive measure. Since by (3.5) $\langle \mathcal{T}^*(x, \Psi)\lambda, 1 \rangle = \langle \lambda, \mathcal{T}(x, \Psi)1 \rangle = 0$, we get (3.17) and we conclude the proof. \square

3.1. The discrete problem and statement of the main result. We consider a particle system of N agents evolving according to

$$\begin{cases} \dot{x}_t^i = v_{\Lambda_t^N}(x_t^i, \lambda_t^i), \\ \dot{\lambda}_t^i = \mathcal{T}^*(x_t^i, \Lambda_t^N)\lambda_t^i, \end{cases} \quad \text{for } i = 1, \dots, N, t \in [0, T], \quad (3.18)$$

where $x^i \in \mathbb{R}^d$, $\lambda^i \in \mathcal{P}(U)$ for each $i \in \{1, \dots, N\}$, and

$$\Lambda_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{(x_t^i, \lambda_t^i)} \quad (3.19)$$

is the empirical measure associated with the system. Recalling the definition of b in (3.1), the evolution (3.18) can be written in compact form as

$$y_t^i = b_{\Lambda_t^N}(y_t^i), \quad \text{for } i = 1, \dots, N, t \in [0, T]. \quad (3.20)$$

We first discuss the well posedness of system (3.18) for every choice of an initial datum $\bar{y}^i = (\bar{x}^i, \bar{\lambda}^i)$, for $i = 1, \dots, N$.

Proposition 3.3. *Assume that for every $y \in Y$ and $\Psi \in \mathcal{P}_1(Y)$ the velocity $v_\Psi: Y \rightarrow \mathbb{R}^d$ satisfies (v1)-(v3) and the operator $\mathcal{T}(x, \Psi): \text{Lip}(U) \rightarrow \text{Lip}(U)$ satisfies (T0)-(T3). Then, for every choice of $\bar{y}^i \in Y$, $i = 1, \dots, N$, the system (3.20) has a unique solution.*

Proof. We introduce the vector-valued variable $\mathbf{y} := (y^1, \dots, y^N) \in Y^N \subset \bar{Y}^N$, which we endow with the norm

$$\|\mathbf{y}\|_{\bar{Y}^N} := \frac{1}{N} \sum_{i=1}^N \|y^i\|_{\bar{Y}}, \quad (3.21)$$

and the associated empirical measure $\Lambda^N := \frac{1}{N} \sum_{i=1}^N \delta_{y^i}$, which belongs to $\mathcal{P}(B_R^Y)$ whenever $\mathbf{y} \in (B_R^Y)^N$. Consider the map $\mathbf{b}^N: Y^N \rightarrow \bar{Y}^N$ whose components are defined through

$$b_i^N(\mathbf{y}) := b_{\Lambda^N}(y^i). \quad (3.22)$$

Then the Cauchy problem associated with (3.20) can be written as

$$\begin{cases} \dot{\mathbf{y}}_t = \mathbf{b}^N(\mathbf{y}_t), \\ \mathbf{y}_0 = \bar{\mathbf{y}}. \end{cases} \quad (3.23)$$

In order to apply Corollary 2.3 to the system above, we first notice that assumption (ii) is automatically satisfied since the system is autonomous. To check the other assumptions, we fix a ball $B_R^{Y^N}$ and notice that $B_R^{Y^N} \subset (B_R^Y)^N$.

Applying (3.15) with $\Psi = \Lambda^N$ to each component y^i of \mathbf{y} , we get that assumption (iii) of Corollary 2.3 is satisfied.

We now show that assumption (i') holds. Fix $\mathbf{y}_1, \mathbf{y}_2 \in B_R^{Y^N} \subset (B_R^Y)^N$, and let Λ_1^N, Λ_2^N be the associated empirical measures. Notice that

$$\mathcal{W}_1(\Lambda_1^N, \Lambda_2^N) \leq \frac{1}{N} \sum_{i=1}^N \|y_1^i - y_2^i\|_{\bar{Y}} = \|\mathbf{y}_1 - \mathbf{y}_2\|_{\bar{Y}^N}. \quad (3.24)$$

With this, by the triangle inequality, (3.13), and (3.14), we estimate

$$\begin{aligned} \|\mathbf{b}^N(\mathbf{y}_1) - \mathbf{b}^N(\mathbf{y}_2)\|_{\bar{Y}^N} &= \frac{1}{N} \sum_{i=1}^N \|b_{\Lambda_1^N}(y_1^i) - b_{\Lambda_2^N}(y_2^i)\|_{\bar{Y}} \\ &\leq L_R \mathcal{W}_1(\Lambda_1^N, \Lambda_2^N) + \frac{L_R}{N} \sum_{i=1}^N \|y_1^i - y_2^i\|_{\bar{Y}} \leq 2L_R \|\mathbf{y}_1 - \mathbf{y}_2\|_{\bar{Y}^N}, \end{aligned} \quad (3.25)$$

which yields (2.7). To see that also (2.8), and therefore assumption (i''), holds, we apply (3.16) and obtain, upon noticing that $m_1(\Lambda^N) = \|\mathbf{y}\|_{\bar{Y}^N}$,

$$\|\mathbf{b}^N(\mathbf{y})\|_{\bar{Y}^N} = \frac{1}{N} \sum_{i=1}^N \|b_{\Lambda^N}(y^i)\|_{\bar{Y}} \leq \frac{M}{N} \sum_{i=1}^N (1 + \|y^i\|_{\bar{Y}} + m_1(\Lambda^N)) = M(1 + 2\|\mathbf{y}\|_{\bar{Y}^N}), \quad (3.26)$$

Therefore we can apply Corollary 2.3, which proves the statement. \square

We are now in a position to state the main result of our paper, concerning the mean-field limit as $N \rightarrow \infty$ of the solutions (y_t^1, \dots, y_t^N) to (3.20), or equivalently the limiting behavior of the associated empirical measures Λ_t^N . In order to do so, we first need to recall the concept of Eulerian solution to the continuity equation.

Definition 3.4 (Eulerian solution). *Let $\Lambda \in C^0([0, T]; (\mathcal{P}_1(Y), \mathcal{W}_1))$ and let $\bar{\Lambda} \in \mathcal{P}_c(Y)$ be a given initial datum. We say that Λ is a Eulerian solution to the initial value problem for the equation*

$$\partial_t \Lambda_t + \operatorname{div}(b_{\Lambda_t} \Lambda_t) = 0 \quad (3.27)$$

starting from $\bar{\Lambda}$ if and only if $\Lambda_0 = \bar{\Lambda}$ and, for every $\phi \in C_b^1([0, T] \times \bar{Y})$,

$$\int_Y \phi(t, y) d\Lambda_t(y) - \int_Y \phi(0, y) d\Lambda_0(y) = \int_0^t \int_Y (\partial_t \phi(s, y) + D\phi(s, y) \cdot b_{\Lambda_s}(y)) d\Lambda_s(y) ds, \quad (3.28)$$

where $D\phi(s, y)$ is the Fréchet differential of ϕ in the y variable.

The main result of our paper is the following theorem, stating the existence of a unique Eulerian solution to (3.27) and its characterization as the mean-field limit of solutions to the discrete problem (3.20).

Theorem 3.5. *Let $r > 0$ and $\bar{\Lambda} \in \mathcal{P}(B_r^Y)$ be a given initial datum. Then*

- (i) *there exists a unique Eulerian solution $t \mapsto \Lambda_t$ to (3.27) starting from $\bar{\Lambda}$;*
- (ii) *if $\bar{\Lambda}^N = \frac{1}{N} \sum_{i=1}^N \delta_{\bar{y}^{N,i}}$ is a sequence of atomic measures in $\mathcal{P}(B_r^Y)$ such that*

$$\lim_{N \rightarrow \infty} \mathcal{W}_1(\bar{\Lambda}^N, \bar{\Lambda}) = 0$$

and, for fixed N , Λ_t^N are the empirical measures associated with the unique solution to (3.20) with initial datum $\bar{y}^{N,i}$, we have

$$\lim_{N \rightarrow \infty} \mathcal{W}_1(\Lambda_t^N, \Lambda_t) = 0 \quad \text{uniformly with respect to } t \in [0, T].$$

The proof of Theorem 3.5 will be based on a fixed point argument and on the notion of Lagrangian solution, which are going to be introduced in the next subsection.

3.2. Lagrangian solutions. We start by proving an auxiliary well posedness result for and ODE in Y of the form

$$\dot{y}_t = b_{\Psi_t}(y_t), \quad y_0 = \bar{y}, \quad (3.29)$$

where $[0, T] \ni t \mapsto \Psi_t \in \mathcal{P}_1(Y)$ is a given continuous curve and $\bar{y} \in Y$.

Proposition 3.6. *Assume that for every $y \in Y$ and $\Psi \in \mathcal{P}_1(Y)$ the velocity $v_\Psi: Y \rightarrow \mathbb{R}^d$ satisfies (v1)-(v3) and the operator $\mathcal{T}(x, \Psi): \operatorname{Lip}(U) \rightarrow \operatorname{Lip}(U)$ satisfies (T0)-(T3). Let $\Psi \in C^0([0, T]; (\mathcal{P}_1(Y), \mathcal{W}_1))$ and assume that there exists $R > 0$ such that, in addition, $\Psi_t \in \mathcal{P}(B_R^Y)$ for all $t \in [0, T]$. Then, for every choice of $\bar{y} \in Y$, the ODE (3.29) has a unique solution.*

Proof. We set $b(t, y) := b_{\Psi_t}(y)$ according to (3.1). Since $t \mapsto \Psi_t$ is continuous, using (3.14) we get that, for any fixed $y \in Y$, $b(\cdot, y)$ is continuous, which is condition (ii) in Corollary 2.3. Condition (iii) of Corollary 2.3 is a direct consequence of (3.15). Furthermore, (3.13) and (3.16) yield (2.7) and (2.8), respectively. The proof is concluded by Corollary 2.3. \square

In view of the previous result, the following definition is justified.

Definition 3.7 (Transition map). *The transition map $\mathbf{Y}_\Psi(t, s, \bar{y})$ associated with the ODE (3.29), replacing the initial condition by $y_s = \bar{y}$, is defined through*

$$\mathbf{Y}_\Psi(t, s, \bar{y}) = y_t,$$

where $t \mapsto y_t$ is the unique solution to (3.29).

We can now proceed to defining the notion of Lagrangian solution to (3.27).

Definition 3.8 (Lagrangian solutions). *Let $\Lambda \in C^0([0, T]; (\mathcal{P}_1(Y), \mathcal{W}_1))$ and let $\bar{\Lambda} \in \mathcal{P}_c(Y)$ be a given initial datum. We say that Λ is a Lagrangian solution to the initial value problem for the equation (3.27) starting from $\bar{\Lambda}$ if and only if it satisfies the fixed point condition*

$$\Lambda_t = \mathbf{Y}_\Lambda(t, 0, \cdot)_{\#} \bar{\Lambda} \quad \text{for every } 0 \leq t \leq T, \quad (3.30)$$

where $\mathbf{Y}_\Lambda(t, s, y)$ are the transition maps associated with the ODE (3.29).

Remark 3.9. *It follows from Definition 2.1 that Lagrangian solutions are also Eulerian solutions.*

Remark 3.10. *For a fixed $N \in \mathbb{N}$, let Λ_t^N be the empirical measures associated with the unique solution to (3.20) with initial datum \bar{y}^i , $i = 1, \dots, N$. If we now set $\bar{\Lambda}^N := \frac{1}{N} \sum_{i=1}^N \delta_{\bar{y}^i}$, by Definition 3.7 there holds*

$$\Lambda_t^N = \mathbf{Y}_{\Lambda^N}(t, 0, \cdot)_{\#} \bar{\Lambda}^N \quad \text{for every } 0 \leq t \leq T. \quad (3.31)$$

Hence, Λ^N is a Lagrangian and Eulerian solution to (3.27) starting from $\bar{\Lambda}^N$.

We now want to show that an infinite-dimensional converse of Proposition 3.6 holds, proving that indeed, in our case, every Eulerian solution is also a Lagrangian solution. This stems out of a general abstract principle known as the *superposition principle*, in the version introduced in [5] (see also [7, Theorem 8.2.1] and [9, Theorem 7.1]). In the statement below, the *evaluation map* ev_t is defined, at a given $t \in [0, T]$, by

$$\text{ev}_t(\gamma) := \gamma(t) \quad \text{for all } \gamma \in C([0, T] : E).$$

Theorem 3.11 (superposition principle). *Let $(E, \|\cdot\|_E)$ be a separable Banach space, let $b : (0, T) \times E \rightarrow E$ be a Borel vector field and let $\mu_t \in \mathcal{P}(E)$, $t \in [0, T]$, be a continuous curve with*

$$\int_0^T \int_E \|b_t\|_E d\mu_t dt < +\infty. \quad (3.32)$$

If

$$\frac{d}{dt} \mu_t + \text{div}(b_t \mu_t) = 0$$

in duality with cylindrical functions $\phi \in C_b^1(E)$, precisely of the form (here $\langle \cdot, \cdot \rangle$ denotes the duality map between E and E')

$$\varphi(\langle y, z'_1 \rangle, \langle y, z'_2 \rangle, \dots, \langle y, z'_N \rangle)$$

with $\varphi \in C_b^1(\mathbb{R}^N)$ and $z'_1, \dots, z'_N \in E'$, then there exists $\eta \in \mathcal{P}(C([0, T]; E))$ concentrated on absolutely continuous solutions to the ODE $\dot{y} = b_t(y)$ and with $(\text{ev}_t)_{\#} \eta = \mu_t$ for all $t \in [0, T]$.

Proof. See [5, Theorem 5.2]. □

Combining the abstract result Theorem 3.11 with the uniqueness granted by Proposition 3.6, we can prove the announced equivalence result. Notice that the proof has an intermediate step, since in order to apply Proposition 3.6 we must first ensure that a Eulerian solution Λ_t has (equi)compact support for all t . We are able to deduce this from Theorem 3.11 and the assumption that the initial datum $\bar{\Lambda} \in \mathcal{P}_c(Y)$.

Theorem 3.12. *Let $\Lambda \in C^0([0, T]; (\mathcal{P}_1(Y), \mathcal{W}_1))$ and let $\bar{\Lambda} \in \mathcal{P}_c(Y)$ be a given initial datum. Assume that Λ is a Eulerian solution to the initial value problem for $\partial_t \Lambda_t + \text{div}(b_{\Lambda_t} \Lambda_t) = 0$ (see (3.27)) starting from $\bar{\Lambda}$, in the sense of (3.28), then there exists $R > 0$ such that $\Lambda_t \in \mathcal{P}(B_R^Y)$ for all $t \in [0, T]$ and*

$$\Lambda_t = \mathbf{Y}_\Lambda(t, 0, \cdot)_{\#} \bar{\Lambda} \quad \text{for every } 0 \leq t \leq T,$$

where $\mathbf{Y}_\Lambda(t, s, y)$ are the transition maps associated with the ODE (3.29).

Proof. Since $\Lambda \in C^0([0, T]; (\mathcal{P}_1(Y), \mathcal{W}_1))$, the map

$$t \mapsto m_1(\Lambda_t) = \int_Y \|y\|_{\bar{Y}} d\Lambda_t(y) \quad (3.33)$$

is continuous, and hence bounded, in $[0, T]$. Set $b_t(y) := b_{\Lambda_t}(y)$, for $y \in Y$, and extend it to zero on $\bar{Y} \setminus Y$. Using (3.16), we have

$$\begin{aligned} \int_0^T \int_{\bar{Y}} \|b_t\|_{\bar{Y}} d\Lambda_t(y) dt &= \int_0^T \int_Y \|b_t\|_{\bar{Y}} d\Lambda_t(y) dt \\ &\leq \int_0^T \int_Y M(1 + \|y\|_{\bar{Y}} + m_1(\Lambda_t)) d\Lambda_t(y) dt \leq TM \left(1 + 2 \max_{t \in [0, T]} m_1(\Lambda_t)\right) < +\infty. \end{aligned} \quad (3.34)$$

Hence, we can apply Theorem 3.11 with $E = \bar{Y}$ and $\mu_t = \Lambda_t$, obtaining that $\Lambda_t = (\text{ev}_t)_\# \eta$ for a suitable $\eta \in \mathcal{P}(C([0, T]; \bar{Y}))$ concentrated on absolutely continuous solutions to ODE

$$\dot{y} = b_{\Lambda_t}(y) \quad \text{in } [0, T]. \quad (3.35)$$

Now, using (3.16) again, we have

$$\|b_{\Lambda_t}(y)\|_{\bar{Y}} \leq M \left(1 + \max_{t \in [0, T]} m_1(\Lambda_t) + \|y\|_{\bar{Y}}\right) \leq M_\Lambda (1 + \|y\|_{\bar{Y}}), \quad (3.36)$$

where we set $M_\Lambda := M(1 + \max_{t \in [0, T]} m_1(\Lambda_t))$. The equality $\bar{\Lambda} = \Lambda_0 = (\text{ev}_0)_\# \eta$, which reads

$$\int_{\mathcal{P}(C([0, T]; \bar{Y}))} \phi(\gamma(0)) d\eta(\gamma) = \int_Y \phi(y) d\bar{\Lambda}(y)$$

for each $\phi \in C_b(Y)$, implies that η is concentrated on the set of solutions to (3.35) satisfying $y(0) \in B_r^Y$, where r is such that $\text{supp } \bar{\Lambda} \subset B_r^Y$. With (3.13) and the Grönwall inequality, each of these solutions must satisfy $y(t) \in B_R^Y$, where R is explicitly given by

$$R := R_{r, M, \Lambda, T} = (r + M_\Lambda T) e^{M_\Lambda T}.$$

From the equality $\Lambda_t = (\text{ev}_t)_\# \eta$ we then deduce that $\Lambda_t \in \mathcal{P}_1(B_R^Y)$ for all $t \in [0, T]$. We can therefore apply Proposition 3.6 and exploit the uniqueness of the solution to the Cauchy problem (3.29) to deduce the representation

$$\gamma(t) = \mathbf{Y}_\Lambda(t, 0, \gamma(0))$$

with $\gamma(0) \in B_r^Y$, for each continuous path $\gamma \in \text{supp } \eta$. With the equality $\Lambda_t = (\text{ev}_t)_\# \eta$, this gives

$$\int_Y \phi(y) d\Lambda_t(y) = \int_{\mathcal{P}(C([0, T]; \bar{Y}))} \phi(\mathbf{Y}_\Lambda(t, 0, \gamma(0))) d\eta(\gamma) = \int_Y \phi(\mathbf{Y}_\Lambda(t, 0, y)) d\bar{\Lambda}(y)$$

for each $\phi \in C_b(Y)$, which implies the conclusion. \square

3.3. Proof of Theorem 3.5. As a preliminary step towards the proof, we need the following lemma, assuring that the size of the support of a Lagrangian solution in the sense of (3.30) can be *a priori* estimated from the data of the problem.

Lemma 3.13. *Assume that for every $y \in Y$ and $\Psi \in \mathcal{P}_1(Y)$ the velocity $v_\Psi: Y \rightarrow \mathbb{R}^d$ satisfies (v1)-(v3) and the operator $\mathcal{T}(x, \Psi): \text{Lip}(U) \rightarrow \text{Lip}(U)$ satisfies (T0)-(T3). Let $\Lambda \in C^0([0, T]; (\mathcal{P}_1(Y), \mathcal{W}_1))$ and let $\bar{\Lambda} \in \mathcal{P}_c(Y)$ be a given initial datum. Fix $r > 0$ such that $\bar{\Lambda}$ has support in B_r^Y , and let M be the constant given by (3.16). Assume that Λ is a Lagrangian solution to the initial value problem for the equation (3.27) starting from $\bar{\Lambda}$ in the sense of (3.30). Then, for $R = (r + MT)e^{2MT}$ we have*

$$\Lambda_t \in \mathcal{P}_1(B_R^Y) \quad \text{for all } t \in [0, T].$$

Proof. For r, R as in the statement, it suffices to show that we have

$$\max_{y \in B_r^Y} \|\mathbf{Y}_{\Lambda_t}(t, 0, y)\|_{\bar{Y}} \leq R \quad (3.37)$$

for all $t \in [0, T]$. Indeed, if this holds the statement immediately follows by (3.30) and elementary properties of the push-forward measure, taking into account that $\bar{\Lambda}$ has support

in B_r^Y . To prove the above claim, we first observe, that by definition of Lagrangian solutions and the fact that $\bar{\Lambda} \in \mathcal{P}(B_r^Y)$ we immediately have

$$m_1(\Lambda_t) \leq \max_{y \in B_r^Y} \|\mathbf{Y}_\Lambda(t, 0, y)\|_{\bar{Y}} \quad (3.38)$$

for all $t \in [0, T]$. We now set $f(s) = \max_{y \in B_r^Y} \|\mathbf{Y}_\Lambda(s, 0, y)\|_{\bar{Y}}$. Then, one has by definition of the transition map, (3.16), and (3.38) that for every choice of $y \in B_r^Y$

$$\|\mathbf{Y}_\Lambda(t, 0, y)\|_{\bar{Y}} \leq r + M \int_0^t (1 + \|\mathbf{Y}_\Lambda(s, 0, y)\|_{\bar{Y}} + m_1(\Lambda_s)) ds \leq r + M \int_0^t (1 + 2f(s)) ds$$

which implies by the Gronwall inequality $f(t) \leq (r + Mt)e^{2Mt}$ for all t , confirming (3.37). \square

Proof of Theorem 3.5. The proof goes through a finite dimensional approximation and involves three steps.

Step 1: stability of Lagrangian solutions. We fix $r > 0$, two initial data $\bar{\Lambda}^1, \bar{\Lambda}^2 \in \mathcal{P}(B_r^Y)$ and assume that two Lagrangian solutions Λ_t^1, Λ_t^2 , starting from $\bar{\Lambda}^1$, and $\bar{\Lambda}^2$, respectively, exist. We fix $R = (r + MT)e^{2MT}$ and the corresponding constant L_R provided by (3.13)-(3.14). We claim that

$$\mathcal{W}_1(\Lambda_t^1, \Lambda_t^2) \leq e^{L_R t + e^{L_R t} - 1} \mathcal{W}_1(\bar{\Lambda}^1, \bar{\Lambda}^2) \quad \text{for all } t \in [0, T]. \quad (3.39)$$

To prove this claim, we fix y_1 and $y_2 \in B_r^Y$ and observe that

$$\|\mathbf{Y}_{\Lambda^i}(t, 0, y_i)\|_{\bar{Y}} \leq R \quad (3.40)$$

for all $t \in [0, T]$ and $i = 1, 2$. This can be proved along similar lines as in the proof of (3.38). With (3.40) and (3.13)-(3.14), the solutions $y^1(t)$ and $y^2(t)$ to the ODE's $\dot{y}^i = b_{\Lambda^i}(y^i)$ with initial data y^1 and y^2 respectively, satisfy

$$\begin{aligned} \frac{d}{dt} \|y^1 - y^2\|(t) &\leq \|b_{\Lambda^1}(y^1(t)) - b_{\Lambda^1}(y^2(t))\| + \|b_{\Lambda^1}(y^2(t)) - b_{\Lambda^2}(y^2(t))\| \\ &\leq L_R \|y^1 - y^2\|(t) + L_R \mathcal{W}_1(\Lambda_t^1, \Lambda_t^2). \end{aligned}$$

This gives, by means of a comparison argument, that

$$\|y^1(t) - y^2(t)\| \leq e^{L_R t} \|y^1 - y^2\| + L_R \int_0^t e^{L_R(t-\tau)} \mathcal{W}_1(\Lambda_\tau^1, \Lambda_\tau^2) d\tau;$$

equivalently,

$$\|\mathbf{Y}_{\Lambda^1}(t, 0, y^1) - \mathbf{Y}_{\Lambda^2}(t, 0, y^2)\| \leq e^{L_R t} \|y^1 - y^2\| + L_R \int_0^t e^{L_R(t-\tau)} \mathcal{W}_1(\Lambda_\tau^1, \Lambda_\tau^2) d\tau \quad (3.41)$$

for all $t \in [0, T]$ and y_1 and $y_2 \in B_r^Y$.

Now, let Π be an optimal coupling between $\bar{\Lambda}^1$ and $\bar{\Lambda}^2$. Then clearly, by the definition of Lagrangian solutions, $(\mathbf{Y}_{\Lambda^1}(t, 0, y^1), \mathbf{Y}_{\Lambda^2}(t, 0, y^2))_{\#} \Pi$ is a coupling between Λ_t^1 and Λ_t^2 . Therefore

$$\begin{aligned} \mathcal{W}_1(\Lambda_t^1, \Lambda_t^2) &\leq \int_{Y \times Y} \|\mathbf{Y}_{\Lambda^1}(t, 0, y^1) - \mathbf{Y}_{\Lambda^2}(t, 0, y^2)\| d\Pi(y^1, y^2) \\ &= \int_{B_r^Y \times B_r^Y} \|\mathbf{Y}_{\Lambda^1}(t, 0, y^1) - \mathbf{Y}_{\Lambda^2}(t, 0, y^2)\| d\Pi(y^1, y^2), \end{aligned}$$

where we also used that $\bar{\Lambda}^1, \bar{\Lambda}^2 \in \mathcal{P}(B_r^Y)$. Hence, using (3.41) we get

$$\begin{aligned} \mathcal{W}_1(\Lambda_t^1, \Lambda_t^2) &\leq e^{L_R t} \int_{B_r^Y \times B_r^Y} \|y^1 - y^2\| d\Pi(y^1, y^2) + L_R \int_0^t e^{L_R(t-\tau)} \mathcal{W}_1(\Lambda_\tau^1, \Lambda_\tau^2) d\tau \\ &= e^{L_R t} \mathcal{W}_1(\bar{\Lambda}^1, \bar{\Lambda}^2) + L_R \int_0^t e^{L_R(t-\tau)} \mathcal{W}_1(\Lambda_\tau^1, \Lambda_\tau^2) d\tau. \end{aligned}$$

With this and the Grönwall inequality, we get (3.39).

Step 2: existence and approximation of Lagrangian solutions. We start by fixing a sequence of atomic measures $\bar{\Lambda}^N \in \mathcal{P}(B_r^Y)$ such that

$$\lim_{N \rightarrow \infty} \mathcal{W}_1(\bar{\Lambda}^N, \bar{\Lambda}) = 0. \quad (3.42)$$

Such a sequence can be for instance constructed as follows: choose $\bar{y}_i(\omega) \in Y$ independent and identically distributed, with law $\bar{\Lambda}$, so that the random measures $\bar{\Lambda}^N(\omega) := \frac{1}{N} \sum_{i=1}^N \delta_{\bar{y}_i(\omega)}$ almost surely converge in $\mathcal{P}_1(Y)$ to $\bar{\Lambda}$, and choose ω such that this happens. Now, let Λ_t^N be the empirical measures associated with the unique solution to (3.20) with initial datum \bar{y}^i , $i = 1, \dots, N$. As noticed in (3.31), Λ_t^N are Lagrangian solutions to (3.27) starting from $\bar{\Lambda}^N$. Hence, (3.39) provides a constant $C := C(M, r, T)$ such that

$$\mathcal{W}_1(\Lambda_t^N, \Lambda_t^M) \leq C \mathcal{W}_1(\bar{\Lambda}^N, \bar{\Lambda}^M)$$

for all $t \in [0, T]$ and $N, M \in \mathbb{N}$. It follows that $\Lambda_t^N \in C([0, T]; (\mathcal{P}_1(B_R^Y), \mathcal{W}_1))$ is a Cauchy sequence. Let then $\Lambda_t \in C([0, T]; (\mathcal{P}_1(B_R^Y), \mathcal{W}_1))$ be the limit of the sequence Λ_t^N .

For a given $\bar{y} \in B_r^Y$, consider now the solutions $y^N(t)$ and $y(t)$ to the ODE's $\dot{y}^N = b_{\Lambda^N}(y^N)$, and $\dot{y} = b_{\Lambda}(y)$, respectively, with initial datum \bar{y} . Let $R' \geq R$ be an upper bound² for $\max_{t \in [0, T]} \|y(t)\|_Y$, which can be taken independent of $\bar{y} \in B_r^Y$. With (3.13)-(3.14) we obtain

$$\frac{d}{dt} \|y^N - y\|(t) \leq \|b_{\Lambda^N}(y^N) - b_{\Lambda}(y^N)\| + \|b_{\Lambda}(y^N) - b_{\Lambda}(y)\| \leq L_{R'} \|y^N - y\| + L_R \mathcal{W}_1(\Lambda_t^N, \Lambda_t).$$

Again by comparison, we deduce that

$$\|\mathbf{Y}_{\Lambda^N}(t, 0, \bar{y}) - \mathbf{Y}_{\Lambda}(t, 0, \bar{y})\| \leq L_R \int_0^t e^{L_{R'}(t-\tau)} \mathcal{W}_1(\Lambda_{\tau}^N, \Lambda_{\tau}) d\tau$$

which entails the uniform convergence of $\mathbf{Y}_{\Lambda^N}(\cdot, 0, \cdot)$ to $\mathbf{Y}_{\Lambda}(\cdot, 0, \cdot)$ in $[0, T] \times B_r^Y$. For each $t \in [0, T]$ this implies, together with (3.42) and the fact that $\mathbf{Y}_{\Lambda}(t, 0, \cdot)$ is a Lipschitz map on B_r^Y , that

$$\Lambda_t^N = \mathbf{Y}_{\Lambda^N}(t, 0, \cdot)_{\#} \bar{\Lambda}^N \rightarrow \mathbf{Y}_{\Lambda}(t, 0, \cdot)_{\#} \bar{\Lambda}$$

in $\mathcal{P}_1(Y)$, which gives (3.30).

Step 3: uniqueness and conclusion. Uniqueness of Lagrangian solutions, given the initial datum, follows now from (3.39). Existence and uniqueness of Eulerian solutions is now a consequence of Remark 3.9, and Theorem 3.12, respectively. The same argument used in the second step of this proof gives also part (ii) of the statement. \square

4. A LEADER-FOLLOWER DYNAMICS

We want to discuss the application of our results to the relevant scenario of two interacting populations, one consisting of “leaders” and the other one of “followers”, with a switching rate between the two. In this setting, the set U consists of two elements, that is $U := \{F, L\}$ and is endowed with a two-valued distance

$$0 = \text{dist}(F, F) = \text{dist}(L, L), \quad 1 = \text{dist}(F, L) = \text{dist}(L, F).$$

The space $\text{Lip}(\{F, L\})$ is a two-dimensional linear space spanned by the two indicator functions $\mathbb{1}_F$ and $\mathbb{1}_L$; accordingly, the space $\mathcal{F}(\{F, L\})$ is the two-dimensional space of signed Borel measures on the discrete set $\{F, L\}$, whose generic element ξ is completely described by the two values $\xi_F := \xi(\{F\})$ and $\xi_L := \xi(\{L\})$. A simple characterization of the operators \mathcal{T} (and \mathcal{T}^*) complying with our set of assumptions (T0)-(T3) is then given in the following proposition.

Proposition 4.1. *Let $U := \{F, L\}$. Then $\mathcal{T}: \mathbb{R}^d \times \mathcal{P}_1(Y) \rightarrow \text{Lip}(\{F, L\})$ satisfies (T0)-(T3) if and only if there exist two functions $\alpha_F, \alpha_L: \mathbb{R}^d \times \mathcal{P}_1(Y) \rightarrow [0, +\infty)$ such that*

($\alpha 0$) *for every $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$ and $\lambda \in \mathcal{P}_1(\{F, L\})$ it holds*

$$\begin{aligned} (\mathcal{T}^*(x, \Psi)\lambda)_F &= -\alpha_F(x, \Psi)\lambda_F + \alpha_L(x, \Psi)(1 - \lambda_F), \\ (\mathcal{T}^*(x, \Psi)\lambda)_L &= \alpha_F(x, \Psi)\lambda_F - \alpha_L(x, \Psi)(1 - \lambda_F); \end{aligned}$$

($\alpha 1$) *for every $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$, there exists $M_{\mathcal{T}} > 0$ such that the*

$$0 \leq \alpha_{\bullet}(x, \Psi) \leq M_{\mathcal{T}}(1 + |x| + m_1(\Psi)), \quad \text{for } \bullet = F, L;$$

²observe that at this point of the proof we cannot a priori exclude that $R' > R$, since we still do not know that Λ_t is a Lagrangian solution, hence we cannot apply (3.40) (which holds instead for Λ^N)

($\alpha 2$) for every $R > 0$ there exists $L_{\mathcal{T},R} > 0$ such that, for every $(x^1, \Psi^1), (x^2, \Psi^2) \in B_R \times \mathcal{P}(B_R^Y)$,

$$|\alpha_{\bullet}(x^1, \Psi^1) - \alpha_{\bullet}(x^2, \Psi^2)| \leq L_{\mathcal{T},R}(|x^1 - x^2| + \mathcal{W}_1(\Psi^1, \Psi^2)), \quad \text{for } \bullet = F, L.$$

Proof. Since $\text{Lip}(\{F, L\})$ is a two-dimensional linear space, we can identify $\mathcal{T}(x, \Psi)$ with its matrix representation with respect to the canonical basis $\{\mathbb{1}_F, \mathbb{1}_L\}$ and endow the space $\mathcal{L}(\text{Lip}(\{F, L\}), \text{Lip}(\{F, L\}))$ with the Frobenius norm of such a matrix representation. Accordingly, the transpose matrix will be the representation of $\mathcal{T}^*(x, \Psi)$ with respect to the canonical basis of $\mathcal{F}(\{F, L\})$ consisting of the two Dirac masses δ_F and δ_L .

Now, condition (T0) is equivalent to the fact that

$$\mathcal{T}(x, \Psi) = \begin{pmatrix} -\alpha_F(x, \Psi) & \alpha_F(x, \Psi) \\ \alpha_L(x, \Psi) & -\alpha_L(x, \Psi) \end{pmatrix} \quad (4.1)$$

with $\alpha_F, \alpha_L: \mathbb{R}^d \times \mathcal{P}_1(Y) \rightarrow [0, +\infty)$. In turn, this is equivalent to ($\alpha 0$) by a direct computation. Again by a direct computation, (T2) is equivalent to ($\alpha 2$), while (T1) is equivalent to the inequalities

$$|\alpha_{\bullet}(x, \Psi)| \leq M_{\mathcal{T}}(1 + |x| + m_1(\Psi)), \quad \text{for } \bullet = F, L.$$

In particular, α_F and α_L are uniformly bounded for $(x, \Psi) \in B_R \times \mathcal{P}_1(B_R^Y)$, hence (T3) holds if and only if the nondiagonal elements $\alpha_F(x, \Psi)$ and $\alpha_L(x, \Psi)$ of the matrix representation (4.1) are nonnegative. This implies that (T1) and (T3) together are equivalent to ($\alpha 1$), which concludes the proof. \square

To each element Ψ of $\mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\}))$ we can associate a followers' and a leaders' distribution in a natural way, as we are going to discuss in the next definition.

Definition 4.2. Let $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\}))$ The followers distribution μ_{Ψ}^F associated to Ψ is the positive Borel measure on \mathbb{R}^d defined by

$$\mu_{\Psi}^F(B) := \int_{B \times \mathcal{P}_1(\{F, L\})} \lambda_F d\Psi(x, \lambda) \quad (4.2a)$$

for each Borel set $B \subset \mathbb{R}^d$. Similarly, the leaders distribution μ_{Ψ}^L associated to Ψ is the positive Borel measure on \mathbb{R}^d defined by

$$\mu_{\Psi}^L(B) := \int_{B \times \mathcal{P}_1(\{F, L\})} (1 - \lambda_F) d\Psi(x, \lambda). \quad (4.2b)$$

Both the measures defined above have a simple interpretation. For instance, $\mu_{\Psi}^F(B)$ is the expected value of the number of leaders in a region B for a probability distribution Ψ on $\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})$. We also observe that the sum of the two measures μ_{Ψ}^F and μ_{Ψ}^L is exactly the x -marginal of $\Psi(x, \lambda)$. From a practical point of view, the integrals appearing in (4.2) can be computed identifying the metric space $(\mathcal{P}_1(\{F, L\}), \mathcal{W}_1)$ with the 1-dimensional simplex $[0, 1]$ endowed with the Euclidean distance (this is indeed an isometry). We also point out the following inequalities, whose proof is straightforward

$$\begin{aligned} m_1(\mu_{\Psi}^F) + m_1(\mu_{\Psi}^L) &\leq m_1(\Psi), \quad \text{for all } \Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})); \\ \|\mu_{\Psi_1}^F - \mu_{\Psi_2}^F\|_{BL} &\leq 2\mathcal{W}_1(\Psi_1, \Psi_2) \quad \text{for all } \Psi_1, \Psi_2 \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})); \\ \|\mu_{\Psi_1}^L - \mu_{\Psi_2}^L\|_{BL} &\leq 2\mathcal{W}_1(\Psi_1, \Psi_2) \quad \text{for all } \Psi_1, \Psi_2 \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})). \end{aligned} \quad (4.3)$$

Using the previous definition, we can also provide some relevant examples of velocity fields v_{Ψ} complying with assumptions (v1)-(v3).

Proposition 4.3. Let $U = \{F, L\}$. For $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\}))$, consider the velocity field

$$v_{\Psi}(x, \lambda) := \lambda_F (K^{FF} \star \mu_{\Psi}^F + K^{LF} \star \mu_{\Psi}^L) + (1 - \lambda_F) (K^{FL} \star \mu_{\Psi}^F + K^{LL} \star \mu_{\Psi}^L) \quad (4.4)$$

where μ_{Ψ}^F and μ_{Ψ}^L are defined in (4.2) and the interaction kernels $K^{ij}: \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfy

$$\begin{aligned} |K^{ij}(x)| &\leq M(1 + |x|), \quad \text{for all } x \in \mathbb{R}^d; \\ |K^{ij}(x_1) - K^{ij}(x_2)| &\leq L_R |x_1 - x_2|, \quad \text{for all } x_1, x_2 \in B_R. \end{aligned}$$

Then, $v_{\Psi}(x, \lambda)$ satisfies (v1)-(v3).

Proof. The result follows by a direct computation. \square

Remark 4.4. *The velocity field (4.4) corresponds to a particle model where each follower experiences a velocity $K^{FF} \star \mu_\Psi^F + K^{LF} \star \mu_\Psi^L$, which combines the action of the overall followers and leaders distribution. Similarly, each leader is under the action of the velocity field $K^{FL} \star \mu_\Psi^F + K^{LL} \star \mu_\Psi^L$. Hence, (4.4) is an average velocity of the system, weighed by the probability λ that a particle at x has of being a leader or a follower.*

In a similar spirit to the previous proposition, also transition rates α_F and α_L depending on μ_Ψ^F and μ_Ψ^L can be considered in our setting. This is for instance the point of view taken in [2, Assumption (H4) and Appendix A], where also some explicit examples were provided. Our assumptions are actually more general than those considered there: in particular, we can allow for an explicit dependence on the space variable x .

Proposition 4.5. *Let $U = \{F, L\}$. Consider two functions $\alpha_F, \alpha_L: \mathbb{R}^d \times \mathcal{M}_+(\mathbb{R}^d) \times \mathcal{M}_+(\mathbb{R}^d) \rightarrow [0, +\infty)$ satisfying the following assumptions:*

- *there exists a constant M such that, for $\bullet = F, L$,*

$$0 \leq \alpha_\bullet(x, \mu, \nu) \leq M(1 + |x| + m_1(\mu) + m_1(\nu)) \quad (4.5)$$

for all $x \in \mathbb{R}^d$ and $(\mu, \nu) \in \mathcal{M}_+(\mathbb{R}^d) \times \mathcal{M}_+(\mathbb{R}^d)$;

- *for all $R > 0$, there exist a constant L_R such that, for $\bullet = F, L$,*

$$|\alpha_\bullet(x_1, \mu_1, \nu_1) - \alpha_\bullet(x_2, \mu_2, \nu_2)| \leq L_R(|x_1 - x_2| + \|\mu_1 - \mu_2\|_{BL} + \|\nu_1 - \nu_2\|_{BL}) \quad (4.6)$$

for all $x_1, x_2 \in B_R$ and $(\mu_1, \nu_1), (\mu_2, \nu_2) \in \mathcal{M}_+(B_R) \times \mathcal{M}_+(B_R)$.

For $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\}))$, define μ_Ψ^F and μ_Ψ^L as in (4.2). Then, the functions

$$\alpha_F(x, \Psi) := \alpha_F(x, \mu_\Psi^F, \mu_\Psi^L) \quad \text{and} \quad \alpha_L(x, \Psi) := \alpha_L(x, \mu_\Psi^F, \mu_\Psi^L) \quad (4.7)$$

satisfy Assumptions $(\alpha 0)$ - $(\alpha 2)$ in Proposition 4.1.

Proof. The result follows from (4.5)-(4.6) by means of the inequalities in (4.3). \square

The main result of this Section is an existence and uniqueness result for the system of equations considered in [2], which we are going to deduce from Theorem 3.5. By doing this, we will extend the result in [2, Proposition 3.2] to the case were the transition rates α_F, α_L are allowed to explicitly depend on x , which was not considered there. The equations we consider are namely

$$\begin{cases} \partial_t \mu_t^F = -\operatorname{div}((K^F \star \mu_t^F + K^L \star \mu_t^L) \mu_t^F) - \alpha_F(x, \mu_t^F, \mu_t^L) \mu_t^F + \alpha_L(x, \mu_t^F, \mu_t^L) \mu_t^L, \\ \partial_t \mu_t^L = -\operatorname{div}((K^F \star \mu_t^F + K^L \star \mu_t^L) \mu_t^L) + \alpha_F(x, \mu_t^F, \mu_t^L) \mu_t^F - \alpha_L(x, \mu_t^F, \mu_t^L) \mu_t^L \end{cases} \quad (4.8)$$

to be solved by two positive Borel measures μ^F and μ^L attaining, for $t = 0$, an initial datum $\bar{\mu}^F$, and $\bar{\mu}^L$, respectively, which we assume to have compact support in \mathbb{R}^d . As customary in this kind of models, we will assume that the initial total population is normalized to 1, i.e.

$$\bar{\mu}^F(\mathbb{R}^d) + \bar{\mu}^L(\mathbb{R}^d) = 1.$$

Remark 4.6. *The velocity field in (4.8) corresponds to the choice $K^{FF} = K^{FL} = K^F$ and $K^{LF} = K^{LL} = K^L$ in (4.4). The key observation is that, in this case, the field $v_\Psi(x, \lambda)$ shows no explicit dependence on λ and is simply given by*

$$v_\Psi(x) = K^F \star \mu_\Psi^F + K^L \star \mu_\Psi^L. \quad (4.9)$$

This will eventually allow us to decouple equation (3.27) into the simpler system (4.8). Such an analysis is not possible if more than two different kernels are considered in (4.4). In that general case, the mean-field limit of the associated particle system must be formulated in terms of a solution Λ to (3.27), defined in the product space $\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})$.

To proceed to the announced result, we need to recall the definition of a solution to (4.8) which has been considered in [2]. Below, the shortcut $\mathcal{M}_c(\mathbb{R}^d)$ is used to denote a positive Borel measure having compact support in \mathbb{R}^d .

Definition 4.7 (Solution of system (4.8)). *Let $(\bar{\mu}^F, \bar{\mu}^L) \in \mathcal{M}_c(\mathbb{R}^d) \times \mathcal{M}_c(\mathbb{R}^d)$ be given, as well as $\mu^F, \mu^L: [0, T] \rightarrow \mathcal{M}_c(\mathbb{R}^d)$. We say that the couple (μ_t^F, μ_t^L) is a solution of system (4.8) with initial datum $(\bar{\mu}^F, \bar{\mu}^L)$ when*

- (i) $\mu_0^F = \bar{\mu}^F$ and $\mu_0^L = \bar{\mu}^L$;
- (ii) for each $i \in \{F, L\}$, the function $t \rightarrow \mu_t^i$ is continuous with respect to the topology of weak convergence of measures;
- (iii) there exists $R_T > 0$ such that $\bigcup_{t \in [0, T]} \text{supp}(\mu_t^i) \subseteq B_{R_T}$ for every $i \in \{F, L\}$;
- (iv) for every $\varphi \in C_c^1(\mathbb{R}^d)$ and $i \in \{F, L\}$ it holds

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) d\mu_t^i(x) &= \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot \left[\sum_{j \in \{F, L\}} (K^j \star \mu_t^j)(x) \right] d\mu_t^i(x) \\ &\quad - \int_{\mathbb{R}^d} \varphi(x) \alpha_i(x, \mu_t^F, \mu_t^L) d\mu_t^i(x) + \int_{\mathbb{R}^d} \varphi(x) \alpha_{-i}(x, \mu_t^F, \mu_t^L) d\mu_t^{-i}(x), \end{aligned}$$

for almost every $t \in [0, T]$, with

$$-i := \begin{cases} L & \text{if } i = F, \\ F & \text{if } i = L. \end{cases}$$

We can now state the existence and uniqueness result for system (4.8).

Theorem 4.8. *Let $U = \{F, L\}$. Consider two functions $\alpha_F, \alpha_L: \mathbb{R}^d \times \mathcal{M}_+(\mathbb{R}^d) \times \mathcal{M}_+(\mathbb{R}^d) \rightarrow [0, +\infty)$ satisfying (4.5)-(4.6) and two kernels $K^F, K^L: \mathbb{R}^d \rightarrow \mathbb{R}^d$ with*

$$\begin{aligned} |K^F(x)| + |K^L(x)| &\leq M(1 + |x|) \quad \text{for all } x \in \mathbb{R}^d; \\ |K^F(x_1) - K^F(x_2)| + |K^L(x_1) - K^L(x_2)| &\leq L_R |x_1 - x_2| \quad \text{for all } x_1, x_2 \in B_R. \end{aligned} \quad (4.10)$$

For $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\}))$, define μ_Ψ^F and μ_Ψ^L as in (4.2). For $x \in \mathbb{R}^d$ and $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\}))$, let $v_\Psi(x)$ and $\mathcal{T}(x, \Psi)$ be given by (4.9) and (4.1) respectively, and consider the corresponding velocity field b_Ψ as in (3.1).

Then, if $\Lambda \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})), \mathcal{W}_1)$ is the unique solution to (3.27) starting from $\bar{\Lambda} \in \mathcal{P}_c(\mathbb{R}^d \times \mathcal{P}_1(\{F, L\}))$, the measures $\mu_t^F := \mu_{\Lambda_t}^F$ and $\mu_t^L := \mu_{\Lambda_t}^L$ are the unique solutions to (4.8) with initial data $\bar{\mu}^F = \mu_{\bar{\Lambda}}^F$ and $\bar{\mu}^L = \mu_{\bar{\Lambda}}^L$.

Proof. We start by observing that, under our assumptions on α_F, α_L, K^F and K^L , the previous results of this section assure that the field b_Ψ complies with the requirements of Theorem 3.5, hence existence and uniqueness of the solution Λ to (3.27) starting from $\bar{\Lambda}$ is guaranteed. We split the proof into two parts, proving first existence and then uniqueness of the solutions.

Existence. For Λ as above, define $\mu_t^F := \mu_{\Lambda_t}^F$ and $\mu_t^L := \mu_{\Lambda_t}^L$. By definition and Lemma 3.13, conditions (i) and (iii) in Definition 4.7 are satisfied. The continuity property (ii) is instead a direct consequence of (4.3) and the continuity of Λ as a function of the time. We therefore only have to check (iv) in Definition 4.7. We perform the required computation only for μ_t^F , since the one for μ_t^L follows along similar lines. We take $\varphi \in C_c^1(\mathbb{R}^d)$ and we define, for all $(x, \lambda) \in \mathbb{R}^d \times \mathcal{F}(\{F, L\})$ the test function

$$\phi(x, \lambda) = \lambda_F \varphi(x).$$

We notice that the $(\mathbb{R}^d)^*$ -component of the Fréchet differential of ϕ at (x, λ) is given by $\lambda_F D\varphi(x)$, while the action of the other component is independent of λ by linearity and is given by

$$\xi \mapsto \varphi(x) \xi_F.$$

We apply the definition (3.28) of Eulerian solution to the above test function ϕ , which does not depend on t , and we get

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})} \lambda_F \varphi(x) d\Lambda_t(x, \lambda) &= \int_{\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})} \lambda_F \nabla \varphi(x) \cdot v_{\Lambda_t}(x) d\Lambda_t(x, \lambda) \\ &\quad + \int_{\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})} \varphi(x) (\mathcal{T}^*(x, \Lambda_t) \lambda)_F d\Lambda_t(x, \lambda), \end{aligned} \quad (4.11)$$

for all $t \in [0, T]$. We observe now that (4.2a)-(4.2b) are equivalent to the duality relationships

$$\begin{aligned} \int_{\mathbb{R}^d} \zeta(x) d\mu_{\Psi}^F(x) &= \int_{\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})} \lambda_F \zeta(x) d\Psi(x, \lambda), \\ \int_{\mathbb{R}^d} \zeta(x) d\mu_{\Psi}^L(x) &= \int_{\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})} (1 - \lambda_F) \zeta(x) d\Psi(x, \lambda) \end{aligned} \quad (4.12)$$

for all $\zeta \in C_b(\mathbb{R}^d)$. Applying the first one to the functions $\varphi(x)$ and $\nabla\varphi(x) \cdot v_{\Lambda_t}(x) \in C_b(\mathbb{R}^d)$ gives³

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})} \lambda_F \varphi(x) d\Lambda_t(x, \lambda) &= \int_{\mathbb{R}^d} \varphi(x) d\mu_t^F(x) \\ \int_{\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})} \lambda_F \nabla\varphi(x) \cdot v_{\Lambda_t}(x) d\Lambda_t(x, \lambda) &= \int_{\mathbb{R}^d} \nabla\varphi(x) \cdot v_{\Lambda_t}(x) d\mu_t^F(x). \end{aligned} \quad (4.13)$$

With Proposition 4.1, (4.13), and applying (4.12) to the functions $-\varphi(x)\alpha_F(x, \mu_t^F, \mu_t^L)$, and $\varphi(x)\alpha_L(x, \mu_t^F, \mu_t^L)$, respectively, we get

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathcal{P}_1(\{F, L\})} \varphi(x) (\mathcal{T}^*(x, \Lambda_t)\lambda)_F d\Lambda_t(x, \lambda) &= \\ - \int_{\mathbb{R}^d} \varphi(x) \alpha_F(x, \mu_t^F, \mu_t^L) d\mu_t^F(x) + \int_{\mathbb{R}^d} \varphi(x) \alpha_L(x, \mu_t^F, \mu_t^L) d\mu_t^L(x). \end{aligned}$$

Hence, also using (4.13) and the explicit expression (4.9) of the field $v_{\Lambda_t}(x)$, (4.11) is equivalent to equality (iv) in Definition 4.7 for $i = F$, as required. This proves existence of a solution to (4.8).

Uniqueness. The proof is divided into two steps.

Uniqueness- Step 1: continuous dependence for an auxiliary equation. For a given $Z > 0$ we fix the class of Carathéodory vector fields

$$\mathcal{V}_Z := \{v \in L^\infty([0, T], \text{Lip}(\mathbb{R}^d)) : \|v\|_{L^\infty([0, T], \text{Lip}(\mathbb{R}^d))} \leq Z\}$$

and we denote as usual by $\mathbf{Y}_v(t, s, \cdot)$ the associated transition maps, which satisfy the equalities

$$\mathbf{Y}_v(s, s, x) = x, \quad \frac{d}{dt} \mathbf{Y}_v(t, s, x) = v_t(\mathbf{Y}_v(t, s, x)) \quad (4.14)$$

for every $0 \leq s \leq t \leq T$ and $x \in \mathbb{R}^d$. From these, we can deduce the Grönwall-type estimates

$$\begin{aligned} |\mathbf{Y}_v(t, s, x_1) - \mathbf{Y}_v(t, s, x_2)| &\leq e^{Z(t-s)} |x_1 - x_2|, \\ |\mathbf{Y}_v(t, s, x) - \mathbf{Y}_w(t, s, x)| &\leq (t-s) e^{Z(t-s)} \sup_{s \in [0, t]} \|w_s - v_s\|_{C_b(\mathbb{R}^d)} \end{aligned} \quad (4.15)$$

for every $0 \leq s \leq t \leq T$, $x, x_1, x_2 \in \mathbb{R}^d$ and $v, w \in \mathcal{V}_Z$. For $v \in \mathcal{V}_Z$ and a given narrowly continuous family ξ_t of signed measures satisfying $|\xi_t| \leq R$ for all $t \in [0, T]$, we consider the inhomogeneous equation

$$\partial_t \mu_t + \text{div}(v_t \mu_t) = \xi_t. \quad (4.16)$$

We claim that, for a given initial datum $\bar{\mu} \in \mathcal{M}(\mathbb{R}^d)$, the unique solution of (4.16) starting from $\bar{\mu}$ is given by the variation of constants formula

$$\mu_t = \mathbf{Y}_v(t, 0, \cdot) \# \bar{\mu} + \int_0^t \mathbf{Y}_v(t, s, \cdot) \# \xi_s ds. \quad (4.17)$$

A direct computation using (4.14) proves indeed that the above formula provides a solution to (4.16), while uniqueness follows by taking the difference of two solutions and using that the comparison principle [7, Proposition 8.1.7] for the continuity equation.

³here it is crucial that the velocity field $v_{\Lambda_t}(x)$ only depends on x .

Now, if σ_t is another narrowly continuous family of signed measures, using (4.15) for all test functions φ with $|\varphi| \leq 1$ and $\text{Lip}(\varphi) \leq 1$ we have

$$\left| \int_0^t \int_{\mathbb{R}^d} \varphi(\mathbf{Y}_v(t, s, x)) d(\xi_s - \sigma_s) ds \right| \leq \left(\int_0^t e^{Z(t-s)} ds \right) \sup_{s \in [0, t]} \|\xi_s - \sigma_s\|_{BL} \leq t C_{T, Z} \sup_{s \in [0, t]} \|\xi_s - \sigma_s\|_{BL}.$$

With this, (4.15), and (4.17), we also deduce that, if ν solves (4.16) for another velocity field $w \in \mathcal{V}_z$, another narrowly continuous family σ_t with $|\sigma_t| \leq R$, and the same initial datum $\bar{\mu}$, the following estimate holds true:

$$\|\mu_t - \nu_t\|_{BL} \leq t C_{\bar{\mu}, T, R, Z} \left(\sup_{s \in [0, t]} \|w_s - v_s\|_{C_b(\mathbb{R}^d)} + \sup_{s \in [0, t]} \|\xi_s - \sigma_s\|_{BL} \right), \quad (4.18)$$

where the constant $C_{\bar{\mu}, T, R, Z}$ is given by

$$C_{\bar{\mu}, T, R, Z} := |\bar{\mu}| e^{ZT} + (1 + RT) C_{T, Z}.$$

Uniqueness- Step 2: conclusion. Consider two solutions (μ_t^F, μ_t^L) and (ν_t^F, ν_t^L) of (4.8) starting from the same initial datum $(\bar{\mu}^F, \bar{\mu}^L)$. Observe that the equation preserves the total mass, that is

$$\mu_t^F(\mathbb{R}^d) + \mu_t^L(\mathbb{R}^d) = \nu_t^F(\mathbb{R}^d) + \nu_t^L(\mathbb{R}^d) = 1$$

for all $t \in [0, T]$. This can be rewritten as

$$|\mu_t^F| + |\mu_t^L| = |\nu_t^F| + |\nu_t^L| = 1, \quad (4.19)$$

since all the above measures are positive. Now, on the one hand μ_t^F solves (4.16) with $v_t = K^F \star \mu_t^F + K^L \star \mu_t^L$ and $\xi_t = -\alpha_F(\cdot, \mu_t^F, \mu_t^L) \mu_t^F + \alpha_L(\cdot, \mu_t^F, \mu_t^L) \mu_t^L$. On the other hand, ν_t^F solves (4.16) with v_t replaced by $w_t = K^F \star \nu_t^F + K^L \star \nu_t^L$ and $\sigma_t = -\alpha_F(\cdot, \nu_t^F, \nu_t^L) \nu_t^F + \alpha_L(\cdot, \nu_t^F, \nu_t^L) \nu_t^L$ in place of ξ_t . Since $\mu_t^F, \mu_t^L, \nu_t^F$, and ν_t^L have compact support contained in a fixed ball B_{R_T} by Definition 4.7, we can assume (up to multiplying v_t and w_t by a suitable cut-off function not affecting equation (4.16)) that v belongs to \mathcal{V}_Z for a constant Z only depending on T, R_T and the constant M in (5.15). Using (4.5), (4.6), and (4.19), we also get the estimates

$$|\xi_t| \leq 2M(1 + 2R_T), \quad |\sigma_t| \leq 2M(1 + 2R_T),$$

$$\|\sigma_t - \xi_t\|_{BL} \leq (M(1 + 2R_T) + 2L_{R_T}) (\|\mu_t^F - \nu_t^F\|_{BL} + \|\mu_t^L - \nu_t^L\|_{BL}).$$

In a similar way, using (5.15) we obtain

$$\|v_t - w_t\|_{C_b(\mathbb{R}^d)} \leq (M(1 + 2R_T) + L_{2R_T}) (\|\mu_t^F - \nu_t^F\|_{BL} + \|\mu_t^L - \nu_t^L\|_{BL}).$$

Combining the previous inequalities with (4.18) and (4.19), we get that there exists a constant C , only depending on T, R_T and M , such that

$$\|\mu_\tau^F - \nu_\tau^F\|_{BL} \leq \tau C \left(\sup_{s \in [0, \tau]} \|\mu_s^F - \nu_s^F\|_{BL} + \sup_{s \in [0, \tau]} \|\mu_s^L - \nu_s^L\|_{BL} \right).$$

for all $0 \leq \tau \leq t \leq T$. A similar estimate also holds for $\|\mu_\tau^L - \nu_\tau^L\|_{BL}$. Taking the supremum in the left-hand sides and summing up the resulting inequalities we obtain

$$\sup_{s \in [0, t]} \|\mu_s^F - \nu_s^F\|_{BL} + \sup_{s \in [0, t]} \|\mu_s^L - \nu_s^L\|_{BL} \leq 2t C \left(\sup_{s \in [0, t]} \|\mu_s^F - \nu_s^F\|_{BL} + \sup_{s \in [0, t]} \|\mu_s^L - \nu_s^L\|_{BL} \right).$$

for all $0 \leq t \leq T$. This implies that $\mu_t^F = \nu_t^F$, as well as $\mu_t^L = \nu_t^L$ for all $0 \leq t \leq \frac{1}{2C}$. As C only depends on T, R_T and M , iterating the argument a finite number of times yields uniqueness on all $[0, T]$. \square

Remark 4.9. For all choices of the initial data $\bar{\mu}^F$ and $\bar{\mu}^L$ whose sum is a probability measure $\bar{\mu}$ with compact support, we can construct $\bar{\Lambda} \in \mathcal{P}_c(Y)$ so that $\mu_{\bar{\Lambda}}^F = \bar{\mu}^F$ and $\mu_{\bar{\Lambda}}^L = \bar{\mu}^L$. This can be done, for instance, as follows: if $g^F(x)$ is the Radon-Nikodym derivative of $\bar{\mu}^F$ with respect to $\bar{\mu}$, we can define, for $\bar{\mu}$ -a.e. $x \in \mathbb{R}^d$, the measure $\lambda_x \in \mathcal{P}_1(\{F, L\})$ via $\lambda_x := g^F(x) \delta_F + (1 - g^F(x)) \delta_L$. Then, the measure $\bar{\Lambda}$, defined by duality through

$$\int_Y \phi(y) d\bar{\Lambda} := \int_{\mathbb{R}^d} \phi(x, \lambda_x) d\bar{\mu}(x)$$

for all $\phi \in C_b(Y)$, satisfies the claim. Hence, Theorem 4.8 is a full existence result for system (4.8).

To conclude the Section, we observe that combining the above Remark with Theorems 3.5 and 4.8 we obtain a mean-field derivation of system (4.8) which extends the results of [2, Section 4].

5. FURTHER APPLICATIONS

In this section we present two situations which fit in the theory presented in Section 3 and are suitable to describe possible applications. In Section 5.1 we present the generalization of the results in Section 4 to the case of a discrete and finite label space U ; in Section 5.2 we focus on a continuum of labels and compare it with the results of [5].

5.1. Discrete and finite spaces of labels U . In this case, we will identify the discrete space of labels $U = \{u_1, \dots, u_H\}$ with the set of the indices of the labels, so that our model will be $U = \{1, \dots, H\}$. We will endow U with the Euclidean distance (restricted to U), namely,

$$\text{dist}(h, k) = |h - k|, \quad \text{for } h, k \in U. \quad (5.1)$$

Then the space $\text{Lip}(U)$ is an H -dimensional space spanned by the indicator functions $\mathbb{1}_h$, for $h \in \{1, \dots, H\}$. Consequently, the free space $\mathcal{F}(U)$ is the space of signed Borel measures on U whose generic element ξ is characterized by the values $\xi_h := \xi(\{h\})$.

Analogously to Proposition 4.1, we have the following characterization of the operators \mathcal{T} (and \mathcal{T}^*) satisfying assumptions (T0)-(T3).

Proposition 5.1. *Let $U = \{1, \dots, H\}$. Then $\mathcal{T}: \mathbb{R}^d \times \mathcal{P}_1(U) \rightarrow \text{Lip}(U)$ satisfies (T0)-(T3) if and only if there exist H^2 functions $\alpha_{hk}: \mathbb{R}^d \times \mathcal{P}_1(U) \rightarrow [0, +\infty)$ such that*

($\bar{\alpha}0$) *for every $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$ and $\lambda \in \mathcal{P}_1(U)$ it holds*

$$(\mathcal{T}^*(x, \Psi)\lambda)_h = -\alpha_{hh}(x, \Psi)\lambda_h + \sum_{k \neq h} \alpha_{kh}(x, \Psi)\lambda_k, \quad \text{for all } h \in U$$

with

$$\alpha_{hh}(x, \Psi) = \sum_{k \neq h} \alpha_{hk}(x, \Psi), \quad \text{for all } h \in U; \quad (5.2)$$

($\bar{\alpha}1$) *for every $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$, there exists $M_{\mathcal{T}} > 0$ such that*

$$0 \leq \alpha_{hk}(x, \Psi) \leq M_{\mathcal{T}}(1 + |x| + m_1(\Psi)), \quad \text{for all } h, k \in U;$$

($\bar{\alpha}2$) *for every $R > 0$ there exists $L_{\mathcal{T}, R} > 0$ such that, for every $(x^1, \Psi^1), (x^2, \Psi^2) \in B_R \times \mathcal{P}(B_R^Y)$,*

$$|\alpha_{hk}(x^1, \Psi^1) - \alpha_{hk}(x^2, \Psi^2)| \leq L_{\mathcal{T}, R}(|x^1 - x^2| + \mathcal{W}_1(\Psi^1, \Psi^2)), \quad \text{for } h, k \in U.$$

Proof. The proof is analogous to that of Proposition 4.1. \square

We notice that a matrix representation of the operator \mathcal{T} analogous to that in (4.1) holds

$$\mathcal{T}(x, \Psi) = \begin{pmatrix} -\alpha_{11}(x, \Psi) & \alpha_{12}(x, \Psi) & \cdots & \alpha_{1H}(x, \Psi) \\ \alpha_{21}(x, \Psi) & -\alpha_{22}(x, \Psi) & \cdots & \alpha_{2H}(x, \Psi) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{H1}(x, \Psi) & \alpha_{H2}(x, \Psi) & \cdots & -\alpha_{HH}(x, \Psi) \end{pmatrix}, \quad (5.3)$$

where, as a consequence of (5.2), the sum of the elements on each row must give zero. In this discrete setting, the operator $\mathcal{T}^*(x, \Psi)$ has a matrix representation given by the inverse matrix of that representing $\mathcal{T}(x, \Psi)$, so that

$$\mathcal{T}^*(x, \Psi) = \begin{pmatrix} -\alpha_{11}(x, \Psi) & \alpha_{21}(x, \Psi) & \cdots & \alpha_{H1}(x, \Psi) \\ \alpha_{12}(x, \Psi) & -\alpha_{22}(x, \Psi) & \cdots & \alpha_{H2}(x, \Psi) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1H}(x, \Psi) & \alpha_{2H}(x, \Psi) & \cdots & -\alpha_{HH}(x, \Psi) \end{pmatrix} \quad (5.4)$$

and condition (5.2) implies that the sum of the elements on each column must vanish.

Definition 4.2 is adapted to the following

Definition 5.2. Let $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(U))$. The distribution μ_Ψ^h of the agents with label $h \in U$ associated to Ψ is the positive Borel measure on \mathbb{R}^d defined by

$$\mu_\Psi^h(B) := \int_{B \times \mathcal{P}_1(U)} \lambda_h \, d\Psi(x, \lambda) \quad (5.5)$$

for each Borel set $B \subset \mathbb{R}^d$.

Also in this case, upon choosing suitable interaction kernels, the measures μ_Ψ^h defined in (5.5) can be used to construct velocity fields (of the type in (4.4)) v_Ψ satisfying (v1)-(v3).

Proposition 5.3. Let $U = \{1, \dots, H\}$. For $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(U))$, consider the velocity field

$$v_\Psi(x, \lambda) := \sum_{h, k \in U} \lambda_k (K^{hk} \star \mu_\Psi^h) \quad (5.6)$$

where μ_Ψ^h , for $h \in U$, are defined in (5.5) and the interaction kernels $K^{hk}: \mathbb{R}^d \rightarrow \mathbb{R}^d$, for $h, k \in U$, satisfy

$$\begin{aligned} |K^{hk}(x)| &\leq M(1 + |x|), \quad \text{for all } x \in \mathbb{R}^d; \\ |K^{hk}(x_1) - K^{hk}(x_2)| &\leq L_R |x_1 - x_2|, \quad \text{for all } x_1, x_2 \in B_R. \end{aligned}$$

Then, $v_\Psi(x, \lambda)$ satisfies (v1)-(v3).

Proof. The result follows by a direct computation. \square

Remark 4.4 applies in this case as well, so that the velocity field v_Ψ defined in (5.3) can be interpreted as an average velocity of the system, weighted by the probability that a particle at x has of having label k .

Proposition 4.5 concerning transition rates α_h 's depending on the μ_Ψ^h 's and explicitly on the space variable x is generalized to the following.

Proposition 5.4. Let $U = \{1, \dots, H\}$. Consider functions $\alpha_{hk}: \mathbb{R}^d \times (\mathcal{M}_+(\mathbb{R}^d))^H \rightarrow [0, +\infty)$, for $h, k \in U$, satisfying the following assumptions:

- there exists a constant M such that, for all $h, k \in U$,

$$0 \leq \alpha_{hk}(x, \mu^1, \dots, \mu^H) \leq M \left(1 + |x| + \sum_{l \in U} m_1(\mu^l) \right) \quad (5.7)$$

for all $x \in \mathbb{R}^d$ and $(\mu^1, \dots, \mu^H) \in (\mathcal{M}_+(\mathbb{R}^d))^H$;

- for all $R > 0$, there exist a constant L_R such that, for all $h, k \in U$,

$$|\alpha_{hk}(x_1, \mu_1^1, \dots, \mu_1^H) - \alpha_{hk}(x_2, \mu_2^1, \dots, \mu_2^H)| \leq L_R \left(|x_1 - x_2| + \sum_{l \in U} \|\mu_1^l - \mu_2^l\|_{BL} \right) \quad (5.8)$$

for all $x_1, x_2 \in B_R$ and $(\mu_1^1, \dots, \mu_1^H), (\mu_2^1, \dots, \mu_2^H) \in (\mathcal{M}_+(B_R))^H$.

For $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(U))$, define μ_Ψ^h as in (5.5), for $h \in U$. Then, the functions

$$\alpha_{hk}(x, \Psi) := \alpha_{hk}(x, \mu_\Psi^1, \dots, \mu_\Psi^H), \quad \text{for } h, k \in U \quad (5.9)$$

satisfy Assumptions $(\bar{\alpha}0)$ - $(\bar{\alpha}2)$ in Proposition 5.1.

Proof. The result follows from (5.7) and (5.8) by means of the following inequalities

$$\sum_{l \in U} m_1(\mu_\Psi^l) \leq m_1(\Psi), \quad \text{for all } \Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(U)), \quad (5.10)$$

$$\|\mu_\Psi^h - \mu_{\Psi_2}^h\|_{BL} \leq 2\mathcal{W}_1(\Psi_1, \Psi_2) \quad \text{for all } \Psi_1, \Psi_2 \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(U)), \quad h \in U,$$

which are a straightforward generalization of those in (4.3). \square

Theorem 4.8 can be generalized to the case of a finite discrete space of labels $U = \{1, \dots, H\}$. To obtain system (4.8) in the current context, we have to assume that the interaction kernels $K^{hk}: \mathbb{R}^d \rightarrow \mathbb{R}^d$, for $h, k \in U$ introduced in Proposition 5.3 are such that

$$K^{hk} = K^h, \quad \text{for all } h \in U, \quad (5.11)$$

for H kernels $K^h: \mathbb{R}^d \rightarrow \mathbb{R}^d$. In this case, analogously to the case with two labels, the velocity field v_Ψ defined in (5.6) does not depend on λ anymore and has the form

$$v_\Psi(x) = \sum_{h \in U} K^h \star \mu_\Psi^h, \quad (5.12)$$

where the μ_Ψ^h 's are defined in (5.5). Then, system (4.8) becomes a set of H equations

$$\partial_t \mu_t^h = -\operatorname{div} \left(\left(\sum_{k \in U} K^k \star \mu_t^k \right) \mu_t^h \right) - \alpha_{hh}(x, \mu_t^1, \dots, \mu_t^H) \mu_t^h + \sum_{k \neq h} \alpha_{kh}(x, \mu_t^1, \dots, \mu_t^H) \mu_t^k, \quad (5.13)$$

to be solved for Borel measures μ_h such that, at the initial time $t = 0$, $\mu_0^h = \bar{\mu}^h$, for $h \in U$, where $\bar{\mu}^1, \dots, \bar{\mu}^H$ are given Borel measures satisfying

$$\sum_{h \in U} \bar{\mu}^h(\mathbb{R}^d) = 1. \quad (5.14)$$

We give the following

Definition 5.5 (Solution to system (5.13)). *Let $(\bar{\mu}^1, \dots, \bar{\mu}^H) \in (\mathcal{M}_c(\mathbb{R}^d))^H$ be given such that (5.14) is satisfied, as well as $\mu^1, \dots, \mu^H: [0, T] \rightarrow \mathcal{M}_c(\mathbb{R}^d)$. We say that $(\mu_t^1, \dots, \mu_t^H)$ is a solution to system (5.13) with initial datum $(\bar{\mu}^1, \dots, \bar{\mu}^H)$ when*

- (i) $\mu_0^h = \bar{\mu}^h$ for each $h \in U$;
- (ii) for each $h \in U$, the function $t \rightarrow \mu_t^h$ is continuous with respect to the topology of weak convergence of measures;
- (iii) there exists $R_T > 0$ such that $\bigcup_{t \in [0, T]} \operatorname{supp}(\mu_t^h) \subseteq B_{R_T}$ for every $h \in U$;
- (iv) for every $\varphi \in \mathcal{C}_c^1(\mathbb{R}^d)$ and $h \in U$ it holds

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) d\mu_t^h(x) &= \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot \left(\sum_{k \in U} (K^k \star \mu_t^k)(x) \right) d\mu_t^h(x) \\ &\quad - \int_{\mathbb{R}^d} \varphi(x) \alpha_{hh}(x, \mu_t^1, \dots, \mu_t^H) d\mu_t^h(x) + \int_{\mathbb{R}^d} \varphi(x) \sum_{k \neq h} \alpha_{kh}(x, \mu_t^1, \dots, \mu_t^H) d\mu_t^k(x), \end{aligned}$$

for almost every $t \in [0, T]$.

We are now in a position to prove the most important result of this section, namely how a solution to (5.13) follows from Theorem 3.5.

Theorem 5.6. *Let $U = \{1, \dots, H\}$. Consider functions $\alpha_{hk}: \mathbb{R}^d \times (\mathcal{M}_+(\mathbb{R}^d))^H \rightarrow [0, +\infty)$ satisfying (5.7) and (5.8) and H kernels $K^h: \mathbb{R}^d \rightarrow \mathbb{R}^d$ with*

$$\begin{aligned} \sum_{h \in U} |K^h(x)| &\leq M(1 + |x|), \quad \text{for all } x \in \mathbb{R}^d; \\ \sum_{h \in U} |K^h(x_1) - K^h(x_2)| &\leq L_R |x_1 - x_2|, \quad \text{for all } x_1, x_2 \in B_R. \end{aligned} \quad (5.15)$$

For $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(U))$, define the μ_Ψ^h 's as in (5.5). For $x \in \mathbb{R}^d$ and $\Psi \in \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(U))$, let $v_\Psi(x)$ and $\mathcal{T}(x, \Psi)$ be given by (5.12), and (5.3), respectively, and consider the corresponding velocity field b_Ψ as in (3.1).

Then, if $\Lambda \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(U), \mathcal{W}_1))$ is the unique solution to (3.27) starting from $\bar{\Lambda} \in \mathcal{P}_c(\mathbb{R}^d \times \mathcal{P}_1(U))$, the measures $\mu_t^h := \mu_{\Lambda_t}^h$, for $h \in U$, are the unique solutions to (5.13) with initial data $\bar{\mu}^h = \mu_{\bar{\Lambda}}^h$, for $h \in U$.

Proof. The proof is analogous to that of Theorem 4.8 □

It is useful to recall the notion of Q -matrix from the literature of Markov chains (see, e.g., [35, Chapter 2]). Let Q be a $H \times H$ matrix; the element q_{hk} represents the transition rate from state h to k . Such a matrix is called Q -matrix satisfies the following conditions

- (Q1) $q_{hk} \geq 0$, for all $h, k \in U, h \neq k$;
- (Q2) $\sum_{k \in U} q_{hk} = 0$, for all $h \in U$.

It is customary, in the literature on Markov chains, to complement these conditions by

- (Q0) $0 \leq -q_{hh} < +\infty$, for all $h \in U$,

even though condition (Q0) is a consequence of (Q1) and (Q2).

Remark 5.7. *It is easy to see that the matrix in (5.3), representing the operator $\mathcal{T}(x, \Psi)$ for all $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$ is a Q -matrix: indeed, condition (5.2) is equivalent to (Q2) and $(\bar{\alpha}1)$ is equivalent to (Q1).*

For $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d \times \mathcal{P}_1(U))$, and for fixed initial conditions $(x^0, \lambda^0) \in Y$, the dynamics described by the vector field b_Ψ defined in (3.1) is

$$\begin{pmatrix} \dot{x} \\ \dot{\lambda} \end{pmatrix} = b_\Psi(x, \lambda) = \begin{pmatrix} v_\Psi(x) \\ \mathcal{T}^*(x, \Psi)\lambda \end{pmatrix}, \quad \begin{pmatrix} x \\ \lambda \end{pmatrix}(0) = \begin{pmatrix} x^0 \\ \lambda^0 \end{pmatrix}. \quad (5.16)$$

Concerning the second equation, $\dot{\lambda} = \mathcal{T}^*(x, \Psi)\lambda$, this is a linear equation in the space of measures, which can be easily integrated to give

$$\lambda(t) = \lambda_t = S_t(x, \Psi)\lambda^0, \quad \text{with } S_t(x, \Psi) := e^{t\mathcal{T}(x, \Psi)}. \quad (5.17)$$

A classical result in the theory of Markov chains, [35, Theorem 2.1.2] grants that the fact that $\mathcal{T}(x, \Psi)$ satisfies (Q1)-(Q2) is equivalent to the matrix $S_t(x, \Psi)$ being a *stochastic matrix* for all times $t \geq 0$, namely a matrix $S \in \mathbb{R}^{H \times H}$ that satisfies

$$(S1) \quad 0 \leq S_{kh} < +\infty \text{ for all } h, k \in U;$$

$$(S2) \quad \sum_{h \in U} S_{kh} = 1 \text{ for all } k \in U.$$

We remark that the explicit solution given by (5.17) is available only for fixed (x, Ψ) , so that the matrix (5.3) is constant and its exponential can be computed. This would restrict the dynamics to constant solutions $x(t) = x^0$ for all t and to a constant distribution $\Psi \in \mathcal{P}(Y)$.

5.2. A continuum of labels. We now turn to the case in which U is a compact metric space, as in the general theory developed in Section 3, and that it is a continuum. We will shortly present some possible expressions of the velocity field b_Ψ defined in (3.1) in which both components feature a two-player game that determines the evolution.

Let $\bar{y} \in Y$ and let $\mathcal{K}: Y \times Y \rightarrow \mathbb{R}^d \times \mathcal{F}(U)$ be an interaction kernel such that, for $\Psi \in \mathcal{P}(Y)$,

$$b_\Psi(y) = \int_Y \mathcal{K}(y, y') d\Psi(y'). \quad (5.18)$$

Our aim is to study system (3.29) where the evolution is driven by the field b_Ψ defined in (5.18). Recalling (3.1), let $\mathcal{K}_x: Y \times Y \rightarrow \mathbb{R}^d$ and $\mathcal{K}_\lambda^*: Y \times Y \rightarrow \mathcal{F}(U)$ be the components of \mathcal{K} . Then (5.18) gives

$$v_\Psi(y) = \int_Y \mathcal{K}_x(y, y') d\Psi(y'), \quad \text{and} \quad \mathcal{T}^*(x, \Psi)\lambda = \int_Y \mathcal{K}_\lambda^*(y, y') d\Psi(y'). \quad (5.19)$$

We furthermore assume that each of \mathcal{K}_x and \mathcal{K}_λ^* accounts for an averaging over all the strategies in U . To this aim, let $V: (\mathbb{R}^d \times U)^2 \rightarrow \mathbb{R}^d$ be such that

$$\mathcal{K}_x(y, y') = \int_U \int_U V(x, u, x', u') d\lambda'(u') d\lambda(u) \quad (5.20)$$

The interpretation of the field V is the following: the quantity

$$V(x, u, x', u') \in \mathbb{R}^d$$

gives the direction that the payer at x should follow if it has label/strategy u when playing a game against a player at x' with label/strategy u' . Keeping the structure (3.29), (5.19), and (5.20) in mind, the full law of motion for the variable x is given by

$$\dot{x} = v_\Psi(y) = \int_U \int_Y \int_U V(x, u, x', u') d\lambda'(u') d\Psi(x', \lambda') d\lambda(u), \quad (5.21)$$

which can be interpreted in the following way: the player at x with label/strategy u plays a game with all the other players at x' with their labels/strategies u' and the velocity \dot{x}_t is determined by averaging over all the possible strategies of the opponent (the integral with respect to $d\lambda'(u')$), over all the possible distributions of opponents with their distributions of strategies (the integral with respect to $d\Lambda(x', \lambda')$), and finally over all the possible strategies that x have at their disposal (the integral with respect to $d\lambda(u)$).

For $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(U)$ define the operator $\mathcal{T}(x, \Psi): \text{Lip}(U) \rightarrow \text{Lip}(U)$ by

$$(\mathcal{T}(x, \Psi)f)(u) := \int_Y \int_U J(x, u, x', u') f(u') d\lambda'(u') d\Psi(x', \lambda') - g(x, \Psi)(u) f(u), \quad (5.22)$$

where $J: (\mathbb{R}^d \times U)^2 \rightarrow \mathbb{R}$, and the function $g(x, \Psi) \in C(U)$ can be interpreted as a *global departure rate* from u .

Proposition 5.8. *Let $\Psi \in \mathcal{P}_1(Y)$, let $V: (\mathbb{R}^d \times U)^2 \rightarrow \mathbb{R}^d$, $J: (\mathbb{R}^d \times U)^2 \rightarrow \mathbb{R}$, and $g(\cdot, \Psi) \in C(U)$ be such that*

- (V1) V is locally Lipschitz with respect to all of its variables;
- (V2) V is sublinear in the spatial variables, namely there exists $C_V > 0$ such that

$$|V(x, u, x', u')| \leq C_V(1 + |x| + |x'|), \quad \text{for all } (x, u), (x', u') \in (\mathbb{R}^d \times U)^2;$$

- (J1) J is locally Lipschitz with respect to all of its variables;
- (J2) J is sublinear in the spatial variables, namely there exists $C_J > 0$ such that

$$|J(x, u, x', u')| \leq C_J(1 + |x| + |x'|), \quad \text{for all } (x, u), (x', u') \in (\mathbb{R}^d \times U)^2.$$

Then, the field $v_\Psi: Y \rightarrow \mathbb{R}^d$ given by (5.19) via (5.20) satisfies (v1)-(v3) and the operator $\mathcal{T}(x, \Psi): \text{Lip}(U) \rightarrow \text{Lip}(U)$ given by (5.22) satisfies (T1)-(T3). If, moreover,

- (g) for $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$,

$$g(x, \Psi)(u) = \int_Y \int_U J(x, u, x', u') d\lambda'(u') d\Psi(x', \lambda'), \quad (5.23)$$

then (T0) is also satisfied. In this case, the adjoint operator $\mathcal{T}^*(x, \Psi): \mathcal{F}(U) \rightarrow \mathcal{F}(U)$ is represented as in (5.19), by means of $\mathcal{K}_\lambda^*: Y \times Y \rightarrow \mathcal{F}(U)$ such that, for $f \in \text{Lip}(U)$,

$$\langle \mathcal{K}_\lambda^*(y, y'), f \rangle = \int_U \int_U J(x, u, x', u') [f(u') - f(u)] d\lambda'(u') d\lambda(u). \quad (5.24)$$

Proof. Let $R > 0$ be fixed and let $L_{V,R}$ be the Lipschitz constant of V on $(B_R \times U)^2$. Properties (v1)-(v2) are obtained from (V1) by standard estimates keeping the definition of BL norm and the structure (5.19)-(5.20) into account; the constant $L_{v,R}$ of (v1)-(v2) is determined by $L_{V,R}$ and $\text{diam } U$. Property (v3) follows from (V1) and (V2) and the constant M_v of (v3) is determined by C_V .

Analogously, let $L_{J,R}$ be the Lipschitz constant of J on $(B_R \times U)^2$. Property (T2) is obtained from (J1) by standard estimates keeping the definition of BL norm and the structure (5.22) into account; the constant $L_{\mathcal{T},R}$ of (T2) is determined by $L_{J,R}$ and $\text{diam } U$. Property (T1) is obtained also using (J2) and the constant $M_{\mathcal{T}}$ of (T1) is determined by C_J . Property (T3) follows from the boundedness of $J(\cdot, \cdot, x', u')$ and $g(\cdot, \Psi)$ on $B_R \times U$ and U , respectively.

Finally, a simple computation shows that (g) and (T0) are equivalent, so that, using (5.23), we obtain the following expression for (5.22)

$$(\mathcal{T}(x, \Psi)f)(u) = \int_Y \int_U J(x, u, x', u') [f(u') - f(u)] d\lambda'(u') d\Psi(x', \lambda'), \quad \text{for } f \in \text{Lip}(U). \quad (5.25)$$

Recalling, from (3.1) and (3.29), that $\dot{\lambda} = \mathcal{T}^*(x, \Psi)\lambda$, for $f \in \text{Lip}(U)$ we have, using (5.19),

$$\begin{aligned} \langle \dot{\lambda}, f \rangle &= \langle \mathcal{T}^*(x, \Psi)\lambda, f \rangle = \langle \lambda, \mathcal{T}(x, \Psi)f \rangle \\ &= \left\langle \lambda, \int_Y \int_U J(x, u, x', u') [f(u') - f(u)] d\lambda'(u') d\Psi(x', \lambda') \right\rangle \\ &= \int_U \int_Y \int_U J(x, u, x', u') [f(u') - f(u)] d\lambda'(u') d\Psi(x', \lambda') d\lambda(u) \\ &= \int_Y \left[\int_U \int_U J(x, u, x', u') [f(u') - f(u)] d\lambda'(u') d\lambda(u) \right] d\Psi(x', \lambda'), \end{aligned}$$

which is (5.24). The proof is complete. \square

Remark 5.9. *Let us assume that the hypotheses of Proposition 5.8 hold; the ODE in (3.29), written component-wise, reads*

$$\begin{pmatrix} \dot{x} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} v_\Psi(y) \\ \mathcal{T}^*(x, \Psi)\lambda \end{pmatrix}, \quad (5.26)$$

for $y = (x, \lambda) \in Y = \mathbb{R}^d \times \mathcal{P}(U)$ and $\Psi \in \mathcal{P}_1(Y)$.

A comparison with [5] is in order. In [5, equations (1.3), (1.4), and (1.8)] the ODE studied is (keeping the notations of [5])

$$\begin{pmatrix} \dot{x} \\ \dot{\sigma} \end{pmatrix} = \begin{pmatrix} a(y) \\ \Delta_{\Sigma,y}\sigma \end{pmatrix}, \quad (5.27)$$

where $y = (x, \sigma) \in C := \mathbb{R}^d \times \mathcal{P}(U)$ and $\Sigma \in \mathcal{P}_1(C)$. The right-hand sides are

$$\begin{aligned} a(y) &= \int_U e(x, u) d\sigma(u), \\ \Delta_{\Sigma,y}\sigma &= \left(\int_C \int_U J(x, u, x', u') d\sigma'(u') d\Sigma(x', \sigma') \right. \\ &\quad \left. - \int_U \int_C \int_U J(x, w, x', u') d\sigma'(u') d\Sigma(x', \sigma') d\sigma(w) \right) \sigma, \end{aligned} \quad (5.28)$$

where $e: \mathbb{R}^d \times U \rightarrow \mathbb{R}^d$ and $J: (\mathbb{R}^d \times U)^2 \rightarrow \mathbb{R}$. The equation for σ in (5.27) is known in the literature of evolutionary games as replicator equation (see, e.g., [28]).

In comparison with (5.28), we notice that the theory developed in Section 3 allows us to deal with a broader class of velocity fields, that include agent interaction. Indeed, the velocity field $a: C \rightarrow \mathbb{R}^d$ in (5.28) is linear in the mixed strategy σ and depends only on the position x of the agent; it does not depend neither on the global distribution Σ of the system nor on any interaction between the agents. The velocity field $v_\Psi: Y \rightarrow \mathbb{R}^d$ in (5.21), instead, additionally takes into account the global distribution Ψ of the system as well as the interaction between players through the kernel V .

The equations for the second component of y , namely λ and σ in (5.26) and (5.27), respectively, share even fewer features. If, on the one hand, they both depend on the global distribution of the system Ψ or Σ , on the other hand the second equation in (5.26) is linear in λ (see (5.24)), whereas the replicator equation is quadratic in σ , through the triple integral in (5.28).

Acknowledgments. The authors thank the hospitality of the departments of mathematics of the Politecnico di Torino, of the Università di Napoli, and of the E. Schrödinger Institute in Vienna, where this research was developed.

The authors are members of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

The work of FS is part of the project "Variational methods for stationary and evolution problems with singularities and interfaces" PRIN 2017 financed by the Italian Ministry of Education, University, and Research.

REFERENCES

- [1] G. Albi, M. Bongini, E. Cristiani, and D. Kalise: *Invisible Control of Self-Organizing Agents Leaving Unknown Environments*. SIAM J. Appl. Math. **76**(4) (2016), 1683-1710.
- [2] G. Albi, M. Bongini, F. Rossi, and F. Solombrino: *Leader formation with mean-field birth and death models*. Math. Models Methods Appl. Sci. **29**(04) (2019), 633-679.
- [3] G. Albi, L. Pareschi, and M. Zanella: *Boltzmann-type control of opinion consensus through leaders*. Phil. Trans. R. Soc. A **372** (2014), 20140138.
- [4] G. Albi, L. Pareschi, and M. Zanella: *Opinion dynamics over complex networks: Kinetic modelling and numerical methods*. Kinetic & Related Models **10**(1) (2017), 1-32.
- [5] L. Ambrosio, M. Fornasier, M. Morandotti, and G. Savaré: *Spatially Inhomogeneous Evolutionary Games*. [arXiv:1805.04027](https://arxiv.org/abs/1805.04027). (Submitted)
- [6] L. Ambrosio and W. Gangbo: *Hamiltonian ODEs in the Wasserstein space of probability measures*. Comm. Pure Appl. Math. **61**(1) (2008), 18-53.
- [7] L. Ambrosio, N. Gigli, and G. Savaré: *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [8] L. Ambrosio and D. Puglisi: *Linear extension operators between spaces of Lipschitz maps and optimal transport*. J. Reine Angew. Math. (2018), to appear. <http://cvgmt.sns.it/paper/3155>.
- [9] L. Ambrosio and D. Trevisan: *Well-posedness of Lagrangian flows and continuity equations in metric measure spaces*. Anal. PDE **7**(5) (2014), 1179-1234.
- [10] R. F. Arens and J. Eells, Jr.: *On embedding uniform and topological spaces*. Pacific J. Math., **6** (1956), 397-403.
- [11] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardinà, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic: *Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study*. PNAS **105**(4) (2008), 1232-1237.

- [12] M. Bongini and G. Buttazzo: *Optimal Control Problems in Transport Dynamics*. Math. Models Methods Appl. Sci. **27**(3) (2017), 427-451.
- [13] M. Bongini, M. Fornasier, M. Hansen, and M. Maggioni: *Inferring interaction rules from observations of evolutive systems I: The variational approach*. Math. Models. Meth. Appl. Sci. **27**(5) (2017), 909-951.
- [14] T. Börgers and R. Sarin: *Learning through reinforcement and replicator dynamics*. Journal of Economic Theory, **77** (1997), 1-14.
- [15] H. Brézis: *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973.
- [16] P. Cardaliaguet and S. Hadikhanloo: *Learning in mean field games: the fictitious play*. ESAIM Control Optim. Calc. Var., **23**(2) (2017), 569-591.
- [17] J. A. Carrillo, J. A. Cañizo and J. Rosado: *A well-posedness theory in measures for some kinetic models of collective behavior*. Math. Models Methods Appl. Sci. **21**(03) (2011), 515-539.
- [18] J. A. Carrillo, Y.-P. Choi, and M. Hauray: *The derivation of swarming models: mean-field limit and Wasserstein distances*. In *Collective dynamics from bacteria to crowds*, pages 1–46. Springer, 2014.
- [19] H. Cartan: *Calcul différentiel*. Hermann, Paris, 1967.
- [20] M. Cirant: *Multi-population mean field games systems with Neumann boundary conditions*. Journal de Mathématiques Pures et Appliquées **103**(5) (2015), 1294-1315.
- [21] J. Dieudonné: *Foundations of modern analysis*. Pure and Applied Mathematics, Vol. X. Academic Press, New York-London, 1960.
- [22] M. Di Francesco and S. Fagioli: *Measure solutions for non-local interaction PDEs with two species*. Nonlinearity **26**(10) (2013), 2777-2808.
- [23] R. Dobrushin. *Vlasov equations*. Funct. Anal. Appl. **13**(2) (1979), 115–123.
- [24] M. Dorigo and C. Blum: *Ant colony optimization theory: A survey*. Theoretical computer science **344**(2-3) (2005), 243-278.
- [25] B. Düring, P. Markowich, J.-F. Pietschmann, and M.-T. Wolfram: *Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders*. Proc. R. Soc. A **465** (2009), 3687-3708.
- [26] M. Fornasier, B. Piccoli, and F. Rossi: *Mean-field Sparse Optimal Control*. Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **372** (2014), 20130400.
- [27] F. Golse. *The mean-field limit for the dynamics of large particle systems*. In: Journées “Équations aux dérivées partielles”, Forges-les-Eaux, France, 2 au 6 juin 2003. Exposés Nos. I-XV pages 1–47. Nantes: Université de Nantes, 2003.
- [28] J. Hofbauer and K. Sigmund: *Evolutionary games and population dynamics*. Cambridge University Press, Cambridge, 1998.
- [29] H. Huang, J.-G. Liu, and J. Lu: *Learning Interacting Particle Systems: Diffusion Parameter Estimation for Aggregation Equations*. Math. Models. Meth. Appl. Sci. **29**(1) (2019), 1-29.
- [30] M. Kac: *Foundations of kinetic theory*. In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 3: Contributions to Astronomy and Physics, pp. 171-197, Berkeley, Calif., University of California Press, 1956.
- [31] J. Kennedy: *Particle swarm optimization*. Encyclopaedia of machine learning, pp. 760-766. Springer, 2011.
- [32] T. S. Lim, Y. Lu, and J. Nolen: *Quantitative Propagation of Chaos in the bimolecular chemical reaction-diffusion model*. [arXiv:1906.01051](https://arxiv.org/abs/1906.01051).
- [33] H. P. McKean. *Propagation of chaos for a class of non-linear parabolic equations*. Lecture Series in Differential Equations, Catholic University, Washington D.C. **7** (1967), 41–57.
- [34] P. Mozgunov, M. Beccuti, A. Horvath, T. Jaki, R. Sirovich, and E. Bibbona: *A review of the deterministic and diffusion approximations for stochastic chemical reaction networks*. Reac. Kinet. Mech. Cat. **123**(2) (2018), 289-312.
- [35] J. R. Norris: *Markov Chains*. Cambridge University Press, 1997.
- [36] K. Oelschläger: *On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes*. Probab. Theory Related Fields **82**(4) (1989), 565-586.
- [37] B. Piccoli and F. Rossi. *Transport equation with nonlocal velocity in Wasserstein spaces: Convergence of numerical schemes*. Acta Appl. Math. **124**(1) (2013), 73–105.
- [38] B. Piccoli and F. Rossi: *Measure-theoretic models for crowd dynamics*. In Crowd Dynamics Volume 1 - Theory, Models, and Safety Problems, pp. 137-165, N. Bellomo and L. Gibelli Eds., Birkhäuser, 2018.
- [39] M.-N. Thai: *Birth and Death process in mean field type interaction*. Bernoulli (2019), to appear.
- [40] G. Toscani: *Kinetic models of opinion formation*. Commun. Math. Sci. **4**(3) (2006), 481-496.
- [41] C. Villani. *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [42] N. Weaver: *Lipschitz algebras. Second edition*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2018.

DIPTAMENTO DI SCIENZE MATEMATICHE “G. L. LAGRANGE”, POLITECNICO DI TORINO, CORSO DUCA DEGLI ABRUZZI, 24, 10129 TORINO, ITALY

Email address, M. Morandotti: marco.morandotti@polito.it

DIPTAMENTO DI MATEMATICA E APPLICAZIONI “RENATO CACCIOPPOLI”, UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II, VIA CINTIA, MONTE S. ANGELO, 80126 NAPOLI, ITALY

Email address, F. Solombrino: francesco.solombrino@unina.it