OPINION FORMATION SYSTEMS VIA DETERMINISTIC PARTICLES APPROXIMATION

SIMONE FAGIOLI[*]

DISIM, Università degli Studi dellAquila
via Vetoio 1 (Coppito), 67100 LAquila (AQ), Italy

EMANUELA RADICI

EPFL SB, Station 8,
CH-1015 Lausanne, Switzerland

(Communicated by the associate editor name)

ABSTRACT. We propose an ODE-based derivation for a generalized class of opinion formation models either for single and multiple species (followers, leaders, trolls). The approach is purely deterministic and the evolution of the single opinion is determined by the competition between two mechanisms: the opinion diffusion and the compromise process. Such deterministic approach allows to recover in the limit an aggregation/(nonlinear)diffusion system of PDEs for the macroscopic opinion densities.

1. **Introduction.** The study of phenomena in social sciences trough mathematical modelling has gained an increasing attention in the scientific community, in particular in the last decades, see [8, 15, 26, 33, 34, 35, 41]. The huge increasing of private and public communications on social networks like Facebook and Twitter speed up the attention on social phenomena, mainly due to a huge amount of data coming from empirical observations. The large information exchange push the study on the analysis of how these interactions influence the process of opinion formation, [4, 9, 11, 29, 31, 40, 42, 47].

Social networks are now particularly important for the political leaders communication, since they give the possibility of driving selected information to potential electors. Actually, the phenomenon of opinion leaders and their possible control strategy on the public opinion dates back to the work of Lazarsfeld, [32], in the study of the USA presidential elactions in 1940, see also [10, 12]. The social networks, however, show an innovative feature: the possibility of *measuring* the popularity of a given leader through factors such as the number of followers or the number of likes, see [16]. The drawback is that these *measure* can be easily falsified to guide the behaviour of real users and persuading them to vote for a specific candidate, see [17, 30]. In particular, in [13] it was estimated that the most followed political US accounts on Twitter posses the 25% of fake followers.

In the present work, we introduce a model general enough to catch such phenomenon.

Among the possible mathematical models, the kinetic formulation of opinion formation, introduced in the seminal paper [43], has gained a lot of attention mainly because of its flexibility to describe the phenomenon at different levels: microscopic, based on the pairwise interaction between agents, mesoscopic for the distribution of the opinions, and last but not least, macroscopic, useful for describing the trend of the opinion density. In this kinetic model, interactions among agents are supposed to be governed by two relevant concepts: compromise describes the way in which pairs of agents reach a compromise after exchanging opinions (its structure and other important features had been intensively studied in [4, 9, 18, 28]), and self-thinking, modelled by a random variable, describes how agents change their opinions in an unpredictable way [9, 43]. As a result of the above considerations, one can consider the following pairwise interactions law between two agents with opinion w and v respectively

$$\begin{aligned} w' &= w - P(w)(w - v) + \eta_w D(w), \\ v' &= v - P(v)(v - w) + \eta_v D(v), \end{aligned} \tag{1}$$

where (w', v') denotes the opinions after the interaction, the functions P and D describe the local relevance of the compromise and the self-thinking (diffusion) for a given opinion, and η_w and η_v are two random variables. After a suitable scaling process, called *quasi-invariant opinion limit*, the author in [43] obtains a Fokker-Planck type equation for the opinion density, precisely of the form

$$\partial_t u(t, w) = \frac{\lambda^2}{2} \partial_{ww} \left(D^2(w) u(t, w) \right) + \partial_w \left(P(w)(w - m) u(t, w) \right), \tag{2}$$

which results to be a good approximation for the stationary profiles of the kinetic equation associated to (1). A similar approach was also used to model more general opinion formation processes, such as the presence of opinion leaders, choice formation, control and networks (see [4, 1, 3, 2, 22, 23, 45]), as well as the different class of trading models for goods and wealth distribution (see [36, 44]).

We propose an alternative derivation which is based on a deterministic many particle limit and allows to consider a generalized version of the above kinetic model.

1.1. **Formal derivation of the generalized opinion formation model.** This section is devoted to introduce and formally derive, via many particle limit, the set of equations we want to investigate in the present paper.

1.1.1. *Basic one-species model.* Consider a population composed of $N + 1$ individuals with given initial opinion W_i^0, for $i = 0, \ldots, N$ and assume that opinions can only range in a bounded set of values, say $I = [-1, 1]$ where $W = \pm 1$ represent the extreme opinions. We further assume that each opinion has a certain *strength* σ_i, for $i = 0, \ldots, N$. According to the kinetic model described above, the i-individual can modify its own opinion depending on two possible mechanisms: the compromise process (interaction) with the others individuals and the diffusion of that given opinion. Therefore, the time evolution of the opinion of the i-individual is described by the following ODE

$$\dot{W}_i = \text{ compromise} + \text{ diffusion} . \tag{3}$$

It is standard to assume that the local or non-local relevance of the compromise depends on the distance between two different opinions. Then, by calling $P(\cdot)$ the

function describing the local relevance of the compromise, the interaction of the $i-$individual with the other individuals can be modelled by

$$\text{compromise} = -\sum_{j=0}^{N} \sigma_j P(W_i, W_j)(W_i - W_j). \tag{4}$$

For what concerns the diffusive part, instead, we assume that the opinions evolve with a speed equal to the *osmotic velocity* associated to the diffusion process, as firstly introduced in [38]. More precisely, if we denote by $D(\cdot)$ the diffusion capacity of a given opinion, then the diffusive process is described by

$$\text{diffusion} \ = \frac{\lambda^2}{2\sigma_i} D^2(W_i) \left(\frac{\sigma_{i-1}}{W_i - W_{i-1}} - \frac{\sigma_i}{W_{i+1} - W_i} \right), \tag{5}$$

where λ is a fixed diffusion coefficient. Let us observe, that, accordingly to the results in [24, 27], it is possible to generalize the diffusion law (5) to more general emphnon linear expressions. Indeed, the quantity

$$u_i = \frac{\sigma_i}{W_{i+1} - W_i}, \tag{6}$$

represents the local density between two consecutive opinions, then one can consider a generic nonlinear non-decreasing real valued function ϕ to rephrase (5) as

$$\text{diffusion} \ = \frac{\lambda^2}{2\sigma_i} D^2(W_i) \left(\phi(u_{i-1}) - \phi(u_i) \right). \tag{7}$$

A further, hopefully realistic, assumption of our model concerns the boundary conditions. More precisely, we impose that the extreme opinions cannot alter during the evolution. As a consequence, we set

$$\dot{W}_0 = 0,$$

$$\dot{W}_i = \frac{\lambda^2}{2\sigma_i} D^2(W_i) \left(\phi(u_{i-1}) - \phi(u_i) \right) - \sum_{j=0}^{N} \sigma_j P(W_i, W_j)(W_i - W_j), \ i \neq 0, N, \quad (8)$$

$$\dot{W}_N = 0.$$

By formally sending $N \to \infty$ and calling $\mathcal{W}(z, t)$ the piecewise linear interpolation of the values W_i, we get

$$\partial_t \mathcal{W}(t, z) = \frac{\lambda^2}{2\sigma_i} D^2(\mathcal{W}) \partial_z \left(\phi \left(\frac{1}{\partial_z \mathcal{W}} \right) \right) - \int_0^\sigma P(\mathcal{W}(z), \mathcal{W}(\zeta)) \left(\mathcal{W}(z) - \mathcal{W}(\zeta) \right) d\zeta.$$

Once here, if $\mathcal{U}(t, w)$ were the inverse of the $\mathcal{W}(t, z)$ and $u(t, w)$ were $\partial_w \mathcal{U}(t, w)$ then, inspired by the results in [19, 24], u would be a weak solutions of the following aggregation-diffusion PDE

$$\partial_t u(t, w) = \partial_w \left(\frac{\lambda^2}{2} D^2(w) \partial_w \phi(u(t, w)) - \mathcal{P}_{uu}\left[u \right](t, w) u(t, w) \right), \tag{9}$$

for $(t, w) \in [0, T] \times I$, endowed with zero flux boundary condition, where the nonlocal operator $\mathcal{P}_{uu}\left[u \right](t, w)$ is defined by

$$\mathcal{P}_{uu}\left[u \right](t, w) = \int_I P(w, v)(v - w) u(t, v) dv.$$

Note that, even in the linear diffusion case $\phi(u) = u$, this derivation leads to a diffusion of Fick type, differently from the one in (2), that is of Laplacian type. However, expanding the inner derivative in (2), we get

$$\partial_{ww}(D^2(w)u(t,w)) = \partial_w(D^2(w)\partial_w u(t,w)) + \partial_w(\partial_w D^2(w)u(t,w)),$$

where the first term on the r.h.s. corresponds to the diffusion in (9), while the second term plays the role of a local nonlinear transport term. A similar particle derivation can be performed in order to reconstruct the this transport, see [21]. Another difference between the two models is the fact that the *mean opinion*

$$m_1(t) = \int_I wu(t,w)dw, \tag{10}$$

is not preserved in time. From a modelling point of view, this fact can be interpreted as a more *realistic* compromise process, see the discussion in section 4.

1.1.2. *Leaders-followers model.* We consider now a situation where the population is divided in subgroups: one group of followers and two (or more) groups of leaders, see [22]. Hence, by denoting with F_i the opinion of the $i-$th follower and with L_h and R_k the opinion of the $h-$th leader in the *left* group and $k-$th leader of the *right* group respectively, we have that the i-follower opinion evolves according to

$$\dot{F}_i = \frac{\lambda_f^2}{2\sigma_{f,i}}D_f^2(F_i)\left(\phi_f(f_{i-1}) - \phi_f(f_i)\right) - \sum_{j=1}^{N_f}\sigma_{f,j}P_{ff}(F_i,F_j)(F_i - F_j)$$
$$- \sum_{h=1}^{N_l}\sigma_{l,j}P_{fl}(F_i,L_h)(F_i - L_h)$$
$$- \sum_{k=1}^{N_r}\sigma_{r,j}P_{fr}(F_i,R_k)(F_i - R_k),$$

where the last two sums concern the interactions between the i-th follower and the different leaders. Similarly, we can write

$$\dot{L}_i = \frac{\lambda_l^2}{2\sigma_{l,i}}D_l^2(L_i)\left(\phi_l(l_{i-1}) - \phi_l(l_i)\right) - \sum_{h=1}^{N_l}\sigma_{l,j}P_{ll}(L_i,L_h)(L_i - L_h)$$
$$- \sum_{k=1}^{N_r}\sigma_{r,j}P_{lr}(L_i,R_k)(L_i - R_k),$$

for the generic i-th left leader and

$$\dot{R}_i = \frac{\lambda_r^2}{2\sigma_{r,i}}D_r^2(R_i)\left(\phi_r(r_{i-1}) - \phi_r(r_i)\right) - \sum_{k=1}^{N_r}\sigma_{r,j}P_{rr}(R_i,R_k)(R_i - R_k)$$
$$- \sum_{h=1}^{N_l}\sigma_{l,j}P_{lr}(R_i,L_h)(R_i - L_h),$$

for the i-th right leader. In the previous equations we are considering *strong* opinion leaders, i.e. the interaction with the followers does not affect the leader's opinion. In other words, we assume that $P_{lf} = P_{rf} = 0$. From the above system of ODEs,

we can formally derive the following system of PDEs

$$\partial_t f = \partial_w \Big(\frac{\lambda_f^2}{2} D_f^2(w) \partial_w \phi_f(f) - f\big(\mathcal{P}_{ff}[f] + \mathcal{P}_{fl}[l] + \mathcal{P}_{fr}[r]\big) \Big),$$

$$\partial_t l = \partial_w \Big(\frac{\lambda_l^2}{2} D_l^2(w) \partial_w \phi_l(l) - l\big(\mathcal{P}_{ll}[l] + \mathcal{P}_{lr}[r]\big) \Big), \tag{11}$$

$$\partial_t r = \partial_w \Big(\frac{\lambda_r^2}{2} D_r^2(w) \partial_w \phi_r(r) - r\big(\mathcal{P}_{rr}[r] + \mathcal{P}_{rl}[l]\big) \Big).$$

1.1.3. *Leaders-followers-trolls model.* As a last example, we consider a case that is of interest for the understanding of opinion formation and opinion clustering in social media. Accordingly to the consideration mentioned in the Introduction, we assume that one of the group of leaders *introduces* a new group of *fake* agents, commonly known as *trolls*. The trolls are indistinguishable from the followers, they only interact with the reference leaders and they cannot diffuse their opinion. It is reasonable to assume that, in this new setting, the leaders opinion evolution is not affected by the presence of the trolls, while the followers have an additional compromise term that models the followers-trolls interaction. Clearly, since the followers cannot distinguish trolls, the compromise function of the follower-troll interaction remains P_{ff}. Therefore, the ODEs for the followers becomes

$$\dot{F}_i = \frac{\lambda_f^2}{2\sigma_{f,i}} D_f^2(F_i) \left(\phi_f(f_{i-1}) - \phi_f(f_i)\right) - \sum_{j=1}^{N_f} \sigma_{f,j} P_{ff}(F_i, F_j)(F_i - F_j)$$

$$- \sum_{h=1}^{N_l} \sigma_{l,j} P_{fl}(F_i, L_h)(F_i - L_h) - \sum_{k=1}^{N_r} \sigma_{r,j} P_{fr}(F_i, R_k)(F_i - R_k)$$

$$- \sum_{j=1}^{N_f} \sigma_{q,j} P_{ff}(F_i, Q_j)(F_i - Q_j).$$

Besides, consistently with the two features that trolls only interact with the corresponding group of leaders and that they cannot diffuse their opinion, we deduce the trolls evolution law

$$\dot{Q}_i = - \sum_{k=1}^{N_r} \sigma_{r,k} P_{qr}(Q_i, R_k)(Q_i - R_k).$$

In the macroscopic limit we then have

$$\partial_t f = \partial_w \Big(\frac{\lambda_f^2}{2} D_f^2(w) \partial_w \phi_f(f) - f\big(\mathcal{P}_{ff}[f] + \mathcal{P}_{ff}[q] + \mathcal{P}_{fl}[l] + \mathcal{P}_{fr}[r]\big) \Big),$$

$$\partial_t l = \partial_w \Big(\frac{\lambda_l^2}{2} D_l^2(w) \partial_w \phi_l(l) - l\big(\mathcal{P}_{ll}[l] + \mathcal{P}_{lr}[r]\big) \Big), \tag{12}$$

$$\partial_t r = \partial_w \Big(\frac{\lambda_r^2}{2} D_r^2(w) \partial_w \phi_r(r) - r\big(\mathcal{P}_{rr}[r] + \mathcal{P}_{rl}[l]\big) \Big),$$

$$\partial_t q = -\partial_w \Big(q \mathcal{P}_{qr}[r] \Big).$$

In the present paper we provide a solid existence theory for the sample models presented so far, nevertheless it can be easily adapted to many other combinations of populations/interactions the readers may wish to consider.

The paper is structured as follows. In section 2 we introduce the main assumptions, the rigorous statement of the discrete setting and we state our main result Theorem 2.3. Section 3 is devoted to the proof of Theorem 2.3. There, we show how a proper piecewise constant reconstruction of density, built from the *microscopic* ODEs, converges to weak a solution of the corresponding PDEs system. In section 4, instead, we study the large-time behaviour of the macroscopic model and we conclude with some numerical simulations based on the particle scheme, which validate the analytical results of the paper.

2. Preliminaries and main result.

2.1. Main assumptions.
In this section we state the setting and the assumptions of our main result. Let w the a generic opinion belonging to the interval $I = [-1, 1]$, and let $u(t, w)$ the macroscopic opinion density at time t and opinion w of a fixed population. Then the equation for u is given by

$$\partial_t u(t, w) = \partial_w \left(\frac{\lambda_u}{2} D_u^2(w) \partial_w \phi_u(u(t, w)) - u(t, w) \sum_{h \in \{f, l, r, q\}} \mathcal{P}_{uh}[h](t, w) \right). \quad (13)$$

where $\lambda_u > 0$ is a diffusion coefficient. Equation (13) is endowed with zero-flux boundary condition and the initial condition $u(0, w) := \bar{u}$, for some $\bar{u}$ satisfying the following assumptions:

(In1) $\bar{u} \in BV(I; \mathbb{R}^+)$ with $\|\bar{u}\|_{L^1(I)} = \sigma_u$, for some $\sigma_u > 0$,

(In2) there exist $m_u, M_u > 0$ such that $m_u \leq \bar{u}(w) \leq M_u$ for every $w \in I$.

We recall that the function D_u models the diffusion of a given opinion and plays the role of a mobility function. We assume that

(D1) $D_u \in C^2(I)$, $0 \leq D_u \leq \|D_u\|_{L^\infty} < \infty$ and $D_u(\pm 1) = 0$.

(D2) $\partial_w D_u^2$ is uniformly bounded in I.

The prototype example for this diffusivity function is

$$D_u(w) = \left(1 - w^2 \right)^{\frac{\alpha}{2}}, \quad \alpha > 0. \quad (14)$$

Observe that, under this choice, condition (D2) corresponds to the values $\alpha \geq 1$, see Figure 1.

We further assume that

(Dif) $\phi_u : [0, \infty) \to \mathbb{R}$ is a nondecreasing Lipschcitz function, with $\phi_u(0) = 0$.

Remember that the nonlocal operator $\mathcal{P}_{uh}[h](t, w)$ is modelling the local relevance of compromise among the same population (case $h = u$) or between different populations (case $h \neq u$), and is defined by

$$\mathcal{P}_{uh}[h](t, w) = \int_I P_{uh}(w, v)(v - w)h(t, v)dv.$$

We will denote with $\partial_i P_{uh}$ the derivative of P_{uh} w.r.t. the i-th variable, for $i = 1, 2$, and we assume that

(P) for every $h \in \{f, l, r, q\}$, $P_{uh}(\cdot, \cdot)$ is a non negative and uniformly bounded function. Moreover, P_{uh} and $\partial_1 P_{uh}$ are Lipschitz w.r.t. to both components, uniformly in the other:

$$|P_{uh}(w_1, w^*) - P_{uh}(w_2, w^*)| + |P_{uh}(w^*, w_1) - P_{uh}(w^*, w_2)| \leq Lip(P_{uh})|w_1 - w_2|,$$
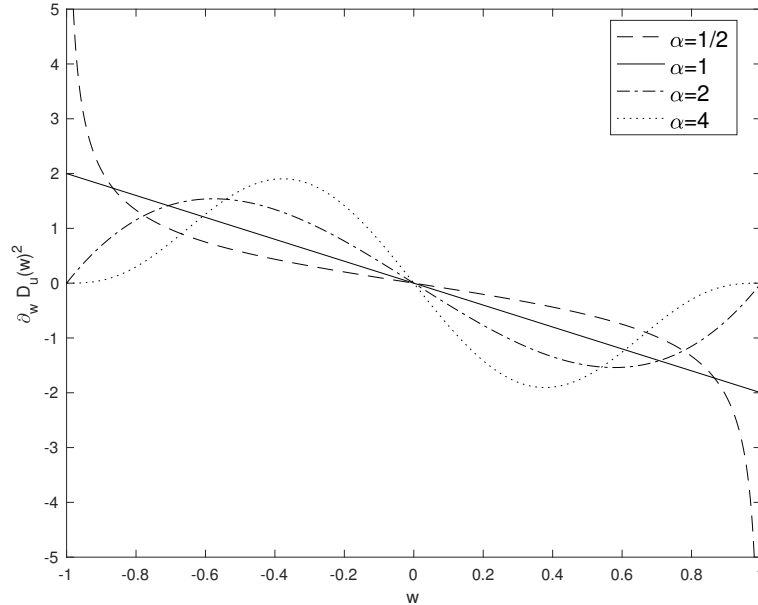
for all $w_1, w_2, w^* \in I$.

FIGURE 1. Behaviours of $\partial_w D_u^2$ for different values of α.

Typical choices for the function P_{uh} are $P_{uh}(w,v) = 1$, that correspond to the Sznajd-Weron model [42], $P_{uh}(w,v) = 1 - |w|$ or $P_{uh}(w,v) = 1 - |w - v|$.

2.2. **Rigorous statement of the discrete setting.** Let us denote with $I_T = [0,T] \times [-1,1]$ and with $\Gamma_T = [0,T] \times \{\pm 1\}$. We notice that we can rewrite both systems (11) and (12) in a compact form as following: let u be either f, l, r or q, then it follows

$$
\begin{cases}
\partial_t u = \partial_w \left(\frac{\lambda_u}{2} D_u^2 \partial_w \phi_u(u) - \sum_{h \in \{f,l,r,q\}} u \mathcal{P}_{uh}[h] \right), & (t,w) \in I_T, \\
\left(\frac{\lambda_u}{2} D_u^2 \partial_w \phi_u(u) - \sum_{h \in \{f,l,r,q\}} u \mathcal{P}_{uh}[h] \right) = 0, & (t,w) \in \Gamma_T, \\
u(0,w) = \bar{u}(w), & w \in I.
\end{cases}
\tag{15}
$$

where P_{uh} and D_u are eventually zero for some $h \in \{f,l,r,q\}$ and $\bar{u} : I \to \mathbb{R}$ is the initial datum for the generic species f, l, r or q, satisfying (In1) and (In2). Fix $N \in \mathbb{N}$, and introduce $\sigma_u^N = \sigma_u/N$ as the infinitesimal mass associated to each opinion. In this way, we will have the same number N of particles for each population. This is just to simplify the computation when we prove the convergence of the scheme. Indeed, our argument works the same even when we have different amounts of particles N^h for different populations, provided that we perform the particle limit at the same time for any all populations. Even if this choice would be more suitable from the modelling point of view, for example in order to catch multi-scale phenomena, taking the same number of particles N is not restrictive since we are interested in recovering the macroscopic equations for all species. Thus, we

atomize $\bar{u}$ in $N+1$ discrete opinions: set $W_0^0 = -1$ and define recursively

$$W_i^0 = \sup\left\{ x \in \mathbb{R} : \int_{W_{i-1}^0}^x \bar{u}(z)dz < \sigma_u^N \right\}, \quad \forall\, i = 1, \ldots, N-1. \qquad (16)$$

By construction we have that $W_N^0 = 1$ and $W_{i+1}^0 - W_i^0 \geq \frac{\sigma_u^N}{M_u}$ for every $i = 0, \ldots, N-1$, where M_u is defined in (In2). Given this initial set of particles, we let the positions W_i evolve according to

$$\begin{cases} \dot{W}_i(t) = \dot{W}_i^d(t) + \dot{W}_i^p(t), & i = 1, \ldots, N-1 \\ \dot{W}_0(t) = \dot{W}_N(t) = 0, \\ W_i(0) = W_i^0, & i = 1, \ldots, N-1 \end{cases} \qquad (17)$$

where

$$\dot{W}_i^d(t) = \frac{\lambda_u}{2\sigma_u^N} D_u^2(W_i(t))\big(\phi_u(u_{i-1}(t)) - \phi_u(u_i(t))\big), \qquad (18)$$

and

$$\dot{W}_i^p(t) = - \sum_{h \in \{f,l,r,q\}} \sigma_h^N \sum_{j=0}^{N+1} P_{uh}(W_i(t), H_j(t))(W_i(t) - H_j(t)). \qquad (19)$$

In (18) and (19), H_j indicates the opinions of the population h and $u_i(t)$ is a local reconstructions for the density

$$u_i(t) := \frac{\sigma_u^N}{(W_{i+1}(t) - W_i(t))}, \quad i = 0, \ldots, N. \qquad (20)$$

Note that

$$\dot{u}_i(t) = -\frac{u_i^2(t)}{\sigma_u^N}(\dot{W}_{i+1}(t) - \dot{W}_i(t)) = \dot{u}_i^d(t) + \dot{u}_i^p(t) \qquad (21)$$

with

$$\dot{u}_i^d(t) = -u_i(t)\frac{\dot{W}_{i+1}^d(t) - \dot{W}_i^d(t)}{W_{i+1}(t) - W_i(t)}, \qquad \dot{u}_i^p(t) = -u_i(t)\frac{\dot{W}_{i+1}^p(t) - \dot{W}_i^p(t)}{W_{i+1}(t) - W_i(t)}. \qquad (22)$$

In order to lighten the notation, we will normalize the diffusion constant $\frac{\lambda_u}{2}$ to 1 and denote by $\mathfrak{H} = \{f, l, r, q\}$ the tag-set of the different populations. We further introduce the constant

$$\Theta_u := \sum_{h \in \mathfrak{H}} \big(Lip[P_{uh}] + \|P_{uh}\|_{L^\infty} + Lip[\partial_1 P_{uh}]\big), \qquad (23)$$

that will be largely used in the following. The following Lemma shows that there is an a priori control from above and below on the mutual distance between particles. In particular, the order of the particles is maintained during the whole evolution.

Lemma 2.1 (Discrete Min-Max Principle). *Let $T > 0$ be fixed and $\bar{u}$ under the assumptions (In1) and (In2). Let c_u be a positive constant satisfying $c_u > \Theta_u$, with Θ_u defined in (23). Then*

$$\frac{\sigma_u^N}{M_u}e^{-c_u T} \leq W_{i+1}(t) - W_i(t) \leq \frac{\sigma_u^N}{m_u}e^{c_u T}, \qquad (24)$$

for every $i = 0, \ldots, N-1$ and $t \in [0, T]$. Moreover

$$e^{-c_u T}m_u \leq u_i(t) \leq e^{c_u T}M_u, \qquad (25)$$

for all $i = 0, ..., N-1$ and $t \in [0, T]$.

Proof. The proof of both inequalities is argued by contradiction. We only prove the lower bound of (24), the proof of the upper bound can be easily obtained by adapting the following steps. Thanks to (In2), we know that

$$W_{i+1}(0) - W_i(0) \geq \frac{\sigma_u^N}{M_u}, \quad i = 0, ..., N.$$

To simplify the notation, let us introduce the *lower bound* function

$$lb_u(t) = \frac{\sigma_u^N e^{-c_u t}}{M_u}.$$

Let now t_1 be

$$t_1 = \inf_{t \in [0,T]} \left\{ W_{i+1}(t) - W_i(t) = lb_u(t), \, W_{j+1}(t) - W_j(t) \geq lb_u(t), j \neq i \right\}. \quad (26)$$

Suppose, by contradiction, that there exists some $0 < t_1 < t_2 \leq T$ such that

$$W_{i+1}(t) - W_i(t) < lb_u(t), \quad t \in (t_1, t_2],$$

then the contradiction follows as soon as we show that

$$\frac{d}{dt} \left[e^{c_u t}(W_{i+1}(t) - W_i(t)) \right]_{|t=t_1} > 0.$$

We can compute the time derivative as follows

$$\frac{d}{dt} \left[e^{c_u t}(W_{i+1}(t) - W_i(t)) \right]_{|t=t_1}$$

$$= e^{c_u t_1} \left[(\dot{W}_{i+1}(t_1) - \dot{W}_i(t_1)) + c_u(W_{i+1}(t_1) - W_i(t_1)) \right]$$

$$= e^{c_u t_1} \left[(\dot{W}_{i+1}^d(t_1) - \dot{W}_i^d(t_1)) + (\dot{W}_{i+1}^p(t_1) - \dot{W}_i^p(t_1)) \right] + \frac{c_u \sigma_u^N}{M_u},$$

with $\dot{W}_i^d$, $\dot{W}_i^p$ defined in (18) and (19) respectively. Thanks to assumption (26),

$$W_{i+1}(t_1) - W_i(t_1) \leq W_{j+1}(t_1) - W_j(t_1), \text{ for all } j \neq i,$$

then $W_{i\pm 1}(t_1) \leq W_i(t_1)$ and this directly implies that $(\dot{W}_{i+1}^d(t_1) - \dot{W}_i^d(t_1)) \geq 0$, using the assumption (Dif). Recalling the definition of $\dot{W}_i^p$ and the assumptions (P) on the generic P_{uh}, it is immediate to get

$$\dot{W}_{i+1}^p(t_1) - \dot{W}_i^p(t_1)$$

$$= -\sum_{h \in \mathfrak{H}} \sigma_h^N \sum_{j=0}^N [P_{uh}(W_{i+1}, H_j)(W_{i+1} - H_j) - P_{uh}(W_i, H_j)(W_i - H_j)]$$

$$= \sum_{h \in \mathfrak{H}} \sigma_h^N \sum_{j=0}^N \left[(P_{uh}(W_{i+1}, H_j) - P_{uh}(W_i, H_j))(W_{i+1} - H_j) \right]$$

$$+ \sum_{h \in \mathfrak{H}} \sigma_h^N \sum_{j=0}^N \left[P_{uh}(W_i, H_j))(W_{i+1} - W_i) \right]$$

$$\geq -\Theta_u(W_{i+1}(t_1) - W_i(t_1)).$$

Finally, by using the above lower bound on the nonlocal part, the 0 lower bound of the diffusive part and by recalling that $c_u > \Theta_u$, we deduce

$$\frac{d}{dt} \left[e^{c_u t}(W_{i+1}(t) - W_i(t)) \right]_{|t=t_1} \geq \frac{\sigma_u^N}{M_u} [c_u - \Theta_u] > 0, \quad (27)$$

which gives the desired contradiction.   □

We are now in position to define the N-discrete density as

$$u^N(t, w) := \sum_{i=0}^{N-1} u_i^N(t) \chi_{[W_i^N(t), W_{i+1}^N(t))}(w), \tag{28}$$

because the intervals $[W_i^N(t), W_{i+1}^N(t))$ are well defined for every $t \in [0, T]$. Moreover,

$$\|u^N(t, \cdot)\|_{L^1(I)} = \sum_{i=0}^{N-1} u_i^N(t) \int_{W_i^N(t)}^{W_{i+1}^N(t)} 1 \, dw = \sum_{i=0}^{N-1} \frac{\sigma_u}{N} = \sigma_u,$$

for every $t \in [0, T]$ and independently on N. As a consequence of (25), we deduce that $(u^N(t, w))_N$ is also uniformly bounded in $L^\infty(I_T)$.

In the following Definition we introduce the notion of weak solutions we are dealing with.

Definition 2.2 (Weak Solution). Let u be either f, l, r or q. The function u is a *weak solution* of (15) if $u \in L^\infty \cap BV(I_T)$, $u(0, \cdot) = \bar{u}$ and satisfies

$$\int_{I_T} \left[u\partial_t\zeta + \phi_u(\rho)\partial_w \left(D_u^2 \partial_w \zeta\right) - u \sum_{h \in \mathfrak{H}} \mathcal{P}_{uh}[h] \, \partial_w \zeta \right] dw \, dt = 0,$$

for all $\zeta \in C_0^\infty(I_T)$.

The main result of this manuscript is stated in the following Theorem.

Theorem 2.3. *For every u, h among $\{f, l, r, q\}$, consider ϕ_u, D_u, P_{uh} under the assumptions (Dif), $(D1)$, $(D2)$ and (P) respectively. Let $\bar{u} : I \to \mathbb{R}$ be as in $(In1)$ and $(In2)$ and let $T > 0$ be fixed. Then, up to a subsequence, the discretized densities u^N defined in (28) strongly converge in $L^1(I_T)$ to a limit in $L^\infty \cap BV(I_T)$ which solves the Initial-boundary value Problem (15) in the sense of Definition 2.2.*

Note that, the well-posedness for equations with a nonlinear *space*-dependent mobility D degenerating at the boundary, as in (9), has not been subject yet of a complete and satisfactory theory, see [5, 7, 25].

2.3. **Tools from Optimal Transport.** The purpose of this section is to collect and present some tools from Optimal Transport that will be useful in the sequel. We refer to [6, 39, 46] for an extensive treatment of the concepts mentioned in this section. In our setting, for arbitrary $t \in [0, T]$, the functions $u^N(t, \cdot)$ are all densities on I with same mass, independently on N. The Wasserstein distance is the right notion of distance that allows us to evaluate how far from each other the two measures $u^N(t, \cdot)$ and $u^N(s, \cdot)$ are at different times $t, s \in [0, T]$. In this section we introduce a well known characterization of the Wasserstein distance in the one-dimensional setting, we refer to [14] for the detailed proof.

For a fixed mass $\sigma > 0$, we consider the space

$$\mathfrak{M}_\sigma = \left\{ \mu \text{ Radon measure on } \mathbb{R} : \mu \geq 0 \text{ and } \mu(\mathbb{R}) = \sigma \right\}.$$

Given $\mu \in \mathfrak{M}_\sigma$, we introduce the pseudo-inverse function $X_\mu \in L^1([0, \sigma]; \mathbb{R})$ as

$$X_\mu(z) = \inf\{x \in \mathbb{R} : \mu((-\infty, x]) > z\}. \tag{29}$$

In particular, if $\sigma = 1$, then $\mathfrak{M}_1$ is the set of non-negative probability densities on $\mathbb{R}$ and it is possible to consider the one-dimensional 1-*Wasserstein distance* between each pair of densities $\rho_1, \rho_2 \in \mathfrak{M}_1$. As shown in [14], in the one dimensional setting

the 1-*Wasserstein distance* can be equivalently defined in terms of the L^1-distance between the respective pseudo-inverse mappings as

$$d_{W^1}(\rho_1, \rho_2) = \|X_{\rho_1} - X_{\rho_2}\|_{L^1([0,1];\mathbb{R})}.$$

For generic $\sigma > 0$, we recall the definition for the *scaled* 1-*Wasserstein distance* between $\rho_1, \rho_2 \in \mathfrak{M}_\sigma$ as

$$d_{W^1,\sigma}(\rho_1, \rho_2) = \|X_{\rho_1} - X_{\rho_2}\|_{L^1([0,\sigma];\mathbb{R})}, \tag{30}$$

see [14, 46].

3. **Proof of the main result.** This section is devoted to the proof of Theorem 2.3. We will first focus on the L^1 compactness for the piecewise constant interpolation $(u^N(t, w))_N$ for $u = f, l, n, g$, defined in (28). The main tool available in this direction is a generalized version of the Aubin-Lions Lemma [37], that we report here in a simplified version adapted to our setting.

Theorem 3.1. *Let $T > 0$ be fixed and $\rho^N(t, \cdot) : [a, b] \to \mathbb{R}$ be a sequence of non negative probability densities for every $t \in [0, T]$ and for every $N \in \mathbb{N}$. Moreover, assume that $\|\rho^N(t, \cdot)\|_{L^\infty} \leq M$ for some constant M independent on t and N. If*

I) $\sup_N \int_0^T TV[\rho^N(t, \cdot)]dt < \infty$,
II) $d_{W^1}(\rho^N(t, \cdot), \rho^N(s, \cdot)) < C|t - s|$ for all $t, s \in [0, T]$, where C is a positive constant independent on N,

then ρ^N is strongly relatively compact in $L^1([0, T] \times [a, b])$.

As a consequence, the desired compactness of $(u^N)_N$ relies on showing a uniform bound of the total variation and the equi-continuity of the $1-$Wasserstein distance with respect to the time variable.

We first focus on the bounds on the Total Variation.

Proposition 1 (BV bound for Systems). *Let $T > 0$ and $\bar{u}$ be under assumptions (In1) and (In2). Then the discrete densities u^N defined in (28) satisfy*

$$TV\left[u^N(t, \cdot)\right] \leq TV\left[\bar{u}\right] C, \qquad \forall\, t \in [0, T], \,\forall\, N \in \mathbb{N}, \tag{31}$$

where the constant C is such that $C = C(D_u, \phi_u, \Theta_u, T)$.

Proof. The proof is based on a Gronwall type estimate for $TV[u^N]$. To shorten the notation in the following computations we introduce the auxiliar functions

$$s_i(t) = \text{sign}(u_i(t) - u_{i+1}(t)) - \text{sign}(u_{i-1}(t) - u_i(t)),$$

for $i = 1, ..., N - 1$. Standard computations lead to the following expression

$$\frac{d}{dt}TV\left[u^N(t, \cdot)\right] = \dot{u}_0(t) + \dot{u}_{N-1}(t) + \sum_{i=1}^{N-2} \text{sign}\left(u_{i+1}(t) - u_i(t)\right)\left(\dot{u}_{i+1}(t) - \dot{u}_i(t)\right)$$

$$= \dot{u}_0(t) + \dot{u}_{N-1}(t) + \text{sign}(u_0(t) - u_1(t))\dot{u}_0^d(t) - \text{sign}(u_{N-2}(t) - u_{N-1}(t))\dot{u}_{N-1}^d(t)$$

$$+ \sum_{i=1}^{N-2} \text{sign}\left(u_{i+1}(t) - u_i(t)\right)\left(\dot{u}_{i+1}^p(t) - \dot{u}_i^p(t)\right) + \sum_{i=1}^{N-2} s_i(t)\dot{u}_i^d(t),$$

where we used (21). We first show that the boundary terms involving $\dot{u}_0$ and $\dot{u}_{N-1}$ are uniformly bounded with respect to N. Consider, for example, the contribution of the left boundary term

$$\dot{u}_0(t) + \text{sign}(u_0(t) - u_1(t))\dot{u}_0^d.$$

When $u_0(t) \leq u_1(t)$, the diffusive part cancels out and only the term $\dot{u}_0^p$ survives. If, instead, $u_1(t) \leq u_0(t)$, from the monotonicity of ϕ_u we deduce that

$$\dot{u}_0 + \text{sign}(u_0 - u_1)\dot{u}_0^d = \dot{u}_0^p + 2\dot{u}_0^d = \dot{u}_0^p - 2u_0\frac{D^2(W_1)(\phi_u(u_0) - \phi_u(u_1))}{\sigma_u^N(W_1 - W_0)} \leq \dot{u}_0^p.$$

Then, in both the cases, it is enough to show that the non-local part of $\dot{u}_0$ is uniformly bounded. From the positivity of P_{uh} and the bounds (24), it is easy to see that

$$\dot{u}_0^p = \frac{u_0}{W_1 - W_0}\sum_{h \in \mathfrak{H}}\sigma_h^N\sum_{j=0}^{N}P_{uh}(W_1, H_j)(W_1 - H_j)$$

$$\leq \frac{u_0}{W_1 - W_0}\sum_{h \in \mathfrak{H}}\sigma_h^N\sum_{j:H_j<W_1}P_{uh}(W_1, H_j)(W_1 - H_j)$$

$$\leq \frac{u_0}{W_1 - W_0}\sum_{h \in \mathfrak{H}}\sigma_h^N\sum_{j:H_j<W_1}P_{uh}(W_1, H_j)(W_1 - W_0)$$

$$\leq u_0\sum_{h \in \mathfrak{H}}[\sigma_h\|P_{uh}\|_{L^\infty}].$$

With a symmetric argument, one can see that the term also satisfies

$$\dot{u}_{N-1} - \text{sign}(u_{N-2} - u_{N-1})\dot{u}_{N-1}^d \leq u_{N-1}\sum_{h \in \mathfrak{H}}[\sigma_h\|P_{uh}\|_{L^\infty}].$$

Let us now prove that

$$\sum_{i=1}^{N-2}s_i(t)\dot{u}_i^d(t) \leq 0. \tag{32}$$

If at time t one has $u_{i+1}(t) \leq u_i(t) \leq u_{i-1}(t)$, or $u_{i+1}(t) \geq u_i(t) \geq u_{i-1}(t)$ then $s_i(t) = 0$. When, instead, $u_{i\pm1}(t)$ are both bigger or both smaller than $u_i(t)$, the monotonicity of ϕ_u implies the desired estimate. Indeed, if we assume for example that $u_{i\pm1}(t) < u_i(t)$, then $s_i(t) = 2$ and, recalling the definition of $\dot{u}_i^d$, we get

$$s_i\dot{u}_i^d = 2\frac{u_i^2}{(\sigma_u^N)^2}\left[D_u^2(W_i)(\phi_u(u_{i-1}) - \phi_u(u_i)) - D_u^2(W_{i+1})(\phi_u(u_i) - \phi_u(u_{i+1}))\right] < 0.$$

The other case $u_{i\pm1}(t) > u_i(t)$ follows analogously. To conclude, we are left to estimate the term concerning the non-local contribution. Using the definition of $\dot{u}_i^p$ in (22), we deduce that

$$\sum_{i=1}^{N-2}\text{sign}(u_{i+1}(t) - u_i(t))(\dot{u}_{i+1}^p(t) - \dot{u}_i^p(t))$$

$$\leq \sum_{i=1}^{N-2}|u_{i+1}(t) - u_i(t)|\frac{|\dot{W}_{i+2}^p(t) - \dot{W}_{i+1}^p(t)|}{W_{i+2}(t) - W_{i+1}(t)}$$

$$+ \sum_{i=1}^{N-2}u_i(t)\left|\frac{\dot{W}_{i+2}^p(t) - \dot{W}_{i+1}^p(t)}{W_{i+2}(t) - W_{i+1}(t)} - \frac{\dot{W}_{i+1}^p(t) - \dot{W}_i^p(t)}{W_{i+1}(t) - W_i(t)}\right|. \tag{33}$$

Thanks to the assumption (P), we can easily bound the term in (33) with the total variation of u^N

$$\sum_{i=1}^{N-2} |u_{i+1}(t) - u_i(t)| \frac{|\dot{W}^p_{i+2}(t) - \dot{W}^p_{i+1}(t)|}{W_{i+2}(t) - W_{i+1}(t)} \leq \Theta_u TV[u^N(t, \cdot)]. \tag{34}$$

On the other hand, since the functions P_{uh} are Lipschitz, there exist some $\bar{W}^{i+2}_{i+1} \in [W_{i+1}, W_{i+2}]$ and $\bar{W}^{i+1}_i \in [W_i, W_{i+1}]$ such that

$$\frac{\dot{W}^p_{i+2} - \dot{W}^p_{i+1}}{W_{i+2} - W_{i+1}} - \frac{\dot{W}^p_{i+1} - \dot{W}^p_i}{W_{i+1} - W_i} =$$

$$= -\sum_{h \in \mathfrak{H}} \sigma^N_h \sum_{j=0}^{N} \frac{P_{uh}(W_{i+2}, H_j)(W_{i+2} - H_j) - P_{uh}(W_{i+1}, H_j)(W_{i+1} - H_j)}{W_{i+2} - W_{i+1}}$$

$$+ \sum_{h \in \mathfrak{H}} \sigma^N_h \sum_{j=0}^{N} \frac{P_{uh}(W_{i+1}, H_j)(W_{i+1} - H_j) - P_{uh}(W_i, H_j)(W_i - H_j)}{W_{i+1} - W_i}$$

$$\leq \sum_{h \in \mathfrak{H}} \sigma^N_h \sum_{j=0}^{N} \left| \partial_1 P_{uh}(\bar{W}^{i+2}_{i+1}, H_j)(\bar{W}^{i+2}_{i+1} - H_j) - \partial_1 P_{uh}(\bar{W}^{i+1}_i, H_j)(\bar{W}^{i+1}_i - H_j) \right|$$

$$+ \sum_{h \in \mathfrak{H}} \sigma^N_h \sum_{j=0}^{N} \left| P_{uh}(\bar{W}^{i+2}_{i+1}, H_j) - P_{uh}(\bar{u}^{i+1}_i, H_j) \right|$$

$$\leq (W_{i+2} - W_i) \sum_{h \in \mathfrak{H}} \left(2Lip[\partial_1 P_{uh}] + \|\partial_1 P_{uh}\|_{L^\infty} + Lip[P_{uh}] \right).$$

Using now the upper bound of (24) in the above estimate, we can handle also the second term of (33) and get

$$\sum_{i=1}^{N-2} u_i(t) \left| \frac{\dot{W}^p_{i+2}(t) - \dot{W}^p_{i+1}(t)}{W_{i+2}(t) - W_{i+1}(t)} - \frac{\dot{W}^p_{i+1}(t) - \dot{W}^p_i(t)}{W_{i+1}(t) - W_i(t)} \right| \leq 4 \frac{e^{c_u T}}{m_u} \Theta_u. \tag{35}$$

Summarizing, (34) and (35), together with (32) and the estimates on the boundary terms, imply that

$$\frac{d}{dt} TV[u^N(t, \cdot)] \leq \Theta_u \left(4 \frac{e^{c_u T}}{m_u} + + TV[u^N(t, \cdot)] \right)$$

and the conclusion follows via a Gronwall type argument. $\qquad \square$

We then show that the sequence $u^N(t, w)$ satisfies and equi-continuity in time with respect to the 1-Wasserstein distance, the proof realise on a, right now, quite standard argument introduced in [20], that we report here for completeness.

Proposition 2. *Let u^N be any of f^N, l^N, r^N, q^N, defined in (28). Let $T > 0$ and $\bar{u}$ under assumptions (In1) and (In2). If D_u, ϕ_u, P_{uh} are under the assumptions (D1), (D2), (Dif) and (P) respectively, then there exists a constant $C > 0$ such that*

$$d_{W^1}(u^N(t, \cdot), u^N(s, \cdot)) \leq C|t - s| \qquad \text{for all } s, t \in [0, T] \text{ and } N \in \mathbb{N}. \tag{36}$$

Proof. The proof of this result is very standard in the one dimensional setting, and we will use the tools introduced in Section 2.3. The pseudo-inverse map associated

to the piecewise constant probability density u^N can be very easily computed and it is given by

$$X_{u^N(t,\cdot)}(x) = \sum_{i=0}^{N-1} \left(W_i(t) + \left(w - \sigma_u^N i \right) \frac{1}{u_i(t)} \right) \chi_{J_u^N(i)}(w),$$

where $J_u^N(i) = \left[\sigma_u^N i, \sigma_u^N(i+1) \right)$. For any $t > s$ we have

$$d_{W^1}(u^N(t,\cdot), u^N(s,\cdot)) = \|X_{u^N(t,\cdot)} - X_{u^N(s,\cdot)}\|_{L^1([0,1])}$$

$$\leq \sum_{i=0}^{N-1} \int_{J_u^N(i)} \left| W_i(t) - W_i(s) + \left(w - \sigma_u^N i \right) \left(\frac{1}{u_i(t)} - \frac{1}{u_i(s)} \right) \right| dw$$

$$\leq \sigma_u^N \sum_{i=0}^{N-1} |W_i(t) - W_i(s)| + \frac{(\sigma_u^N)^2}{2} \sum_{i=0}^{N-1} \int_s^t \left| \frac{d}{d\tau} \frac{1}{u_i(\tau)} \right| d\tau$$

$$\leq 3\sigma_u^N \sum_{i=0}^{N-1} \int_s^t |\dot{W}_i(\tau)| d\tau$$

$$\leq 3\|D_u^2\|_{L^\infty} Lip[\phi_u](t-s) \sum_{i=0}^{N-1} |W_{i+1} - W_i| + 3(t-s)\Theta_u$$

$$\leq C(D_u, \phi_u, \bar{u}, P_{uh})(t-s).$$

where we have used the assumption on the uniform bound of both on the interaction potentials P_{uh} and the diffusivity D_u, and the estimate (31). $\qquad\square$

Gathering together Theorem 3.1 with Propositions 1 and 2, we conclude the following.

Corollary 1. *Let u^N be any of f^N, l^N, r^N, q^N, defined in (28). Then there exists some $u \in L^1 \cap L^\infty(I_T)$ such that $\|u^N - u\|_{L^1} \to 0$ as $N \to \infty$.*

We now focus on the identification of the limit given by Corollary 1 as weak solution of the system (15). This will be, indeed, a straightforward consequence of Propositions 3 and 4.

Proposition 3. *Let $T > 0$ and $\bar{u}$ under assumptions $(In1)$ and $(In2)$. Let D_u, ϕ_u, P_{uh} be under the assumptions $(D1)$, $(D2)$, (Dif) and (P) respectively. Consider u^N and h^N be any two sequences among $\{f^N, l^N, r^N, q^N\}$ and the respective L^1-strong limits u, h provided by Corollary 1. Then for every $\zeta \in C_0^\infty((0,T) \times (-1,1))$ we get*

$$\lim_{N\to\infty} \int_0^T \int_I u^N \partial_t \zeta = \int_0^T \int_I u \partial_t \zeta,$$

$$\lim_{N\to\infty} \int_0^T \int_I \phi_u(u^N) \partial_w (D_u^2(w)\partial_w \zeta) = \int_0^T \int_I \phi_u(u) \partial_w (D_u^2(w)\partial_w \zeta), \qquad (37)$$

$$\lim_{N\to\infty} \int_0^T \int_I u^N \sum_{h \in \mathfrak{H}} \mathcal{P}_{uh}[\hat{h}^N] \partial_w \zeta = \int_0^T \int_I u \sum_{h \in \mathfrak{H}} \mathcal{P}_{uh}[h] \partial_w \zeta,$$

where $\hat{h}^N(t,w) := \frac{1}{N} \sum_{j=0}^N \delta_{H_j(t)}(w)$ indicates the sequence of empirical measures.

Proof. The first two of (37) are follow as direct consequence of the strong L^1-compactness obtained in Corollary 1, together with the Lipschitz regularity of the

nonlinear diffusion ϕ_u and the uniform bound of $\partial_w D^2$ on I. Concerning the third part, we need to first show that the empirical measures $\hat{h}^N$ and the piecewise constant densities h^N share the same limit h with respect to a suitable topology. Indeed, it is possible to prove that

$$d_{W^1}\left(h(t,\cdot),\hat{h}^N(t,\cdot)\right) \longrightarrow 0 \quad \text{as } N \to \infty \text{ for all } t \in [0,T]. \tag{38}$$

This is follows by the identity $d_{W^1}(\mu,\nu) = \|X_\mu - X_\nu\|_{L^1}$ between probability measures. Observing that the pseudo-inverse mapping of an empirical measure is piecewise constant, it is easy to see that

$$\|X_{h^N(t,\cdot)} - X_{h\hat{N}(t,\cdot)}\|_{L^1([-1,1])} \leq \sum_{i=0}^{N-1} \int_{\sigma_h^N i}^{\sigma_h^N(i+1)} \left|(w - \sigma_h^N)\frac{1}{h_i(t)}\right| dw$$

$$= \frac{\sigma_h^N}{2} \sum_{i=0}^{N-1}(H_{i+1}(t) - H_i(t)) \leq \sigma_h^N.$$

Recalling that the sequence h^N also converges to h with respect to the 1-Wasserstein distance, we deduce

$$d_{W_1}(h(t,\cdot),\hat{h}^N(t,\cdot)) \leq d_{W_1}(h(t,\cdot),h^N(t,\cdot)) + d_{W_1}(h^N(t,\cdot),\hat{h}^N(t,\cdot)) \leq C\sigma_h^N,$$

for some positive geometric constant C, and then (38) follows. We are now in position to prove the third limit in (37), which is equivalent to show the following

$$\int_0^T \int_I (u^N - u) \sum_{h \in \mathfrak{H}} \mathcal{P}_{uh}[\hat{h}^N]\partial_w \zeta \to 0,$$

and

$$\int_0^T \int_I u \sum_{h \in \mathfrak{H}} (\mathcal{P}_{uh}[\hat{h}^N] - \mathcal{P}_{uh}[h])\partial_w \zeta \to 0.$$

The first limit is immediate because of the boundedness assumption on any of the P_{uh}. On the other hand, for the second limit, we can always consider an optimal plan $\gamma_h^N(t)$ between the probability measures $\hat{h}^N(t,\cdot)$ and $h^N(\cdot)$ with respect to the cost $|\cdot|$. Then for every $t \geq 0$ fixed, calling $C_1 = \|\partial_w \zeta\|_{L^\infty} M_u e^{c_u T}$ we get

$$\int_I \left|u \sum_{h \in \mathfrak{H}} (\mathcal{P}_{uh}[\hat{h}^N] - \mathcal{P}_{uh}[h])\partial_w \zeta\right| dw$$

$$\leq C_1 \sum_{h \in \mathfrak{H}} \int_I \left|\int_I P_{uh}(w,v)(v - w)dh^N(t,v) - \int_I P_{uh}(w,v')(v' - w)d\hat{h}^N(t,v')\right| dw$$

$$= C_1 \sum_{h \in \mathfrak{H}} \int_I \left|\int_{I \times I} \left(P_{uh}(w,v)(v - w) - P_{uh}(w,v')(v' - w)\right)d\gamma_h^N(t)(v,v')\right| dw$$

$$\leq C_1 \sum_{h \in \mathfrak{H}} (Lip[P_{uh}] + \|P_{uh}\|_{L^\infty}) \int_I \int_{I \times I} |v - v'|d\gamma_h^N(t)(v,v')dw$$

$$\leq C_1 \Theta_u d_{W_1}(h^N(t,\cdot),\hat{h}^N(t,\cdot)) \leq C_1 \Theta_u \sigma_h^N.$$

Then,

$$\int_0^T \int_I \left|u \sum_{h \in \mathfrak{H}} (\mathcal{P}_{uh}[\hat{h}^N] - \mathcal{P}_{uh}[h])\partial_w \zeta\right| dw dt \leq C_1 T \Theta_u \frac{1}{N} \sum_{h \in \mathfrak{H}} \sigma_h$$

and this concludes the proof. $\square$

The remaining part of this section is devoted to prove that the sequence u^N satisfies the the weak formulation of (15) in the limit as $N \to \infty$.

Proposition 4. *Under the same assumptions of Proposition 3, for every $\zeta \in C_c^\infty((0,T) \times (-1,1))$, one has*

$$\lim_{N \to \infty} \int_0^T \int_I u^N \partial_t \zeta + \phi_u(u^N) \partial_w \left(D_u^2(w) \partial_w \zeta \right) - u^N \sum_{h \in \mathfrak{H}} \mathcal{P}_{uh} \left[\hat{h}^N \right] \partial_w \zeta \, dw dt = 0. \quad (39)$$

Proof. For simplicity we will denote the three contributions on the r.h.s. of (39) as follows

$$I := \int_0^T \int_I u^N \partial_t \zeta \, dw dt, \quad II := \int_0^T \int_I \phi(u^N) \partial_w (D^2(w) \partial_w \zeta) \, dw dt,$$
$$III := -\int_0^T \int_I u^N \sum_{h \in \mathfrak{H}} \mathcal{P}_{uh} [\hat{h}^N] \partial_w \zeta \, dw dt. \quad (40)$$

Recalling the definition of u^N and performing some standard computations, as discrete integration by parts and reconstruction of the derivative, it is easy to rewrite

$$I = \sum_{i=0}^{N-1} \int_0^T \int_{W_i(t)}^{W_{i+1}(t)} u_i^N(t) \partial_t \zeta(t,w) dw \, dt$$
$$= \sum_{i=0}^{N-1} \int_0^T u_i^N(t) \dot{W}_{i+1}(t) \int_{W_i(t)}^{W_{i+1}(t)} \left(\fint_{W_i(t)}^{W_{i+1}(t)} \zeta(t,w) - \zeta(t,W_{i+1}(t)) \right) dw dt$$
$$- \sum_{i=0}^{N-1} \int_0^T u_i^N(t) \dot{W}_i(t) \int_{W_i(t)}^{W_{i+1}(t)} \left(\fint_{W_i(t)}^{W_{i+1}(t)} \zeta(t,w) - \zeta(t,W_i(t)) \right) dw dt,$$

thus, expanding $\zeta(t,w)$ at first order with respect to $W_{i+1}(t)$ and $W_i(t)$ respectively, we obtain that for some $\alpha_w^i \in [w, W_{i+1}(t)]$ and $\beta_w^i \in [W_i(t), w]$ for which

$$I = \frac{\sigma_u^N}{2} \sum_{i=0}^{N-1} \int_0^T [-\partial_w \zeta(t, W_{i+1}(t)) \dot{W}_{i+1}(t) - \partial_w \zeta(t, W_i(t)) \dot{W}_i(t)] dt$$
$$+ \frac{1}{2} \sum_{i=0}^{N-1} \int_0^T u_i^N(t) \dot{W}_{i+1}(t) \int_{W_i(t)}^{W_{i+1}(t)} \partial_{ww} \zeta(t, \alpha_w^i)(w - W_{i+1}(t))^2 dw dt$$
$$- \frac{1}{2} \sum_{i=0}^{N-1} \int_0^T u_i^N(t) \dot{W}_i(t) \int_{W_i(t)}^{W_{i+1}(t)} \partial_{ww} \zeta(t, \beta_w^i)(w - W_i(t))^2 dw dt.$$

Recalling that $\dot{W}_i(t) = \dot{W}_i^d(t) + \dot{W}_i^p(t)$, we can split the contribution of the first term of the r.h.s. in a diffusive part and a non local part. Then we can rewrite I

as the sum of the following three terms

$$I_1 := -\frac{\sigma_u^N}{2} \sum_{i=0}^{N-1} \int_0^T [\partial_w \zeta(t, W_{i+1}(t))\dot{W}_{i+1}^d(t) + \partial_w \zeta(t, W_i(t))\dot{W}_i^d(t)]dt$$

$$I_2 := -\frac{\sigma_u^N}{2} \sum_{i=0}^{N-1} \int_0^T [\partial_w \zeta(t, W_{i+1})\dot{W}_{i+1}^p(t) + \partial_w \zeta(t, W_i(t))\dot{W}_i^p(t)]dt$$

$$I_3 := \frac{1}{2} \sum_{i=0}^{N-1} \int_0^T u_i^N(t)\dot{W}_{i+1}(t) \int_{W_i(t)}^{W_{i+1}(t)} \partial_{ww}\zeta(t, \alpha_w^i)(w - W_{i+1}(t))^2$$

$$- \frac{1}{2} \sum_{i=0}^{N-1} \int_0^T u_i^N(t)\dot{W}_i(t) \int_{W_i(t)}^{W_{i+1}(t)} \partial_{ww}\zeta(t, \beta_w^i)(w - W_i(t))^2.$$

In the sequel we do not explicit the dependence on N or on the variable t whenever this is clear from the context. For sake of clarity, we divide the rest of the proof in three steps.

Step 1. We first prove that $I_1 + II = 0$. Standard algebraic computations and discrete integration by parts lead to

$$I_1 + II = \sum_{i=2}^{N-2} \int_0^T \partial_w \zeta(t, W_i)D_u^2(W_i)[\phi_u(u_i) - \phi_u(u_{i-1})]dt$$

$$+ \sum_{i=1}^{N-1} \partial_w \zeta(t, W_i)D_u^2(W_i)(\phi_u(u_{i-1}) - \phi_u(u_i))$$

$$+ \frac{1}{2} \int_0^T [\partial_w \zeta(t, W_N)D_u^2(W_N)\phi_u(u_{N-1}) - \partial_w \zeta(t, W_0)D_u^2(W_0)\phi_u(u_0)]dt$$

$$- \int_0^T \partial_w \zeta(t, W_{N-1})D_u^2(W_{N-1})(\phi_u(u_{N-2}) - \phi_u(u_{N-1}))dt$$

$$- \int_0^T \partial_w \zeta(t, W_1)D_u^2(W_1)(\phi_u(u_0) - \phi_u(u_1))dt$$

$$= \frac{1}{2} \int_0^T \partial_w \zeta(t, Z_N)D^2(Z_N)\phi_u(W_{N-1}) - \partial_w \zeta(t, W_0)D_u^2(W_0)\phi_u(u_0)dt = 0,$$

where the last equality holds because $\partial_w \zeta(t, W_0) = \partial_w \zeta(t, W_N) = \partial_w \zeta(t, \pm 1) = 0$ for all $t \in [0, T]$.

Step 2. In this step we show that $I_2 + III$ vanishes as $N \to \infty$. A discrete integration by parts, together with the fact that $\partial_w \zeta(t, W_0) = \partial_w \zeta(t, W_N) = 0$, imply

$$I_2 + III = \sum_{i=1}^{N-1} \sum_{h \in \mathfrak{H}} \sum_{j=0}^N \sigma_u^N \sigma_h^N \int_0^T \partial_w \zeta(t, W_i)P_{uh}(W_i, H_j)(W_i - H_j)dt$$

$$- \sum_{i=0}^{N-1} \sum_{h \in \mathfrak{H}} \sum_{j=0}^N \sigma_h^N \int_0^T u_i \int_{W_i}^{W_{i+1}} \partial_w \zeta(t, w)P_{uh}(w, H_j)(w - H_j)dt.$$

Then, expanding $\partial_w\zeta$ at first order with respect to W_i and still using that ζ has compact support in $(-1,1)$, there is $\gamma_w^i \in [W_i, w]$ such that

$$
\begin{aligned}
I_2 + III = & \\
= & \sum_{i=1}^{N-1}\sum_{h\in\mathfrak{H}}\sum_{j=0}^{N}\sigma_h^N\int_0^T u_i\partial_w\zeta(t,W_i)\int_{W_i}^{W_{i+1}} P_{uh}(W_i,H_j)(W_i-H_j)dwdt \\
& - \sum_{i=1}^{N-1}\sum_{h\in\mathfrak{H}}\sum_{j=0}^{N}\sigma_h^N\int_0^T u_i\partial_w\zeta(t,W_i)\int_{W_i}^{W_{i+1}} P_{uh}(w,H_j)(w-H_j)dwdt \\
& - \sum_{i=1}^{N-1}\sum_{h\in\mathfrak{H}}\sum_{j=0}^{N}\sigma_h^N\int_0^T u_i\int_{W_i}^{W_{i+1}} P_{uh}(w,H_j)(w-H_j)(w-W_i)\partial_{ww}\zeta(t,\gamma_w^i)dwdt.
\end{aligned}
$$

Since P_{uh} is Lipschitz in the first component, we obtain

$$
\begin{aligned}
|I_2 + III| & \leq C\sum_{h\in\mathfrak{H}} Lip[P_{uh}]\sum_{i=1}^{N-1}\int_0^T \left[u_i(W_{i+1}-W_i)^2 + u_i(W_{i+1}-W_i)^2\right]dt \\
& \leq 2CT\,\sigma_u^N\sum_{h\in\mathfrak{H}} Lip[P_{uh}]
\end{aligned}
$$

where $C = \left(3\|\zeta_w\|_\infty + \|\zeta_{ww}\|_\infty\right)$, thus the term $I_2 + III \to 0$ as $N \to \infty$.

Step 3. In this last step we prove that I_3 also vanishes as $N \to \infty$. For simplicity we introduce the rest functions

$$
\mathcal{R}_{i,\alpha}(t) = \int_{W_i}^{W_{i+1}} \partial_{ww}\zeta(t,\alpha_w^i)(w-W_{i+1})^2 dw,
$$

$$
\mathcal{R}_{i,\beta}(t) = \int_{W_i}^{W_{i+1}} \partial_{ww}\zeta(t,\beta_w^i)(w-W_{i+1})^2 dw,
$$

in particular, thanks to upper bound of (24), we observe

$$
|\mathcal{R}_{i,\alpha}|, |\mathcal{R}_{i,\beta}| < \|\partial_{ww}\zeta\|_{L^\infty}(W_{i+1}-W_i)^3 < \|\partial_{ww}\zeta\|_{L^\infty}\left(\frac{e^{c_uT}}{m_u}\right)^3(\sigma_u^N)^3. \tag{41}
$$

Using $\mathcal{R}_{i,\alpha/\beta}(t)$ we can rewrite I_3 as

$$
\begin{aligned}
I_3 = & \frac{1}{2}\sum_{i=0}^{N-1}\int_0^T u_i[(\dot{W}_{i+1}^d - \dot{W}_i^d) + (\dot{W}_{i+1}^p - \dot{W}_i^p)]\mathcal{R}_{i,\alpha} + u_i\dot{W}_i[\mathcal{R}_{i,\alpha} - \mathcal{R}_{i,\beta}]dt \\
& + \frac{1}{2}\sum_{i=0}^{N-1}\int_0^T u_i\dot{W}_i\int_{W_i}^{W_{i+1}}\partial_{ww}\zeta(t,\beta_w^i)[(w-W_{i+1})^2 - (w-W_i)^2]dwdt
\end{aligned}
$$

and by replacing the expression of $\dot{W}_i^d$ and $\dot{W}_i^p$, we obtain

$$I_3 = \frac{1}{2\sigma_u^N} \sum_{i=1}^{N-2} \int_0^T u_i (D_u^2(W_{i+1}) - D_u^2(W_i))(\phi_u(u_i) - \phi_u(u_{i+1}))\mathcal{R}_{i,\alpha} dt$$

$$+ \frac{1}{2\sigma_u^N} \sum_{i=1}^{N-2} \int_0^T u_i D_u^2(W_i)[(\phi_u(u_i) - \phi_u(u_{i+1})) - (\phi_u(u_{i-1}) - \phi_u(u_i))]\mathcal{R}_{i,\alpha} dt$$

$$+ \frac{1}{2} \sum_{i=1}^{N-2} \sum_{h\in\mathfrak{H}} \sum_{j=0}^{N} \sigma_h^N \int_0^T u_i [P_{uh}(W_{i+1}, H_j) - P_{uh}(W_i, H_j))(H_j - W_{i+1}]\mathcal{R}_{i,\alpha} dt$$

$$+ \frac{1}{2} \sum_{i=1}^{N-2} \sum_{h\in\mathfrak{H}} \sum_{j=0}^{N} \sigma_h^N \int_0^T u_i P_{uh}(W_i, H_j)(W_i - H_j)\mathcal{R}_{i,\alpha}(t) dt$$

$$+ \frac{1}{2} \int_0^T \left[u_0 \dot{W}_1 \mathcal{R}_{0,\alpha} - u_{N-1} \dot{W}_{N-1} \mathcal{R}_{N-1,\alpha} + \sum_{i=0}^{N-1} u_i \dot{W}_i (\mathcal{R}_{i,\alpha} - \mathcal{R}_{i,\beta}) \right] dt$$

$$+ \frac{1}{2} \sum_{i=0}^{N-1} \int_0^T u_i \dot{W}_i \int_{W_i}^{W_{i+1}} \partial_{ww}\zeta(t, \beta_w^i)[(w - W_{i+1})^2 - (w - W_i)^2] dw\, dt.$$

Thanks to estimates (41) and (31), and the fact that

$$|\dot{W}_i| < 2\frac{Lip[D_u]}{\sigma_u^N} + \sum_{h\in\mathfrak{H}} \|P_{uh}\|_{L^\infty} \qquad \text{for all } i \text{ in} \{0, \ldots, N\},$$

we deduce that all the terms of I_3 except the last one can be estimated from above by $C\sigma_u^N$, where C is some positive constant depending on ζ, T and on the proper bounds on D_u, ϕ_u, P_{uh} provided by the assumptions (D1), (D2), (Dif) and (P). Concerning the last term of I_3, standard computations, the above estimate on $|\dot{W}_i|$ and (24) and (25) bring to

$$\frac{1}{2} \sum_{i=0}^{N-1} \int_0^T u_i \dot{W}_i \int_{W_i}^{W_{i+1}} \partial_{ww}\zeta(t, \beta_w^i)[(w - W_{i+1})^2 - (w - W_i)^2] dw\, dt$$

$$\leq 2 \sum_{i=0}^{N-1} \int_0^T u_i \dot{W}_i \int_{W_i}^{W_{i+1}} \partial_{ww}\zeta \big[(w - W_{i+1})(W_{i+1} - W_i) + (W_{i+1} - W_i)^2\big] dw\, dt$$

$$\leq 2\|\partial_{ww}\zeta\|_\infty \sum_{i=0}^{N-1} \int_0^T u_i |\dot{W}_i|(W_{i+1} - W_i)^3 dt$$

$$\leq \|\partial_{ww}\zeta\|_\infty \frac{2M_u e^{4c_u T}}{m_u^3} (\sigma_u^N)^3 \sum_{i=0}^{N-1} \int_0^T |\dot{W}_i| dt$$

$$\leq \|\partial_{ww}\zeta\|_\infty \frac{2T\sigma_u M_u e^{4c_u T}}{m_u^3} \left(2Lip[D_u] + \sum_{h\in\mathfrak{H}} \|P_{uh}\|_{L^\infty} \right) \sigma_u^N.$$

Finally, there exists some constant C depending on the data of the problem and on the function ζ so that

$$|I_3| \leq C\sigma_u^N,$$

and this concludes the proof of the Step 3 and of Proposition 4. $\qquad\square$

4. **Large-time behaviour.** This last section is devoted to the study of the large-time behaviour for the model (13). The main difficulty of such study concerns the evolution of the mean opinion which is, unfortunately, not close in general. The mean opinion corresponds to the first moment function

$$m_1(t) = \int_I v u(t,v) dv.$$

Consider, for example, the easiest case of one species, with $P(w,v) = 1$ in (13). Then the compromise part corresponds to

$$\mathcal{P}_{uu}[u](t,w) = \int_I (v-w)u(t,v)dv = m_1(t) - \sigma w, \tag{42}$$

and $m_1(t)$ evolves according to

$$\frac{d}{dt}m_1(t) = \int_I w\partial_w \left(\frac{\lambda^2}{2}D^2(w)\partial_w\phi(u(t,w)) - (m_1(t) - \sigma w)\, u(t,w) \right) dw$$

$$= -\int_I \left(\frac{\lambda^2}{2}D^2(w)\partial_w\phi(u(t,w)) - (m_1(t) - \sigma w)\, u(t,w) \right) dw$$

$$= -\int_I \frac{\lambda^2}{2}D^2(w)\partial_w\phi(u(t,w))dw.$$

It is then evident that the evolution of m_1 is independent on higher order moments of u and it strongly depend on the choice of the mobility D. As a consequence, the exact evaluation of the limit

$$m_1^\infty = \lim_{t\to\infty} m_1(t), \tag{43}$$

is quite difficult to investigate, we refer to [5] for a more detailed discussion on this topic.

4.1. **Stationary states for the single species model.** We start investigating the stationary states for equation (13) under the choice $P = 1$, as in (42). We assume that the limit value m_1^∞ exist, than the equation for the stationary states writes as

$$\partial_w \left(\frac{\lambda^2}{2}D^2(w)\partial_w\phi(u(w)) - (m_1^\infty - \sigma w)\, u(w) \right) = 0. \tag{44}$$

Moreover, to simplify the following study, we introduce the quantity

$$\mathfrak{D}_\alpha(w, m_1^\infty) = \int \frac{(m_1^\infty - \sigma w)}{D^2(w)} d \tag{45}$$

and we distinguish between the two sub-cases of linear and nonlinear diffusion.

4.1.1. *Stationary states in the case of linear diffusion.* We consider here the case of linear diffusion $\phi(u) = u$, thus (44) reduces to

$$\frac{\lambda^2}{2}D^2(w)\partial_w u(w) = (m_1^\infty - \sigma w)u(w).$$

By standard separation of variable, we obtain

$$\frac{\lambda^2}{2}\log(u) = \mathfrak{D}_\alpha(w, m_1^\infty),$$

which, in turn, gives

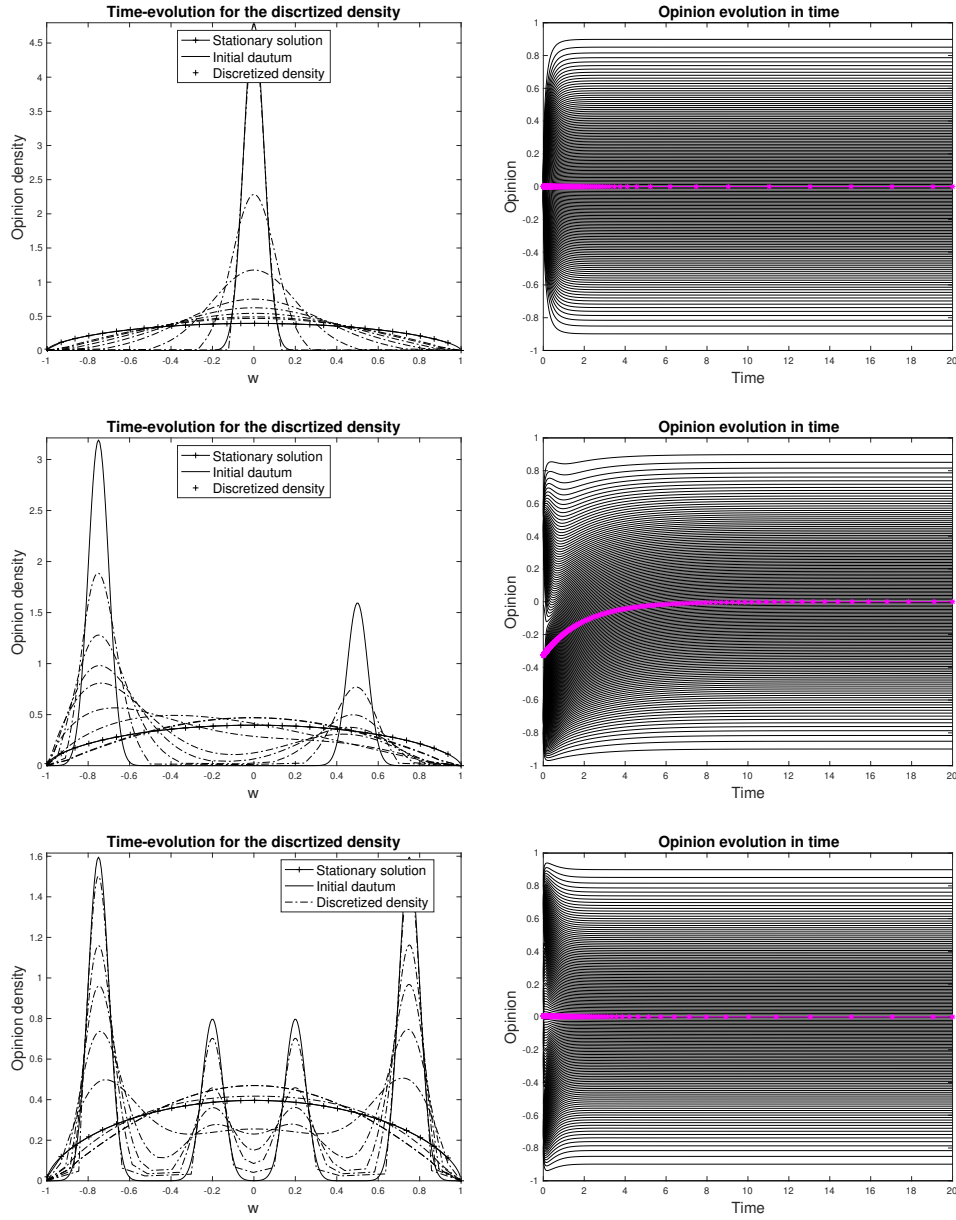$$u_\infty(w) = C_\infty \exp\left(\frac{2}{\lambda^2}\mathfrak{D}_\alpha(w, m_1^\infty) \right), \tag{46}$$

FIGURE 2. Convergence for different initial data to the stationary state (47). The left column shows the evolution in time for the reconstructed density, where the initial data are (48), (49) and (50) respectively, while the right column shows the opinions evolution in time. The (magenta) stars-line represent the mean opinion $m_1(t)$ in each case. Note that, in the second simulation, m_1 is not conserved in time but still converges to zero.
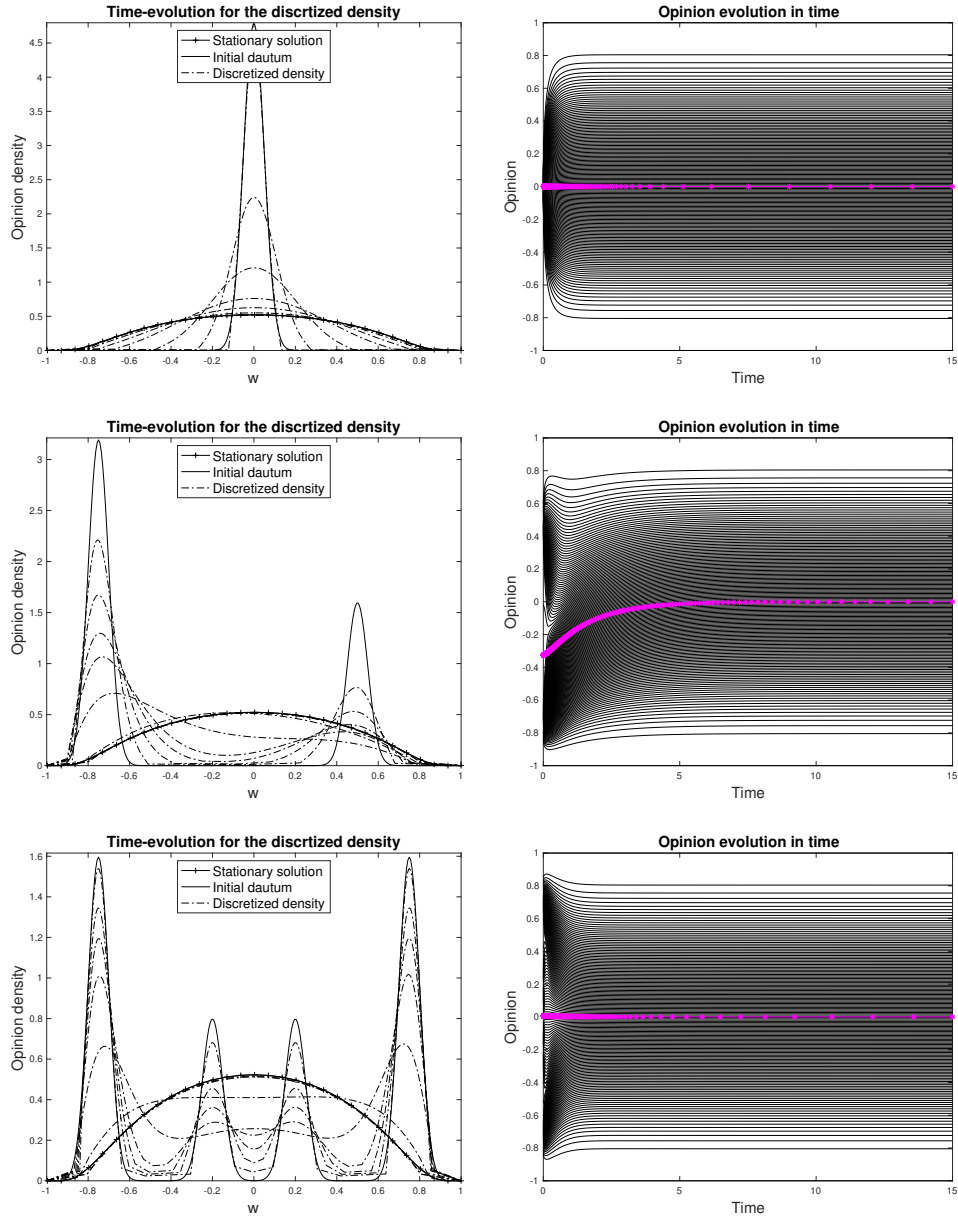
FIGURE 3. Convergence for different initial data to the stationary state (51), where the initial data are (48), (49) and (50) respectively. Note that also in this case the mean opinion m_1 (star magenta line in the right column) converges to zero asymptotically.
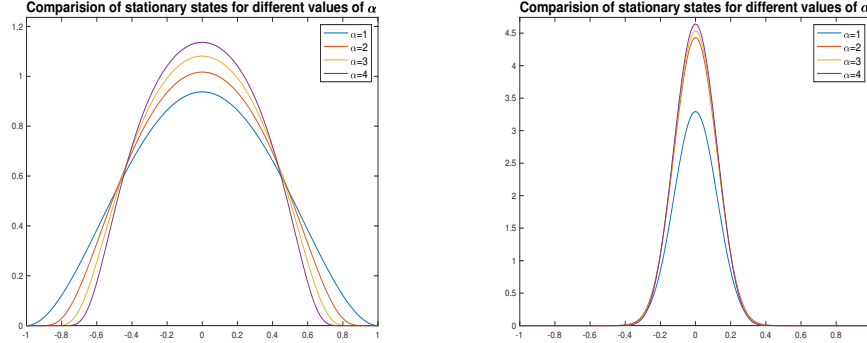
FIGURE 4. Comparison of different stationary states (46) for different values of α. On the left the diffusion coefficient $\lambda^2 = 0.5$, on the right $\lambda^2 = 0.03$.

where C_∞ is the normalisation constant such that $\|u_\infty\|_{L^1} = \sigma$. In order to get a better understanding of the solutions, we consider below some explicit cases for the function D.

- Let us start by taking D as in (14) for $\alpha = 1$

$$D(w) = (1 - w^2)^{\frac{1}{2}}.$$

In this case, the integral in (45) becomes

$$\mathfrak{D}_\alpha(w, m_1^\infty) = \frac{m_1^\infty}{2} \log\left(\frac{1+w}{1-w}\right) + \frac{\sigma}{2} \log(1 - w^2),$$

and

$$u_\infty(w) = C_\infty (1 + w)^{\frac{m_1^\infty + \sigma}{\lambda^2}} (1 - w)^{\frac{\sigma - m_1^\infty}{\lambda^2}}. \tag{47}$$

The well-posedness of the stationary state (47) in I, is guaranteed provided that

$$m_1^\infty + \sigma > 0, \text{ and } \sigma - m_1^\infty > 0,$$

but, according to the consideration at the beginning of this section, we have that

$$\frac{d}{dt} m_1(t) = -\int_{-1}^{1} \left(1 - w^2\right) \partial_w u \, dw = -2m_1(t).$$

In particular, $m_1(t) \to m_1^\infty = 0$ as $t \to +\infty$, and (47) reduces to

$$u_\infty(w) = C_\infty (1 + w)^{\frac{\sigma}{\lambda^2}} (1 - w)^{\frac{\sigma}{\lambda^2}}.$$

In Figure 2, we provide some numerical evidence of the converge to the above stationary state from different initial data. We set the mass of opinion $\sigma = 0.6$ and we choose

1. a single spike centred in the origin, miming a population with opinions symmetrically distributed around the average opinion $w = 0$

$$\bar{u}(w) = \frac{\sigma}{\sqrt{2\pi(0.05)^2}} e^{\frac{-w^2}{2(0.05)^2}}, \tag{48}$$

2. two spikes with different weights $\sigma_1 = 0.4 \ \sigma_2 = 0.2$, non symmetric around the origin,

$$\bar{u}(w) = \frac{\sigma_1}{\sqrt{2\pi(0.05)^2}} e^{\frac{-(w+0.75)^2}{2(0.05)^2}} + \frac{\sigma_2}{\sqrt{2\pi(0.05)^2}} e^{\frac{-(w-0.5)^2}{2(0.05)^2}}, \tag{49}$$

3. a combination of four spikes symmetrically distributed around the origin, with weight $\sigma_1 = 0.2 \ \sigma_2 = 0.1$

$$\begin{aligned} \bar{u}(w) = &\frac{\sigma_1}{\sqrt{2\pi(0.05)^2}} e^{\frac{-(w+0.75)^2}{2(0.05)^2}} + \frac{\sigma_2}{\sqrt{2\pi(0.05)^2}} e^{\frac{-(w+0.2)^2}{2(0.05)^2}} \\ &+ \frac{\sigma_2}{\sqrt{2\pi(0.05)^2}} e^{\frac{-(w-0.2)^2}{2(0.05)^2}} + \frac{\sigma_1}{\sqrt{2\pi(0.05)^2}} e^{\frac{-(w-0.75)^2}{2(0.05)^2}}. \end{aligned} \tag{50}$$

All the simulations are performed using the deterministic particle approximation introduced in the previous sections as a numerical scheme for (13). More precisely, given one of the above initial data, we construct an initial distribution of opinions according to (16), then we solve the ODEs system (17) and we reconstruct the density as in (28).

- We now consider the mobility function as in (14) for $\alpha = 2$

$$D(w) = (1 - w^2).$$

In this case, the stationary solution reduces to

$$u_\infty(w) = C_\infty (1+w)^{\frac{m_1^\infty}{2\lambda^2}} (1-w)^{\frac{-m_1^\infty}{2\lambda^2}} e^{\frac{m_1^\infty w - \sigma}{\lambda^2(1-w^2)}}, \tag{51}$$

where C_∞ is the usual normalisation constant. The well-posedness of the steady state (51) is not guaranteed *a priori* and it seem to be strongly dependent on the relation between m_1^∞ and σ. As mentioned above, the evolution equation for the second moment is not closed, here it is corresponds to

$$\frac{d}{dt} m_1(t) = -2\lambda^2 m_1(t) + 2\lambda^2 m_3(t),$$

where m_3 denotes the third order moment of u. Let us observe, that for any $k \geq 2$, the $k-$th moment evolves according to

$$\begin{aligned} \frac{d}{dt} m_k(t) = &\frac{\lambda^2}{2}(k-1)m_{k-2}(t) + m_1(t)m_{k-1}(t) \\ &- k\left[\lambda^2(k+1) + \sigma\right] m_k(t) + \frac{\lambda^2}{2}(k+3)m_{k+2}(t), \end{aligned}$$

that is a dynamical system with infinite dimension, whose study requires deeper investigations which exceed the scope of this paper. However, by introducing the variance function

$$\text{Var}\,[u]\,(t) = m_2(t) - (m_1(t))^2,$$

it is easy to see that $(m_1(t))^2 \leq 1 - \text{Var}\,[u]\,(t)$, with $\text{Var}\,[u]\,(t) \in [0, 1]$. Since also $|m_3| \leq 1$, this suggests the decay in time of the mean opinion. This is supported by the numerical results in Figure 3.

- The argument for generic values of α is more involved, but does not present further complications. For this reason, we decided to omit the proof in the present paper. Nonetheless, we highlight that the parameter α plays an important modelling role: as shown in Figure 4 (left), the support of the stationary state shrinks when alpha increases, thus providing an higher *consensus*

around the limiting mean opinion (in the previous examples the mean opinion is $w = 0$). A similar effect can be reached while decreasing the diffusion coefficient λ. Indeed, for smaller values of λ, the contribution of the reaction part is stronger and the stationary states are more concentrated, see Figure 4 (right).

4.1.2. *Nonlinear diffusion.* Here we discuss the case of nonlinear diffusion of a *porous medium* type, namely

$$\phi(u) = \frac{u^{\gamma}}{\gamma}, \quad \gamma > 1.$$

The stationary states, then, correspond to the solutions of

$$\partial_w \left(\frac{\lambda^2}{2} D^2(w) u \partial_w \frac{u(w)^{\gamma-1}}{\gamma - 1} - (m_1^{\infty} - \sigma w) u(w) \right) = 0, \tag{52}$$

which can be rewritten as

$$\partial_w \left(\frac{\lambda^2}{2} \frac{u(w)^{\gamma-1}}{\gamma - 1} - \mathfrak{D}_\alpha(w, m_1^{\infty}) \right) = 0.$$

Since the physical solutions are non negative, we deduce

$$u_\infty(w) = \left[\frac{2(\gamma - 1)}{\lambda^2} \left(C + \mathfrak{D}_\alpha(w, m_1^{\infty}) \right)_+ \right]^{\frac{1}{\gamma-1}}, \tag{53}$$

where C is a suitable normalization constant. In Figure 5 we show the convergence towards the stationary state in the case $D(w) = (1 - w^2)^{(1/2)}$ and $\gamma = 2$, with the same initial data and diffusion coefficient of the previous examples. Observe, that the nonlinear diffusion induces a stronger *consensus* around the compromise value $w = 0$. In Figure 6 we plot the stationary states corresponding to $\alpha = 1$ and $\alpha = 2$.

4.2. **Simulations with many species.** We now explore the large-time behaviour for the many species case. The analytical study in this case became more complicated, since it requires to check the solvability of a system of coupled equations in the form (44). We here focus our attention on the numerical comparison among the large time behaviours of systems (11) and (12).

In Figure 7, we show the opinions dynamics in the Follower-Leader system (11), that we rewrite below for the reader convenience,

$$\partial_t f = \partial_w \left(\frac{\lambda_f^2}{2} D_f^2(w) \partial_w \phi_f(f) - f \left(\mathcal{P}_{ff}[f] + \mathcal{P}_{fl}[l] + \mathcal{P}_{fr}[r] \right) \right)$$

$$\partial_t l = \partial_w \left(\frac{\lambda_l^2}{2} D_l^2(w) \partial_w \phi_l(l) - l \left(\mathcal{P}_{ll}[l] + \mathcal{P}_{lr}[r] \right) \right),$$

$$\partial_t r = \partial_w \left(\frac{\lambda_r^2}{2} D_r^2(w) \partial_w \phi_r(r) - r \left(\mathcal{P}_{rr}[r] + \mathcal{P}_{rl}[l] \right) \right).$$

In particular, we assume that there are two groups of Leaders with equal mass and symmetric centres of mass. Moreover, we assume that the initial followers' opinions are symmetrically distributed, see Figure 7 top-left. In this simulation, we fix the diffusion coefficients $\frac{\lambda_u^2}{2} = 0.03$ and the mobility functions $D_u^2(w) = (1 - w^2)$ for $u \in \{f, l, r\}$, taking $\sigma_f = 1$ and $\sigma_l = \sigma_r = 0.6$. The compromise functions are the following: for the interactions among the same species we set
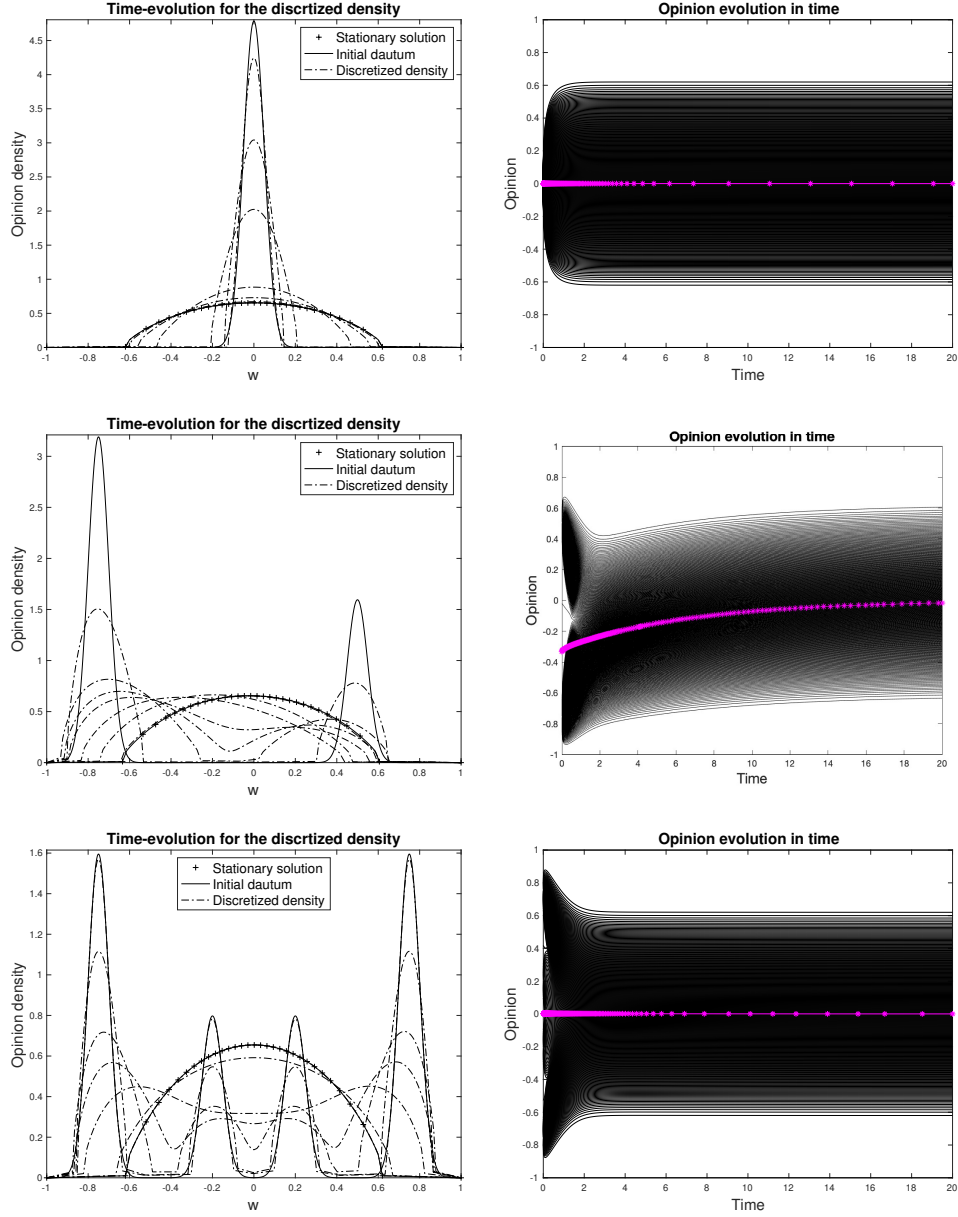
$$P_{ff}(w, v) = P_{ll}(w, v) = P_{rr}(w, v) = 1,$$

FIGURE 5. Convergence for the initial data (48), (49) and (50) to the stationary state (53), with $\alpha = 1$. Note that the nonlinearity in the diffusion produces stationary solutions with supports that are smaller than the one we have seen in the linear diffusion case.
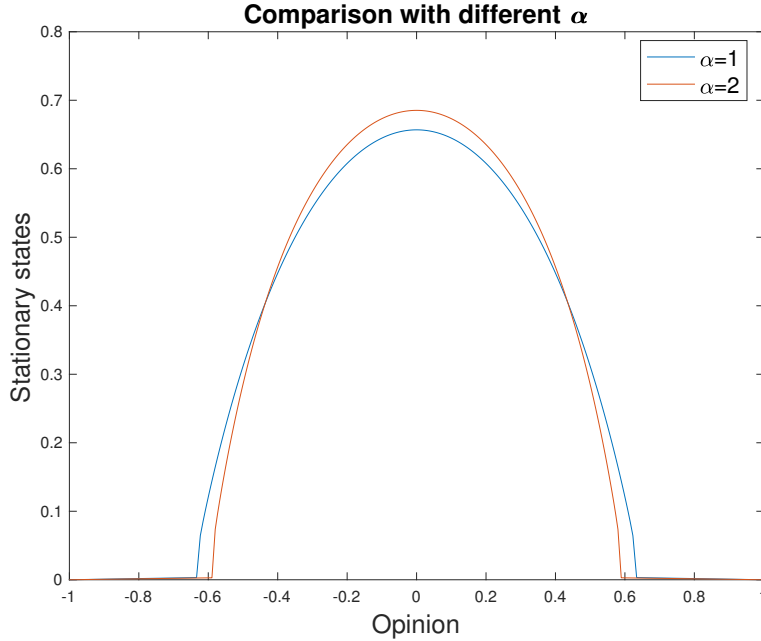
FIGURE 6. Stationary states in (53) for different values of α.

while the follower-leaders interactions are given by

$$P_{fr}(w,v) = P_{fl}(w,v) = 1 - w^2.$$

In this way, we are modelling the situation where a follower agent with an extreme opinion is less likely to revise its own opinion. According to this, we set

$$P_{lr}(w,v) = P_{rl}(w,v) = 0.001(1 - w^2).$$

As expected, the opinions evolution is symmetric and solution converges to a compromise (the opinion $w = 0$).

In Figure 8 we consider the same setting as before, but this time one leader group is *stronger* than the other, namely $\sigma_l = 0.6$ and $\sigma_r = 0.2$. This difference strongly effects the evolution of the followers in the short period, see Figure 8 top-right, while for long time we still observe convergence to the compromise value.

We finally consider system (12), where a small species of *fake* agents is present. The agents of this species are called *trolls*, they are indistinguishable from the followers and they interact with only one group of leaders. In this case, we assume that the fake species has mass $\sigma_q = 0.3$ and that it interacts only with leaders sharing the right opinion ($w = +1$). The fact that trolls are not perceived by the followers is modelled by

$$P_{ft}(w,v) = P_{ff}(w,v).$$

On the other hand, trolls cannot diffuse their opinion, so, according to (12), the evolution of their opinion is only driven by the compromise part. In this example we set

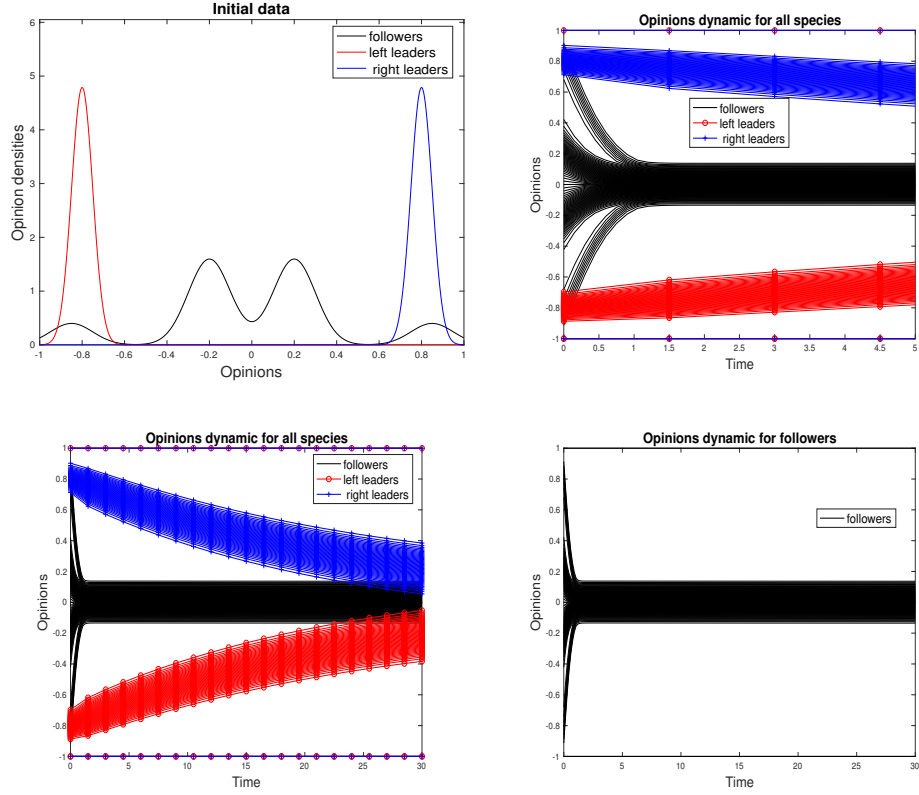$$P_{ft}(w,v) = 1, \qquad P_{tr}(w,v) = \left(1 - \frac{1}{4}(w - v)^2\right).$$

FIGURE 7. Opinion dynamics in presence of *equally-strong* leaders. Top-left initial data for followers(black), left leaders (red) and right leaders (blue). Top-right and bottom-left opinions evolutions for all the groups in short and long term-respectively. Bottom-right opinions evolutions for the followers.

In Figure 9 we show how the presence of trolls affects the behaviour of the population both in the short and the long term.

REFERENCES

[1] G. Albi, L. Pareschi and M. Zanella, Boltzmann-type control of opinion consensus through leaders, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **372** (2014), 20140138.

[2] G. Albi, L. Pareschi and M. Zanella, *On the Optimal Control of Opinion Dynamics on Evolving Networks*, vol. 494, 58–67, Springer, Cham, 2016.

[3] G. Albi, L. Pareschi and M. Zanella, Opinion dynamics over complex networks: Kinetic modelling and numerical methods, *Kinetic and Related Models*, **10** (2017), 1.

[4] G. Albi, P. Pareschi, G. Toscani and M. Zanella, *Recent advances in opinion modeling: control and social influence*, 49–98, Birkhuser-Springer, 2017.

[5] G. Aletti, G. Naldi and G. Toscani, First-order continuous models of opinion formation, *SIAM J. Appl. Math.*, **67** (2007), 837–853.
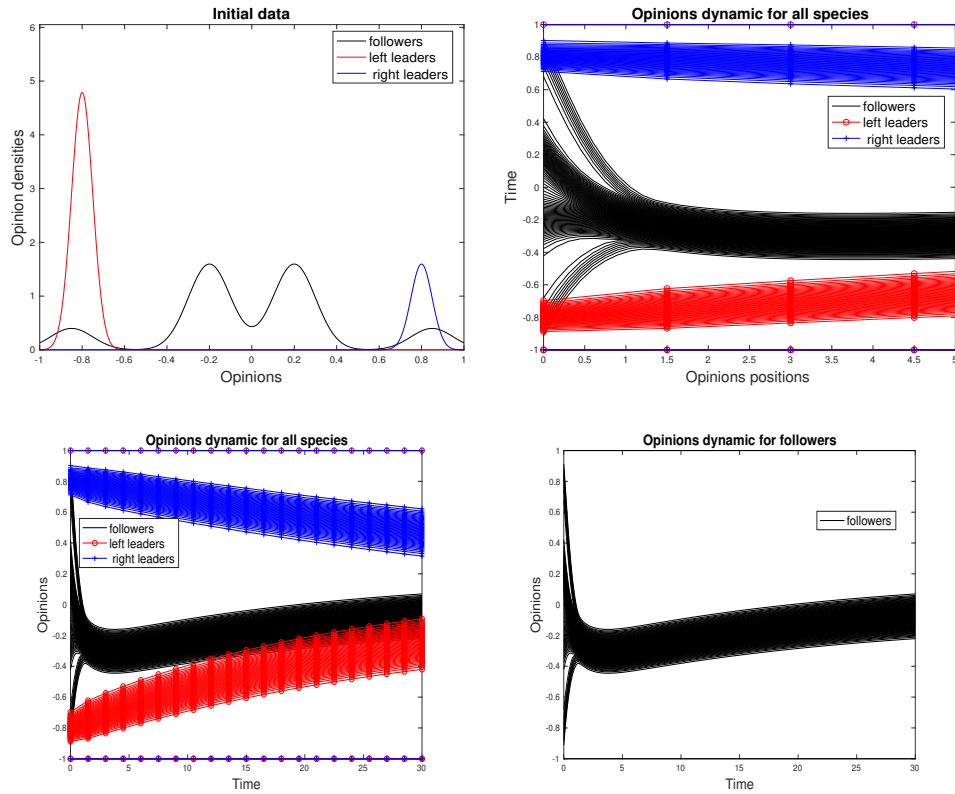
FIGURE 8. Opinion dynamics in presence of one *strong* group of leaders and one *weak* group of leaders. Top-left initial data for followers(black), left leaders (red) and right leaders (blue). Top-right and bottom-left opinions evolutions for all the groups in short and long term-respectively. Bottom-right opinions evolutions for the followers.

[6] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd edition, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.

[7] N. Ansini and S. Fagioli, Nonlinear diffusion equations with degenerate fast-decay mobility by coordinate transformation, *to appear on Communications in Mathematical Sciences*, 23.

[8] N. Bellomo, G. Ajmone Marsan and A. Tosin, *Complex Systems and Society. Modeling and Simulation. SpringerBriefs in Mathematics*, Springer, 2013.

[9] E. Ben-Naim, Opinion dynamics: rise and fall of political parties, *Europhysics Letters*, **69** (2005), 671.

[10] S. Biswas and P. Sen, Critical noise can make the minority candidate win: The u.s. presidential election cases, *Phys. Rev. E*, **96** (2017), 032303.

[11] D. Borra and T. Lorenzi, A hybrid model for opinion formation, *Zeitschrift für angewandte Mathematik und Physik*, **64** (2013), 419–437.

[12] L. Boudin and F. Salvarani, Opinion dynamics: Kinetic modelling with mass media, application to the scottish independence referendum, *Physica A: Statistical Mechanics and its Applications*, **444** (2016), 448 – 457.
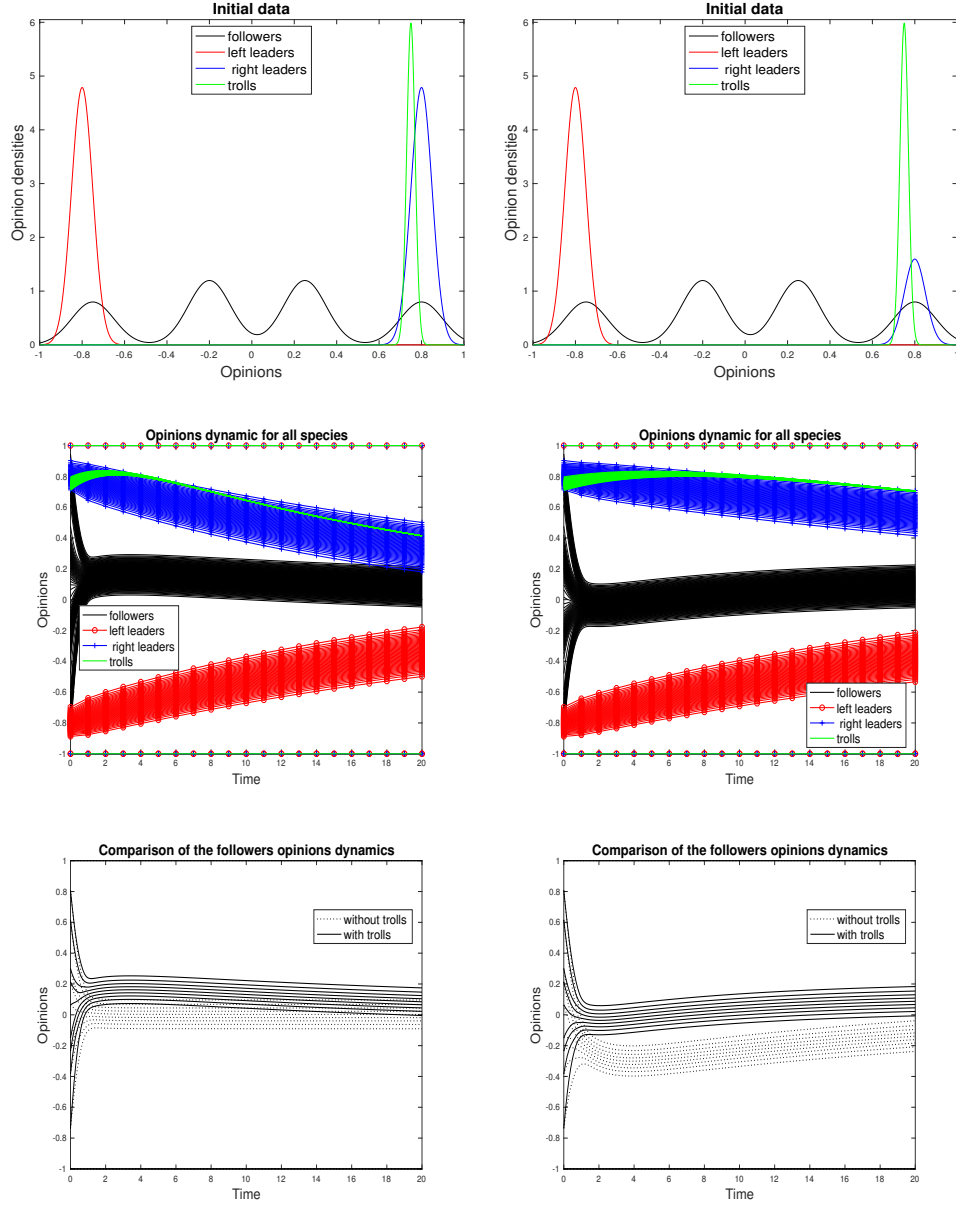
FIGURE 9. Evolution of system (12). The left column concerns the case of two equally strong groups of leaders, the right column instead describes the situation where the right leader is *weaker*. Trolls are plotted in green and are associated to the right leaders. Top: initial data. Centre: opinions dynamic for all species. Bottom: comparison between the followers paths with or without trolls.

[13] G. R. e. a. Boynton, The reach of politics via twitter? can that be real?, *Open Journal of Political Science*, **3** (2013), 91–97.

[14] J. A. Carrillo and G. Toscani, Wasserstein metric and large–time asymptotics of nonlinear diffusion equations, *New Trends in Mathematical Physics, (In Honour of the Salvatore Rionero 70th Birthday)*, 234–244.

[15] C. Castellano, S. Fortunato and V. Loreto, Statistical physics of social dynamics., *Review of Modern Physics*, **81** (2009), 591–646.

[16] S. Cresci, M. N. La Polla and M. Tesconi, *Il fenomeno dei Fake Follower in Twitter*, 151–191, Pisa University Press, Pisa, 2017.

[17] E. De Cristofaro, A. Friedman, G. Jourjon, M. A. Kaafar and M. Z. Shafiq, Paying for likes?: Understanding facebook like fraud using honeypots., *2014 ACM 14th Internet Measurement Conference (IMC)*, 129–136.

[18] G. Deffuant, F. Amblard, G. Weisbuch and T. Faure, How can extremism prevail? a study based on the relative agreement interaction model, *Journal of Artificial Societies and Social Simulation*, **5**.

[19] M. Di Francesco, S. Fagioli and E. Radici, Deterministic particle approximation for nonlocal transport equations with nonlinear mobility, *Journal of Differential Equations*, **266** (2019), 2830–2868.

[20] M. Di Francesco, S. Fagioli and M. D. Rosini, Deterministic particle approximation of scalar conservation laws, *Bollettino dell'Unione Matematica Italiana*, **10** (2017), 487–501.

[21] M. Di Francesco and G. Stivaletta, Convergence of the follow-the-leader scheme for scalar conservation laws with space dependent flux, *Discrete & Continuous Dynamical Systems - A*, **40** (2020), 233–266.

[22] 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. Lond. Ser. A Math. Phys. Eng. Sci.*, **465** (2009), 3687–3708.

[23] B. During and M. T. Wolfram, Opinion dynamics: inhomogeneous boltzmann-type equations modelling opinion leadership and political segregation, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **471** (2015), 20150345.

[24] S. Fagioli and E. Radici, Solutions to aggregationdiffusion equations with nonlinear mobility constructed via a deterministic particle approximation, *Math. Mod. and Meth. in App. Sci.*, **28** (2018), 1801–1829.

[25] G. Furioli, A. Pulvirenti, E. Terraneo and G. Toscani, Wright-fisher-type equations for opinion formation, large time behavior and weighted logarithmic-sobolev inequalities, *Annales de l'Institut Henri Poincar C, Analyse non linaire*, **36** (2019), 2065 – 2082.

[26] S. Galam, *Sociophysics: a physicists modeling of psycho-political phenomena (understanding complex systems)*, Springer, 2012.

[27] L. Gosse and G. Toscani, Identification of asymptotic decay to self-similarity for one-dimensional filtration equations, *SIAM J. Numer. Anal.*, **43** (2006), 2590–2606 (electronic).

[28] R. Hegselmann and U. Krause, Opinion dynamics and bounded confidence, models, analysis and simulation, *Journal of Artificial Societies and Social Simulation*, **5**.

[29] A. Klein, H. Ahlf and V. Sharma, Social activity and structural centrality in online social networks, *Telematics and Informatics*, **32** (2015), 321–332.

[30] A. D. I. Kramer, J. E. Guillory and J. T. Hancock, Experimental evidence of massive scale emotional contagion through social networks, *Proceedings of the National Academy of Sciences*, **11** (2014), 8788–8789.

[31] H. Lavenant and B. Maury, Opinion propagation on social networks: a mathematical standpoint, *preprint*, 53.

[32] P. F. Lazarsfeld, B. R. Berelson and H. Gaudet, *The people's choice: how the voter makes up his mind in a presidential campaign*, Duell, Sloan & Pierce, New York, 1944.

[33] S. Motsch and E. Tadmor, Heterophilious dynamics enhances consensus, *SIAM Review*, **56** (2014), 577–621.

[34] G. Naldi, L. Pareschi and G. Toscani, *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, Birkhäuser, Boston, 2010.

[35] L. Pareschi and G. Toscani, *Interacting Multiagent Systems. Kinetic Equations and Monte Carlo Methods.*, Oxford University Press, 2013.

[36] L. Pareschi and G. Toscani, Wealth distribution and collective knowledge: a boltzmann approach, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **372** (2014), 20130396.

[37] R. Rossi and G. Savaré, Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, **2** (2003), 395–431.

[38] G. Russo, Deterministic diffusion of particles, *Comm. on Pure and Applied Mathematics*, **43** (1990), 697–733.

[39] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, vol. 86 of Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser Verlag, Basel, 2015.

[40] F. Slanina and H. Lavička, Analytical results for the sznajd model of opinion formation., *Eur.Phys. J. B*, **35** (2003), 279?288.

[41] S. H. Strogatz, Exploring complex networks, *Nature*, **410** (2001), 268–276.

[42] K. Sznajd-Weron and J. Sznajd, Opinion evolution in closed community., *Int. J. Mod. Phys. C*, **11** (2000), 1157?1165.

[43] G. Toscani, Kinetic models of opinion formation, *Comm. Math. Sci.*, **4** (2006), 481–496.

[44] G. Toscani, C. Brugna and S. Demichelis, Kinetic models for the trading of goods, *Journal of Statistical Physics*, **151** (2013), 549–566.

[45] G. Toscani, A. Tosin and M. Zanella, Opinion modeling on social media and marketing aspects, *Phys. Rev. E*, **98** (2018), 022315.

[46] C. Villani, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.

[47] S. Yardi, D. Romero and G. Schoenebeck, Detecting spam in a twitter network, *First Monday*, **15**.

E-mail address: simone.fagioli@univaq.it

E-mail address: emanuela.radici@epfl.ch