European Journal of Applied Mathematics (2024), **35**, pp. 748–775
doi:10.1017/S0956792524000093

CAMBRIDGE UNIVERSITY PRESS

PAPER

Opinion formation on evolving network: the DPA method applied to a nonlocal cross-diffusion PDE-ODE system

Simone Fagioli and Gianluca Favre

Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Università degli Studi dell'Aquila, Coppito, L'Aquila, Italy
Corresponding author: Simone Fagioli; Email: simone.fagioli@univaq.it

Received: 16 June 2023; **Revised:** 12 December 2023; **Accepted:** 22 February 2024; **First published online** 21 May 2024

Keywords: deterministic particle approximation; opinion formation on networks; cross-diffusion; nonlocal transport equations

2020 Mathematics Subject Classification: 35Q91, 35Q70 (Primary); 35A35, 91D30, 65M75, 35R09 (Secondary)

Abstract
We study a system of nonlocal aggregation cross-diffusion PDEs that describe the evolution of opinion densities on a network. The PDEs are coupled with a system of ODEs that describe the time evolution of the agents on the network. Firstly, we apply the Deterministic Particle Approximation (DPA) method to the aforementioned system in order to prove the existence of solutions under suitable assumptions on the interactions between agents. Later on, we present an explicit model for opinion formation on an evolving network. The opinions evolve based on both the distance between the agents on the network and the 'attitude areas', which depend on the distance between the agents' opinions. The position of the agents on the network evolves based on the distance between the agents' opinions. The goal is to study radicalisation, polarisation and fragmentation of the population while changing its open-mindedness and the radius of interaction.

1. Introduction

The study of social phenomena through mathematical modelling has gained significant attention in the scientific community, especially in recent decades [6, 17, 28, 35, 37, 41, 50]. The exchange of information on these platforms has sparked research in understanding how social interactions shape the process of opinion formation [3, 7, 10, 30, 33, 48, 51, 57].

In social interactions, the relationships between individuals are often structured as networks that co-evolve with the individuals themselves [49]. A prominent example of this is the formation of opinions or norms within social networks, where interactions only occur between connected agents. However, the network connections are dynamic, and this change influences the states of the individuals. For example, opinions can be influenced by connections, such as followers reacting to posts, while individuals tend to follow others with closer opinions.

The network structure of social interactions plays a vital role and is commonly represented using random networks. However, there are two natural levels in examining opinion formation processes on network: the microscopic and macroscopic scales. The microscopic models have been employed to simulate phenomena such as opinion formation, knowledge networks, social norm formation and biological transport networks [1, 2, 8, 31, 39, 54]. In considering processes with a huge number of agents, a natural question arises in considering a limit procedure between the two scales. However, the specific details of the network structure can be lost, and only few general characteristics are incorporated into the models [19, 20].

From a mathematical standpoint, it is natural to apply methods from statistical physics or kinetic theory to bridge the gap between microscopic interactions and macroscopic models, see [12, 52].

This approach involves formulating partial differential equations for distributions and using well-established asymptotic methods to simplify the equations and analyse pattern formation [11, 18]. These mathematical approaches have successfully explained macroscopic distributions in socio-economic interactions and various aspects of opinion formation and polarisation [3, 15, 36, 42]. Using these tools has significantly contributed to understanding the emergence of macroscopic behaviour from microscopic interactions in a wide range of social phenomena.

1.1. Modelling motivation: the social context

During the last two decades, the diffusion of smartphones and the increasing use of social networks have changed how people interact and form their beliefs. There are two main aspects that have been drastically disrupted: the **number of connections** and the **frequency of interactions**.

Due to the *hyperconnectivity* of the globalised world, each individual can get in contact with a wide range of opinions. Delocalising the place of interaction from physical space to the digital realm has destroyed the local cultural bias of interactions. This implies that each individual can come into contact with cultures and ideas they do not know and cannot deeply understand. This aspect has resulted in a change in the epistemic processes, specifically altering the dynamics that govern the formation of beliefs when individuals are exposed to new inputs (news, visual art, songs, posts, tweets, chats, etc.).

Moreover, the increasing amount of inputs and their high frequency create a physical upper bound on the processing capacity of the human brain. An individual walking through a mall cannot process all the inputs coming from screens, speakers, billboards, smartphones and so on. The same situation occurs while scrolling through social networks or digital social media. As a result, there is a need to filter the inputs, both on the physical side through the network and on the rational side, by selectively processing only a few of them and disregarding others at a peripheral level of thinking.

All these aspects are well understood by scientists from social sciences and social epistemologists (a non-exhaustive list includes references such as [5, 9, 32, 38, 44, 56]). However, it is challenging to fit them into a unified mathematical description. In this paper, we propose an approach that mainly focuses on two tools: attitude areas and the Euclidean network. In Section 4, we introduce and simulate a new model to investigate the evolution of the network and opinion distributions for agents interacting on social networks and social media.

1.2. Modelling tools

We consider a generalisation of the model introduced by Burger in [13, 14]. In these works, the author derives a kinetic description of an opinion formation process on networks. Inspired by these studies, we associate two variables with each individual: the *network position* and the *opinion distribution*.

Opinion Distribution

The description of an individual's opinion during the opinion formation process can be represented by a distribution rather than a single value. The motivation lies in the fact that an individual may be indecisive, lacking a clear and definite opinion. Instead, their thoughts span multiple possibilities. An example can be related to political elections, where voters loyal to smaller parties often face the dilemma of the *strategic vote*. This term refers to the choice of voting for a party other than their own, typically one similar but with a higher chance of surpassing the threshold percentage. In such a case, the individual's opinion is described by a distribution with two peaks, one at the position of their affiliated party and another at the position associated with the *strategic vote*. Now, let's consider a finite population $\mathscr{M} = \{1, \ldots, M \in \mathbb{N}\}$ with M individuals. To each individual, we associate a distribution $\rho^{i} : (x, t) \in \Omega \times [0, \infty) \to [0, \infty)$, where $i \in \mathscr{M}$ and $\Omega \in \mathbb{R}$ is bounded.

The Social Strength

Every individual possesses a distinct social strength denoted as $\sigma^{\imath}$. This corresponds to the integral of their opinion, i.e. $\sigma^{\imath} = \int_{\Omega} \rho^{\imath}(y)\,dy$. Mathematically, this quantity coincides with mass. However, from a modelling perspective, it represents an individual's inclination to uphold their own opinion. In the scenario of an attractive binary interaction between two agents with different social strengths, the average of their opinions tends to align more closely with the original opinion of the agent with the higher social strength. Conversely, in the case of a repulsive interaction, the agent with the lower social strength will deviate further from the average of the original opinions.

Euclidean Network

The network position is described by the function $a^{\imath} : t \in [0, \infty) \to \mathbb{R}^2$ with $\imath \in \mathscr{M}$. In this way, the network is based on a Euclidean space and its natural distance. Instead of considering interactions mediated by weights between individuals $\imath$ and $\jmath$, we use their distance in $\mathbb{R}^2$. This has a direct consequence on the modelling interpretation of interactions. Specifically, cases are excluded where two agents are close, but a third agent is close to one but not the other. If we take three individuals $\imath$, $\jmath$ and $\mathscr{k}$ in $\mathscr{M}$, then if $\imath$ is close to $\jmath$, it implies that $\mathscr{k}$ cannot be distant from one but close to the other.

The PDE-ODE System

The model under consideration describes the temporal evolution of the opinion distribution and the network position of each agent. Given the vectors of positions $\mathbf{a}$ and opinion distributions $\boldsymbol{\rho}$, their evolution is described by the following system:

$$\partial_t \rho^{\imath}(t,x) = \partial_x\big(\beta^{\imath}(\boldsymbol{\rho},\mathbf{a};x)\partial_x\rho^{\imath}(t,x)\big) - \partial_x\left(\rho^{\imath}(t,x)\theta^{\imath}(\boldsymbol{\rho},\mathbf{a};x)\right), \tag{1.1a}$$

$$\partial_t a^{\imath}(t) = \sum_{\jmath\in\mathscr{M}} \mathbf{V}(\mu_{\rho^{\imath}}(t), \mu_{\rho^{\jmath}}(t), a^{\imath\jmath}), \tag{1.1b}$$

where $\beta^{\imath}$ and $\theta^{\imath}$ are defined by:

$$\beta^{\imath}(\boldsymbol{\rho},\mathbf{a};x) = \sum_{\jmath=1}^{M}\int_{\Omega}\mathbf{A}^{\imath\jmath}(x,y,a^{\imath\jmath})\rho^{\jmath}(y,t)\,dy, \tag{1.2}$$

$$\theta^{\imath}(\boldsymbol{\rho},\mathbf{a};x) = \sum_{\jmath=1}^{M}\int_{\Omega}\mathbf{K}^{\imath\jmath}(x,y,a^{\imath\jmath})\rho^{\jmath}(y,t)\,dy, \tag{1.3}$$

where $\mathbf{A}^{\imath\jmath}$ and $\mathbf{K}^{\imath\jmath}$ are interaction potentials, to be defined later. In (1.1), $\mu_{\rho^{\imath}}$ indicates the mean of $\rho^{\imath}$. The term $a^{\imath\jmath}$ represents the Euclidean distance between agents $\imath$ and $\jmath$ in the network space, i.e. $a^{\imath\jmath} = | a^{\imath} - a^{\jmath} |$, coinciding with $a^{\jmath\imath}$ by construction. In (1.1) we are not imposing boundary conditions. However, we are interested in the case where individual opinions do not exit the considered interval. We will take care of the proper boundary conditions to be imposed on (1.1) in the following.

The Interaction Potentials: Diffusion Mobility and Transport

The operators $\beta^{\imath}$ and $\theta^{\imath}$ appearing in the evolution of the opinion distribution are related to the terms of diffusion mobility and transport, respectively. The resulting equation combines these terms in an aggregation-diffusion manner. Both operators have a nonlocal nature. The diffusion mobility term depends on the vector $\boldsymbol{\rho}$ and introduces a cross-diffusion mechanism affecting the system's opinion evolution. The terms $\mathbf{A}^{\imath\jmath}$ and $\mathbf{K}^{\imath\jmath}$ represent the interaction kernels, applicable to both social networks and social media descriptions. Importantly, they depend on both the distance between opinions and the distance between agents in the network space.

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

In Section 4, we define the kernels explicitly and we give the interpretation in relation to social network's and social media's structures.

Furthermore, if $\beta^{\imath} \equiv 1$ in (1.1) for all $\imath$, the diffusion terms become a one-dimensional Laplacian. The natural question is whether we can consider more general types of nonlinear (possibly degenerate) diffusion. In what follows, we consider the general diffusion terms below

$$\partial_x\big(\beta^{\imath}(\boldsymbol{\rho}, \mathbf{a}; x)\partial_x\Phi^{\imath}(\rho^{\imath})\big)$$

Here, $\Phi^{\imath}$ represents nonlinear functions, refer to hypothesis (Diff) below for a precise statement. The minimal assumptions on $\Phi^{\imath}$ include Lipschitz regularity and nondecreasing monotonicity. This class of diffusion includes the classical porous medium equations (one-point degeneracy), the two-phase reservoir flow equations (two-point degeneracy) and the so-called strongly degenerate diffusion equations where $\Phi^{\imath}(s) = 0$ for $s \in [s_1, s_2]$.

1.3. The goal of the model

This model aims to describe common and general phenomena emerging in the last two decades due to the increasingly widespread use of social networks and social media. Specifically, it focuses on the changes in the epistemic process caused by such technology. Through the DPA scheme and numerical simulations, we aim to observe the segregation of the population or network into distinct opinion bubbles. Additionally, maintaining a highly general model based on attraction and repulsion principles, we aim to observe the radicalisation and polarisation of opinions based on network connectivity. In other words, we seek to examine the tendency of the hyperconnected society towards a more polarised distribution of opinions and a network divided into large opinion clusters.

Polarisation, Radicalisation and Segregation

Polarisation is the inclination of an individual or a group of individuals to adopt very extreme opinions, generally in stark opposition to the majority who do not share the same opinion. Another phenomenon often confused with polarisation is radicalisation. Radicalisation involves the tendency to become more entrenched in one's opinion and less prone to change. For example, a highly radicalised person is not subject to polarisation, but what often happens is that polarisation occurs first, followed by radicalisation, leading to the creation of extremist groups. Segregation, on the other hand, is not related to opinion but to the network. In this work, segregation refers to the tendency of the network to develop groups that share the same opinion. However, it is important to note that sharing the same opinion does not imply belonging to the same group since the network can split into multiple opinion balls not connected in terms of interactions.

Observing these phenomena without imposing them from a modelling perspective but maintaining a very general description based on the epistemic process of the individual rather than population behaviour is the applied goal of this work.

1.4. Structure and results

The paper pursues a dual objective. Firstly, we aim to establish the existence of solutions for the system (1.1) within an appropriate functional framework. Secondly, we seek to numerically investigate solutions to the system in order to determine its ability to replicate processes such as polarisation, radicalisation, fragmentation and clustering of the population. Similar inquiries have been recently explored in [40].

To establish the existence of solutions for system (1.1), we draw inspiration from the deterministic particle approximation (DPA) developed for similar equations in [25–27]. The method, along with its various modifications, traces back to seminal works [29, 46]. It demonstrates the convergence of the

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

resulting equation and has been applied in various contexts such as traffic flow [22, 23] and local or nonlocal transport equations [21, 24, 27].

The DPA is then employed for numerical simulations of (1.1). While the numerical scheme can be connected to moving mesh schemes applied in diverse contexts, its limitations should be acknowledged, especially in one-dimensional applications.

The paper is organised as follows. In Section 2, we introduce the rigorous Deterministic Particle Approximation (DPA). Preliminaries on optimal transportation theory are presented along with the main assumptions. The section concludes with the statement of the main result in Theorem 2.2. Section 3 focuses on the proof of the main theorem by providing fundamental a priori estimates, allowing the deduction of convergence of properly reconstructed piecewise constant densities to weak solutions. Finally, in Section 4, we utilise the DPA numerical scheme to simulate an explicit model for opinion formation on an evolving network. The goal is to study radicalisation, polarisation and fragmentation of the population while altering its open-mindedness and the radius of interaction.

2. Rigorous formulation, assumptions and main result

2.1. Deterministic Particle Approximation (DPA)

We begin this section with the rigorous formulation of the particle evolution already sketched in the Introduction. We consider in $\mathbb{R}^d$ a network of M nodes and we locate an agent $a^{\imath} \in \mathbb{R}^d$ with $\imath \in \mathcal{M} = \{1, \ldots, M\}$ in each node. Assume that each agent may have opinion ranging on a compact set $\Omega \subset \mathbb{R}$, without loss of generality we consider $\Omega = [-1, 1]$. To each agent we associate a finite *opinion strength* $\sigma^{\imath}$ and an initial opinion density $\bar{\rho}^{\imath}(x) \in L^1(\Omega)$ such that

$$\sigma^{\imath} = \int_{\Omega} \bar{\rho}^{\imath}(y)\, dy\,, \qquad \forall \imath \in \mathcal{M}\,.$$

Given $N \in \mathbb{N}$, we consider the strength fractions $\sigma_N^{\imath} = \sigma^{\imath}/N$, and we introduce for each $\imath \in \mathcal{M}$ the $\{\bar{x}_k^{\imath}\}$ partition of Ω with $k \in \mathcal{N} = \{0, \ldots, N\}$ given by

$$\begin{aligned} &\bar{x}_0^{\imath} = -1\,, \\ &\bar{x}_k^{\imath} = \inf\left\{x \in \Omega : \int_{\bar{x}_{k-1}}^{x} \bar{\rho}^{\imath}(y)\, dy = \sigma_N^{\imath}\right\}\,, \qquad k \in \mathcal{N} \setminus \{0, N\} \\ &\bar{x}_N^{\imath} = 1\,. \end{aligned} \tag{2.1}$$

Note that $\bar{x}_k^{\imath} < \bar{x}_{k+1}^{\imath}$, for any $\imath \in \mathcal{M}$ and $k \in \mathcal{N} \setminus \{N\}$. This procedure allows to associate with each agent a finite number of time-evolving opinions $x_k^{\imath}(t)$. Assume that initially all the nodes $\bar{a}^{\imath} = a^{\imath}(t = 0)$ are located in a certain smooth and bounded domain $\Lambda \in \mathbb{R}^d$. We then let the nodes evolve in time depending on the distances $a^{\imath\jmath}$ between the agents $a^{\imath}$ and $a^{\jmath}$ and the mean opinion of the agents.

We define the *discrete opinion densities* for the $\imath$-th agent as

$$\rho_k^{\imath}(t) = \frac{\sigma_N^{\imath}}{|I_k^{\imath}(t)|}, \qquad \textit{with} \qquad I_k^{\imath}(t) = [x_k^{\imath}(t), x_{k+1}^{\imath}(t)) \tag{2.2}$$

with $k \in \mathcal{N} \setminus \{N\}$, and the *discrete mean opinions* by

$$\mu_{x^{\imath}}^{\mathcal{N}}(t) = \frac{1}{N+1} \sum_{k \in \mathcal{N}} x_k^{\imath}(t)\,, \tag{2.3}$$

where $a_{\mathcal{N}}^{\imath\jmath}$ is the Euclidean distance between the agents after the opinion discretisation. In the following, we may denote with $\mathbf{x}^{\imath,\mathcal{N}}(t) := \left(x_0^{\imath}(t), \ldots, x_N^{\imath}(t)\right)$, for all $\imath \in \mathcal{M}$.

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

Thus, we consider the following system of ODEs

$$\dot{x}_k^{\imath}(t) = \frac{\beta_k^{\imath}}{\sigma_N^{\imath}} \left(\Phi^{\imath}(\rho_{k-1}^{\imath}) - \Phi^{\imath}(\rho_k^{\imath}) \right) + \theta_k^{\imath} \tag{2.4a}$$

$$\dot{a}_{\mathcal{N}}^{\imath}(t) = \sum_{\jmath \in \mathcal{M}} \mathbf{V}(\mu_{x^{\imath}}^{\mathcal{N}}, \mu_{x^{\jmath}}^{\mathcal{N}}, a_{\mathcal{N}}^{\imath\jmath}), \tag{2.4b}$$

for $k \in \mathcal{N} \setminus \{0, N\}$ and $\imath \in \mathcal{M}$, endowed with the *boundary conditions*

$$\dot{x}_0^{\imath}(t) = 0 \quad , \quad \dot{x}_N^{\imath}(t) = 0 \qquad \text{for all } \imath \in \mathcal{M}, \tag{2.5}$$

and initial conditions

$$x_k^{\imath}(0) = \bar{x}_k^{\imath} \quad , \quad a_{\mathcal{N}}^{\imath}(0) = \bar{a}^{\imath} \qquad \text{for all } \imath \in \mathcal{M}, \ k \in \mathcal{N}. \tag{2.6}$$

In (2.4), we have denoted with $\beta_k^{\imath}(t)$ the *discrete diffusion mobilities*

$$\beta_k^{\imath}(t) = \sum_{\jmath \in \mathcal{M}} \sum_{l \in \mathcal{N}} \sigma_N^{\jmath} \mathbf{A}^{\imath\jmath}(x_k^{\imath}, x_l^{\jmath}, a_{\mathcal{N}}^{\imath\jmath}), \tag{2.7}$$

and with $\Phi^{\imath}$ the nonlinear diffusion for the agent $\imath$, see assumption (Dif) below. The contribution of the diffusion at the particle level can be interpreted assuming that opinions evolve with a speed equal to the osmotic velocity associated with the diffusion process, see [25, 26, 46].

Functions $\theta_k^{\imath}(t)$ describe the *discrete transports* and are given by

$$\theta_k^{\imath}(t) = \sum_{\jmath \in \mathcal{M}} \sum_{l \in \mathcal{N}} \sigma_N^{\jmath} \mathbf{K}^{\imath\jmath}(x_k^{\imath}, x_l^{\jmath}, a_{\mathcal{N}}^{\imath\jmath}). \tag{2.8}$$

We briefly comment on the boundary conditions in (2.5). As mentioned earlier, in (1.1) we did not introduce any boundary conditions. However, we are interested in the case where individual opinions do not exit the considered interval. Condition (2.5) enforces zero velocity for extreme opinions, which, along with the results of Lemma 3.2 on the preservation of the order of opinions, implies zero-flux boundary conditions for the opinion densities, as specified in equation (2.11) below.

2.2. Preliminaries and assumptions

We now present some tools from optimal transport that will be useful in the sequel. The Wasserstein distance is the right notion of distance for the opinions since it allows to measure the distances between measures (densities) with same mass. For a fixed mass $\sigma > 0$, we consider the space

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

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

$$X_\mu(z) = \inf\left\{ x \in \mathbb{R} \colon \mu((-\infty, x]) > z \right\}. \tag{2.9}$$

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 [16], in the one-dimensional setting the *p-Wasserstein distance* can be equivalently defined in terms of the L^1-distance between the respective pseudo-inverse mappings as

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

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

$$d_{W_\sigma^1}(\rho_1, \rho_2) = \|X_{\rho_1} - X_{\rho_2}\|_{L^1([0,\sigma];\mathbb{R})}, \tag{2.10}$$

We refer to [4, 47, 55] for a complete presentation of the subject.

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

We assume that the **initial densities** are under the following assumptions:

(In1) $\bar{\rho}^{\imath} \in BV(\Omega;\mathbb{R}^+)$ with $\|\bar{\rho}^{\imath}\|_{L^1(\Omega)} = \sigma^{\imath}$, for some $\sigma^{\imath} > 0$,

(In2) there exists $m^{\imath}, M^{\imath} > 0$ such that $m^{\imath} \le \bar{\rho}^{\imath}(x) \le M^{\imath}$ for every $x \in \Omega$.

Due to technical constraints, initial data with a vacuum region cannot be considered, as evident in the proof of Proposition 3.6. The lower bound is employed to control the time derivatives of extreme opinion densities $\dot{\rho}_0^i$ and $\dot{\rho}_N^i$. The essential point is the need for either a uniform control over these quantities or an estimation at an appropriate rate of N. We believe that this technical issue can be resolved, especially in the case of nonlinear diffusion, where finite speed of propagation is known and not utilised in the numerical Section. However, addressing this matter is currently beyond our capabilities. We now introduce the assumptions for the **diffusive** and **transport operators**.

(A) We assume that $\mathbf{A}^{\imath\jmath} : \Omega \times \Omega \times \mathbb{R} \to \mathbb{R}$ is a C^1 non-negative function w.r.t. the first variable and for all pairs $(\imath, \jmath)$ it exists $c_{\mathbf{A}} > 0$ such that

$$|\mathbf{A}^{\imath\jmath}(x^i_{k^*}, x^j_{s^*}, a^{\imath\jmath}) - \mathbf{A}^{\imath\jmath}(x^i_k, x^j_s, a^{\imath\jmath})| \le c_{\mathbf{A}} \left(|x^i_{k^*} - x^i_k| + |x^j_{s^*} - x^j_s|\right) ,$$

and it exists $c_{1,\mathbf{A}} > 0$ such that

$$|\partial_1\mathbf{A}^{\imath\jmath}(x^i_{k^*}, x^j_{s^*}, a^{\imath\jmath}) - \partial_1\mathbf{A}^{\imath\jmath}(x^i_k, x^j_s, a^{\imath\jmath})| \le c_{1,\mathbf{A}}|x^j_{s^*} - x^j_s|,$$

for all (k, k^*) and (s, s^*) pairs of indexes in $\mathcal{N} \times \mathcal{N}$, where $\partial_1\mathbf{A}^{\imath\jmath}$ denotes the derivatives with respect to the first entrance.

(K) We assume that $\mathbf{K}^{\imath\jmath} : \Omega \times \Omega \times \mathbb{R} \to \mathbb{R}$ is bounded, continuous, and for all pairs $(\imath, \jmath)$, it exists $c_{\mathbf{K}} > 0$ such that

$$|\mathbf{K}^{\imath\jmath}(x^i_{k^*}, x^j_{s^*}, a^{\imath\jmath}) - \mathbf{K}^{\imath\jmath}(x^i_k, x^j_s, a^{\imath\jmath})| \le c_{\mathbf{K}} \left(|x^i_{k^*} - x^i_k| + |x^j_{s^*} - x^j_s|\right) ,$$

for all (k, k^*) and (s, s^*) pairs of indexes in $\mathcal{N} \times \mathcal{N}$. We further assume that

$$\mathbf{K}^{\imath\imath}(x, x, a^{\imath\imath}) = 0.$$

(Dif) $\Phi^{\imath} : [0, \infty) \to \mathbb{R}$ is a nondecreasing Lipschitz function, with $\Phi^{\imath}(0) = 0$.

(V) The network velocity $\mathbf{V}$ is a C^1 bounded function on $\Omega \times \Omega \times \mathbb{R}^+$.

2.3. Continuous reconstruction and main result

Given the preliminary assumptions, we give the definition of weak solutions to equation (1.1) together with the statement of the main result.

By setting $\Omega_T = [0, T] \times \Omega$ and $\partial\Omega_T = [0, T] \times \{-1, 1\}$, and considering $\bar{\rho}^{\imath} \in L^1 \cap L^\infty(\Omega)$ and $\bar{a}^{\imath} \in \Lambda \subset \mathbb{R}^d$, we are going to deal with the following PDE-ODE system

$$\begin{cases} \partial_t\rho^{\imath} = \partial_x\big(\beta^{\imath}(\boldsymbol{\rho}, \mathbf{a}; x)\partial_x\Phi^{\imath}(\rho^{\imath})\big) - \partial_x\left(\rho^{\imath}\theta^{\imath}(\boldsymbol{\rho}, \mathbf{a}; x)\right), & (t, x) \in \Omega_T, \\ \partial_t a^{\imath}(t) = \sum\limits_{\jmath \in \mathcal{M}} \mathbf{V}(\mu_{\rho^{\imath}}(t), \mu_{\rho^{\jmath}}(t), a^{\imath\jmath}), & t \in [0, T], \\ \beta^{\imath}(\boldsymbol{\rho}, \mathbf{a}; x)\partial_x\Phi^{\imath}(\rho^{\imath}(t, x)) - \rho^{\imath}(t, x)\theta^{\imath}(\boldsymbol{\rho}, \mathbf{a}; x) = 0, & (t, x) \in \partial\Omega_T, \\ \rho^{\imath}(0, \cdot) = \bar{\rho}^{\imath}, & x \in \Omega, \\ a^{\imath}(0) = \bar{a}^{\imath}, \end{cases} \tag{2.11}$$

for all $\imath \in \mathcal{M}$, where the bold notation refers to the vectors $\boldsymbol{\rho} = (\rho^1, \ldots, \rho^M)$ and the set $\mathbf{a} = (a^1, \ldots, a^M)$.

We state the notion of weak solution for the system (2.11) as follows

Definition 2.1 (Weak solution). *We say that the couple* $(\boldsymbol{\rho}, \mathbf{a})$ *is a weak solution of* (1.1) *in the formulation* (2.11) *if*

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

- $\rho^{\imath} \in L^\infty \cap BV(\Omega_T)$, *with* $\rho^{\imath}(0,\cdot) = \bar{\rho}^{\imath}$, *for all* $\imath \in \mathscr{M}$
- $a^{\imath} \in C^2\left([0,T]\,;\mathbb{R}^d\right)$,

and taken $\zeta \in C_0^\infty(\Omega_T)$, *for all* $\imath \in \mathscr{M}$ *it satisfies*

$$\begin{aligned}\int_{\Omega_T} \rho^{\imath}(t,x)\partial_t\zeta(t,x) + \rho^{\imath}(t,x)\theta^{\imath}(\boldsymbol{\rho},\mathbf{a};x)\partial_x\zeta(x)\,dx\,dt \\ + \int_{\Omega_T} \Phi^{\imath}\left(\rho^{\imath}(t,x)\right)\left(\partial_x\beta^{\imath}(\boldsymbol{\rho},\mathbf{a};x)\partial_x\zeta(t,x) + \beta^{\imath}(\boldsymbol{\rho},\mathbf{a};x)\partial_{xx}\zeta(t,x)\right)\,dx = 0\end{aligned} \tag{2.12}$$

and

$$a^{\imath}(t) = \bar{a}^{\imath} + \sum_{\jmath\in\mathscr{M}} \int_0^t \mathbf{V}\left(\mu_{\rho^{\imath}}(\tau), \mu_{\rho^{\jmath}}(\tau), a^{\imath\jmath}(\tau)\right)\,d\tau \tag{2.13}$$

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

Given the discrete opinion densities defined in (2.2), we consider the following piecewise constant density reconstructions

$$\rho^{\imath,\mathscr{N}}(t,x) = \sum_{k\in\mathscr{N}\setminus\{N\}} \rho_k^{\imath}(t)\,\chi_{I_k^{\imath}(t)}(x)\,. \tag{2.14}$$

The main result of the paper reads as follows

Theorem 2.2. *Given* $M \in \mathbb{N}$ *and* $T > 0$ *fixed, consider* $\mathbf{a}^{\imath\jmath}$, $\mathbf{K}^{\imath\jmath}$, $\mathbf{V}$ *and* $\Phi^{\imath}$ *under assumptions* (**A**), (**K**), (*V*) *and* (*Dif*), *respectively, for all* $\imath,\jmath \in \mathscr{M}$. *Let* $\bar{\rho}^{\imath} : \Omega \to \mathbb{R}$ *under assumptions* (*In*1) *and* (*In*2) *and* $\bar{a}^{\imath} \in \Lambda \subset \mathbb{R}^d$ *for all* $\imath \in \mathscr{M}$. *Then, for all* $\imath \in \mathscr{M}$, *when* $N \to \infty$ *the density* $\rho^{\imath,\mathscr{N}}$ *introduced in equation* (2.14) *converges (up to subsequence) strongly to a non-negative function* $\rho^{\imath} \in L^\infty \cap BV(\Omega_T)$ *such that* $\|\rho^{\imath}\| = \sigma^{\imath}$ *and the solution to equation* (2.4b) $a^{\imath}_{\mathscr{N}}$ *converges to* $a^{\imath} \in C^2\left([0,T]\,;\mathbb{R}^d\right)$ *where the couple* $(\boldsymbol{\rho},\mathbf{a})$ *is a solution to equation* (1.1) *in the sense of Definition* 2.1.

Note that non-negativeness of the limit functions ρ easily follows by construction, see also (3.4) below, as well as the mass preservation. To avoid lack of notation, we highlight that from now on while considering the limit $N \to \infty$ we refer to the limit to infinity of the cardinality of the set of indexes $\mathscr{N} = \{0,\ldots,N\}$.

3. Proof of the main result

3.1. Basic estimates

We start providing some fundamental estimates that allow to deduce the well-posedness of (2.4) and of the discrete densities (2.2). Let us recall that in (2.2), we introduced the intervals

$$I_k^{\imath}(t) = \left[x_k^{\imath}(t), x_{k+1}^{\imath}(t)\right), \qquad |I_k^{\imath}|(t) = |x_{k+1}^{\imath} - x_k^{\imath}|, \tag{3.1}$$

for $\imath \in \mathscr{M}$ and $k \in \mathscr{N}\setminus\{N\}$. The first step is to prove that such intervals are well-defined. We start with the following auxiliary lemma, that follows directly from Assumption (**K**)

Lemma 3.1. *With the setting of the Main Theorem* 2.2, *and referring to the previous definitions, given* $\mathbf{K}^{\imath\jmath}$ *under Assumption* (**K**), *then it exists* $C > 0$ *such that the following inequality holds*

$$|\theta_{k+1}^i(t) - \theta_k^i(t)| \le C|x_{k+1}^i(t) - x_k^i(t)| \quad \forall \imath \in \mathscr{M},\ \forall t \in [0,T]\,, \tag{3.2}$$

with $C = c_{\mathbf{K}}\sigma^{\mathscr{M}}$, *where* $\sigma^{\mathscr{M}} = \sum_{j\in\mathscr{M}} \sigma^{\jmath}$.

Lemma 3.2 (Ordering preservation). *Assume* $\mathbf{A}^{\imath\jmath}$ *and* $\mathbf{K}^{\imath\jmath}$ *under assumptions* (**A**) *and* (**K**), *respectively, for all* $\imath,\jmath \in \mathscr{M}$. *Let us consider the DPA system described by* (2.4) *with initial conditions* $\bar{x}^{\imath}$

constructed in (2.1)*, for* $i \in \mathcal{M}$*, and a finite time* $T > 0$*. Then, for all* $t \in [0, T)$ *there is a positive constant* μ *independent from* $\mathcal{N}$ *– and so from* N *too – such that the distance between two adjacent opinions* $x_k^i, x_{k+1}^i \in C[0, T]$ *is bounded from below by*

$$\left(x_{k+1}^i - x_k^i\right)(t) \geq \min_{i,k}(\bar{x}_{k+1}^i - \bar{x}_k^i)\, e^{-\mu T},$$

for all $k \in \mathcal{N} \setminus \{N\}$ *and* $i \in \mathcal{M}$*.*

Proof. Given $T > 0$ and $i \in \mathcal{M}$, we define τ_1 as

$$\tau_1 = \inf\left\{s \in (0, T) : \exists\, k \in \mathcal{N}\setminus\{N\} \text{ s.t. } (x_{k+1}^i - x_k^i)(s) = (\bar{x}_{k+1}^i - \bar{x}_k^i)\, e^{-\mu s}\right\},$$

then the same index k corresponds also to the one of the maximum ρ_k^i at time τ_1 because I_k^i is the minimum interval of the $\mathcal{N}$ partition of Ω for the i-th agent at time τ_1.

At this point, let assume that exists $\tau_2 \in (\tau_1, T)$ such that

$$\left(x_{k+1}^i - x_k^i\right)(s) < (\bar{x}_{k+1}^i - \bar{x}_k^i)e^{-\mu s} \quad \forall s \in (\tau_1, \tau_2).$$

We show that the existence of τ_2 would bring to a contradiction. Let us consider the evolution of the interval I_k^i

$$\begin{aligned}\frac{d}{dt}\left[e^{\mu t}\left(x_{k+1}^i - x_k^i\right)(t)\right]_{|t=\tau_1} = e^{\mu\tau_1}\frac{\beta_{k+1}^i(\tau_1)}{\sigma_N^i}&\left[\Phi^i(\rho_k^i(\tau_1)) - \Phi^i(\rho_{k+1}^i(\tau_1))\right] + e^{\mu\tau_1}\theta_{k+1}^i(\tau_1)\\ - e^{\mu\tau_1}\frac{\beta_k^i(\tau_1)}{\sigma_N^i}&\left[\Phi^i(\rho_{k-1}^i(\tau_1)) - \Phi^i(\rho_k^i(\tau_1))\right] - e^{\mu\tau_1}\theta_k^i(\tau_1)\\ + \mu e^{\mu\tau_1}&(x_{k+1}^i(\tau_1) - x_k^i(\tau_1)).\end{aligned}$$

At time $t = \tau_1$, as highlighted before, by construction of the discrete densities in (2.2) we have $\rho_k^i \geq \rho_l^i$ for all $l \neq k$.

Then, the monotonicity of Φ^i gives

$$\frac{d}{dt}\left[e^{\mu t}\left(x_{k+1}^i - x_k^i\right)(t)\right]_{|t=\tau_1} \geq e^{\mu\tau_1}\left[\mu\ \left(x_{k+1}^i(\tau_1) - x_k^i(\tau_1)\right) + \left(\theta_{k+1}^i(\tau_1) - \theta_k^i(\tau_1)\right)\right].$$

Thanks to (3.2) we get

$$\frac{d}{dt}\left[e^{\mu t}\left(x_{k+1}^i - x_k^i\right)(t)\right]_{|t=\tau_1} \geq e^{\mu\tau_1}(\mu - C)(x_{k+1}^i(\tau_1) - x_k^i(\tau_1)) \geq 0,$$

which holds while choosing $\mu \geq C$. At this point, we fix $t^* \in (\tau_1, \tau_1 + \delta < \tau_2)$ with δ as small as we want, and we get

$$e^{\mu t^*}\left(x_{k+1}^i - x_k^i\right)(t^*) = e^{\mu\tau_1}\left(x_{k+1}^i - x_k^i\right)(\tau_1) + \int_{\tau_1}^{t^*}\frac{d}{ds}\left[e^{\mu t}\left(x_{k+1}^i - x_k^i\right)(s)\right] ds,$$

due to the positiveness of the last term we get the wished result which show the absurd,

$$\left(x_{k+1}^i - x_k^i\right)(t^*) \geq e^{-\mu t^*}\left(\bar{x}_{k+1}^i - \bar{x}_k^i\right),$$

this proves that $I_k^i(t)$ cannot decrease faster than $I_k^i(0)e^{-\mu t}$. Nevertheless, this does not deny the existence of an index $\tilde{k}$ such that the interval $I_{\tilde{k}}^i(t)$, satisfying $I_{\tilde{k}}^i(\tau_1) > I_{\tilde{k}}^i(0)$, decreases faster than $I_k^i(t)$. This means that could exists $\tilde{\tau}_1 > \tau_1$ for which $I_{\tilde{k}}^i(\tilde{\tau}_1) = I_{\tilde{k}}^i(0)$, with $\tilde{\tau}_1 < T$. At this point, we should prove that there exists $\tilde{\mu}$ such that $I_{\tilde{k}}^i(t) \geq e^{-\tilde{\mu}t}I_{\tilde{k}}^i(0)$ for all $t \in (\tilde{\tau}_1, T)$. To prove it, we repeat the same procedure explained before but defining $\tilde{\tau}_1$ considering the set of indexes $\mathcal{N}/\{k\}$. The final exponential rate will be the largest μ among those considered.

Let us also consider the case with k not unique, i.e. for τ_1 there are several intervals satisfying the definition of τ_1, the set of these indexes is denoted by $\mathcal{J} = \{k_j\}$. If there is at least one k_{j^*} not adjacent to other indexes of $\{k_j\}$, then we take that index and we follow the proof above. In the event that there are three indexes of $\{k_j\}$ in a row, we take without distinction that one with the fastest decrease in time of the interval $I_{k_j}^i$. At this point, we are back to the steps shown above. This concludes the proof. □

Remark 3.3. *A similar procedure of the one in Lemma 3.2 allows to produce the upper bound on discrete opinions*

$$\left(x_{k+1}^{\imath}-x_k^{\imath}\right)(t)\leq\max_k(\bar{x}_{k+1}^{\imath}-\bar{x}_k^{\imath})\,e^{CT}\qquad\forall t\in[0,T]\,. \tag{3.3}$$

This estimate, together with the one in Lemma 3.2, allows to deduce the following bounds on the discrete densities in (2.2)

$$m^{\imath}e^{-CT}\leq\rho_k^{\imath}\leq M^{\imath}e^{\mu T}\,,\ \text{for all}\ \ \imath\in\mathscr{M},\ \ k\in\mathscr{N}\setminus\{N\}\,, \tag{3.4}$$

with $m^{\imath}$ and $M^{\imath}$ in assumption (In2).

Lemma 3.4 (Velocity boundedness). *Fix $T>0$ and assume $\mathbf{A}^{\imath\jmath}$ and $\mathbf{K}^{\imath\jmath}$, respectively, under assumptions ($\mathbf{A}$) and ($\mathbf{K}$) for all $(\imath,\jmath)\in\mathscr{M}\times\mathscr{M}$. Then, solutions to system (2.4) satisfy*

$$\sup_{t\in[0,T]}\|\dot{\mathbf{x}}^{\imath,\mathscr{N}}(t)\|_\infty<+\infty,\ \ \text{for all}\ \ \imath\in\mathscr{M}.$$

Proof. Using the equation for the evolution of the partitioning, we get

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\left(\dot{x}_k^{\imath}(t)\right)^2&\leq x_k^{\imath}\left(\frac{\beta_k^{\imath}}{\sigma_N^{\imath}}\left(\Phi^{\imath}(\rho_{k-1}^{\imath})-\Phi^{\imath}(\rho_k^{\imath})\right)-\theta_k^{\imath}\right)\\&\leq|x_k^{\imath}|\left(\frac{1}{\sigma_N^{\imath}}\|\beta_k^{\imath}\|_\infty Lip\left(\Phi^{\imath}\right)\left|\rho_k^{\imath}-\rho_{k-1}^{\imath}\right|+\left|\theta_k^{\imath}\right|\right)\\&\leq c_1|x_k^{\imath}|^2+c_2\left|\rho_k^{\imath}-\rho_{k-1}^{\imath}\right|^2+c_3\left|\theta_k^{\imath}\right|^2,\end{aligned}$$

for some constants $c_1,c_2,c_3>0$. Then, the thesis follows from equation (3.4), assumption ($\mathbf{K}$) and the fact that $x_k^{\imath}\in\Omega$ because (2.5) and Lemma 3.2. □

We now have all the tools needed to prove the convergence in some strong sense for the piecewise constant densities of equation (2.14). We based our strategy on the ones proposed in the context of DPA, see for instance the proofs in [25, 27], that are using the generalised Aubin-Lions lemma version in [45], that we report here in a simplified version adapted to our setting.

Theorem 3.5. *Let $T>0$ be fixed, and $\rho^{\mathscr{N}}(t,\cdot):[a,b]\to\mathbb{R}$ be a sequence of non-negative probability densities for every $t\in[0,T]$ and for every $N\in\mathbb{N}$, where $\mathscr{N}=\{0,\ldots,N\}$. Moreover, assume that $\|\rho^{\mathscr{N}}(t,\cdot)\|_{L^\infty}\leq M$ for some constant M independent on t and N. If*

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

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

The result reads as follows

Proposition 3.6. *Let $\rho^{\imath,\mathscr{N}}$ be defined as in equation (2.14) for $\imath\in\mathscr{M}$. Then, there exists $\rho^{\imath}\in L^1\cap L^\infty(\Omega_T)$ such that $\|\rho^{\imath,\mathscr{N}}-\rho^{\imath}\|_{L^1}\to 0$ as $N\to\infty$.*

Proof. The proof reduces to the application of Theorem 3.5, in order to show that we can apply that result we first prove that

$$\sup_N\int_0^T TV[\rho^{\imath,\mathscr{N}}(t,\cdot)]\,dt\leq\infty\,. \tag{3.5}$$

To show this result, we look for a Grönwall type inequality. We start performing the following preliminary computation

$$\begin{aligned}
\dot\rho_k^i &= -\frac{\rho_k^i}{|I_k^i|}\left(\dot x_{k+1}^i(t)-\dot x_k^i(t)\right)\\
&= -\frac{\rho_k^i}{|I_k^i|}\Big(\underbrace{\frac{\beta_{k+1}^i}{\sigma_N^i}\left(\Phi^i\left(\rho_k^i\right)-\Phi^i\left(\rho_{k+1}^i\right)\right)-\frac{\beta_k^i}{\sigma_N^i}\left(\Phi^i\left(\rho_{k-1}^i\right)-\Phi^i\left(\rho_k^i\right)\right)}_{:=B_k^i}+\underbrace{(\theta_{k+1}^i-\theta_k^i)}_{:=|I_k^i|\Theta_k^i}\Big)\\
&= -\frac{\rho_k^i}{|I_k^i|}B_k^i+\rho_k^i\Theta_k^i.
\end{aligned}$$

The first step consists of proving the Lipschitz continuity in time of $t\to TV[\rho_t^{i,\mathcal{N}}(t,\cdot)]$. The total variation can be explicitly computed as

$$TV[\rho^{i,\mathcal{N}}(t,\cdot)]=\rho_0^i(t)+\sum_{k\in\mathcal{N}\setminus\{0,N\}}|\rho_k^i(t)-\rho_{k-1}^i(t)|+\rho_{N-1}^i(t).$$

From the boundedness in Lemmas 3.2 and 3.4, for all $s,t\in(0,T)$, we estimate

$$\begin{aligned}
|\rho_k^i(t)-\rho_k^i(s)| \le \left|\int_s^t \dot\rho_k^i(\tau)\,d\tau\right| &\le \int_s^t \frac{\rho_k^i}{|I_k^i|}\left|\dot x_{k+1}^i(t)-\dot x_k^i(t)\right|\,d\tau\\
&\le 2\frac{M^i}{\min\limits_{k,i}|\bar x_{k+1}^i-\bar x_k^i|}\sup_{\tau\in[0,T]}\|\dot{\mathbf{x}}^{i,\mathcal{N}}(\tau)\|_\infty|t-s|.
\end{aligned}$$

Then, it follows

$$\begin{aligned}
\left|TV[\rho^{i,\mathcal{N}}(t,\cdot)]-TV[\rho^{i,\mathcal{N}}(s,\cdot)]\right| \le& |\rho_0^i(t)-\rho_0^i(s)|+|\rho_{N-1}^i(t)-\rho_{N-1}^i(s)|\\
&+\sum_{k\in\mathcal{N}\setminus\{0,N\}}\left(|\rho_k^i(t)-\rho_k^i(s)|+|\rho_{k-1}^i(t)-\rho_{k-1}^i(s)|\right)\\
&\le 2\max_{k,i}\frac{M^i}{\min\limits_{k,i}|\bar x_{k+1}^i-\bar x_k^i|}\sup_{\tau\in[0,T]}\|\dot{\mathbf{x}}^{i,\mathcal{N}}(\tau)\|_\infty|t-s|,
\end{aligned}$$

that proves the asserted Lipschitz continuity.

We now consider the time derivative

$$\frac{d}{dt}TV[\rho^{i,\mathcal{N}}(t,\cdot)]=\dot\rho_0^i(t)+\sum_{k\in\mathcal{N}\setminus\{0,N\}}\operatorname{sign}\left(\rho_k^i(t)-\rho_{k-1}^i(t)\right)\left(\dot\rho_k^i(t)-\dot\rho_{k-1}^i(t)\right)+\dot\rho_{N-1}^i(t)\,.$$

Rearranging the sum and defining the operator

$$\mathrm{s}_k=\begin{cases}1-\operatorname{sign}\left(\rho_1^i(t)-\rho_0^i(t)\right) & k=0,\\ \operatorname{sign}\left(\rho_k^i(t)-\rho_{k-1}^i(t)\right)-\operatorname{sign}\left(\rho_{k+1}^i(t)-\rho_k^i(t)\right), & k=1,\dots,N-2,\\ 1+\operatorname{sign}\left(\rho_{N-1}^i(t)-\rho_{N-2}^i(t)\right), & k=N-1,\end{cases}$$

with we can rewrite the previous equation as

$$\frac{d}{dt}TV[\rho_t^{i,\mathcal{N}}(t,\cdot)]=\mathrm{s}_0\dot\rho_0^i(t)+\sum_{k=1}^{N-2}\mathrm{s}_k\dot\rho_k^i(t)+\mathrm{s}_{N-1}\dot\rho_{N-1}^i(t)\,. \tag{3.6}$$

At this point, we show that the terms involving the diffusion are always negative, i.e.

$$-\mathrm{s}_k\frac{\rho_k^i}{|I_k^i|}B_k^i\le 0\quad k=2,\dots,N-2.$$

In order to prove this statement, we should distinguish different cases. First, we observe that s_k is always zero if $\rho_{k-1}^i<\rho_k^i<\rho_{k+1}^i$ or $\rho_{k+1}^i<\rho_k^i<\rho_{k-1}^i$. In the other two cases, namely $\rho_k^i<\rho_{k-1}^i$ and $\rho_k^i<\rho_{k+1}^i$, or $\rho^{i}_{k-1},\ \rho^{i}_{k+1} < \rho^{i}_{k}$ and $\rho^{i}_{k-1} < \rho^{i}_{k}$, the monotonicity of Φ^{i} implies, respectively, $s_k \le 0$ and $B^{i}_{k} \le 0$, and $s_k \ge 0$ and $B^{i}_{k} \ge 0$, and hence, the negativity of the diffusion contribution is proved.

Concerning the term involving Θ^{i}_{k}, we can rearrange the sum as follows

$$
\begin{aligned}
\sum_{k=1}^{N-2} s_k \rho^{i}_{k} \Theta^{i}_{k} = \operatorname{sign}(\rho^{i}_{1} - \rho^{i}_{0})\rho^{i}_{1}\Theta^{i}_{1} &- \operatorname{sign}(\rho^{i}_{N-1} - \rho^{i}_{N-2})\rho^{i}_{N-2}\Theta^{i}_{N-2} \\
&+ \sum_{k=2}^{N-2} \operatorname{sign}(\rho^{i}_{k} - \rho^{i}_{k-1})(\rho^{i}_{k} - \rho^{i}_{k-1})\Theta^{i}_{k} \\
&+ \sum_{k=2}^{N-2} \operatorname{sign}(\rho^{i}_{k} - \rho^{i}_{k-1})(\Theta^{i}_{k} - \Theta^{i}_{k-1})\rho^{i}_{k-1}.
\end{aligned}
$$

Observing that $|\Theta^{i}_{k}| \le C$ because of equation (3.2) and that

$$
\begin{aligned}
\left|\Theta^{i}_{k} - \Theta^{i}_{k-1}\right| \le& \left| \frac{1}{|I^{i}_{k}|} \left[\left(\theta^{i}_{k+1} - \theta^{i}_{k}\right) - \left(\theta^{i}_{k} - \theta^{i}_{k-1}\right) \right] \right| \\
&+ \left| \left(\frac{1}{|I^{i}_{k}|} - \frac{1}{|I^{i}_{k-1}|} \right) \left(\theta^{i}_{k} - \theta^{i}_{k-1}\right) \right| \\
\le& C \frac{\rho^{i}_{k}}{\sigma^{i}_{N}} \left(|I^{i}_{k}|^2 + |I^{i}_{k-1}|^2 + \left| |I^{i}_{k}| - |I^{i}_{k-1}| \right| \right) + \frac{C}{\sigma^{i}_{N}} \left| \rho^{i}_{k} - \rho^{i}_{k-1} \right| |I^{i}_{k-1}|,
\end{aligned}
$$

we can bound

$$
\left| \sum_{k=1}^{N-2} s_k \rho^{i}_{k} \theta^{i}_{k} \right| \le 2M^{i} C \left(1 + |\Omega|\right) + 3C\, TV[\rho^{i,\mathcal{N}}(t)].
$$

We can finally estimate

$$
\frac{d}{dt} TV[\rho^{i,\mathcal{N}}_{t}(t,\cdot)] \le 2C \frac{M^{i}}{m^{i}} \left(\rho^{i}_{0}(t) + \rho^{i}_{N-1}(t) \right) + 2M^{i} C \left(1 + |\Omega|\right) + 3C\, TV[\rho^{i,\mathcal{N}}(t)]\,,
$$

and thus, equation (3.5) follows by Grönwall type argument.

We now prove that the second requirement of Theorem 3.5 holds, namely there exists a positive constant C such that

$$
d_{W^1}\left(\rho^{i,\mathcal{N}}(t,\cdot), \rho^{i,\mathcal{N}}(s,\cdot)\right) \le C|t-s| \qquad \forall\, s,t \in [0,T]\,. \tag{3.7}
$$

In order to do this, we use the isometry of equation (2.10), where the *pseudo-inverse* function for $\rho^{i,\mathcal{N}}$ is given by

$$
X_{\rho^{i,\mathcal{N}}}(m,t) = \sum_{k \in \mathcal{N} \setminus \{N\}} \left(x^{i}_{k}(t) + \frac{m - k\sigma^{i}_{N}}{\rho^{i}_{k}(t)} \right) \chi_{\left[k\sigma^{i}_{N},(k+1)\sigma^{i}_{N}\right]}(m)\,.
$$

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

Then, for any $t > s$, we have

$$\begin{aligned} d_{W^1}\left(\rho^{\iota,\mathcal{N}}(t,\cdot),\rho^{\iota,\mathcal{N}}(s,\cdot)\right) &\le \sum_{k\in\mathcal{N}\setminus\{N\}} \int_{k\sigma_N^\iota}^{(k+1)\sigma_N^\iota} \left|x_k^\iota(t)-x_k^\iota(s)\right| dm \\ &\quad + \sum_{k\in\mathcal{N}\setminus\{N\}} \int_{k\sigma_N^\iota}^{(k+1)\sigma_N^\iota} \left|(m-k\sigma_N^\iota)\left(\frac{1}{\rho_k^\iota(t)}-\frac{1}{\rho_k^\iota(s)}\right)\right| dm \\ &\le \sum_{k\in\mathcal{N}\setminus\{N\}} \sigma_N^\iota \int_s^t \left|\dot{x}_k^\iota(\tau)\right| d\tau + \frac{(\sigma_N^\iota)^2}{2}\int_s^t \left|\frac{d}{d\tau}\frac{1}{\rho_k^\iota(\tau)}\right| d\tau \\ &\le 3\sigma_N^\iota \sum_{k\in\mathcal{N}\setminus\{N\}} \int_s^t \left|\dot{x}_k^\iota(\tau)\right| d\tau \\ &\le C|t-s|, \end{aligned}$$

in view of Lemma 3.4.

Once proved the bounds described by equations (3.5) and (3.7), we can apply Theorem 3.5, which concludes the proof. □

Lemma 3.7 (Convergence of momenta). *Given $\mu_{\rho^\iota}(t)$ and $\mu_{x^\iota}^{\mathcal{N}}(t)$, respectively,*

$$\mu_{\rho^\iota}(t) = \frac{1}{\sigma^\iota|\Omega|}\int_\Omega \rho^\iota(y,t)y\,dy, \qquad \mu_{x^\iota}^{\mathcal{N}}(t) = \frac{1}{N+1}\sum_{k=0}^{N} x_k^\iota(t),$$

we have that

$$\lim_{N\to\infty}\left(\mu_{\rho^\iota}(t)-\mu_{x^\iota}^{\mathcal{N}}(t)\right)=0$$

for all $t > 0$.

Proof. We recall that $x_0^i + x_N^i = 0$, and $\Omega = [-1, 1]$, from the definitions we have

$$\begin{aligned} \mu_{x^\iota}^{\mathcal{N}}(t) &= \frac{1}{2(N+1)}\sum_{k\in\mathcal{N}\setminus\{N\}}(x_{k+1}^\iota + x_k^\iota) = \frac{1}{2(N+1)}\sum_{k\in\mathcal{N}\setminus\{N\}}\frac{(x_{k+1}^\iota+x_k^\iota)(x_{k+1}^\iota-x_k^\iota)}{x_{k+1}^\iota-x_k^\iota} \\ &= \frac{N}{(N+1)}\sum_{k\in\mathcal{N}\setminus\{N\}}\frac{\rho_k^\iota}{\sigma^\iota}\frac{(x_{k+1}^\iota)^2-(x_k^\iota)^2}{2} = \frac{N}{(N+1)}\sum_{k\in\mathcal{N}\setminus\{N\}}\frac{\rho_k^\iota(t)}{\sigma^\iota}\int_{-\infty}^{\infty} x\,\chi_{I_k^\iota}\,dx \\ &= \frac{N}{(N+1)}\,\mu_{\rho^{\iota,N}}(t), \end{aligned}$$

where we used equation (2.14). We conclude that there exists a constant depending only on the domain Ω such that

$$\begin{aligned} |\mu_{\rho^\iota}(t)-\mu_{x^\iota}^{\mathcal{N}}(t)| &\le \int_\Omega |\rho^\iota - \frac{N}{N+1}\rho^{\iota,N}|\,|x|\,dx \\ &\le C(|\Omega|)\|\rho^\iota-\rho^{\iota,N}\|_{L^1} + o\left(\frac{C}{N}\right), \end{aligned}$$

which concludes the proof. □

The following Proposition concerns the convergence of the approximated nodes $a_{\mathcal{N}}^\iota$.

Proposition 3.8. *Let $T > 0$ be fixed and consider* **V** *under assumption* (**V**). *Then, for any $\iota \in \mathcal{M}$, there exists $a^\iota \in C([0,T])$ such that $a_{\mathcal{N}}^\iota \to a^\iota$ as $N \to \infty$ uniformly in $[0,T]$. Moreover, the limits a^ι satisfy* (2.13) *for all $t \in [0,T]$.*

Proof. We first notice that from (2.4b) and the boundedness of $\mathbf{V}$, we have the uniform bound

$$|a^{\imath}_{\mathcal{N}}(t)| \leq |\bar{a}^{\imath}| + T\|\mathbf{V}\|_{\infty}, \qquad \textit{with } t \in [0, T],$$

for all $\imath \in \mathscr{M}$ uniformly in N. Thus, there exist $a^{\imath}$ such that $a^{\imath}_{\mathcal{N}}(t)$ pointwise converges (up to subsequences) to $a^{\imath}(t)$ as $N \to \infty$. Consider now $N_1, N_2 \in \mathbb{N}$, then

$$\begin{aligned}
\sum_{\imath\in\mathscr{M}} |a^{\imath}_{\mathcal{N}_1}(t) - a^{\imath}_{\mathcal{N}_2}(t)| \leq & \sum_{\imath,\jmath\in\mathscr{M}} \int_0^t \left|\mathbf{V}(\mu^{\mathcal{N}_1}_{x^{\imath}}, \mu^{\mathcal{N}_1}_{x^{\jmath}}, a^{\imath\jmath}_{\mathcal{N}_1}) - \mathbf{V}(\mu^{\mathcal{N}_2}_{x^{\imath}}, \mu^{\mathcal{N}_2}_{x^{\jmath}}, a^{\imath\jmath}_{\mathcal{N}_2})\right| d\tau \\
\leq & C \sum_{\imath,\jmath\in\mathscr{M}} \int_0^t \left|\mu^{\mathcal{N}_1}_{x^{\imath}}(\tau) - \mu^{\mathcal{N}_2}_{x^{\imath}}(\tau)\right| + \left|\mu^{\mathcal{N}_1}_{x^{\jmath}}(\tau) - \mu^{\mathcal{N}_2}_{x^{\jmath}}(\tau)\right| \, d\tau \\
& + C \sum_{\jmath\in\mathscr{M}} \int_0^t \left|a^{\imath\jmath}_{\mathcal{N}_1}(\tau) - a^{\imath\jmath}_{\mathcal{N}_2}(\tau)\right| \, d\tau \\
\leq & 2C \sum_{\imath\in\mathscr{M}} \sum_{l=1,2} \|\rho^{\imath} - \rho^{\imath,\mathcal{N}_l}\|_{L^1(\Omega_T)} + 2C \int_0^t \sum_{\imath\in\mathscr{M}} |a^{\imath}_{\mathcal{N}_1}(t) - a^{\imath}_{\mathcal{N}_2}(t)| \, d\tau,
\end{aligned}$$

where we used the estimate from the proof of Lemma 3.7 and the fact that by straightforward manipulations we have

$$\begin{aligned}
\sum_{\imath,\jmath\in\mathscr{M}} \left|a^{\imath\jmath}_{\mathcal{N}_1}(\tau) - a^{\imath\jmath}_{\mathcal{N}_2}(\tau)\right| = & \sum_{\imath,\jmath\in\mathscr{M}} \left|\|a^{\imath}_{\mathcal{N}_1}(\tau) - a^{\jmath}_{\mathcal{N}_1}(\tau)\| - \|a^{\imath}_{\mathcal{N}_2}(\tau) - a^{\jmath}_{\mathcal{N}_2}(\tau)\|\right| \\
= & \sum_{\imath,\jmath\in\mathscr{M}} \left|\|a^{\imath}_{\mathcal{N}_1}(\tau) \pm a^{\imath}_{\mathcal{N}_2}(\tau) \pm a^{\jmath}_{\mathcal{N}_2}(\tau) - a^{\jmath}_{\mathcal{N}_1}(\tau)\| - \|a^{\imath}_{\mathcal{N}_2}(\tau) - a^{\jmath}_{\mathcal{N}_2}(\tau)\|\right| \\
\leq & \sum_{\imath,\jmath\in\mathscr{M}} \left|\|a^{\imath}_{\mathcal{N}_1}(\tau) - a^{\imath}_{\mathcal{N}_2}(\tau)\| + \|a^{\imath}_{\mathcal{N}_2} - a^{\jmath}_{\mathcal{N}_2}(\tau)\|\right. \\
& \left. - \|a^{\imath}_{\mathcal{N}_2}(\tau) - a^{\jmath}_{\mathcal{N}_2}(\tau)\| + \|a^{\jmath}_{\mathcal{N}_2}(\tau) - a^{\jmath}_{\mathcal{N}_1}(\tau)\|\right| \\
\leq & \sum_{\imath,\jmath\in\mathscr{M}} \left|a^{\imath}_{\mathcal{N}_1}(\tau) - a^{\imath}_{\mathcal{N}_2}(\tau)\right| + \left|a^{\jmath}_{\mathcal{N}_2}(\tau) - a^{\jmath}_{\mathcal{N}_1}(\tau)\right| \\
\leq & 2 \sum_{\imath\in\mathscr{M}} |a^{\imath}_{\mathcal{N}_1}(t) - a^{\imath}_{\mathcal{N}_2}(t)| \, .
\end{aligned}$$

Thus, by using the integral version of Gronwall's inequality we can deduce

$$\sup_{t\in[0,T]} \sum_{\imath\in\mathscr{M}} |a^{\imath}_{\mathcal{N}_1}(t) - a^{\imath}_{\mathcal{N}_2}(t)| \leq 2C \sum_{\imath\in\mathscr{M}} \sum_{l=1,2} \|\rho^{\imath} - \rho^{\imath,\mathcal{N}_l}\|_{L^1(\Omega_T)} e^{2CT}.$$

and then

$$\sup_{t\in[0,T]} |a^{\imath}_{\mathcal{N}_1}(t) - a^{\imath}_{\mathcal{N}_2}(t)| \leq 2C \sum_{\imath\in\mathscr{M}} \sum_{l=1,2} \|\rho^{\imath} - \rho^{\imath,\mathcal{N}_l}\|_{L^1(\Omega_T)} e^{2CT},$$

that ensure the uniform convergences of $a^{\imath}_{\mathcal{N}_1}$ to $a^{\imath}$, for all $\imath \in \mathscr{M}$.

In order to show that $a^{\imath}$ satisfies (2.13), it is enough to observe that we can invoke the dominated convergence theorem since by the continuity of $\mathbf{V}$ and the uniform converges proved we have that

$$\mathbf{V}(\mu^{\mathcal{N}}_{x^{\imath}}, \mu^{\mathcal{N}}_{x^{\jmath}}, a^{\imath\jmath}_{\mathcal{N}}) \to \mathbf{V}(\mu_{\rho^{\imath}}, \mu_{\rho^{\jmath}}, a^{\imath\jmath}) \quad \text{a.e. in} \quad t \in [0, T],$$

and $\mathbf{V}(\mu^{\mathcal{N}}_{x^{\imath}}, \mu^{\mathcal{N}}_{x^{\jmath}}, a^{\imath\jmath}_{\mathcal{N}})$ is uniformly bounded w.r.t. N. Thus,

$$\int_0^t \mathbf{V}(\mu^{\mathcal{N}}_{x^{\imath}}, \mu^{\mathcal{N}}_{x^{\jmath}}, a^{\imath\jmath}_{\mathcal{N}}) \, d\tau \to \int_0^t \mathbf{V}(\mu_{\rho^{\imath}}, \mu_{\rho^{\jmath}}, a^{\imath\jmath}) \, d\tau,$$

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

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

We now prove that the empirical measures associated with the solution of equation (2.4a) and the piecewise constant densities in equation (2.14) share the same limit with respect to a suitable topology.

Lemma 3.9. *For any $T > 0$, the empirical measures associated with the solution of equation* (2.4a)*, defined by*

$$\tilde{\rho}^{i,\mathcal{N}}(t,x) = \sigma_N^i \sum_{k\in\mathcal{N}} \delta_{x_k^i(t)}(x), \qquad i \in \mathcal{M}, \tag{3.8}$$

satisfy

$$d_{W^1}\left(\tilde{\rho}^{i,\mathcal{N}}(t,\cdot), \rho^i(t,\cdot)\right) \to 0, \ \ \text{as} \ \ N \to \infty,$$

for all $t \in [0, T]$, where ρ^i is the limit obtained in Proposition 3.6.

Proof. We take again advantage of the isometry between the $1-$Wasserstain space for probability measures and the L^1 space in the space pseudo-inverse functions by noticing that the pseudo-inverse of an empirical measure is piecewise constant and then

$$\begin{aligned} d_{W^1}\left(\tilde{\rho}^{i,\mathcal{N}}(t,\cdot), \rho^{i,\mathcal{N}}(t,\cdot)\right) =& \|X_{\tilde{\rho}^{i,\mathcal{N}}}(t,\cdot) - X_{\rho^{i,\mathcal{N}}}(t,\cdot)\|_{L^1([0,\sigma^i])} \\ \leq& \sum_{k\in\mathcal{N}\setminus\{N\}} \int_{k\sigma_N^i}^{(k+1)\sigma_N^i} \left| (m - k\sigma_N^i)\frac{1}{\rho_k^i(t)} \right| dm \\ =& \frac{\sigma_N^i}{2}|\Omega|. \end{aligned}$$

The statement then follows from a triangulation argument. □

3.2. Convergence to weak solutions

With $\mathbf{a}^{\mathcal{N}}$ we denoted the vector $(a_{\mathcal{N}}^i)_{i\in\mathcal{M}}$, of agents with piecewise constant opinion distribution $\rho^{i,\mathcal{N}}$, while $\mathbf{a}$ is related to the vector of continuous distributions $\boldsymbol{\rho}$. This distinction is not stressed in the rest of the paper where the context does not allow misunderstanding.

Remark 3.10. *We notice that evaluating the operator β^i in* (3.8) *we have*

$$\begin{aligned} \beta^i(\tilde{\boldsymbol{\rho}}^N, \boldsymbol{a}^N; x_k^i) =& \sum_{j\in\mathcal{M}} \int_\Omega \mathbf{A}^{ij}(x_k^i, y, a_N^{ij})\tilde{\rho}^{j,N}(y,t)\,dy \\ =& \sum_{j\in\mathcal{M}} \sum_{l\in\mathcal{N}} \sigma_N^j \mathbf{A}^{ij}(x_k^i, x_l^j, a_N^{ij}) \\ =& \beta_k^i. \end{aligned} \tag{3.9}$$

Moreover,

$$\partial_x \beta^i\left(\tilde{\boldsymbol{\rho}}^N, \boldsymbol{a}^N; x\right) = \sum_{j\in\mathcal{M}} \int_\Omega \partial_1 \mathbf{A}^{ij}(x, y; a_N^{ij})\tilde{\rho}^{j,N}(t,y)\,dy, \tag{3.10}$$

Lemma 3.11. *Let $T > 0$, and consider the kernels $\mathbf{A}^{ij}$, $\mathbf{K}^{ij}$ under assumptions* (**A**) *and* (**K**)*, respectively. Let $\rho^{i,\mathcal{N}}$ and $\tilde{\rho}^{i,\mathcal{N}}$ be the sequences defined in equations* (2.14) *and* (3.8)*, respectively, and their limits ρ^i given by Proposition 3.6 and Lemma 3.9, for all $i \in \mathcal{M}$. Then, for every $\zeta \in C_0^\infty(\Omega_T)$ we have*

$$\int_{\Omega_T} \rho^{i,\mathcal{N}} \partial_t \zeta \, dx\, dt \to \int_{\Omega_T} \rho^i \partial_t \zeta \, dx\, dt \tag{3.11}$$

$$\int_{\Omega_T} \Phi^i\left(\rho^{i,\mathcal{N}}\right) \partial_x \beta^i\left(\tilde{\boldsymbol{\rho}}^{\mathcal{N}}, \mathbf{a}^{\mathcal{N}}; x\right) \partial_x \zeta \, dx\, dt \to \int_{\Omega_T} \Phi^i\left(\rho^i\right) \partial_x \beta^i\left(\boldsymbol{\rho}, \mathbf{a}; x\right) \partial_x \zeta \, dx\, dt \tag{3.12}$$

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

$$\int_{\Omega_T} \Phi^{\imath}\left(\rho^{\imath,\mathcal{N}}\right) \beta^{\imath}\left(\tilde{\boldsymbol{\rho}}^{\mathcal{N}}, \mathbf{a}^{\mathcal{N}}; x\right) \partial_{xx}\zeta \, dx\, dt \to \int_{\Omega_T} \Phi^{\imath}\left(\rho^{\imath}\right) \beta^{\imath}\left(\boldsymbol{\rho}, \mathbf{a}; x\right) \partial_{xx}\zeta \, dx\, dt \tag{3.13}$$

$$\int_{\Omega_T} \rho^{\imath,\mathcal{N}} \theta^{\imath}\left(\tilde{\boldsymbol{\rho}}^{\mathcal{N}}, \mathbf{a}^{\mathcal{N}}; x\right) \partial_{x}\zeta \, dx\, dt \to \int_{\Omega_T} \rho^{\imath} \theta^{\imath}\left(\boldsymbol{\rho}, \mathbf{a}; x\right) \partial_{x}\zeta \, dx\, dt \tag{3.14}$$

as $N \to \infty$.

Proof. We only prove equation (3.12), since equations (3.13) and (3.14) follow from similar argument, and equation (3.11) is a direct consequence of the L^1 strong compactness proved in Proposition 3.6. We first split the terms as following

$$\begin{aligned}
&\left|\int_{\Omega_T} \left(\Phi^{\imath}\left(\rho^{\imath,\mathcal{N}}\right) \partial_x\beta^{\imath}\left(\tilde{\boldsymbol{\rho}}^{\mathcal{N}}, \mathbf{a}^{\mathcal{N}}; x\right) - \Phi^{\imath}\left(\rho^{\imath}\right) \partial_x\beta^{\imath}\left(\boldsymbol{\rho}, \mathbf{a}; x\right)\right) \partial_x\zeta(t,x) dx\, dt\right| \\
&\leq \left|\int_{\Omega_T} \left(\Phi^{\imath}\left(\rho^{\imath,\mathcal{N}}\right) - \Phi^{\imath}\left(\rho^{\imath}\right)\right) \partial_x\beta^{\imath}\left(\tilde{\boldsymbol{\rho}}^{\mathcal{N}}, \mathbf{a}^{\mathcal{N}}; x\right) \partial_x\zeta(t,x)\, dx\, dt\right| \\
&\quad + \left|\int_{\Omega_T} \Phi^{\imath}\left(\rho^{\imath}\right) \left(\partial_x\beta^{\imath}\left(\tilde{\boldsymbol{\rho}}^{\mathcal{N}}, \mathbf{a}^{\mathcal{N}}; x\right) - \partial_x\beta^{\imath}\left(\boldsymbol{\rho}, \mathbf{a}; x\right)\right) \partial_x\zeta(t,x) dx\, dt\right| \\
&= |I| + |II|.
\end{aligned}$$

We now treat the two terms separately. Assumption (**A**) and equation (3.10) ensure the following bound

$$\begin{aligned}
|I| \leq &\sum_{\jmath\in\mathcal{M}} \int_{\Omega_T} \left|\Phi^{\imath}\left(\rho^{\imath,\mathcal{N}}\right) - \Phi^{\imath}\left(\rho^{\imath}\right)\right| \int_{\Omega} \left|\partial_1\mathbf{A}^{\imath\jmath}(x,y;a^{\imath\jmath})\tilde{\rho}^{\jmath,\mathcal{N}}(t,y)\, dy\right| \left|\partial_x\zeta(t,x)\right| \, dx\, dt \\
\leq &C\|\rho^{\imath,\mathcal{N}} - \rho^{\imath}\|_{L^1(\Omega_T)}.
\end{aligned}$$

where C is a constant depending on $\left\|\partial_1\mathbf{A}^{\imath\jmath}\right\|_{\infty}$, $\|\partial_x\zeta\|_{\infty}$, $Lip\left(\Phi^{\imath}\right)$ and $\sigma^{\mathcal{M}}$. In order to bound the second integral, let us introduce $\Pi^{\imath,\mathcal{N}}$ an optimal transport plan between $\tilde{\rho}^{\imath,\mathcal{N}}$ and $\rho^{\imath}$. Then, we have

$$\begin{aligned}
|II| \leq &\sum_{\jmath\in\mathcal{M}} \int_{\Omega_T} \Phi^{\imath}\left(\rho^{\imath}\right) \int_{\Omega^2} \left|\partial_1\mathbf{A}^{\imath\jmath}(x,y;a^{\imath\jmath}_{\mathcal{N}}) - \partial_1\mathbf{A}^{\imath\jmath}(x,z;a^{\imath\jmath})\right| \, d\Pi^{\imath,\mathcal{N}}(y,z)|\partial_x\zeta|\, dx\, dt \\
\leq &C\sum_{\jmath\in\mathcal{M}} \int_{\Omega_T}\int_{\Omega\times\Omega} |y-z| \, d\Pi^{\imath,\mathcal{N}}(y,z)\, dx\, dt \\
\leq &CM|\Omega| \int_0^T d_{W^1}(\tilde{\rho}^{\imath,\mathcal{N}}(t,\cdot), \rho^{\imath}(t,\cdot))\, dt,
\end{aligned}$$

where in this case C is a constant depending on $\|\rho^{\imath}\|_{\infty}$, $\|\partial_x\zeta\|_{\infty}$, and the constant $c_{1,A}$ from Assumption (**A**). The convergences in Proposition 3.6 and Lemma 3.9 ensure that equation (3.12) holds. □

We are now in the position of proving that the limit densities and nodes satisfy the weak formulation in the sense of Definition 2.1. More precisely, we are going to show that for $N \to +\infty$ we have

$$\begin{aligned}
&\int_{\Omega_T} \rho^{\imath,\mathcal{N}}\partial_t\zeta + \rho^{\imath,\mathcal{N}}\theta^{\imath}(\tilde{\boldsymbol{\rho}}^{\mathcal{N}}, \mathbf{a}^{\mathcal{N}}; x)\partial_x\zeta\, dx\, dt \\
&\quad + \int_{\Omega_T} \Phi^{\imath}\left(\rho^{\imath,\mathcal{N}}\right)\left(\partial_x\beta^{\imath}(\tilde{\boldsymbol{\rho}}^{\mathcal{N}}, \mathbf{a}^{\mathcal{N}}; x)\partial_x\zeta + \beta^{\imath}(\tilde{\boldsymbol{\rho}}^{\mathcal{N}}, \mathbf{a}^{\mathcal{N}}; x)\partial_{xx}\zeta\right) \, dx\, dt \to 0
\end{aligned} \tag{3.15}$$

that combined with the convergences in Lemma 3.11 gives the assertion. We state the following

Proposition 3.12. *Given* $M \in \mathbb{N}$, *and* $T > 0$ *fixed, for all* $\imath, \jmath \in \mathcal{M}$ *consider* $\mathbf{A}^{\imath\jmath}$, $\mathbf{K}^{\imath\jmath}$, $\mathbf{V}$ *and* $\Phi^{\imath}$ *under assumptions* (**A**), (**K**), (*V*) *and* (*Dif*), *respectively. Let* $\bar{\rho}^{\imath} : \Omega \to \mathbb{R}$ *under assumptions* (*In*1) *and* (*In*2) *for all* $\imath \in \mathcal{M}$.

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

Then, for all $\imath \in \mathscr{M}$ the densities $\rho^{\imath,\mathscr{N}}$ introduced in equation (2.14) and $\tilde{\rho}^{\imath,\mathscr{N}}$ introduced in equation (3.8) satisfy the condition given by the limit (3.15) as $N \to \infty$, for all $\zeta \in C_0^\infty(\Omega_T)$.

Proof. We start considering the term involving the time derivative. By definition of $\rho^{\imath,\mathscr{N}}$ in equation (2.14), a discrete integration by parts and Fundamental Theorem of Calculus give

$$\begin{aligned}\int_{\Omega_T} \rho^{\imath,\mathscr{N}} \partial_t \zeta &= \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T \rho_k^\imath(t) \int_{I_k^\imath(t)} \partial_t \zeta(t,x)\,dx\,dt \\ &= \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T \rho_k^\imath(t)\dot{x}_{k+1}^\imath(t) \left(\int_{I_k^\imath(t)} \zeta(t,x)\,dx - \zeta(t,x_{k+1}^\imath(t)) \right) dt \\ &\quad - \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T \rho_k^\imath(t)\dot{x}_k^\imath(t) \left(\int_{I_k^\imath(t)} \zeta(t,x)\,dx - \zeta(t,x_k^\imath(t)) \right) dt.\end{aligned}$$

A second-order expansion of ζ around $x_{k+1}^\imath(t)$ in the first average integral and around $x_k^\imath(t)$ in the second average integral produces

$$\begin{aligned}\int_{\Omega_T} \rho^{\imath,\mathscr{N}} \partial_t \zeta = &-\frac{\sigma_N^\imath}{2} \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T \dot{x}_{k+1}^\imath(t) \partial_x \zeta(t,x_{k+1}^\imath)\,dt \\ &+ \frac{1}{2} \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T \rho_k^\imath(t)\dot{x}_{k+1}^\imath(t) \int_{I_k^\imath(t)} \partial_{xx}\zeta(t,\hat{x}_{k+1}^\imath) \left(x - x_{k+1}^\imath(t)\right)^2 dx\,dt \\ &- \frac{\sigma_N^\imath}{2} \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T \dot{x}_k^\imath(t) \partial_x \zeta(t,x_k^\imath)\,dt \\ &- \frac{1}{2} \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T \rho_k^\imath(t)\dot{x}_k^\imath(t) \int_{I_k^\imath(t)} \partial_{xx}\zeta(t,\hat{x}_k^\imath) \left(x - x_k^\imath(t)\right)^2 dx\,dt,\end{aligned}$$

where $\hat{x}_k^\imath$ and $\hat{x}_{k+1}^\imath$ are points in $\left[x_k^\imath, x\right]$ and equation $\left[x, x_{k+1}^\imath\right]$, respectively. We now combine the first and third terms on the r.h.s. above and use (2.4) in order to obtain

$$-\frac{\sigma_N^\imath}{2} \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T \left(\dot{x}_k^\imath(t)\partial_x\zeta(t,x_k^\imath) + \dot{x}_{k+1}^\imath(t)\partial_x\zeta(t,x_{k+1}^\imath) \right) dt = A_1 + A_2,$$

where

$$\begin{aligned}A_1 = -\frac{1}{2} \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T &\left(\beta_k^\imath \left(\Phi^\imath(\rho_{k-1}^\imath) - \Phi^\imath(\rho_k^\imath) \right) \partial_x \zeta(t,x_k^\imath) \right. \\ &\left. + \beta_{k+1}^\imath \left(\Phi^\imath(\rho_k^\imath) - \Phi^\imath(\rho_{k+1}^\imath) \right) \partial_x \zeta(t,x_{k+1}^\imath) \right) dt,\end{aligned}$$

and

$$A_2 = -\frac{\sigma_N^\imath}{2} \sum_{k\in\mathscr{N}\setminus\{N\}} \int_0^T \left(\theta_k^\imath(t)\partial_x\zeta(t,x_k^\imath) + \theta_{k+1}^\imath(t)\partial_x\zeta(t,x_{k+1}^\imath) \right) dt.$$

We now combine the integral A_1 with the two terms involving the diffusion in equation (3.15) in order to show that

$$A_1 + \int_{\Omega_T} \Phi^\imath\left(\rho^{\imath,\mathscr{N}}\right) \partial_x \left(\beta^\imath(\tilde{\boldsymbol{\rho}}^{\mathscr{N}}, \mathbf{a}^{\mathscr{N}}; x) \partial_x \zeta(t,x) \right) dx\,dt = 0.$$

Invoking equation (3.9) and the fact that $\partial_x\zeta(t,x_0^i)=\partial_x\zeta(t,x_N^i)=0$, we can compute

$$
\begin{aligned}
&\int_{\Omega_T}\Phi^i\left(\rho^{i,\mathcal{N}}\right)\partial_x\left(\beta^i(\tilde{\boldsymbol{\rho}}^{\mathcal{N}},\mathbf{a}^{\mathcal{N}};x)\partial_x\zeta(t,x)\right)\,dx\,dt\\
&=\sum_{k\in\mathcal{N}\setminus\{N\}}\int_0^T\Phi^i\left(\rho_k^i\right)\int_{I_k^i(t)}\partial_x\left(\beta^i(\tilde{\boldsymbol{\rho}}^{\mathcal{N}},\mathbf{a}^{\mathcal{N}};x)\partial_x\zeta(t,x)\right)\,dx\,dt\\
&=\sum_{k\in\mathcal{N}\setminus\{N\}}\int_0^T\Phi^i\left(\rho_k^i\right)\left(\beta^i(\tilde{\boldsymbol{\rho}}^{\mathcal{N}},\mathbf{a}^{\mathcal{N}};x_{k+1}^i)\partial_x\zeta(t,x_{k+1}^i)-\beta^i(\tilde{\boldsymbol{\rho}}^{\mathcal{N}},\mathbf{a}^{\mathcal{N}};x_k^i)\partial_x\zeta(t,x_k^i)\right)\,dt\\
&=\sum_{k\in\mathcal{N}\setminus\{N\}}\int_0^T\Phi^i\left(\rho_k^i\right)\left(\beta_{k+1}^i\partial_x\zeta(t,x_{k+1}^i)-\beta_k^i\partial_x\zeta(t,x_k^i)\right)\,dt\\
&=\sum_{k\in\mathcal{N}\setminus\{0,N\}}\int_0^T\beta_k^i\left(\Phi^i(\rho_{k-1}^i)-\Phi^i(\rho_k^i)\right)\partial_x\zeta(t,x_k^i)\,dt,
\end{aligned}
$$

which can be combined with A_1 by shifting the indexes and using again the fact that the test function vanishes at the boundary.

Recalling the definition of θ_k^i in equation (2.8), rearranging the indexes in A_2, using the fact that $\partial_x\zeta$ vanishes at the boundary of Ω and summing with the second term in equation (3.15) we obtain

$$
\begin{aligned}
&\sum_{k\in\mathcal{N}\setminus\{0,N\}}\sum_{i\in\mathcal{M}}\sum_{l\in\mathcal{N}\setminus\{N\}}\sigma_N^i\sigma_N^j\int_0^T\mathbf{K}^{ij}(x_k^i,x_l^j;a_N^{ij})\partial_x\zeta(t,x_k^i)\,dt\\
&+\sum_{k\in\mathcal{N}\setminus\{N\}}\sum_{i\in\mathcal{M}}\sum_{l\in\mathcal{N}\setminus\{N\}}\sigma_N^j\int_0^T\rho_k^i(t)\int_{I_k^i}\mathbf{K}^{ij}(x,x_l^j;a_N^{ij})\partial_x\zeta(t,x)\,dx\,dt.
\end{aligned}
$$

A first-order expansion on $\partial_x\zeta$ around x_k^i for $\hat{\hat{x}}_k^i\in\left[x_k^i,x\right]$, together with the definition of ρ_k^i and assumption (**K**), yields

$$
\begin{aligned}
&\Bigg|\sum_{k\in\mathcal{N}\setminus\{0,N\}}\sum_{i\in\mathcal{M}}\sum_{l\in\mathcal{N}\setminus\{N\}}\sigma_N^j\int_0^T\rho_k^i\partial_x\zeta(t,x_k^i)\int_{I_k^i}\mathbf{K}^{ij}(x_k^i,x_l^j;a_N^{ij})-\mathbf{K}^{ij}(x,x_l^j;a_N^{ij})\,dx\,dt\\
&+\sum_{k\in\mathcal{N}\setminus\{N\}}\sum_{i\in\mathcal{M}}\sum_{l\in\mathcal{N}\setminus\{N\}}\sigma_N^j\int_0^T\rho_k^i(t)\int_{I_k^i}\mathbf{K}^{ij}(x,x_l^j;a_N^{ij})(x-x_k^i)\partial_{xx}\zeta(t,\hat{\hat{x}}_k^i)\,dx\,dt\Bigg|\\
&\le\sum_{k\in\mathcal{N}\setminus\{0,N\}}\sum_{i\in\mathcal{M}}\sum_{l\in\mathcal{N}\setminus\{N\}}\sigma_N^j\left(c_{\mathbf{K}}\|\partial_x\zeta\|_\infty+\|\mathbf{K}^{ij}\|_\infty\|\partial_{xx}\zeta\|_\infty\right)\int_0^T\rho_k^i\int_{I_k^i}|x_k^i-x|\,dx\,dt\\
&=\sum_{k\in\mathcal{N}\setminus\{0,N\}}\sum_{i\in\mathcal{M}}\frac{\sigma^j}{2}\left(c_{\mathbf{K}}\|\partial_x\zeta\|_\infty+\|\mathbf{K}^{ij}\|_\infty\|\partial_{xx}\zeta\|_\infty\right)\int_0^T\sigma_N^i|I_k^i|(t)\,dt\\
&\le\sigma_N^i\sum_{i\in\mathcal{M}}\frac{\sigma^j}{2}\left(c_{\mathbf{K}}\|\partial_x\zeta\|_\infty+\|\mathbf{K}^{ij}\|_\infty\|\partial_{xx}\zeta\|_\infty\right)T|\Omega|,
\end{aligned}
$$

that vanishes as $N\to\infty$ together with σ_N^i. We are now left in showing that the remainder, i.e.

$$
\begin{aligned}
R_{k,k+1}^i=&\frac{1}{2}\sum_{k\in\mathcal{N}\setminus\{N\}}\int_0^T\rho_k^i(t)\dot{x}_{k+1}^i(t)\int_{I_k^i(t)}\partial_{xx}\zeta(t,\hat{x}_{k+1}^i)\left(x-x_{k+1}^i(t)\right)^2\,dx\,dt\\
&-\frac{1}{2}\sum_{k\in\mathcal{N}\setminus\{N\}}\int_0^T\rho_k^i(t)\dot{x}_k^i(t)\int_{I_k^i(t)}\partial_{xx}\zeta(t,\hat{x}_k^i)\left(x-x_k^i(t)\right)^2\,dx\,dt,
\end{aligned}
$$

goes to zero. We first notice that for all $h, k \in \mathcal{N} \setminus \{N\}$, we have

$$\int_{I_k^i(t)} \partial_{xx}\zeta(t, \hat{x}_h^i)\left(x - x_k^i(t)\right)^2 \, dx \le \|\partial_{xx}\zeta\|_\infty \frac{|I_k^i|^3(t)}{3} \le C$$

for some constant $C > 0$, then

$$\begin{aligned}|R_{k,k+1}^i| \le & \frac{\|\partial_{xx}\zeta\|_\infty}{2} \sum_{k \in \mathcal{N} \setminus \{N\}} \int_0^T \rho_k^i(t) \left(|\dot{x}_{k+1}^i(t)| + |\dot{x}_k^i(t)|\right) \frac{|I_k^i|^3(t)}{3} \, dt \\ \le & C\|\partial_{xx}\zeta\|_\infty \sum_{k \in \mathcal{N} \setminus \{N\}} \int_0^T \rho_k^i(t) |I_k^i|^3(t) \, dt \\ = & \sigma_N^i C \|\partial_{xx}\zeta\|_\infty |\Omega| T,\end{aligned}$$

where the second inequality holds in view of Lemma 3.4 and the BV bound on ρ_k^i. □

4. Modelling, simulation and interpretation

4.1. Modelling

As stated in the introduction, the objective of this model is to elucidate various aspects inherent in interactions on social networks and social media. Specifically, it focuses on elements related to network evolution, homophily and heterophobia. The aim is to examine whether polarisation of opinions and fragmentation of the population can be observed solely by modelling the processes governing network rewiring and the epistemic process. It is essential to note that we do not explicitly model polarisation; instead, we adhere to describing the dynamics of opinion formation that underlie the formation of echo chambers and epistemic bubbles.

In recent years, epistemologists have placed particular emphasis on studying and understanding the formation and evolution processes of the aforementioned social structures. Following Nguyen's definition [38], echo chambers and epistemic bubbles are social epistemic structures based on two distinct types of filtering. In the case of epistemic bubbles, filtering occurs at the network level, while for echo chambers, it takes place at the epistemic process level – specifically, in the attitude towards an interlocutor. In this work, we consider an attitude guided by the distance between the opinions of two interacting agents. Nevertheless, recent studies suggest that it is important to also take into account individuals' biases for a more comprehensive analysis of the social dynamics based on distrust leading to social filtering, as demonstrated in the work by Pederneschi [43].

In this work, we introduce the concepts of **attitude areas** and a **Euclidean network** to describe these dynamics.

Attitude Areas – Open-Mindedness

In this section, we model and simulate a possible choice of opinion dynamic. In particular, we focus our attention on a model that takes into account the heterophobia and homophilia dynamics. The model that we propose is based on five attitude areas: *attraction/homophilia*, *curiosity*, *indifference*, *mistrust*, *repulsion/heterophobia*. When two agents interact, their attitude depends on the distance between their respective opinions. We do not consider the attitude depending locally on the opinion itself, this could be a possible improvement that describes the low propensity of changing extreme beliefs. The interaction can be attractive, i.e. the agents move towards a position of consensus looking for a compromise, or can be repulsive, i.e. each agent changes its own opinion moving farther from the one of the other agent. We consider the DPA structure of equations (2.4a), (2.4b), (2.7), (2.8). The operator describing this phenomenon has the following structure

$$\mathbf{K}^{i\,j}(w, v, a^{i\,j}) = \omega(a^{i\,j})\,\zeta(\mu^i - \mu^j)\,(v - w) \tag{4.1}$$

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

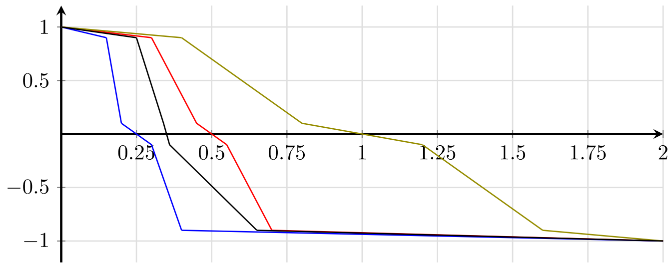


Figure 1. *Attraction/repulsion function.*
The positive values of the function coincide with the attraction, on the other hand, while the function has negative values it describes repulsion between the agents' opinion, which could bring to radicalisation or polarisation. Due to the choice of the domain, we have that $s \in [0, 2]$. *The different colours coincide with the following definitions of the attitude intervals:*
Blue: $r_f = 0.15$, $r_a = 0.20$, $r_r = 0.30$, $r_l = 0.40$.
Black: $r_f = 0.25$, $r_a = 0.34$, $r_r = 0.36$, $r_l = 0.65$.
Red: $r_f = 0.30$, $r_a = 0.45$, $r_r = 0.55$, $r_l = 0.70$.
Olive: $r_f = 0.40$, $r_a = 0.80$, $r_r = 1.20$, $r_l = 1.60$.

where the function ω depends on the network connections, and ζ is the *attitude function* that measures the distance between the agents' mean opinion. Being s the distance between the mean opinions, the five attitude areas coincide with the following intervals

$s \in (0, r_f)$	strong attraction, homophilia	$(r_{friends})$
$s \in (r_f, r_a)$	curiosity	$(r_{attraction})$
$s \in (r_a, r_r)$	indifference	
$s \in (r_r, r_l)$	mistrust	$(r_{repulsion})$
$s \in (r_l, \lvert\Omega\rvert)$	repulsion, heterophobia	(r_{limit}) .

The function ζ is given by

$$\zeta(s) = \begin{cases} 1 - \dfrac{1}{10}\dfrac{|s|}{r_f} & \textit{if} \quad |s| < r_f \\ 0.1 + \dfrac{8}{10}\left[1 - \dfrac{|s| - r_f}{r_a - r_f}\right] & \textit{if} \quad r_f \le |s| < r_a \\ -0.1 + \dfrac{2}{10}\left[1 - \dfrac{|s| - r_a}{r_r - r_a}\right] & \textit{if} \quad r_a \le |s| < r_r \\ -0.9 + \dfrac{8}{10}\left[1 - \dfrac{|s| - r_r}{r_l - r_r}\right] & \textit{if} \quad r_r \le |s| < r_l \\ -0.9 - \dfrac{1}{10}(|s| - r_l) & \textit{if} \quad r_l \le |s|\,, \end{cases}$$

depending on the values of the extreme of the intervals it looks like those in Figure 1. The term **open-mindedness** refers to the individual's inclination to positively consider opinions that differ from their own. In this article, the quantitative description of this term is associated with the value $s \in [0, 2]$ such that $\zeta(s) = 0$. In other words, open-mindedness corresponds to the maximum distance within which an agent regards another opinion favourably.

Diffusion

Moreover, the opinion dynamic is not ruled only by the direct interaction with the connected agents in the network. We continuously get inputs from all the media, this phenomenon is described in this model by the following operator

$$\mathbf{A}^{\imath\jmath}(w, v, a^{\imath\jmath}) = |\mu^{\jmath} - w|^2 . \tag{4.2}$$

In this case, the interaction is not filtered by the network connections, i.e. by ω, and it is not affected by the attitude. This operator plays the role of the diffusion mobility, the opinion tends to diffuse the more the inputs are far from it, and it is not affected by the diffusion when the inputs coincide with the opinion itself. The underlying idea is that we cannot process actively – i.e. through ζ – all the information that we get. Those inputs that we cannot elaborate they influence our opinion distribution smoothing it depending on the distance between the input and our mean opinion.

Evolution of the Network

As anticipated in the introduction, the Euclidean network relies on the distance between agents in $\mathbb{R}^2$. The choice of employing $\mathbb{R}^d$ with $d = 2$ instead of another dimension lacks specific applicative rationale in this case. Should a concrete application be considered, the dimensions of the network space might depend on factors such as biases, as discussed in [43]. In other words, the network space could potentially capture factors of social segregation correlated with the subject of opinion evolution under consideration.

In this paper, the distance would play a singular role, specifically in defining the interaction radius between agents through the function ζ. This function is associated with an interaction resembling that of social networks and epistemic bubbles, where filtering occurs at the network level, and interaction in the network space takes place locally rather than globally. We consider r_{loc} the radius defining the ball of the local interaction,

$$\omega(a^{\imath\jmath}) = \begin{cases} 1 & if \quad a^{\imath\jmath} \leq \rho_{loc} \\ 0 & if \quad a^{\imath\jmath} > \rho_{loc} \end{cases} . \tag{4.3}$$

The agents $a^{\imath}$ and $a^{\jmath}$ interact if their distance on the network is smaller or equal to ρ_{loc}. The global interaction coincides with the radius $\rho_{loc} = +\infty$. The initial condition used for our simulation is the one given in Figure 2.

Considering 3 different radii of interaction the network connections change as in Figure 3.

The evolution of the agents is ruled by the following operator

$$\mathbf{V}(\mu_{\rho^{\imath}}, \mu_{\rho^{\jmath}}, a^{\imath\jmath}) = \sum_{l=1}^{2} \sum_{j} \zeta\, |\mu_{\rho^{\imath}} - \mu_{\rho^{\jmath}}|\, \omega(a^{\imath\jmath})\, (a^{\imath} - a^{\jmath})\mathbf{e}_l\, , \tag{4.4}$$

with $\mathbf{e}_l$ being the j-th normal vector. In this scenario, we have three terms to consider. The function ω evaluates the distance between two agents in the network space; if this distance is not greater than ρ_{loc}, interaction occurs between the agents. The function ζ determines whether there is attraction or repulsion and assesses the respective intensity by evaluating the distance between the opinion means of the two agents. The last term defines both the intensity and the direction of movement in the network space. Both the last term and the attitude function play a role in determining the interaction intensity. However, their roles should not be confused: the attitude function pertains to the distance between opinions, while the third term depends on the network.

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

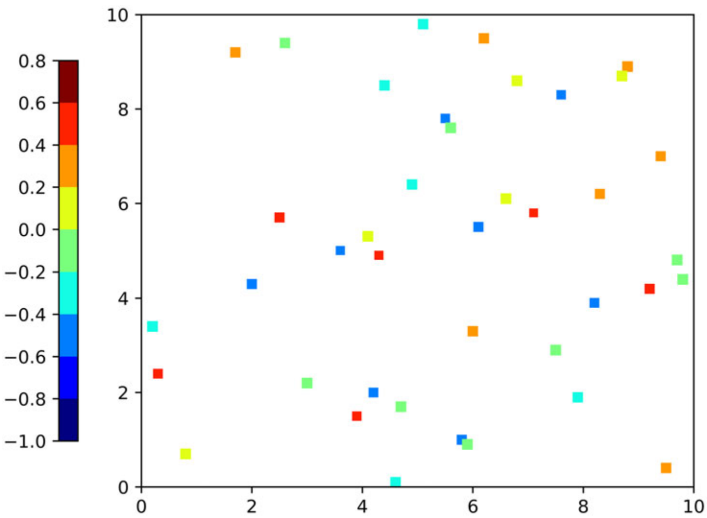


Figure 2. *Initial network condition.*
In this figure, the initial agent's coordinates belong to $[0, 10]^2$. *The colours describe the mean opinion of each agent, which belongs to the interval* $[-1, 1]$. *The number of agents is* $N = 40$. *The agent's coordinates are uniformly randomly distributed on each axis. The dimension of each square is proportional to the social strength* σ^i *of each agent, and in this case, they all almost coincide.*

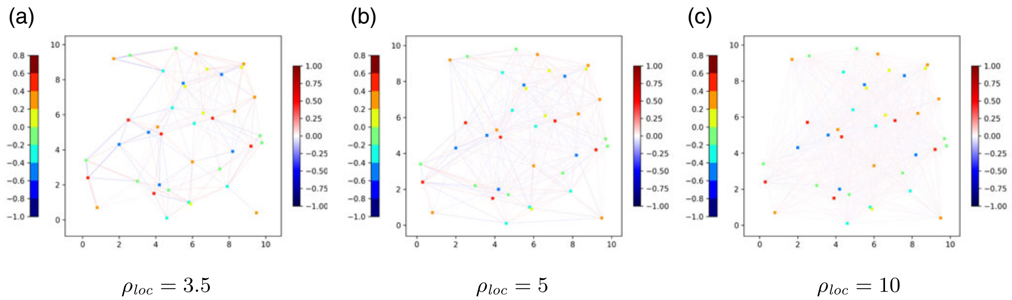


Figure 3. *Local initial network interaction.*
The agents interact if they are connected by a link. The magnitude and the sign of the connection range from $[-1, 1]$ *and are described by the legend on the right of the pictures. In this case, the attitude areas are given by the parameters* $r_f = 0.25$, $r_a = 0.34$, $r_r = 0.36$, $r_l = 0.65$, *i.e. the black function in Figure 1.*

4.2. Simulation

Radicalisation

The term radicalisation denotes an agent's inclination to anchor their opinion within a defined range of opinions. In our model, this phenomenon is observed when the opinion distribution increases within an interval and decreases sharply outside of it.

Remark 4.1. *Note that this dynamics is not observable if one opts for a description that does not consider the opinion distribution but only a singular value.*

We consider the attitude areas given by the following intervals: $r_f = 0.25$, $r_a = 0.34$, $r_r = 0.36$, $r_l = 0.65$. This setting coincides with the ζ function described by the black line in Figure 1.

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

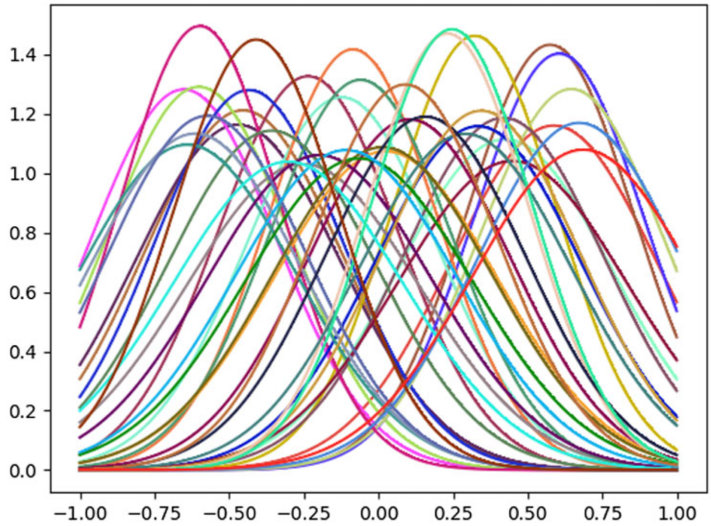


Figure 4. *Initial opinion distribution.*
In this picture are represented the opinion distributions at initial time of the 40 agents considered for the simulation. Each distribution is described by a truncated Gaussian function, mean and variance of the Gaussian functions are independently uniform random distributed respectively in the intervals $[-0.7, 0.7]$ *and* $[0.07, 0.15]$.

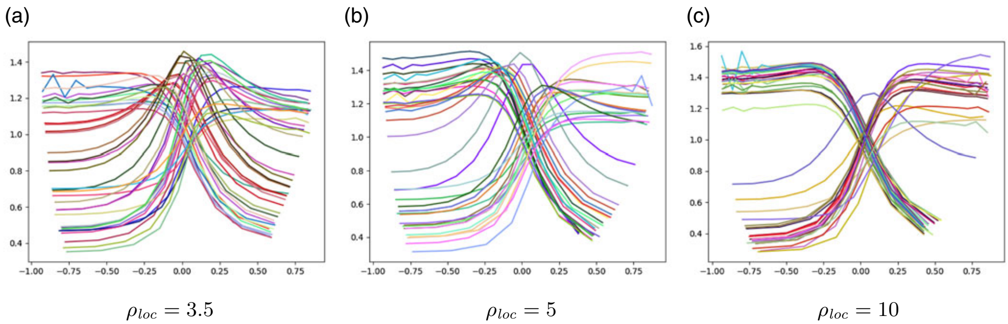


Figure 5. *Final opinion distribution.*
We observe how the distributions are more and more concentrated either on the positive or negative side as the radius increases. Due to the diffusion, the distributions tend to flatten once that they are concentrated on one of the two sides.
The sharp oscillations close to the extreme values are due to the low resolution of the numerical partition of Ω.

Given the initial opinion distribution as in Figure 4, we can observe that the distribution evolves towards a more and more radicalised society as the radius increases. As evident from Figure 5, contrary to common intuition, increasing the interaction radius—thus engaging with a greater number of agents—leads to distributions becoming more concentrated around either positive or negative values. As the interaction radius expands, fewer agents exhibit distributions encompassing both positive and negative values.

The same phenomenon is described by the histogram of the mean opinions distribution in Figure 6.

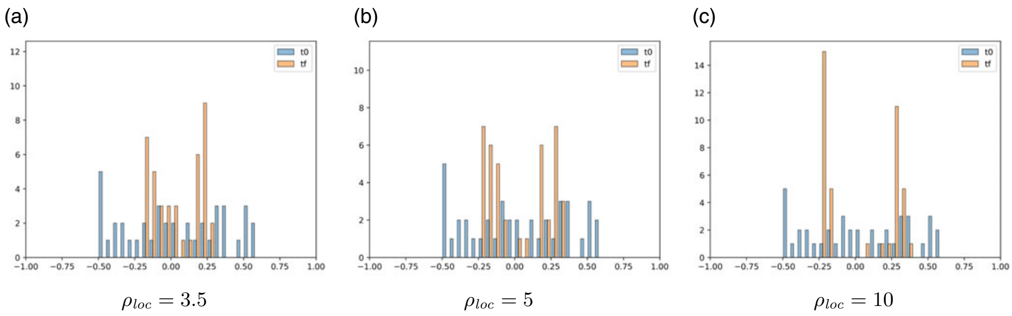


Figure 6. *Initial and final mean opinions distribution.*
In this figure, in blue the distribution of the mean opinions at time $t = 0$, and in orange the mean opinions' distribution at final time (which corresponds to the time showing a quasi-stable status of the simulation result). We observe that the final distribution tends to have two peaks, which means that the opinions of the population are more and more split into two opinion groups. However, they are also more close to the centre. This means that we observe a sort of fragmentation and radicalisation, but there is no polarisation.

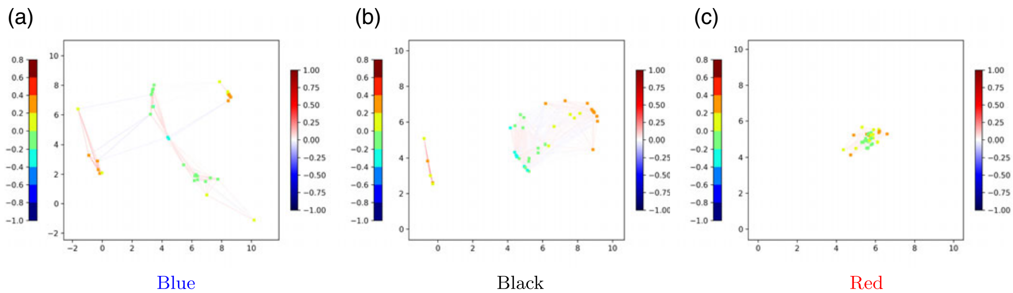


Figure 7. *Final network distribution.*
The olive function has not been plotted because it describes an extreme behaviour, all the agents collapse very fast into a unique point.
Blue: $r_f = 0.15,\ r_a = 0.20,\ r_r = 0.30,\ r_l = 0.40.$
Black: $r_f = 0.25,\ r_a = 0.34,\ r_r = 0.36,\ r_l = 0.65.$
Red: $r_f = 0.30,\ r_a = 0.45,\ r_r = 0.55,\ r_l = 0.70.$

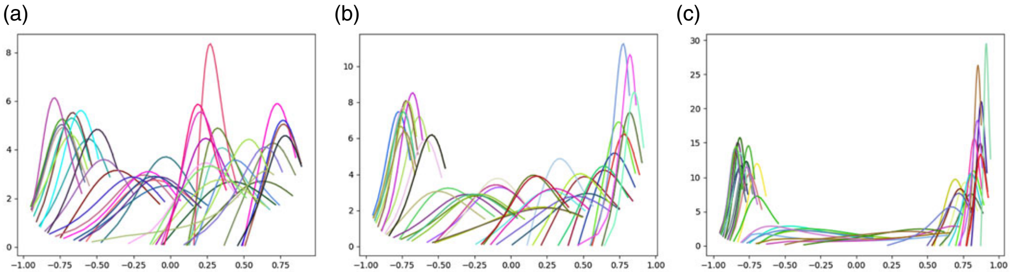
Opinion Polarisation and Network Fragmentation

It is intriguing to examine the final network while varying the parameters of the attitude areas. We set the interaction radius as $\rho_{loc} = 5$, and we compare the first three functions depicted in Figure 1. In this scenario, we note that a more open-minded society corresponds to a less fragmented network, as illustrated in Figure 7.

While focusing solely on the operator $\mathbf{K}^{ij}$ by setting $\mathbf{A}^i = 0$, we observe the phenomenon of polarisation, i.e. the tendency of agents towards an opinion distribution centred on more extreme values, in our case ± 1.

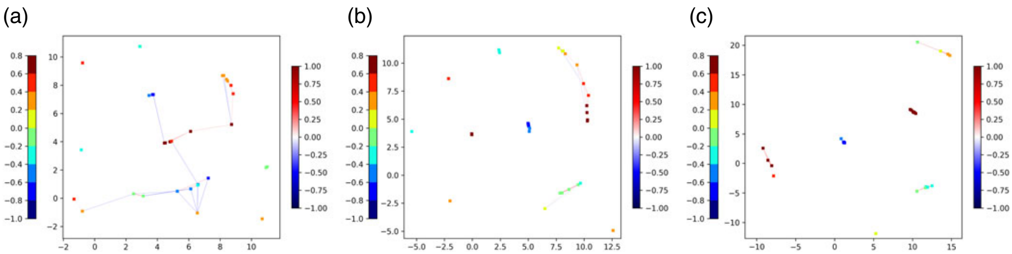
It is interesting to observe how the connectivity of the population is the real driver of the polarisation, instead of the *open-mindedness*. We now compare the results of the simulation with fixed open-mindedness. We chose the most close-minded population, i.e. the one described by the blue function in Figure 1. Increasing the radius of the interaction, the opinion gets more and more polarised. In Figure 8, we notice that a more connected society tends to cluster into extreme opinions, while a less

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press



$\rho_{loc} = 3.5$ $\rho_{loc} = 5$ $\rho_{loc} = 10$

Figure 8. *Polarisation while increasing the radius of interaction. Attitude function Blue: $r_f = 0.15$, $r_a = 0.20$, $r_r = 0.30$, $r_l = 0.40$.*



$\rho_{loc} = 3.5$ $\rho_{loc} = 5$ $\rho_{loc} = 10$

Figure 9. *Network fragmentation and opinion homogeneity.*
Network while increasing the radius of interaction and keeping the same attitude function, i.e. Blue: $r_f = 0.15$, $r_a = 0.20$, $r_r = 0.30$, $r_l = 0.40$.

connected society keeps a more sparse distribution of opinions. This is a not expected behaviour, usually a large or global interaction is related to a higher consensus. In the work by Tucker et al. [53], it is emphasised how in the literature on opinion polarisation, the crucial role of social networks in accelerating and intensifying the polarisation of opinions on highly debated topics online is now well-established. This dynamic aligns with the outcomes of our simulations, given that the operator $\mathbf{K}^{ij}$ describes the interaction process on social networks.

If we observe the evolution of the network, it seems that increasing the radius of interaction does not really affect the fragmentation of the population, but it plays a role in the opinion homogeneity of the network clusters. In Figure 9, at the same fixed time, a larger radius of interaction brings to a wider network space, and the groups of connected agents show a stronger homogeneity of the opinion.

4.3. Conclusions, interpretations and possible follow-up

The goal of this model is to introduce the study of the processes describing the interaction on social networks and social media. We mainly focused on the role of the attitude areas and on that of the radius of interaction. The dynamics ruling the opinion formation of agents interacting on social platforms is different from the one described by models based on *alignment*, *averaged consensus* or Cucker-Smale with positive communication rate.

https://doi.org/10.1017/S0956792524000093 Published online by Cambridge University Press

Recently, sociologists and philosophers described the epistemic processes of the hyperconnected society typical of the last two decades. The high amount of interactions and notions, together with their high frequency, modified the way how we create and reinforce our beliefs. Authors like Nguyen, see [38], explain how the network and the opinion distance are the discriminant for different epistemic processes. In our model, we describe these two aspects through the definition of the attitude areas and through the dynamic of the network, which takes into account the distance on the network and the distance of the opinions.

Drawing on the literature, e.g. [5, 8, 34, 53], we observe that describing interactions through attitude areas and Euclidean network suggests that the approach is heading in the right direction to capture the dynamics of polarisation, radicalisation and fragmentation that arise through interactions on social networks and social media. A particularly intriguing phenomenon to observe is the fragmentation of society into numerous 'mono-opinion' groups.

Future steps could involve assuming that the social strength σ^i of each agent evolves over time based on connections. Additionally, in our case, we used a single function ζ for the entire population; however, it would be interesting to consider a different function for each agent, reflecting varying levels of open-mindedness. Numerous possible avenues exist, but caution should be exercised to avoid overexposing the problem to parameters and external choices. The authors' intention is to continue in the direction of general models based on epistemological theories and interacting group theories. Certainly, an intriguing insight for future work comes from Pederneschi [43] concerning the possibility of introducing biases to explain the dynamics of trust and distrust.

Financial support. The research of SF is supported by the Ministry of University and Research, Italy under the grant PRIN 2020 – Project N. 20204NT8W4, Nonlinear Evolutions PDEs, fluid dynamics and transport equations: theoretical foundations and applications. The research of SF and GF is supported by the Italian INdAM project N. E55F22000270001 'Fenomeni di trasporto in leggi di conservazione e loro applicazioni'. SF is also supported by the University of L'Aquila 2021 project 04ATE2021 – 'Mathematical Models For Social Innovations: Vehicular And Pedestrian Traffic, Opinion Formation And Seismology'.

Competing interests. There are no competing interests.

References

[1] Albi, G., Pareschi, L. & Zanella, M. (2016) On the optimal control of opinion dynamics on evolving networks. In: *System Modeling and Optimization. CSMO 2015. IFIP Advances in Information and Communication Technology*, Vol. **494**, Springer, Cham, pp. 58–67.

[2] Albi, G., Burger, M., Haskovec, J., Markowich, P. & Schlottbom, M. (2016) Continuum modeling of biological network formation. In: *System Modeling and Optimization. CSMO 2015. IFIP Advances in Information and Communication Technology*, Vol. **494**, Springer, Cham, pp. 58–67.

[3] Albi, G., Pareschi, L., Toscani, G. & Zanella, M. (2017) Recent advances in opinion modeling: Control and social influence. In: Bellomo, N., Degond, P. & Tadmor, E. (eds) *Active Particles, Volulme 1. Advances in Theory, Models, and Applications*, Springer, Birkhäuser, pp. 49–98.

[4] Ambrosio, L., Gigli, N. & Savaré, G. (2008) *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed., Birkhäuser Verlag, Basel, Lectures in Mathematics ETH Zurich.

[5] Begby, E. (2022) From belief polarization to echo chambers: A rationalizing account. In: Episteme, pp. 1–21.

[6] Bellomo, N., Ajmone Marsan, G. & Tosin, A. (2013) *Complex Systems and Society. Modeling and Simulation*, Springer, New York, NY, SpringerBriefs in Mathematics.

[7] Ben-Naim, E. (2005) Opinion dynamics: Rise and fall of political parties. *Europhys. Lett.* **69**(5), 671–677.

[8] Benatti, A., de Arruda, H. F., Silva, F. N., Comin, C. H. & da Fontoura Costa, L. (2020) Opinion diversity and social bubbles in adaptive Sznajd networks. J. Stat. Mech. Theory Exp. 2020(2), 023407.

[9] Bernecker, S. (2021) An epistemic defense of news abstinence. In: *The Epistemology of Fake News*, Oxford University Press. ISBN: 9780198863977.

[10] Borra, D. & Lorenzi, T. (2013) A hybrid model for opinion formation. *Z. Angew. Math. Phys.* **64**(3), 419–437.

[11] Bouchut, F., Golse, F. & Pulvirenti, M. (2000) *Kinetic equations and asymptotic theory*. Series in Applied Mathematics, Vol. **4**, Gauthier-Villars, Paris

[12] Boudin, L. & Salvarani, F. (2009) A kinetic approach to the study of opinion formation. *ESAIM Math. Modell. Numer. Anal.* **43**, 507–522.

[13] Burger, M. (2022) Kinetic equations for processes on co-evolving networks. *Kinet. Relat. Models* **15**(2), 187–212.

[14] Burger, M. (2021) Network structured kinetic models of social interactions. *Vietnam J. Math.* **49**(3), 937–956.

[15] Burger, M., Caffarelli, L. & Markowich, P. A. (2014) Partial differential equation models in the socioeconomic sciences. *Philos. Trans. R. Soc.* **372**, 20130406.

[16] Carrillo, J. A. & Toscani, G. (2005) Wasserstein metric and large–time asymptotics of non-linear diffusion equations. In: *New Trends in Mathematical Physics, (In Honour of the Salvatore Rionero 70th Birthday)*, pp. 234–244.

[17] Castellano, C., Fortunato, S. & Loreto, V. (2009) Statistical physics of social dynamics. *Rev. Modern Phys.* **81**(2), 591–646.

[18] Cercignani, C. (1969) *Mathematical Methods in Kinetic Theory*, Plenum Press, New York.

[19] Coppini, F., Dietert, H. & Giacomin, G. (2020) A law of large numbers and large deviations for interacting diffusions on Erd osRenyi graphs. *Stoch. Dyn.* **20**, 2050010.

[20] Delattre, S., Giacomin, G. & Lucon, E. (2016) Lucon a note on dynamical models on random graphs and FokkerPlanck equations. *J. Stat. Phys.* **165**(4), 785–798.

[21] Di Francesco, M., Fagioli, S. & Radici, E. (2019) Deterministic particle approximation for nonlocal transport equations with nonlinear mobility. *J Differ. Equations* **266**(5), 2830–2868.

[22] Di Francesco, M., Fagioli, S. & Rosini, M. D. (2017) Deterministic particle approximation of scalar conservation laws. *Boll. Unione Mat. Ital.* **10**(3), 487–501. ISSN: 1972-6724.

[23] Di Francesco, M. & Rosini, M. D. (2015) Rigorous derivation of nonlinear scalar conservation laws from follow-the-leader type models via many particle limit. *Arch. Ration. Mech. Anal.* 217, 831–871.

[24] Di Francesco, M. & Stivaletta, G. (2020) Convergence of the follow-the-leader scheme for scalar conservation laws with space dependent flux. *Discrete Contin. Dyn. Syst. A* **40**(1), 233–266.

[25] Fagioli, S. & Radici, E. (2020) Opinion formation systems via deterministic particles approximation. *Kinet. Relat. Models* **14**(1), 45–76.

[26] Fagioli, S. & Radici, E. (2018) Solutions to aggregation-diffusion equations with nonlinear mobility constructed via a deterministic particle approximation. *Math. Mod. and Meth. in App. Sci.* **28**(09), 1801–1829.

[27] Fagioli, S. & Tse, O. (2022) On gradient flow and entropy solutions for nonlocal transport equations with nonlinear mobility. *Nonlinear Anal.* **221**, 112904.

[28] Galam, S. (2012) *Sociophysics: A Physicists Modeling of Psycho-Political Phenomena (under-Standing Complex Systems)*, New York, NY, Springer.

[29] Gosse, L. & Toscani, G. (2006) Identification of asymptotic decay to self-similarity for one-dimensional filtration equations. *SIAM J. Numer. Anal.* **43**(6), 2590–2606.

[30] Klein, A., Ahlf, H. & Sharma, V. (2015) Social activity and structural centrality in online social networks. *Telemat. Inform.* **32**(2), 321–332.

[31] Kohne, J., Gallagher, N., Kirgil, Z. M., Paolillo, R., Padmos, L. & Karimi, F. (2020) The role of network structure and initial group norm distributions in norm conflict. In: *Computational Conflict Research*, Springer, Cham, pp. 113–140.

[32] Lackey, J. (2021) Echo Chambers, Fake News, and Social Epistemology.

[33] Lavenant, H. & Maury, B. (2019) Opinion propagation on social networks: A mathematical standpoint. In: preprint, p. 53.

[34] Lim, S. & Bentley, P. J. (2022) Opinion amplification causes extreme polarization in social networks. *Sci. Rep.-UK* **12**(1), 18131. doi: 10.1038/s41598-022-22856-z.

[35] Motsch, S. & Tadmor, E. (2014) Heterophilious dynamics enhances consensus. *SIAM Rev.* **56**(4), 577–621.

[36] Naldi, G., Pareschi, L. & Toscani, G. (2010) *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, Springer, New York.

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

[38] Thi Nguyen, C. (2020) Echo chambers and epistemic bubbles. *Episteme* **17**(2), 141–161.

[39] Nigam, A., Shin, K., Bahulkar, A., Hooi, B., Hachen, D., Szymanski, B. K., Faloutsos, C. & Chawla, N. V. (2018) ONE-M: modeling the co-evolution of opinions and network connections. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 122–140.

[40] Nugent, A. J., Gomes, S. N. & Wolfram, M.-T. (2023) On evolving network models and their influence on opinion formation. *Phys. D Nonlinear Phenom.* **456**, 133914. doi: 10.1016/j.physd.2023.133914.

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

[42] Pareschi, L. & Toscani, G. (2014) Wealth distribution and collective knowledge: A Boltzmann approach. *Philos. Trans. R. Soc. A.* **372**(2028), 20130396. doi: 10.1098/rsta.2013.0396.

[43] Pederneschi, A. (2023) An analysis of bias and distrust in social hinge epistemology. *Philos. Psychol.*, 1–20. doi: 10.1080/09515089.2023.2247010.

[44] Rosenstock, S., Bruner, J. & O'Connor, C. (2017) In epistemic networks, is less really more? *Philos. Sci.* **84**(2), 234–252.

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

[46] Russo, G. (1990) Deterministic diffusion of particles. *Commun. Pure Appl. Math.* **43**(6), 697–733.

[47] Santambrogio, F. (2015) *Optimal Transport for Applied Mathematicians*, Birkhäuser Verlag, Basel, Progress in Nonlinear Differential Equations and Their Applications, Vol. **86**.

[48] Slanina, F. & Lavička, H. (2003) Analytical results for the Sznajd model of opinion formation. *Eur. Phys. J. B* **35**(2), 279–288.

[49] Thurner, S., Klimek, P. & Hanel, R. (2018) *Introduction to the Theory of Complex Systems*, Oxford, online edn, Oxford Academic. doi: 10.1093/oso/9780198821939.001.0001.

[50] Strogatz, S. H. (2001) Exploring complex networks. *Nature* **410**(6825), 268–276.

[51] Sznajd-Weron, K. & Sznajd, J. (2000) Opinion evolution in closed community. *Int. J. Mod. Phys. C* **11**(06), 1157–1165.

[52] Toscani, G. (2006) Kinetic models of opinion formation. *Comm. Math. Sci.* **4**(3), 481–496.
[53] Tucker, J. A., Guess, A. M., Barberá, P., Vaccari, C., Siegel, A. A., Sanovich, S., Stukal, D. K. & Nyhan, B. (2019) Social media, political polarization, and political disinformation: A review of the scientific literature. *SSRN Electron. J.* Available at https://ssrn.com/abstract=3144139
[54] Tur, E. M. & Azagra-Caro, J. M. (2018) The coevolution of endogenous knowledge networks and knowledge creation. *J. Econ. Behav. Organ.* **145**, 424–434.
[55] Villani, C. (2003) *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. **58**, American Mathematical Society, Providence, RI
[56] Weatherall, J. O. & O'Connor, C. (2021) Endogenous epistemic factionalization. *Synthese* **198**(Suppl 25), 6179–6200.
[57] Yardi, S., Romero, D., Schoenebeck, G. & boyd, d. (2009) Detecting spam in a Twitter network. *First Monday* **15**(1). doi: 10.5210/fm.v15i1.2793

Cite this article: Fagioli S. and Favre G. (2024). Opinion formation on evolving network: the DPA method applied to a nonlocal cross-diffusion PDE-ODE system. *European Journal of Applied Mathematics*, **35**, 748–775. https://doi.org/10.1017/S0956792524000093