Feedback Loops in Opinion Dynamics of Agent-Based Models with Multiplicative Noise

Nataša Djurdjevac Conrad*[a], Jonas Köppl[a,c], and Ana Djurdjevac[b]

[a]Zuse Institute Berlin, Berlin, Germany
[b]Freie Universität Berlin, Institut für Mathematik und Informatik, Berlin, Germany
[c]Weierstrass Institute for Applied Analysis and Stochastics Berlin, Germany

Abstract

We introduce an agent-based model for co-evolving opinion and social dynamics, under the influence of multiplicative noise. In this model, every agent is characterized by a position in a social space and a continuous opinion state variable. Agents' movements are governed by positions and opinions of other agents and similarly, the opinion dynamics is influenced by agents' spatial proximity and their opinion similarity. Using numerical simulations and formal analysis, we study this feedback loop between opinion dynamics and mobility of agents in a social space. We investigate the behavior of this ABM in different regimes and explore the influence of various factors on appearance of emerging phenomena such as group formation and opinion consensus. We study the empirical distribution and in the limit of infinite number of agents we derive a corresponding reduced model given by a partial differential equation (PDE). Finally, using numerical examples we show that a resulting PDE model is a good approximation of the original ABM.

Keywords: opinion dynamics, feedback loop, agent-based modeling, multiplicative noise, group formation, mean-field limit, empirical distribution, stochastic partial differential equations.

1 Introduction

Opinion dynamics is one of the most important processes of our society, as our opinions influence not only our individual actions and behavior; but can also shape collective dynamics governing societal change and social movement. Complex interaction patterns between individuals and coupled social mechanisms in different environments are shown to be the crucial drivers of opinion dynamics [1]. With the introduction of online social media, the way people interact and share their opinions has drastically changed. For example, physical proximity is now not anymore a constrain for communication, everybody can engage in information transmission and express their opinions to a large number of people in different social, political and cultural environment. Additionally, large amounts of data became available about how people influence and are getting influenced in their opinions [2], which provided new insights into the social mechanisms and emerging phenomena such as formation of echo chambers and opinion consensus.

During the last decades, extensive research has been carried out in order to understand how people shape their opinions in their social space, see recent reviews [3, 4]. Governed by an increasing amount of available large-scale social data and fast computational advances, the topic of opinion dynamics gathered an interdisciplinary research community [3, 5]. Existing work ranges from the studies on (1) *model-driven approaches* that produce formal models for opinion dynamics that can be analysed using theories from mathematics and statistical physics to (2) *data-driven approaches* that are used to explore empirical data using knowledge from social sciences. Using computer simulations and computational analysis, opinion dynamics models can be used as a tool for understanding social mechanisms, uncovering social interaction patterns and exploring influences of various factors on e.g. group formation and opinion consensus [6, 7, 3, 8]. Furthermore, mathematical description of models is a starting point that enables the use of analytical tools. In such a way we can get theoretical predictions of the models, such as: long time behaviour, limiting behaviour of the system when the number of agents tends to infinity (macroscale) and description of the fluctuations in the case when the number of agents is very big but still finite (mesoscale). In addition, this type of analysis is a basis for developing rigorous numerical analysis which implies error estimates that should be expected by numerical computations and which determine the choice of parameters in the model that should be used in the experiments. However, most existing models are rather simple and rarely connect to empirical studies and available real-world data [9]. In order to close this gap between model- and data-driven approaches, new formal models need to be introduced that can better represent real-world social systems and capture

---

*natasa.conrad@zib.de

complex mechanisms that govern how people shape their opinions.

The largest group of formal models for studying opinion dynamics are agent-based-models (ABMs), where a process of opinion formation takes place through interactions between individual agents. One example of such ABMs are *Voter models* [10], that describe opinion change between agents with discrete opinion states, where agents interaction dynamics is defined through an underlying social network. In the DeGroot model [11], agents have continuous-valued opinions that are formed as the average opinions of all other agents. Further mathematical literature mostly focuses on *bounded confidence models* [12], where the dynamics of opinions depends only on interactions among agents that have similar opinions, without assuming an underlying interaction network, but rather all-to-all possible interactions. Different extensions of these models have been considered in order to account for more realistic scenarios. For example, introducing stochastic effects [13, 14, 15, 16, 6, 17], complex interaction mechanisms [18, 19] and different types of agents (e.g. stubborn agents, influencers, campaigners) [20, 21]. Additionally, in the context of complex social networks, the co-evolution of the opinion and the network dynamics has been studied, where the changes in the network structure influence the opinion dynamics and vice versa. It has been shown that this co-evolution process in network models governs the appearance of emerging structures, e.g. echo chambers [22]. However, most of the existing ABMs for opinion dynamics do not include the dynamics of agents in a social space, despite that this process determines agents' interaction patterns. Extending on the models for epidemic spreading [23] and cultural dissemination [24, 25], recently the so-called *mobile agents* [26, 7] were introduced to account for the feedback between the spatial movement of agents and the social contagion dynamics. While simulation results on co-evolving dynamics (both for network models and ABMs) provided many useful insights into the social mechanisms behind opinion formation, theoretical considerations of such models are still largely missing.

In this article, we introduce a mathematical ABM, that includes feedback loops in opinion dynamics of agents moving in a social space, influencing and being influenced in their opinions. The focus of this manuscript is on the non-trivial two-way interaction between the agents' movement in a social space and their opinion dynamics. More precisely, agents' spatial movements can induce changes in opinion states over time and additionally, opinion dynamics can influence the spatial position of agents and opinion states of agents in their vicinity. This feedback loop between spatial and opinion changes is at the core of the systems co-evolving dynamics. We consider that both spatial and opinion dynamics are governed by stochastic dynamics with multiplicative noise, which generalizes the case of constant, additive noise that is usually considered [14, 15, 16, 6, 17]. We explore the impact the feedback has on the behavior of the system and in particular on the grouping of agents in opinion and/or social space. From the mathematical point of view, our ABM can be seen as an interacting particle system. In particular, on the microscopic level (level of agents) we formulate our model as a system of coupled stochastic differential equations (SDEs) with multiplicative noise. From the application point of view, the interest is to have weak regularity assumptions on the drift and diffusion coefficients. Next, we study the corresponding limiting equation in the case when the number of agents tends to infinity, i.e. the so called McKean-Vlasov equation. Furthermore, we show the well-posedness results for the limiting system, which is a very popular and challenging topic in the field of stochastic analysis. Nowadays, there is extensive literature about these results, the standard results are [27, 28, 29] and see the references therein. Another interesting class of results that is discussed in the literature and that we also consider, is the so-called propagation of chaos, meaning that one wants to prove the convergence of the microscopic model to McKean-Vlasov SDEs. The challenge in these proofs lies in the weak assumptions on the coefficients. Since the topic of this article is not the theoretical investigation of the weak regularity assumption, in order to illustrate our message we concentrate on the simple case of Lipschitz bounded coefficients with multiplicative noise case. Furthermore, due to the high computational cost of ABM simulations when the number of agents is large, we suggest the standard model reduction approach that considers instead the empirical density rather than each agent individually. We derive the formal equations of the empirical density and its so called hydrodynamic limit. These results are in the spirit of the standard, so called Dean-Kawasaki equation [30]. In the setting of the social dynamics, this model reduction for the uncoupled system has already been considered in [31]. Using a numerical example, we illustrate the expected behaviour of the system on the macroscopic scale, that is given by the partial differential equation.

The article is organized as follows. Our agent-based model for opinion dynamics with feedback loops is introduced and studied through numerical simulations in Section 2. Next, we develop a theoretical framework for studying the system at the macroscopic level by an mean-field approach in Section 3 and we present the well-posedness result of the McKean-Vlasov SDE system with Lipschitz coefficients and the convergence results of propagation of chaos. In Section 4 we present formal derivation of the equation that describes the dynamics of the empirical measure and its hydrodyanamic limit. We illustrate the limiting behaviour of the system on the macroscopic scale, by a numerical example. Finally, we derive our conclusions and possible future directions in Section 5.

2 Model Description

We consider a closed system of N interacting agents and agents' co-evolving opinion and social dynamics. At time $t \in [0,T]$, every agent k, $k = 1, \ldots, N$ has a position state $X_t^k \in \mathbb{R}^d$ and an opinion state $\Theta_t^k \in \mathbb{R}$. The position state X_t^k of an agent is a point in an abstract social space, such that a distance between two agents refers to their social proximity that is described by their social similarity. In real-world social systems, information about a position in social space may be inferred from e.g. online social media. The opinion Θ_t^k of an agent k is considered to be a continuous variable. For more generality, this model can be extended to incorporate several opinion entities, such that $\Theta_t^k \in \mathbb{R}^m$. However, for technical simplicity in this paper we will assume that $m = 1$. The state of the system at time $t \geq 0$ for the set of N agents is given by

$$Z_t = (X_t, \Theta_t) \in (\mathbb{R}^d)^N \times \mathbb{R}^N,$$

where the k-th row of the systems' state corresponds to the state $Z_t^k = (X_t^k, \Theta_t^k)$ of the $k-$th agent. All agents follow the same rules that describe how their positions and opinions change. More precisely, agents move in a social space governed by the position of other agents and their opinions. Similarly, the opinion states of agents are influenced by both agents' spatial proximity and their opinions. This feedback loop between spatial and opinion changes determines the systems adaptive dynamics. Additionally, to be able to account for external influences on agents and the sometimes seemingly random nature of human interactions, we model this system via a coupled system of stochastic differential equations with multiplicative noise of the form

$$\begin{cases} dX_t & = \tilde{U}(X_t, \Theta_t)dt + \sigma_{sp}(X_t, \Theta_t)dB_t^{sp}, \\ d\Theta_t & = \tilde{V}(X_t, \Theta_t)dt + \sigma_{op}(X_t, \Theta_t)dB_t^{op}, \end{cases} \tag{1}$$

where:

- $\tilde{U} : (R^d)^N \times \mathbb{R}^N \to (\mathbb{R}^d)^N$ is a *spatial interaction map* that models how the positions and opinions of the agents influence the spatial movement of the agents,

- $\tilde{V} : (R^d)^N \times \mathbb{R}^N \to (\mathbb{R}^d)^N$ is an *opinion interaction map* that models how the positions and opinions of the agents influence the opinion states of the agents,

- B^{sp} and B^{op} are independent Brownian motions starting in 0,

- $\sigma_{sp}(X_t, \Theta_t), \sigma_{op}(X_t, \Theta_t)$ are diffusion coefficients for spatial and opinion dynamics respectively.

Dynamics given by the SDEs in (1) is rather abstract and in this form it doesn't provide much intuition on how it can be adapted to known social mechanisms coming from real-world systems. Thus, in the following we will focus on how complex interaction patterns and stochastic influences can be enforced in our model.

2.1 Pairwise interactions

We start by exploring the simplest type of interaction dynamics, namely pairwise (or 2-body) interactions that only take into account interactions between pairs of agents. More precisely, we consider the case where the interaction maps $\tilde{U}, \tilde{V}$ are linear functions of simpler interaction maps U, V, i.e. we define $\tilde{U} = (\tilde{U}_1, ..., \tilde{U}_N) : (R^d)^N \times \mathbb{R}^N \to (\mathbb{R}^d)^N$ with

$$\tilde{U}_k(X, \Theta) = \frac{1}{N} \sum_{j=1}^N U(X^k, X^j, \Theta^k, \Theta^j), \tag{2}$$

for some pair-interaction map $U : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^d$ and analogously for $\tilde{V}$. In this model, agents shape their opinions based on the mean opinion of other agents through pairwise interactions, which is the setting of many of the classical models for opinion dynamics [11, 32, 12]. Thus, the dynamics of the k-th agent are given by

$$\begin{cases} dX_t^k = \frac{1}{N} \sum_{j=1}^N U(X_t^k, X_t^j, \Theta_t^k, \Theta_t^j)dt + \sigma_{sp}(X_t, \Theta_t)dB_t^{sp,k}, \\ d\Theta_t^k = \frac{1}{N} \sum_{j=1}^N V(X_t^k, X_t^j, \Theta_t^k, \Theta_t^j)dt + \sigma_{op}(X_t, \Theta_t)dB_t^{op,k}. \end{cases} \tag{3}$$

Note that in order to simplify the notation, we do not write $X_t^{k,N}$, $\Theta_t^{k,N}$, but instead X_t^k, Θ_t^k respectively. As a more concrete example one can consider the following model that can be seen as an extension of a classical model by DeGroot [11], in a sense that it does not assume interactions between all agents, but instead assumes that two agents can only interact if their positions in a social space are closer than a certain interaction radius R, like in the bounded confidence models. The reason for such modeling decision is that agents that are further away in a social space, i.e. are having a low social similarity, may have conflicting attitudes and social norms, and thus lack motivation to interact with each other. An opinion interaction map that models this idea is

$$V(x_1, x_2, \theta_1, \theta_2) := \alpha \cdot 1_{[0,R]} (\|x_1 - x_2\|) \cdot (\theta_2 - \theta_1), \tag{4}$$

where α is the *opinion strength* parameter and $\|\cdot\|$ refers to the Euclidean distance. In our model, α regulates the strength of the social influence on the agent's opinion, i.e. the higher the α the more influence the pairwise

opinion difference has on an agent's opinion. As given by (4), this model formalizes the idea that two agents can only interact if their positions in a social space are close enough and if they interact, then their opinions become more similar. Since we do not distinguish between individual agent types, we assume that α is a constant, i.e. it is equal for all agents in the system. In the literature there are different variations of this classical interaction dynamics [33, 19]. As discussed above, our model extends these by introducing the movements of agents in a social space and its feedback loop with the opinion dynamics. An example of such dynamics is given by the spatial interaction map

$$U(x_1, x_2, \theta_1, \theta_2) := \beta \cdot 1_{[0,R]}(\|x_1 - x_2\|) \cdot \text{sgn}(\theta_1 \cdot \theta_2) \cdot (x_2 - x_1), \tag{5}$$

where β denotes the *spatial strength* parameter regulating how much agents compromise when updating their positions. The third term in (5) introduces a direction of agents' spatial motion based on their opinions, i.e. agents can either attract or repel each other, depending on whether their opinions are similar or different to each other. Thus, movement of an agent in a social space is determined by its spatial closeness and opinion similarity with other agents in the system.

2.2 Multi-body interactions

Most existing models of opinion dynamics only take pairwise interactions into account, ignoring any higher order interaction dynamics. While this simplifies mathematical considerations significantly, this is a very rough approximation of real life interactions and does not allow to model group effects such as peer pressure, see e.g. [34, 35]. Recently there has been an increasing interest in going beyond pairwise interactions to create more realistic models for social dynamics, e.g. [36]. Our model can be extended with such effects by including the spatial interaction dynamics as introduced in the previous section by (5) and considering multi-body interactions for the opinion dynamics proposed in [36]. In particular, the effect of peer pressure within groups of agents in spatial proximity can be modeled by the opinion interaction map

$$V(x_1, x_2, x_3, \theta_1, \theta_2, \theta_3) := \alpha \cdot \left(\prod_{i,j=1}^{3} 1_{[0,R]}(\|x_i - x_j\|) \right) s(|\theta_2 - \theta_3|)[(x_2 - x_1) + (x_3 - x_1)], \tag{6}$$

where $s : [0, \infty) \to \mathbb{R}$ is a non-increasing positive-definite function, e.g. $s(x) = \exp(\lambda x)$, for some $\lambda < 0$. This means, that the influence of agents 2 and 3 on the opinion of agent 1 is stronger if they have similar opinions. Different from the classical DeGroot model, in a case when multi-body interactions are defined on hypergraphs, shifts in the average opinion of the system have been observed [36]. Similar results are to be expected in our model and will be the topic of our future work. Here, in order to keep our analytical results more trackable, we will focus on the case of pairwise interactions.

2.3 Stochastic influence: Multiplicative noise

Many of the classical models of opinion dynamics are of a deterministic nature [11, 12, 32]. Historically, the first stochastic versions of these models have always used additive noise, i.e. stochastic noise with constant (in time, status and over the population) strength [37, 38]. Having a constant noise coefficient also implies that the driving noise of agents i and j are independent for all $i, j \in \{1, \ldots, N\}$. This makes the model mathematically easier, but such a modelling choice is questionable from a social sciences perspective. The usual role of the noise is to account for external influences, that are not already incorporated into the model through agents interactions, such as randomly occurring environmental changes or some other significant events. However, these events influence all (or at least most) agents, introducing non-trivial correlation between the individual noises. The choice of additive noise also doesn't account for more complex social mechanisms observed in real world systems [7]. Namely, the strength of the noise can additionally depend on how homogeneous the opinions of agent's peers are. In a highly polarised group of individuals, one might be more susceptible to random influences than in a group where everyone has roughly the same opinion.
Recently, a few models considered such complex mechanisms by introducing a *multiplicative noise*, e.g. in the context of opinion dynamics [7, 39], animal movement [40] and flocking of a Cucker-Smale system [41].

Extending on the ideas from [7], one concrete example of a multiplicative noise for our model is defined by

$$\sigma_{op}^i(X_t, \Theta_t) = \min_{j: \|X_t^i - X_t^j\| \leq R} \left| \Theta_t^i - \Theta_t^j \right| \cdot \text{Id}, i = 1, \ldots, N. \tag{7}$$

Thus, the noise in this model is characterized by a similarity bias, i.e. the closer the opinions of agents' peers are to the opinion of that agent, the smaller the random fluctuations. Similarly, agents have a higher probability to move away from agents with very different opinions. These effects have been shown to yield fast formation of stable spatial clusters in which a local consensus is reached [7].

2.4 Numerical simulations of the ABM

In this section, we show the main properties of our proposed model for opinion dynamics with feedback loops given by (3), (4) and (5). Using numerical simulations, we study how different parameters influence grouping of agents into social and opinion clusters. Additionally, we demonstrate a difference between additive

and multiplicative noise and how this affects the stability of clusters and opinion distribution within clusters.

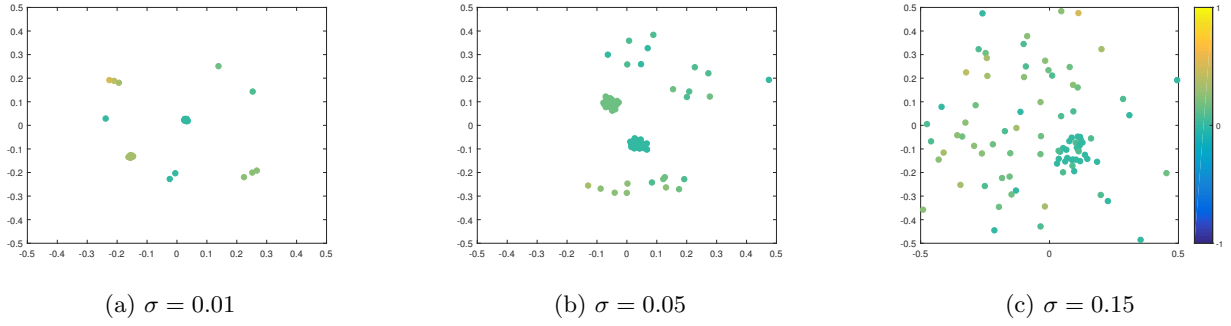(a) $\sigma = 0.01$ (b) $\sigma = 0.05$ (c) $\sigma = 0.15$

Figure 1: Snapshots from numerical simulations at final time $T = 2.5$ for different influences of additive noise: (a) $\sigma = 0.01$, (b) $\sigma = 0.05$ and (c) $\sigma = 0.15$. Position of agents indicated their position in a social space. Color of agents denotes their opinions according to the color-bar. Other parameters are fixed to $R = 0.015$ and $\alpha = \beta = 20$.

(a) $\sigma = 0.01$ (b) $\sigma = 0.05$ (c) $\sigma = 0.15$
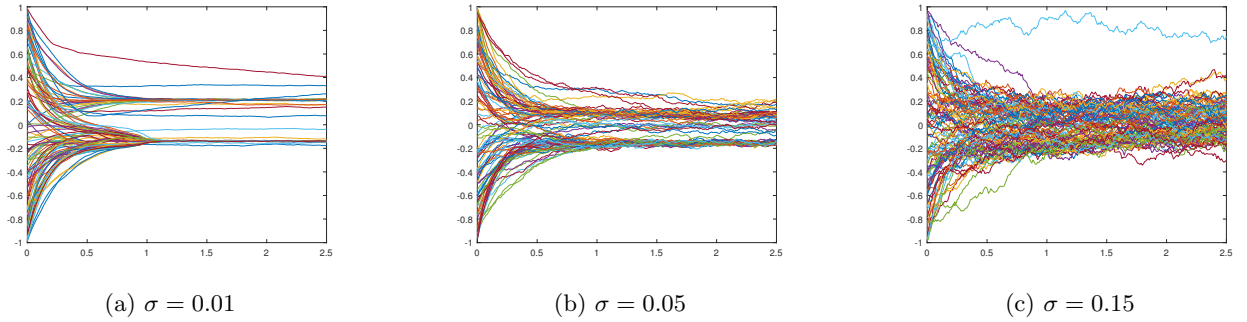
Figure 2: Opinion trajectories of agents' over time-period $[0, 2.5]$ for different influences of additive noise: (a) $\sigma = 0.01$, (b) $\sigma = 0.05$ and (c) $\sigma = 0.15$. Other parameters are fixed to $R = 0.015$ and $\alpha = \beta = 20$.

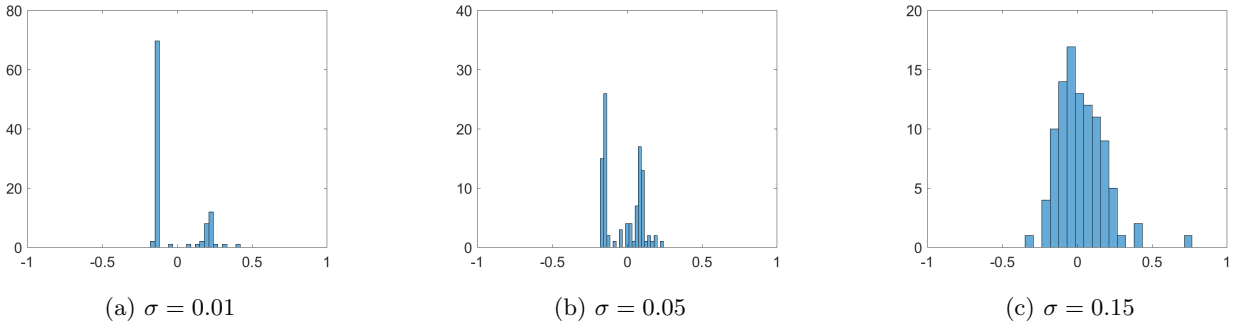(a) $\sigma = 0.01$ (b) $\sigma = 0.05$ (c) $\sigma = 0.15$

Figure 3: Distribution of agents' opinions at final time $T = 2.5$ for different influences of additive noise: (a) $\sigma = 0.01$, (b) $\sigma = 0.05$ and (c) $\sigma = 0.15$. Other parameters are fixed to $R = 0.15$ and $\alpha = \beta = 20$.

We run stochastic simulations for $N = 100$ agents, where at $t = 0$ all agents are placed uniformly at random inside $[-0.25, 0.25]^2$ and their initial opinions are distributed uniformly in the interval $[-1, 1]$. We run each simulation for 250 time-steps, i.e. until $T = 2.5$ where $\Delta t = 0.01$. In every time-step, movement of agents and their opinions are obtained using a Euler–Maruyama scheme [42, 43].

First, we consider the case of additive noise, where we assume that $\sigma := \sigma_{op} = \sigma_{sp}$ and we distinguish between different noise strengths, namely $\sigma = 0.01$, $\sigma = 0.05$ and $\sigma = 0.15$. We fix the interaction radius $R = 0.15$ and include strong opinion influence $\alpha = 20$ and strong spatial influence $\beta = 20$. In Figure 1 we plot simulation snapshots at the final time $T = 2.5$ for these different values of σ, where agents' positions correspond to their position in a social space and agents' colours indicate their opinion. We observe a strong influence of noise on the cluster formation process. Namely, the higher the σ, the denser the inter-cluster connections are, such that for $\sigma = 0.15$ there is no clear separation into different clusters neither in social nor in opinion space. This can be seen in Figure 2 of opinion trajectories of individual agents. Additionally,

stronger noise in the system leads to more diversified opinion distributions when reaching consensus, as can be observed in the final distribution of agents' opinions in Figure 3. For small values of noise, i.e. $\sigma = 0.01$, the system reaches a stable state and we see several spatially separated clusters of agents that are stable and within each cluster agents have similar opinions. This behavior has been observed in previous studies of network models [22, 44].

(a) $\alpha = \beta = 50$ (b) $\alpha = \beta = 5$ (c) $\alpha = 50, \beta = 5$
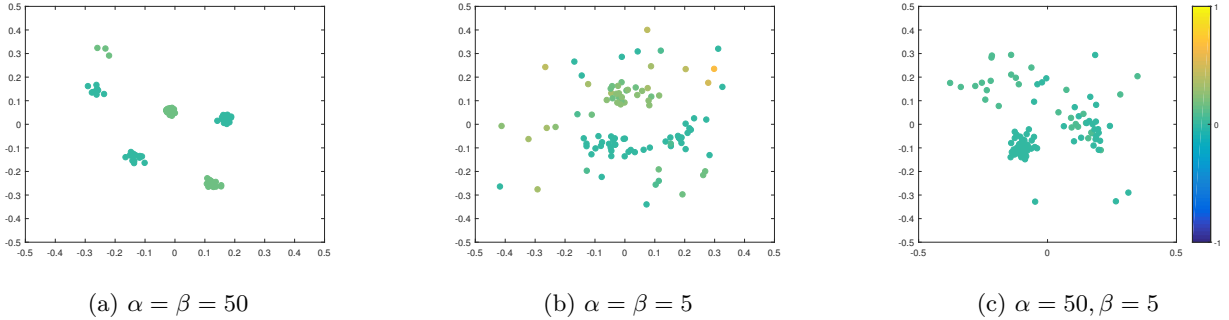
Figure 4: Snapshots from numerical simulations at final time $T = 2.5$ for different influences of opinion and spatial strength: (a) $\alpha = \beta = 50$, (b) $\alpha = \beta = 5$, (c) $\alpha = 50, \beta = 5$. Position of agents indicated their position in a social space. Color of agents denotes their opinions according to the color-bar. Other parameters are fixed to $R = 0.15$ and $\sigma = 0.05$.

(a) $\alpha = 50$ (b) $\alpha = 5$ (c) $\alpha = 50, \beta = 5$
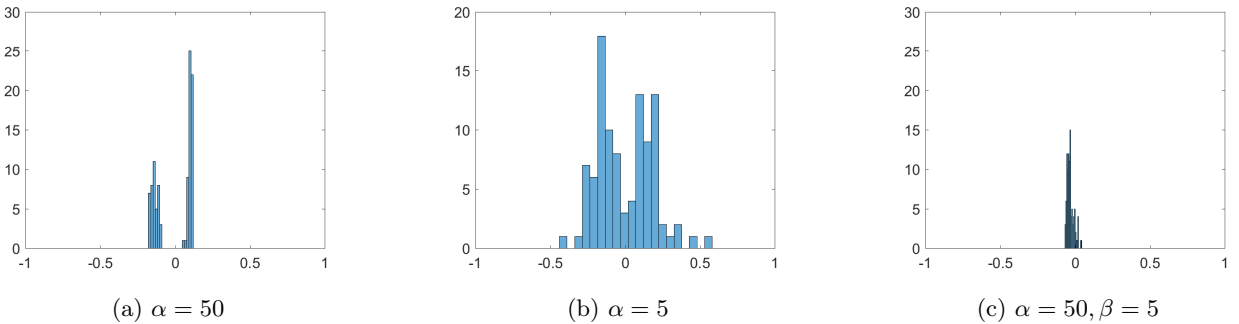
Figure 5: Distribution of agents' opinions at final time $T = 2.5$ for different influences of opinion and spatial strength: (a) $\alpha = \beta = 50$, (b) $\alpha = \beta = 5$. Position of agents indicated their position in a social space. Color of agents denotes their opinions according to the color-bar. Other parameters are fixed to $R = 0.15$ and $\sigma = 0.05$.

Next, we study the influence of opinion strength α and spatial strength β on the cluster formation process. Additional to results shown in Figure1b for $\alpha = \beta = 20$, in Figure 4 we show the snapshots after 250 time-steps for $\alpha = \beta = 50$, $\alpha = \beta = 5$ and $\alpha = 50, \beta = 5$. Large values of α, β mean that agents are strongly influenced by their peers, such that they compromise heavily towards their neighbours when updating their opinions and positions. This leads to faster formation of stable clusters where within the clusters agents have similar opinions, see Figure 4 and Figure 5. Because of this effect, in standard bounded confidence models, these parameters are often called 'convergence parameters' as they affect the speed of convergence [45]. Comparing the three scenarios in this experiment and in particular the cases when $\alpha = \beta$ to $\alpha \neq \beta$, we see that the spatial strength is important for formation of spatial clusters, but also in the case when spatial strength is small $\beta = 5$ and opinion strength is large $\alpha = 50$, because of the feedback loop agents tend to form large, loose groups. This effect emphasises the importance of the feedback loop in this system.

Finally, we examine the effects of multiplicative noise on the dynamics of our ABM given by (3), (4) and (5). To this end, we consider multiplicative noise $\sigma_{op} = \sigma_{sp}$, as introduced in (7). In Figure 6 we see a fast formation of 8 spatially well separated clusters. In each of these clusters local consensus is reached, as discussed in Section 2.3. These clusters are stable in their spatial position and opinion distribution. In comparison to the case of additive noise, see Figure 1, clusters obtained through the influence of multiplicative noise are constructed faster and are more stable.

3   Theoretical analysis: Coupled mean-field limit

In this section we consider the theoretical setting that describes the feedback loop dynamics of the system and its limit when $N \to \infty$. In particular, we start by briefly motivating the mean-field equation, and then
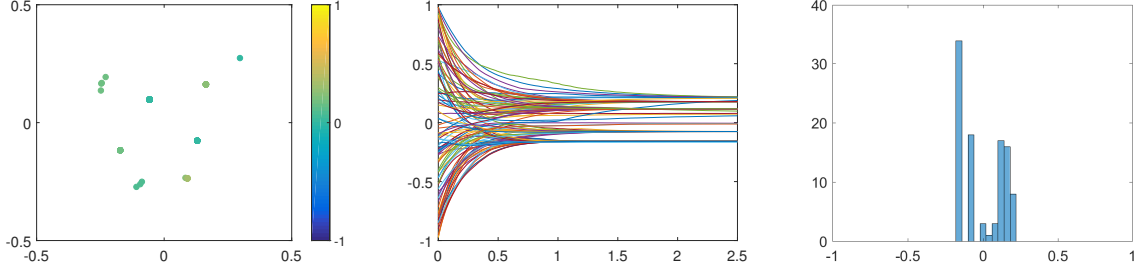
6

Figure 6: Results of one simulation of the ABM with multiplicative noise, for $\alpha = \beta = 20$ and $R = 0.15$. Left: Snapshot of the dynamics at $T = 2.5$. Middle: Opinion trajectories during $[0, 2.5]$. Right: Distribution of agents' opinions at $T = 2.5$. Position of agents indicated their position in a social space. Color of agents denotes their opinions according to the color-bar. Other parameters are fixed to $R = 0.015$ and $\sigma = 0.05$.

we state its well-posedness and the convergence of ABM to this mean-field equation.

From the theoretical point of view, one of the main challenges of these results lies in the regularity assumptions of drift and diffusion coefficients. The standard result is considering Lipschitz coefficients and additive noise, as presented in [27]. On the one hand, one can consider the dynamics just for one agent with singular interactions and additive noise, such as in [29] and its extensions to multiplicative noise case as in [46]. However, as noted in [47], these results can not be trivially extended for N particle systems. Thus, different techniques need to be used to obtain results for the singular interactions, such as mixed L^p drifts. The literature on the weaker assumptions on drift and diffusion than the standard global Lipschitz assumption is tremendous as it is a very popular topic, see [48, 49, 50, 51, 52].

However, in order to illustrate the main message of this article that is the consideration of the feedback loop dynamics with the multiplicative noise, it is enough to consider the standard Lipschitz setting. Since the case of the multiplicative noise in this setting can not be found in an easily accessible form, for the completeness of results, we state in the appendix the proofs of the well-posedness and convergence results, without relying on the sophisticated analysis results. This result can be seen as an extension of the standard result in [27].

In order to simplify the notation, we will write $\int$ for $\int_{\mathbb{R}^d \times \mathbb{R}} = \int_{\mathbb{R}^{d+1}}$ and we define $\mathcal{C} := C([0, T]; \mathbb{R}^{d+1})$. Hence, for the rest of the article, we will assume that potentials U and V are real valued Lipschitz functions on $\mathbb{R}^{2d} \times \mathbb{R}^2$. More precisely, we will always assume the following on the regularity of initial data:

Assumption 1 *Let (B^{sp}, B^{op}) be a Brownian motion in $\mathbb{R}^d \times \mathbb{R}$ with respect to a filtration $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$, the initial conditions X_0 and Θ_0 are $\mathcal{F}_0$-measurable and square-integrable, and the maps*

$$U : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^d, \quad V : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \tag{8}$$

$$\sigma_{sp} : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{d \times d}, \quad \sigma_{op} : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \tag{9}$$

are bounded and Lipschitz continuous.

3.1 Motivation for the limiting equations

In this section we will present a general motivation on how one can derive the so-called mean-field limit of coupled SDEs on $[0, T]$ for a fixed number $N \in \mathbb{N}$ of interacting agents which is a special case of our ABM (3). The mean-field limit is the coupled system of SDEs that describes the averaged dynamics of the system when the number of agents tends to infinity.

In particular, from now on the coupled SDE system that we will analyse is given by

$$\begin{cases} dX_t^{k,N} &= \frac{1}{N} \sum_{j=1}^{N} U(X_t^k, X_t^j, \Theta_t^k, \Theta_t^j) dt + \frac{1}{N} \sum_{j=1}^{N} \sigma_{sp}(X_t^k, X_t^j, \Theta_t^k, \Theta_t^j) dB_t^{sp,k}, \\ d\Theta_t^{k,N} &= \frac{1}{N} \sum_{j=1}^{N} V(X_t^k, X_t^j, \Theta_t^k, \Theta_t^j) dt + \frac{1}{N} \sum_{j=1}^{N} \sigma_{op}(X_t^k, X_t^j, \Theta_t^k, \Theta_t^j) dB_t^{op,k}, \end{cases} \tag{10}$$

for $k = 1, ..., N$, an independent family $(B^{sp,k}, B^{op,k})_{k \in \mathbb{N}}$ of $(d+1)$-dimensional Brownian motions.

Let us define the empirical density by

$$\mu_t^N := \frac{1}{N} \sum_{j=1}^{N} \delta_{(X_t^{j,N}, \Theta_t^{j,N})} \tag{11}$$

with δ being the Dirac distribution. Now the previous system can alternatively be written as

$$\begin{cases} dX_t^{k,N} &= \int_{\mathbb{R}^d \times \mathbb{R}} U(X_t^k, y, \Theta_t^k, \theta) \mu_t^N(dyd\theta) dt + \int_{\mathbb{R}^d \times \mathbb{R}} \sigma_{sp}(X_t^k, y, \Theta_t^k, \theta) \mu_t^N(dyd\theta) dB_t^{sp,k}, \\ d\Theta_t^{k,N} &= \int_{\mathbb{R}^d \times \mathbb{R}} V(X_t^k, y, \Theta_t^k, \theta) \mu_t^N(dyd\theta) dt + \int_{\mathbb{R}^d \times \mathbb{R}} \sigma_{op}(X_t^k, y, \Theta_t^k, \theta) \mu_t^N(dyd\theta) dB_t^{op,k}, \\ \mu_t^N &= \frac{1}{N} \sum_{j=1}^{N} \delta_{(X_t^{j,N}, \Theta_t^{j,N})}. \end{cases} \tag{12}$$

Utilizing the law of large numbers we will show that the mean-field limit of the system is of the form

$$
\begin{cases}
d\overline{X}_t &= \left(\int_{\mathbb{R}^d \times \mathbb{R}} U(\overline{X}_t, y, \overline{\Theta}_t, \theta) \mu_t(dyd\theta)\right) dt + \left(\int_{\mathbb{R}^d \times \mathbb{R}} \sigma_{sp}(\overline{X}_t, y, \overline{\Theta}_t, \theta) \mu_t(dyd\theta)\right) dB_t^{sp}, \\
d\overline{\Theta}_t &= \left(\int_{\mathbb{R}^d \times \mathbb{R}} V(\overline{X}_t, y, \overline{\Theta}_t, \theta) \mu_t(dyd\theta)\right) dt + \left(\int_{\mathbb{R}^d \times \mathbb{R}} \sigma_{op}(\overline{X}_t, y, \overline{\Theta}_t, \theta) \mu_t(dyd\theta)\right) dB_t^{op}, \\
\overline{\mu}_t &= \mathrm{Law}(\overline{X}_t, \overline{\Theta}_t), \\
\overline{X}_{t=0} &= X_0, \overline{\Theta}_{t=0} = \Theta_0,
\end{cases}
\tag{13}
$$

for suitable initial values X_0, Θ_0 and a $(d+1)$-dimensional Brownian motion (B_t^{sp}, B_t^{op}).

Remark 1 *One can write the system more compactly by summarizing the state of each agent into a single $\mathbb{R}^{d+1}$-valued random variable $Y = (X, \Theta)$. In order to do this, we define the combined interaction map to be*

$$
W : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^{d+1}, \quad (x_1, \theta_1, x_2, \theta_2) \mapsto (U(x_1, x_2, \theta_1, \theta_2), V(x_1, x_2, \theta_1, \theta_2)).
$$

With this notation we can write the system (13) *as*

$$
\begin{cases}
dY_t &= \left(\int_{\mathbb{R}^{d+1}} W(Y_t, y, \theta) \mu_t(d(y, \theta))\right) dt + \left(\int_{\mathbb{R}^{d+1}} \tilde{\sigma}(Y_t, y, \theta) \mu_t(d(y, \theta))\right) d\tilde{B}_t, \\
\mu_t &= Law(Y_t),
\end{cases}
\tag{14}
$$

where $\tilde{B}$ is a $(d+1)$-dimensional standard Brownian motion and $\tilde{\sigma}$ is a $(d+1) \times (d+1)$ diagonal matrix with $\mathrm{diag}(\sigma_{sp}(\tilde{X}_t, \Theta_t)) = (\sigma_{sp}(X_t, \Theta_t), ..., \sigma_{sp}(X_t, \Theta_t), \sigma_{op}(X_t, \Theta_t))$.
In the case of additive noise, if U, V are assumed to be Lipschitz continuous, then W is also a Lipschitz function and one can apply the classical existence and uniqueness and convergence results from [27] directly to (14). *With the focus on the different interpretation of processes X_t and Θ_t from the application point of view, we will keep the separate notion which enables us to directly see the limiting equations for position and opinion.*

3.2 Well-posedness result of the coupled mean-field SDE

Here, we will study the well-posedness of the limiting stochastic differential equations given by (13).

The proof is based on the classical results on mean-field theory (see e.g. for additive noise [27]) and the classical existence and uniqueness theory for SDEs via the standard fixed-point argument.

In order to set up our fixed point argument we, let us define

$$
\mathcal{P}^2 := \mathcal{P}^2(C([0, T]; \mathbb{R}^{d+1}))
$$
$$
:= \left\{ \mu : \mu \text{ is a probability measure on } C([0, T]; \mathbb{R}^{d+1}) \text{ s.t. } \int \|x\|_\infty^2 \mu(dx) < \infty. \right\}.
$$

Furthermore, we need to equip this space with a notion of distance, that turns $\mathcal{P}^2$ into a complete metric space. To make use of a common tool for this type of proof, namely Gronwall's inequality, we need to consider a whole family of distances (which are not necessarily metrics due to a lack of definiteness). To be precise we define the *truncated 2-Wasserstein distance* for $t \in [0, T]$ by

$$
D_t^2(\mu_1, \mu_2) := \inf_\mu \int_{\mathcal{C} \times \mathcal{C}} \left(\sup_{s \le t} (\|X_s(\omega_1) - X_s(\omega_2)\|^2 + |\Theta_s(\omega_1) - \Theta_s(\omega_2)|^2) \right) d\mu(\omega_1, \omega_2),
\tag{15}
$$

where we take the infimum over all couplings μ of μ_1 and μ_2 and $X_s(\cdot), \Theta_s(\cdot)$ are the projections onto the time s marginal of the d-dimensional component and the 1-dimensional component respectively. Note that for $t = T$, i.e. if we don't truncate, we obtain the standard 2-Wasserstein metric $\mathcal{W}_2^2(\cdot, \cdot)$ on $\mathcal{P}^2$.

Remark 2 *Observe that the Lipschitz Assumption 1 on the coefficients also implies, that the induced maps defined by*

$$
\hat{U} : \mathbb{R}^d \times \mathbb{R} \times \mathcal{P}^2(\mathbb{R}^{d+1}) \to \mathbb{R}^d, \quad (x, \theta, \mu) \mapsto \int_{\mathbb{R}^{d+1}} U(x, y, \theta, \eta) \mu(dyd\theta),
\tag{16}
$$

and analogously $\hat{V}, \hat{\sigma}_{sp}, \hat{\sigma}_{op}$, satisfy a Lipschitz-type inequality (17) with respect to the product metric, when we equip $\mathcal{P}^2$ with the 2-Wasserstein metric. Indeed, for $(x_1, \theta_1, \mu_1), (x_2, \theta_2, \mu_2)$ we have by Jensen's inequality

$$
\left\| \hat{U}(x_1, \theta_1, \mu_1) - \hat{U}(x_2, \theta_2, \mu_2) \right\|^2 \le \int \|U(x_1, y_1, \theta_1, \eta_1) - U(x_2, y_2, \theta_2, \eta_2)\|^2 \pi(d(y_1, \theta_1)d(y_2, \theta_2)),
$$

where π is an arbitrary coupling of μ_1, μ_2. We can now use the Lipschitz assumption on U to get

$$
\int \|U(x_1, y_1, \theta_1, \eta_1) - U(x_2, y_2, \theta_2, \eta_2)\|^2 \pi(d(y_1, \theta_1)d(y_2, \theta_2))
$$
$$
\le 2L^2 \|x_1 - x_2\| + 2L^2 \|\theta_1 - \theta_2\| + 2L^2 \int (\|y_1 - y_2\|^2 + \|\eta_1 - \eta_2\|^2) \pi(d(y_1, \theta_1)d(y_2, \theta_2)).
$$

Now we can take the infimum over all couplings π of μ_1 and μ_2 on the right hand side to obtain

$$
\left\| \int_{\mathbb{R}^{d+1}} U(x_1, y_1, \theta_1, \eta_1) \mu_1(dy_1 d\eta_1) - \int_{\mathbb{R}^{d+1}} U(x_2, y_2, \theta_2, \eta_2) \mu_2(dy_2 d\eta_2) \right\|^2
$$
$$
\le 2L^2(\|x_1 - x_2\|^2 + \|\theta_1 - \theta_2\|^2 + \mathcal{W}_2^2(\mu_1, \mu_2)).
$$

As already announced, to show the well-posedness of (13), we rely on Banach's classical fixed-point theorem. Therefore we need to make sure that the metric space $(\mathcal{P}^2, D_T^2)$ is sufficiently regular.

Lemma 1 $(\mathcal{P}^2, D_T^2)$ *is a complete metric space.*

The proof of this lemma is standard and is given in Appendix A.1.

In addition, for the proof of the next theorem we need the following a priori estimate.

Lemma 2 *Assume that the coefficients* $U, V, \sigma_{op}, \sigma_{sp}$ *are bounded by some positive constant* $K > 0$ *and that the initial values* X_0, Θ_0 *satisfy* $\mathbb{E}[\|X_0\|^2] < \infty$ *and* $\mathbb{E}[|\Theta_0|^2] < \infty$. *Then for any* $T > 0$ *there exists a constant* $C = C(T, K) > 0$ *such that every solution* (X, Θ) *to (13) satisfies*

$$\mathbb{E}\left[\sup_{t \in [0,T]} \|X_t\|^2 + |\Theta_t|^2\right] \leq C\left(\mathbb{E}\left[\|X_0\|^2 + |\Theta_0|^2\right] + K^2(T^2 + T)\right),$$

In particular, it holds that $Law(X, \Theta) \in \mathcal{P}^2$.

The proof of this lemma is classic and it is based on the Burkholder-Davis-Gundy inequality. For the completeness we wrote the proof in Appendix A.1. Now we can state the well-posedness result:

Theorem 1 *Let* X_0 *and* Θ_0 *be* $\mathbb{R}^d$*-valued, respectively* $\mathbb{R}$*-valued, random variables with finite second moment i.e.*

$$\mathbb{E}\left[\|X_0\|^2\right] < \infty, \quad \mathbb{E}\left[|\Theta_0|^2\right] < \infty.$$

Under Assumption 1, there exists a unique (pathwise and in law) solution to equation (13).

As already announced, the proof of the existence is based on the fixed point argument and it can be found in a more general setting. Nevertheless, for the completeness we sketch in the Appendix 1.1 the basic idea of the proof in our setting that is simpler than in the general setting, and hence easier accessible.

3.3 Convergence of the microscopic model to the mean-field equation

In this section we will prove that the the system of coupled SDEs for fixed $N \in \mathbb{N}$ given by (10) indeed converges to the mean-field limit, i.e. to show the so called propagation of chaos. For $N \in \mathbb{N}$ we denote the (measure-valued) empirical process of this system by $\mu^N = (\mu_t^N)_{t \geq 0}$, i.e. for $t \geq 0$ we set

$$\mu_t^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{(X_t^{i,N}, \Theta_t^{i,N})}. \tag{17}$$

For fixed i we consider the process $(\overline{X}_t^i, \overline{\Theta}_t^i)_{t \geq 0}$ that solves

$$\begin{cases} d\overline{X}_t^i &= \left(\int_{\mathbb{R}^{d+1}} U(\overline{X}_t^i, y, \overline{\Theta}_t^i, \theta)\mu_t(dy\, d\theta)\right) dt + \left(\int_{\mathbb{R}^{d+1}} \sigma_{sp}(\overline{X}_t^i, y, \overline{\Theta}_t^i, \theta)\mu_t(dy\, d\theta)\right) dB_t^{i,sp}, \\ d\overline{\Theta}_t^i &= \left(\int_{\mathbb{R}^{d+1}} V(\overline{X}_t^i, y, \overline{\Theta}_t^i, \theta)\mu_t(dy\, d\theta)\right) dt + \left(\int_{\mathbb{R}^{d+1}} \sigma_{op}(\overline{X}_t^i, y, \overline{\Theta}_t^i, \theta)\mu_t(dy\, d\theta)\right) dB_t^{i,op}, \\ \mu_t &= \text{Law}(\overline{X}_t^i, \overline{\Theta}_t^i). \end{cases}$$

The existence of $(\overline{X}_t^i, \overline{\Theta}_t^i)_{t \geq 0}$ follows from Theorem 1 and $\mu = (\mu_t)_{t \geq 0}$ does not depend on i. From now on we will denote the law of $(\overline{X}_t^i, \overline{\Theta}_t^i)_{t \geq 0}$ by μ.

Next, we will study the case when N tends to ∞. The key idea for the proof is to use the law of large numbers (LLN) for empirical measures $\overline{\mu}^N$ of i.i.d. copies of the process $(\overline{X}^i, \overline{\Theta}^i)$, cf. [53, Theorem 11.4.1]. We will see that by a uniform integrability argument, LLN also implies that

$$\mathbb{E}\left[D_T^2(\mu, \overline{\mu}^N)\right] \to 0 \quad \text{as } N \to \infty.$$

The rest of the needed estimates are then the same as in the proof of Theorem 1 and are mainly done for the purpose of showing that

$$\mathbb{E}\left[D_T^2(\mu, \mu^N)\right] \leq C\mathbb{E}\left[D_T^2(\mu, \overline{\mu}^N)\right]$$

for some constant $C > 0$.

Theorem 2 *For any* $i \in \mathbb{N}$ *we have*

$$\mathbb{E}\left[\sup_{0 \leq s \leq t} \left\|X_s^{i,N} - \overline{X}_s^i\right\|^2 + \left|\Theta_s^{i,N} - \overline{\Theta}_s^i\right|^2\right] \to 0 \quad \text{as } N \to \infty.$$

Moreover, we have

$$\lim_{N \to \infty} \mathbb{E}\left[D_T^2(\mu, \mu^N)\right] = 0. \tag{18}$$

The details of the proof are presented in Appendix 1.2.

4 Characterization of the empirical measure and its limit

In this section we derive the formal stochastic partial differential equation (SPDE) for the empirical measure μ^N, that is the so-called Dean–Kawasaki type equation [30, 54] with multiplicative noise and we derive the McKean-Vlasov type PDE for the limiting measure μ. Recall that if $X = (X_t)_{t \geq 0}$ is a Markov process with generator $\mathcal{L}$, then its law $\mu = (\mu_t)_{t \geq 0}$ is a (weak) solution to the linear PDE

$$\partial_t \mu = \mathcal{L}^* \mu, \tag{19}$$

also known as the Kolmogorov forward equation. A similar equation still holds true for solutions (X, Θ, μ) of the type of equation (13) and empirical measure for (10). However, the PDE that we will derive will be non-linear.

4.1 Derivation of the PDE for the law of the coupled mean-field SDEs

We consider a system of coupled mean-field SDEs given by (13) and we want to derive equation for the hydrodynamic limit $\mu_t = \mathrm{Law}(X_t, \Theta_t)$.

Let $\phi : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ be a smooth and compactly supported function. Then by Itô's formula we have

$$
\begin{aligned}
d\phi(X_t, \Theta_t) =& \\
& \left(\int U(X_t, y, \Theta_t, \theta) \cdot \nabla_x \phi(X_t, \Theta_t) \mu_t(dyd\theta) \right) dt + \left(\int V(X_t, y, \Theta_t, \theta) \frac{\partial}{\partial \theta} \phi(X_t, \Theta_t) \mu_t(dyd\theta) \right) dt \\
&+ \frac{1}{2} \left(\sum_{i,j=1}^d \left(\int \sigma_{sp}(X_t, y, \Theta_t, \theta) \mu_t(dyd\theta) \cdot \left(\int \sigma_{sp}(X_t, y, \Theta_t, \theta) \mu_t(dyd\theta) \right)^\top \right)_{i,j} \times \frac{\partial^2}{\partial x_i \partial x_j} \phi(X_t, \Theta_t) \right) dt \\
&+ \frac{1}{2} \left(\left(\int \sigma_{op}(X_t, y, \Theta_t, \theta) \mu_t(dyd\theta) \right)^2 \frac{\partial^2}{\partial \theta^2} \phi(X_t, \Theta_t) \right) dt \\
&+ \left((\nabla_x \phi(X_t, \Theta_t))^\top \cdot \int \sigma_{sp}(X_t, y, \Theta_t, \theta) \mu_t(dyd\theta) \right) dB_t^{sp} \\
&+ \left(\int \sigma_{op}(X_t, y, \Theta_t, \theta) \mu_t(dyd\theta) \cdot \frac{\partial}{\partial \theta} \phi(X_t, \Theta_t) \right) dB_t^{op}.
\end{aligned}
$$

Taking the expectation the martingale part vanishes, differentiating in t and recalling that $\mu_t = \mathrm{Law}(X_t, \Theta_t)$ we can rewrite the previous equation as

$$
\begin{aligned}
\frac{d}{dt} \int \phi(z, \eta) \mu_t(dzd\eta) =& \\
\int \bigg[& \left(\int U(z, y, \eta, \theta) \cdot \nabla_z \phi(z, \eta) \mu_t(dyd\theta) \right) + \left(\int V(z, y, \eta, \theta) \frac{\partial}{\partial \eta} \phi(z, \eta) \mu_t(dyd\theta) \right) \\
&+ \frac{1}{2} \left(\sum_{i,j=1}^d \left(\int \sigma_{sp}(X_t, y, \Theta_t, \theta) \mu_t(dyd\theta) \cdot \left(\int \sigma_{sp}(X_t, y, \Theta_t, \theta) \mu_t(dyd\theta) \right)^\top \right)_{i,j} \right. \\
&\left. \times \frac{\partial^2}{\partial x_i \partial x_j} \phi(X_t, \Theta_t) \right) + \frac{1}{2} \left(\left(\int \sigma_{op}(X_t, y, \Theta_t, \theta) \mu_t(dyd\theta) \right)^2 \frac{\partial^2}{\partial \theta^2} \phi(X_t, \Theta_t) \right) \bigg] \mu_t(dzd\eta).
\end{aligned}
$$

Assuming that μ_t is absolutely continuous with respect to Lebesgue measure and that its density $\mu_t(z, \eta)$ is sufficiently regular, we can apply Fubini's theorem and integration by parts. Since ϕ is an arbitrary sufficiently regular test-function, we see that $\mu = (\mu_t)_{t \geq 0}$ is a distribution-valued solution of the PDE

$$
\begin{aligned}
\partial_t \mu_t(z, \eta) =& -\mathrm{div}_z \left(\mu_t(z, \eta) \cdot U(z, \eta, \mu_t) \right) - \mathrm{div}_\eta \left(\mu_t(z, \eta) \cdot V(z, \eta, \mu_t) \right) \\
&+ \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial z_i \partial z_j} \left(\mu_t(z, \eta) \cdot \sigma_{sp}(z, \eta, \mu_t) \cdot \sigma_{sp}(z, \eta, \mu_t)^\top \right)_{i,j} \\
&+ \frac{1}{2} \frac{\partial^2}{\partial \eta^2} \left(\mu_t(z, \eta) \sigma_{op}(z, \eta, \mu_t)^2 \right),
\end{aligned} \tag{20}
$$

where we use the shorthand notation

$$U(z, \eta, \mu_t) := \int_{\mathbb{R}^d \times \mathbb{R}} U(z, y, \eta, \theta) \mu_t(dyd\theta), \tag{21}$$

$$V(z, \eta, \mu_t) := \int_{\mathbb{R}^d \times \mathbb{R}} V(z, y, \eta, \theta) \mu_t(dyd\theta), \tag{22}$$

$$\sigma_{space}(z, \eta, \mu_t) := \int_{\mathbb{R}^d \times \mathbb{R}} \sigma_{space}(z, y, \eta, \theta) \mu_t(dyd\theta), \tag{23}$$

$$\sigma_{op}(z, \eta, \mu_t) := \int_{\mathbb{R}^d \times \mathbb{R}} \sigma_{op}(z, y, \eta, \theta) \mu_t(dyd\theta). \tag{24}$$

10

Remark 3 *Note that in the case of the additive noise this equation becomes*

$$\partial_t \mu_t(z, \eta) = - div_z\Big(U(z, \eta, \mu_t)\mu_t(z, \eta)\Big) - div_\eta\Big(V(z, \eta, \mu_t)\mu_t(z, \eta)\Big)$$
$$+ \frac{1}{2}\sigma_1^2\Delta_z\mu_t(z, \eta) + \frac{1}{2}\sigma_2^2\Delta_\eta\mu_t(z, \eta),$$

which coincides with the standard result from [27]. Hence, as expected the only difference are the coefficients of the second order part of the differential operator on the right hand side, which are either constants or functions.

4.2 SPDE description for the empirical measure μ^N

The goal is to derive a formal SPDE for the empirical measure μ^N, i.e. the analogue to the Dean-Kawasaki equation, but with multiplicative noise. More precise, for $N \in \mathbb{N}$ let $(X_t^N, \Theta_t^N)_{t\geq 0}$ be such that its components $(X_t^{k,N}, \Theta_t^{k,N})_{t\geq 0}$ solve the system (10).

In the following, we want to (formally) derive the SPDE that is solved by the empirical measure $\mu^N = (\mu_t^N)_{t\geq 0}$ defined by (11).

Let $\phi : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ be a smooth and compactly supported function. By definition of the empirical measure we have

$$\int_{\mathbb{R}^d \times \mathbb{R}} \phi(z, \eta)\mu_t^N(dzd\eta) = \frac{1}{N}\sum_{i=1}^{N}\phi(X_t^{i,N}, \Theta_t^{i,N}).$$

Hence, we will first calculate each of the summands individually. For fixed $i = 1, \ldots, N$ we denote by $B^{sp,i} = (B^{sp,i,1}, \ldots, B^{sp,i,d})$ the components of the driving Brownian motion of the i-th agent. By Itô's formula we have

$$d\phi(X_t^{i,N}, \Theta_t^{i,N}) = \int \mu_t^{i,N}(dzd\eta)\Bigg[\Big(\int U(z, \eta, \mu_t^N)\cdot\nabla_x\phi(z, \eta) + V(z, \eta, \mu_t^N)\frac{\partial}{\partial\theta}\phi(z, \eta)\Big)dt$$

$$+ \frac{1}{2}\Bigg(\sum_{i,j=1}^{d}\Big(\sigma_{sp}(z, \eta, \mu_t^N)\cdot\sigma_{sp}(z, \eta, \mu_t^N)^\top\Big)_{i,j}\frac{\partial^2}{\partial z_i\partial z_j}\phi(z, \eta)\Bigg)dt$$

$$+ \frac{1}{2}\Big(\sigma_{op}(z, \eta, \mu_t^N)^2\frac{\partial^2}{\partial\eta^2}\phi(z, \eta)\Big)dt$$

$$+ \Big((\nabla_z\phi(z, \eta))^\top\sigma_{sp}(z, \eta, \mu_t^N)\Big)dB_t^{sp,i} + \Big(\sigma_{op}(z, \eta, \mu_t^N)\frac{\partial}{\partial\theta}\phi(z, \eta)\Big)dB_t^{op,i}\Bigg],$$

where we again used the shorthand notation introduced in (21)-(24).

Assuming that $\mu_t^{i,N}$ is absolutely continuous with respect to Lebesgue measure with a sufficiently smooth density $\mu_t^{i,N}(z, \eta)$ and applying integration by parts with respect to the Lebesgue integral, we obtain

$$d\phi(X_t^{i,N}, \Theta_t^{i,N}) =$$

$$\int dzd\eta\phi(z, \eta)\Bigg[-\mathrm{div}_z\Big[U(z, \eta, \mu_t^N)\mu_t^{i,N}(z, \eta)\Big]dt - \mathrm{div}_\eta\Big[V(z, \eta, \mu_t^N)\mu_t^{i,N}(z, \eta)\Big]dt$$

$$+ \frac{1}{2}\Bigg(\sum_{i,j=1}^{N}\frac{\partial^2}{\partial z_i\partial z_j}\Big(\Big(\sigma_{sp}(z, \eta, \mu_t^N)\cdot\sigma_{sp}(z, \eta, \mu_t^N)^\top\Big)_{i,j}\mu_t^{i,N}(z, \eta)\Big)\Bigg)dt$$

$$+ \frac{1}{2}\frac{\partial^2}{\partial\eta^2}\Big(\sigma_{op}(z, \eta, \mu_t^N)^2\mu_t^{i,N}(z, \eta)\Big)dt - \sum_{k,l=1}^{d}\frac{\partial}{\partial z_k}\Big(\sigma_{sp}^{l,k}(z, \eta, \mu_t^N)\mu_t^{i,N}(z, \eta)dB_t^{sp,i,k}\Big)$$

$$- \frac{\partial}{\partial\eta}\Big(\sigma_{op}(z, \eta, \mu_t^N)\mu_t^{i,N}(z, \eta)dB_t^{op,i}\Big)\Bigg].$$

Utilizing that for each $i = 1, \ldots, N$ we have

$$\frac{d}{dt}\phi(X^{i,N}, \Theta^{i,N}) = \int_{\mathbb{R}^{d+1}} dzd\eta\phi(z, \eta)\Big[\frac{\partial}{\partial t}\mu_t^{i,N}(z, \eta)\Big],$$

and after the summation over i we conclude that the empirical measure $(\mu_t^N)_{t \geq 0}$ solves the SPDE

$$
\begin{aligned}
\partial_t \mu_t^N = & - \operatorname{div}_z \left[U(z, \eta, \mu_t^N) \mu_t^N(z, \eta) \right] - \operatorname{div}_\eta \left[V(z, \eta, \mu_t^N) \mu_t^N \right] \\
& + \frac{1}{2} \sum_{i,j=1}^N \left(\frac{\partial^2}{\partial z_i \partial z_j} \left(\sigma_{sp}(z, \eta, \mu_t^N) \cdot \sigma_{sp}(z, \eta, \mu_t^N)^\top \right)_{i,j} \mu_t^N(z, \eta) \right) \\
& + \frac{1}{2} \frac{\partial^2}{\partial \eta^2} \left(\sigma_{op}(z, \eta, \mu_t^N)^2 \mu^N(z, \eta) \right) \\
& - \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^q \sum_{l=1}^d \frac{\partial}{\partial z_k} \left(\sigma_{sp}^{l,k}(z, \eta, \mu_t^N) \mu_t^{i,N}(z, \eta) dB_t^{sp,i,k} \right) \\
& - \frac{1}{N} \sum_{i=1}^N \frac{\partial}{\partial \eta} \left(\sigma_{op}(z, \eta) \mu_t^{i,N}(z, \eta) dB_t^{op,i} \right).
\end{aligned}
$$

However, this is not yet a closed equation for the empirical measure, because it still depends on the individual trajectories through the noise term. As in the case of additive noise [30], we will first calculate the covariance of this noise term and then replace it by a statistically identical term. For this, let's denote the two noise terms by

$$
\xi^{sp}(t, z, \eta) := \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^d \sum_{l=1}^d \frac{\partial}{\partial z_k} \left(\sigma_{sp}^{l,k}(z, \eta, \mu_t^N) \mu_t^{i,N}(z, \eta) dB_t^{sp,i,k} \right),
$$

$$
\xi^{op}(t, z, \eta) := \frac{1}{N} \sum_{i=1}^N \frac{\partial}{\partial \eta} \left(\sigma_{op}(z, \eta) \mu_t^{i,N}(z, \eta) dB_t^{op,i} \right).
$$

Then, the covariances are given by

$$
\begin{aligned}
\mathbb{E} \left[\xi^{sp}(t, z, \eta) \xi^{sp}(s, y, \theta) \right] = \delta(t - s) \frac{1}{N^2} \sum_{i=1}^N \sum_{k=1}^d \sum_{l,g=1}^d \frac{\partial}{z_k} & \left(\sigma_{sp}^{l,k}(z, \eta, \mu_t^N) \mu_t^{i,N}(z, \eta) \right) \\
& \times \frac{\partial}{x_k} \left(\sigma_{sp}^{g,k}(x, \theta, \mu_s^N) \mu_s^{i,N}(x, \theta) \right),
\end{aligned}
$$

$$
\mathbb{E} \left[\xi^{sp}(t, z, \eta) \xi^{op}(s, x, \theta) \right] = 0,
$$

$$
\begin{aligned}
\mathbb{E} \left[\xi^{op}(t, z, \eta) \xi^{op}(s, x, \theta) \right] = \delta(t - s) \frac{1}{N^2} \sum_{i=1}^N \frac{\partial}{\partial \eta} & \left(\sigma_{op}(z, \eta, \mu_t^N) \mu_t^{i,N}(z, \eta) \right) \\
& \times \frac{\partial}{\partial \theta} \left(\sigma_{op}(x, \theta, \mu_s^N) \mu_s^{i,N}(x, \theta) \right).
\end{aligned}
$$

First note that by the definition of $\mu_t^{i,N}$ as a Dirac delta distribution, we have for each $i = 1, \ldots, N$

$$
\mu_t^{i,N}(z, \eta) \mu_t^{i,N}(x, \theta) = \delta(z - x) \delta(\eta - \theta) \mu^{i,N}(z, \eta) = \delta(z - x) \delta(\eta - \theta) \mu^{i,N}(x, \theta).
$$

Therefore we can rewrite the non-trivial covariances as

$$
\begin{aligned}
& \mathbb{E} \left[\xi^{sp}(t, z, \eta) \xi^{sp}(s, y, \theta) \right] \\
& = \delta(t - s) \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^d \sum_{l,g=1}^d \frac{\partial}{\partial z_k} \frac{\partial}{\partial x_k} \left(\delta(\eta - \theta) \delta(z - x) \sigma_{sp}^{l,k}(z, \eta, \mu_t^N) \sigma_{sp}^{g,k}(z, \eta, \mu_t^N) \mu_t^N(z, \eta) \right),
\end{aligned}
$$

and

$$
\mathbb{E} \left[\xi^{op}(t, z, \eta) \xi^{op}(s, x, \theta) \right] \qquad = \delta(t - s) \frac{1}{N} \frac{\partial}{\partial \eta} \frac{\partial}{\partial \theta} \left(\delta(z - x) \delta(\eta - \theta) \sigma_{op}(z, \eta, \mu_t^N)^2 \mu_t^N(z, \eta) \right).
$$

Furthermore, let ζ be a $(d+1)$-dimensional Gaussian random field indexed by $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}$ with covariance given by

$$
\mathbb{E} \left[\zeta^i(t, z, \eta) \zeta^j(s, x, \theta) \right] = \delta^{i,j} \delta(t - s) \delta(z - x) \delta(\eta - \theta).
$$

We define two noise fields $\xi^{sp'}, \xi^{op'}$ via

$$
\xi^{sp'}(t, z, \eta) := \frac{1}{\sqrt{N}} \sum_{k=1}^d \sum_{l,g=1}^d \frac{\partial}{\partial z_k} \frac{\partial}{\partial x_k} \left(\sigma_{sp}^{l,k}(z, \eta, \mu_t^N) \sigma_{sp}^{g,k}(z, \eta, \mu_t^N) \mu_t^N(z, \eta)^{\frac{1}{2}} \zeta(t, z, \eta) \right)
$$

$$
\xi^{op'}(t, z, \eta) := \frac{1}{\sqrt{N}} \frac{\partial}{\partial \eta} \frac{\partial}{\partial \theta} \left(\sigma_{op}(z, \eta, \mu_t^N) \mu_t^N(z, \eta)^{\frac{1}{2}} \zeta(t, z, \eta) \right).
$$

Note that these noise fields are statistically equivalent to ξ^{sp}, ξ^{op}. Altogether, we conclude that the dynamics of the empirical measure μ^N of the feedback-loop dynamics with multiplicative noise given by (1) is described by the following formal SPDE

$$
\begin{aligned}
\partial_t \mu_t^N = & - \operatorname{div}_z \left[U(z, \eta, \mu_t^N) \mu_t^N(z, \eta) \right] - \operatorname{div}_\eta \left[V(z, \eta, \mu_t^N) \mu_t^N(z, \eta) \right] \\
& + \frac{1}{2} \sum_{i,j=1}^N \left(\frac{\partial^2}{\partial z_i \partial z_j} \left(\sigma_{sp}(z, \eta, \mu_t^N) \cdot \sigma_{sp}(z, \eta, \mu_t^N)^\top \right)_{i,j} \mu_t^N(z, \eta) \right) \\
& + \frac{1}{2} \frac{\partial^2}{\partial \eta^2} \left(\sigma_{op}(z, \eta, \mu_t^N)^2 \mu^N(z, \eta) \right) \\
& - \frac{1}{\sqrt{N}} \sum_{k=1}^d \sum_{l,g=1}^d \frac{\partial}{\partial z_k} \frac{\partial}{\partial x_k} \left(\sigma_{sp}^{l,k}(z, \eta, \mu_t^N) \sigma_{sp}^{g,k}(z, \eta, \mu_t^N) \mu_t^N(z, \eta)^{\frac{1}{2}} \zeta^k(t, z, \eta) \right) \\
& - \frac{1}{\sqrt{N}} \frac{\partial}{\partial \eta} \frac{\partial}{\partial \theta} \left(\sigma_{op}(z, \eta, \mu_t^N) \mu_t^N(z, \eta)^{\frac{1}{2}} \zeta^{d+1}(t, z, \eta) \right).
\end{aligned}
\tag{25}
$$

Remark 4 *The previous equation (25) is the generalization of the standard Dean-Kawasaki equation [30] to the dynamics with the multiplicative case. In particular, in the case of the additive noise this SPDE becomes*

$$
\begin{aligned}
\partial_t \mu_t^N(x, \theta) = & - div_x \left[\mu_t^N(x, \theta) \cdot \int U(x, y, \theta, \eta) \mu_t^N(y, \eta) dy d\eta \right] \\
& - div_\theta \left[\mu_t^N(x, \theta) \cdot \int V(x, y, \theta, \eta) \mu_t^N(y, \eta) dy d\eta \right] \\
& + \frac{1}{2} \sigma_1^2 \Delta_x \mu_t^N(x, \theta) + \frac{1}{2} \sigma_2^2 \Delta_\theta \mu_t^N(x, \theta) \\
& - \frac{1}{N} \sigma_1 div_x \left[\mu_t^N(x, \theta)^{\frac{1}{2}} \cdot \zeta(t, x, \theta) \right] - \frac{1}{N} \sigma_2 div_\theta \left[\mu_t^N(x, \theta)^{\frac{1}{2}} \cdot \zeta(t, x, \theta) \right]
\end{aligned}
$$

which is the standard type of Dean-Kawasaki equation. Concerning the SPDE formulations of an ABM, observe that the diffusion part that appears in the dynamics of the empirical measure (see [31]) is exactly of this type and the case with multiplicative noise is its generalization.

4.3 Numerical experiment

We will here illustrate the behaviour of the feedback loop system at the macroscopic level, i.e. when the number of agents tend to infinity. As we showed in the previous Section, this dynamics is described by the PDE (20). We will simulate it using finite difference method with one spatial dimension and one opinion dimension. The considered domain is $[-2, 2]^2$ with the grid size 0.05 and the time interval is $[0, 1]$ with time step $dt = 0.0001$. We use no-flux boundary conditions. The initial conditions are randomly chosen four clusters with normal distribution. Other parameters are chosen in a similar way as in the experiments made in Section 2.4. In particular, the potentials U and V are given by (4) and (5) respectively, with the additional scaling parameter that is taken to be 0.5 and represents scaled space/opinion interaction strength. Note that these potentials do not satisfy our regularity Assumption (1), however, as already explained, from the literature is expected that the PDE equation for the empirical density has the same form as (20). As in the ABM, the interaction radius is taken to be 0.15 and we consider the additive noise with strength 0.01 for both space and opinion dynamics. In Figure 7 we show the empirical density of agents from the numerical discretization of the equation (20) at the initial time $t = 0$, intermediate time $t = 0.5$ and final time $t = 1$. We observe that the behavior of the ABM shown in 2.4 agrees with the emerging dynamics of the PDE model. Namely, the agents empirical density show the cluster formation that is in agreement with the ABM and reflects the feedback loop dynamics of the system. Note that the diffusive behaviour depends on the choice of the scaled space/opinion strength parameter. In particular, it is expected to have stronger clustering effects with the increasing of space/opinion strength parameter. More detailed investigation of this, and the effect of the boundary conditions will be the topic for the future research.

5 Conclusion

Literature on opinion dynamics is very rich and diverse, spanning from empirical approaches, that rely on real-world data, to mathematical models, that mostly consider simple social rules and allow for rigorous analysis. However, there is a large gap between data- and model- driven approaches, that asks for introducing new formal models that can better represent complex social mechanisms governing how people shape and share their opinions. As a step towards closing this gap, in this paper we introduced an agent-based model for studying the feedback loop between opinion and social dynamics. This co-evolving dynamics governs how agents position in a social space (e.g. online social media) is influencing and is being influenced by other agents opinions. Additionally, unlike most existing models for opinion dynamics that consider only deterministic dynamics or additive noise, in order to account for more realistic scenarios, we introduce the influence of a multiplicative noise. In order to explore how these different factors influence the appearance

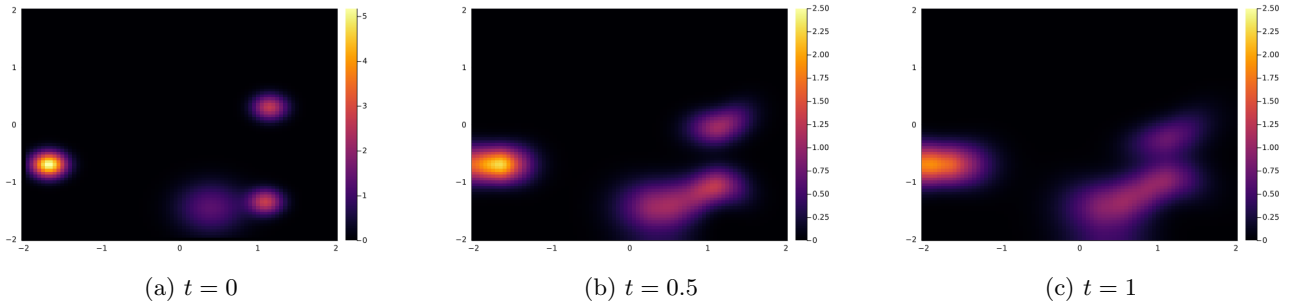|(a) $t = 0$ | (b) $t = 0.5$ | (c) $t = 1$ |

Figure 7: Empirical density of agents in mean-field limit given by equation (20) at initial time $t = 0$, intermediate time $t = 0.5$ and final time $t = 1$.

of emerging phenomena in the system, we tested our ABM model on several toy examples. In particular, we simulated the model for different parameter choices and we showed how these govern grouping of agents into spatial clusters, within which agents hold similar opinions. Our experiments have shown that cluster formation with respect to agents' social and/or opinion states is strongly influenced by the feedback loop. Further investigations of this model and possible effects the feedback loop could have in real-world social systems, will be the topic of future research.

Additionally, we formulated the feedback loop model in a rigorous mathematical framework and considered its behaviour in the case when the number of agents tends to infinity. Although the well-posedness propagation of chaos results have been proved for these type of systems in various weak regularity assumptions scenarios, we stated the proofs of these results in the case of the Lipschitz assumption that are easier to access and present the extension of the work presented in [27]. Finally, we derived formal equation for the empirical density of the ABM that is in the spirit of the so-called Dean-Kawasaki equation, and the equation of its hydrodynamic limit. Motivated by applications, considering the feedback loop dynamics with more singular interaction potentials and its numerical analysis would be a natural next step and will be considered in the future.

Acknowledgments

6 Appendix

Appendix A.1.

proof 1 (Proof of Lemma 1) *This follows from Theorem [55, Theorem 6.18] and the fact that*

$$d((x,\theta),(y,\eta)) := \sup_{s \leq T} \|(x_s, \theta_s) - (y_s, \eta_s)\|, \quad (x,\theta),(y,\eta) \in C([0,T];\mathbb{R}^{d+1}),$$

where $\|(x_s, \theta_s)\| := \left(\sum_{k=1}^{d} |x_s^k|^2 + |\theta_s|^2 \right)^2$, *defines a metric on* $C([0,T];\mathbb{R}^{d+1})$.

proof 2 (Proof of Lemma 2) *For* $n \in \mathbb{N}$ *consider the stopping time*

$$\tau_n := \min \left\{ \inf \left\{ t \geq 0 : |X_t| \geq n \right\}, \inf \left\{ t \geq 0 : |\Theta_t| \geq n \right\} \right\}.$$

Then for the stopped process $(X_t^{\tau_n}, \Theta_t^{\tau_n})_{t \geq 0}$ *it holds that*

$$\mathbb{E} \left[\sup_{t \in [0,T]} \|X_t^{\tau_n}\|^2 + |\Theta_t^{\tau_n}|^2 \right] \leq \mathbb{E} \left[\|X_0\|^2 + |\Theta_0|^2 \right] + 2n^2.$$

Via Cauchy-Schwarz and the Burkholder-Davis-Gundy inequality one can now derive the upper bound

$$\mathbb{E} \left[\sup_{t \in [0,T]} \|X_t^{\tau_n}\|^2 + |\Theta_t^{\tau_n}|^2 \right] \leq C \left(\mathbb{E} \left[\|X_0\|^2 + |\Theta_0|^2 \right] + K^2(T^2 + T) \right),$$

where the constant C *only depends on* T *and* K. *The right hand side of this inequality does not depend on* n, *so we can conclude that the unstopped process* X *also satisfies this inequality by letting* n *tend to infinity.*

proof 3 (Proof of Theorem 1) *As already announced, the proof of the existence is based on the fixed point argument. Let us define a map $\Phi : \mathcal{P}^2 \to \mathcal{P}^2$ by mapping $\mu \in \mathcal{P}^2$ to the law of the solution of the coupled system of SDEs*

$$
\begin{cases}
dX_t^{\mu} &= \left(\int_{\mathcal{C}} U(X_t^{\mu}, y, \Theta_t^{\mu}, \theta)\mu(dyd\theta)\right) dt + \left(\int_{\mathcal{C}} \sigma_{sp}(X_t^{\mu}, y, \Theta_t^{\mu}, \theta)\mu(dyd\theta)\right) dB_t^{sp}, \\
d\Theta_t^{\mu} &= \left(\int_{\mathcal{C}} V(X_t^{\mu}, y, \Theta_t^{\mu}, \theta)\mu(dyd\theta)\right) dt + \left(\int_{\mathcal{C}} \sigma_{op}(X_t^{\mu}, y, \Theta_t^{\mu}, \theta)\mu(dyd\theta)\right) dB_t^{op}, \\
X_{t=0}^{\mu} &= X_0, \\
\Theta_{t=0}^{\mu} &= \Theta_0.
\end{cases}
$$

The existence and uniqueness of a solution to this system follows from the fact that for fixed μ the drift and diffusion coefficient are Lipschitz continuous and bounded. Moreover, from standard a priori estimates for classical SDEs it follows that $\Phi(\mu) \in \mathcal{P}^2$.

It remains to show, that Φ is indeed a contraction on $\mathcal{P}^2$. To see this, fix two probability measures $\mu^1, \mu^2 \in \mathcal{P}^2$ and let $(X^1, \Theta^1), (X^2, \Theta^2)$ be the corresponding solutions, i.e. for $i = 1, 2$ we have

$$
\begin{cases}
dX_t^i &= \left(\int_{\mathcal{C}} U(X_t^i, y, \Theta_t^i, \theta)\mu^i(dyd\theta)\right) dt + \left(\int_{\mathcal{C}} \sigma_{sp}(X_t^i, y, \Theta_t^i, \theta)\mu^i(dyd\theta)\right) dB_t^{sp}, \\
d\Theta_t^i &= \left(\int_{\mathcal{C}} V(X_t^i, y, \Theta_t^i, \theta)\mu^i(dyd\theta)\right) dt + \left(\int_{\mathcal{C}} \sigma_{op}(X_t^i, y, \Theta_t^i, \theta)\mu^i(dyd\theta)\right) dB_t^{op}.
\end{cases}
$$

To estimate $D_T^2(\Phi(\mu^1), \Phi(\mu^2))$, we will use that the law of $(X^1, \Theta^1, X^2, \Theta^2)$ is actually a coupling of the two probability measures $\Phi(\mu^1)$ and $\Phi(\mu^2)$. Therefore, by definition of the 2-Wasserstein distance as the infimum over all couplings we have

$$
D_T^2(\Phi(\mu^1), \Phi(\mu^2)) \leq \mathbb{E}\left[\sup_{s \leq T} \left\|X_s^1 - X_s^2\right\|^2 + \left|\Theta_s^1 - \Theta_s^2\right|^2\right].
$$

We will now set up a Gronwall-type argument to obtain an upper bound of the right-hand side of this inequality in terms of the 2-Wasserstein distance of μ^1 and μ^2. As a first step, we use Hölder's and Young's inequality to see that for fixed $s \leq T$ we have

$$
|X_s^1 - X_s^2|^2 \leq 2s \int_0^s \left|\int_{\mathcal{C}} U(X_r^1, y_r, \Theta_r^1, \theta_r)\mu^1(dyd\theta) - \int_{\mathcal{C}} U(X_r^2, y_r, \Theta_r^2, \theta_r)\mu^2(dyd\theta)\right|^2 ds
$$
$$
+ 2\left|\int_0^s \left(\int_{\mathcal{C}} \sigma_{sp}(X_r^1, y_r, \Theta_r^1, \theta_r)\mu^1(dyd\theta) - \int_{\mathcal{C}} \sigma_{sp}(X_r^2, y_r, \Theta_r^2, \theta_r)\mu^2(dyd\theta)\right) dB_r^{sp}\right|^2
$$

where y_r, θ_r are the projections onto the time r coordinate of the d-dimensional spatial component and the 1-dimensional opinion component respectively. We obtain the analogoue estimate for $\left|\Theta_s^1 - \Theta_s^2\right|^2$.

Utilizing then Doob's L^2-inequality to estimate the stochastic part and Lipschitz-type inequality (17), enable us to apply Gronwall's lemma which implies that for all $t \in [0, T]$ we have

$$
\mathbb{E}\left[\sup_{s \leq t} \left\|X_s^1 - X_s^2\right\|^2 + \left|\Theta_s^1 - \Theta_s^2\right|^2\right] \leq C \int_0^t \mathcal{W}_{2, \mathbb{R}^{d+1}}^2(\mu_r^1, \mu_r^2) dr, \tag{26}
$$

where the constant C is given explicitly by

$$
C := 2(8 + 2T)L^2 \exp(2(8 + 2T)L^2)
$$

and $\mathcal{W}_{2, \mathbb{R}^{d+1}}$ is the usual 2-Wasserstein distance on $\mathcal{P}^2(\mathbb{R}^{d+1})$. Next observe that for two measures $\mu, \nu \in \mathcal{P}^2((C([0, T]; \mathbb{R}^{d+1}))$ and fixed $r \in [0, T]$ we have

$$
\mathcal{W}_{2, \mathbb{R}^{d+1}}^2(\mu_r, \nu_r) \leq D_r^2(\mu, \nu). \tag{27}
$$

Applying (27) to (26) gives us

$$
\mathbb{E}\left[\sup_{s \leq t} \left\|X_s^1 - X_s^2\right\|^2 + \left|\Theta_s^1 - \Theta_s^2\right|\right] \leq C \int_0^t D_r^2(\mu^1, \mu^2) dr.
$$

Furthermore, since the law of $(X^1, \Theta^1, X^2, \Theta^2)$ is a coupling of $\Phi(\mu^1), \Phi(\mu^2)$, the definition of the 2-Wasserstein distance (15) implies

$$
D_T^2(\Phi(\mu^1), \Phi(\mu^2)) \leq \mathbb{E}\left[\sup_{s \leq T} \left\|X_s^1 - X_s^2\right\|^2 + \left|\Theta_s^1 - \Theta_s^2\right|\right] \leq C \int_0^T D_r^2(\mu_r^1, \mu_r^2) dr.
$$

After establishing this inequality, the well-posedness follows from the usual application of Banach's fixed point theorem. By Lemma 2 this shows that the solution is unique in law. Now if we let μ be the unique law of the solution, and let it be the drift and diffusion coefficient of (13), then we have a standard SDE with bounded and Lipschitz continuous drift and diffusion coefficients. For such systems it is well known, that pathwise uniqueness holds. Therefore we can conclude that the solution to (13) is not only unique in law but also pathwise unique.

Appendix A.2.

proof 4 (Proof of Theorem 2) *For fixed $t \in [0, T]$ we have by Hölder's inequality*

$$\left\| X_t^{i,N} - \overline{X}_t^i \right\|^2$$

$$\leq 2t \int_0^t \left\| \int_{\mathcal{C}} U(X_s^{i,N}, y, \Theta^{i,N}, \theta) \nu_N(dyd\theta) - \int_{\mathcal{C}} U(\overline{X}_s^i, y, \overline{\Theta^{i,N}}, \theta) \mu(dyd\theta) \right\|^2 ds$$

$$+ 2 \left\| \int_0^t \left(\int_{\mathcal{C}} \sigma_{sp}(X_s^{i,N}, y, \Theta^{i,N}, \theta) \nu_N(dyd\theta) - \int_{\mathcal{C}} \sigma_{sp}(\overline{X}_s^i, y, \overline{\Theta^{i,N}}, \theta) \mu(dyd\theta) \right) dB_s^{sp} \right\|^2 .$$

We obtain the analogue estimate for $\left| \Theta_t^{i,N} - \overline{\Theta}_t^i \right|^2$.

Similarly as before, taking the supremum of previous estimates and integrating with respect to $\mathbb{P}$ and Doob's L^2-inequality and the Lipschitz type inequality (17) we obtain

$$\mathbb{E}\left[\sup_{0 \leq s \leq t} \left\| X_s^{i,N} - \overline{X}_s^i \right\|^2 + \left| \Theta_s^{i,N} - \overline{\Theta}_s^i \right|^2 \right]$$

$$\leq 2(2t+8)L^2 \mathbb{E}\left[\int_0^t \left\| X_s^{i,N} - \overline{X}_s^i \right\|^2 + \left| \Theta_s^{i,N} - \overline{\Theta}_s^i \right|^2 + \mathcal{W}_{2,\mathbb{R}^{d+1}}^2(\mu_s^N, \mu_s) ds. \right],$$

where $\mathcal{W}_{2,\mathbb{R}^{d+1}}(\cdot, \cdot)$ is the usual 2-Wasserstein metric on the set of probability measures on $\mathbb{R}^{d+1}$ with finite second moment. The application of Gronwall's inequality and the analogue arguments as in the proof of Theorem 1, yield the estimate

$$\mathbb{E}\left[\sup_{0 \leq s \leq t} \left\| X_s^{i,N} - \overline{X}_s^i \right\|^2 + \left| \Theta_s^{i,N} - \overline{\Theta}_s^i \right|^2 \right] \leq C\mathbb{E}\left[\int_0^t D_s^2(\mu, \mu^N) ds \right].$$

By construction we know that $(\overline{X}^i, \overline{\Theta}^i)_{i \in \mathbb{N}}$ is an i.i.d. family with $Law(\overline{X}^i, \overline{\Theta}^i) = \mu$ for all $i \in \mathbb{N}$. We define the empirical measure of the $(\overline{X}^i, \overline{\Theta}^i)$'s by

$$\overline{\mu}^N := \frac{1}{N} \sum_{i=1}^N \delta_{(\overline{X}^i, \overline{\Theta}^i)}. \tag{28}$$

Note that the random measure $\frac{1}{N} \sum_{i=1}^N \delta_{(X^{i,N}, \Theta^{i,N}, \overline{X}^i, \overline{\Theta}^i)}$ is a random coupling of the random measures μ^N and $\overline{\mu}^N$. Therefore the definition of the truncated 2-Wasserstein distance gives us the pathwise inequality

$$D_t^2(\mu^N, \overline{\mu}^N) \leq \frac{1}{N} \sum_{i=1}^N \sup_{0 \leq s \leq t} \left(\left\| X_s^{i,N} - \overline{X}_s^i \right\|^2 + \left| \Theta_s^{i,N} - \overline{\Theta}_s^i \right|^2 \right). \tag{29}$$

Hence, by our previous considerations we have

$$\mathbb{E}\left[D_t^2(\mu^N, \overline{\mu}^N) \right] \leq C\mathbb{E}\left[\int_0^t D_s^2(\mu, \mu^N) ds \right]. \tag{30}$$

Moreover, combining this with the triangle inequality for the truncated 2-Wasserstein distance (15), we obtain

$$\mathbb{E}\left[D_t^2(\mu, \mu^N) \right] \leq 2\mathbb{E}\left[D_t^2(\mu, \overline{\mu}^N) \right] + 2C\mathbb{E}\left[\int_0^t D_s^2(\mu, \mu^N) ds \right].$$

Now we can apply Gronwall's inequality and take $t = T$ to get

$$\mathbb{E}\left[D_T^2(\mu, \mu^N) \right] \leq 2e^{2CT} \mathbb{E}\left[D_T^2(\mu, \overline{\mu}^N) \right],$$

where the right hand side tends to 0 as $N \to \infty$ by the LLN for empirical measures of i.i.d. random variables, cf. [53, Theorem 11.4.1].
Since this LLN only gives us the almost sure convergence of the random measures $\mu^N \to \mu$, it remains to show that the family $(D_T^2(\mu, \mu^N))_{N \in \mathbb{N}}$ is uniformly integrable.

This can be shown by observing that

$$D_T^2(\mu, \overline{\mu}^N) \leq 2D_T^2(\mu, \delta_0) + 2\mathcal{W}_2^2(\delta_0, \overline{\mu}^N) \leq 2D_T^2(\mu, \delta_0) + 2\frac{1}{N} \sum_{j=1}^N \sup_{s \leq T} \left\| (\overline{X}_s^j, \overline{\Theta}_s^j) \right\|^2, \tag{31}$$

where δ_0 is the Dirac zero measure on $C([0, T]; \mathbb{R}^{d+1})$. The right hand side of this inequality is uniformly integrable, since the first term does not depend on N and the family $(\frac{1}{N} \sum_{j=1}^N \sup_{s \leq T} \left\| (\overline{X}_s^j, \overline{\Theta}_s^j) \right\|^2)_{N \in \mathbb{N}}$ is uniformly integrable as the empirical average of i.i.d. copies of an integrable random variable. Then the uniform integrability of the empirical averages follows directly from the definition.

References

[1] Stephan Lewandowsky, Laura Smillie, David Garcia, Ralph Hertwig, Jim Weatherall, Stefanie Egidy, Ronald E Robertson, Cailin O'Connor, Anastasia Kozyreva, Philipp Lorenz-Spreen, et al. Technology and democracy: Understanding the influence of online technologies on political behaviour and decision-making. Technical report, 2020.

[2] Pablo Porten-Cheé and Christiane Eilders. The effects of likes on public opinion perception and personal opinion. *Communications*, 45(2):223–239, 2020.

[3] Antonio F. Peralta, János Kertész, and Gerardo Iñiguez. Opinion dynamics in social networks: From models to data, 2022.

[4] Alina Sîrbu, Vittorio Loreto, Vito D. P. Servedio, and Francesca Tria. *Opinion Dynamics: Models, Extensions and External Effects*, pages 363–401. Springer International Publishing, Cham, 2017.

[5] Richard A Holley and Thomas M Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *The annals of probability*, pages 643–663, 1975.

[6] F. Schweitzer and J. A. Hołyst. Modelling collective opinion formation by means of active brownian particles. *The European Physical Journal B - Condensed Matter and Complex Systems*, 15(4):723–732, Jun 2000.

[7] Michele Starnini, Mattia Frasca, and Andrea Baronchelli. Emergence of metapopulations and echo chambers in mobile agents. *Scientific Reports*, 6(1):31834, October 2016.

[8] Unchitta Kan, Michelle Feng, and Mason A. Porter. An Adaptive Bounded-Confidence Model of Opinion Dynamics on Networks. preprint, SocArXiv, December 2021.

[9] Dietrich Stauffer. *Opinion Dynamics and Sociophysics*, pages 6380–6388. Springer, New York, 2009.

[10] PETER CLIFFORD and AIDAN SUDBURY. A model for spatial conflict. *Biometrika*, 60(3):581–588, 12 1973.

[11] Morris H. Degroot. Reaching a Consensus. *Journal of the American Statistical Association*, 69(345):118–121, March 1974.

[12] Rainer Hegselmann and Ulrich Krause. Opinion dynamics and bounded confidence: models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.

[13] Frank Schweitzer and J Doyne Farmer. *Brownian agents and active particles: collective dynamics in the natural and social sciences*, volume 1. Springer, 2003.

[14] M. Pineda, R. Toral, and E. Hernández-García. The noisy Hegselmann-Krause model for opinion dynamics. *Eur. Phys. J. B*, 86(12):1–10, 2013.

[15] Benjamin D Goddard, Beth Gooding, H Short, and GA Pavliotis. Noisy bounded confidence models for opinion dynamics: the effect of boundary conditions on phase transitions. *IMA Journal of Applied Mathematics*, 87(1):80–110, 2022.

[16] Chu Wang, Qianxiao Li, Bernard Chazelle, et al. Noisy hegselmann-krause systems: Phase transition and the 2r-conjecture. *Journal of Statistical Physics*, 166(5):1209–1225, 2017.

[17] Susana N Gomes, Grigorios A Pavliotis, and Urbain Vaes. Mean field limits for interacting diffusions with colored noise: phase transitions and spectral numerical methods. *Multiscale Modeling & Simulation*, 18(3):1343–1370, 2020.

[18] Nuno Crokidakis and Celia Anteneodo. Role of conviction in nonequilibrium models of opinion formation. *Phys. Rev. E*, 86:061127, Dec 2012.

[19] Pavlin Mavrodiev and Frank Schweitzer. The ambigous role of social influence on the wisdom of crowds: An analytic approach. *Physica A: Statistical Mechanics and its Applications*, 567:125624, 2021.

[20] Letizia Milli. Opinion dynamic modeling of news perception. *Applied Network Science*, 6(1):1–19, 2021.

[21] R. Hegselmann and U. Krause. Opinion dynamics under the influence of radical groups, charismatic leaders, and other constant signals: A simple unifying model. *Netw. Heterog. Media*, 10(3):477, 2015.

[22] Y. Yu, G. Xiao, G. Li, W. P. Tay, and H. F. Teoh. Opinion diversity and community formation in adaptive networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(10):103115, 2017.

[23] Arturo Buscarino, Luigi Fortuna, Mattia Frasca, and Alessandro Rizzo. Local and global epidemic outbreaks in populations moving in inhomogeneous environments. *Phys. Rev. E*, 90:042813, Oct 2014.

[24] Damon Centola, Juan Carlos González-Avella, Víctor M. Eguíluz, and Maxi San Miguel. Homophily, cultural drift, and the co-evolution of cultural groups. *Journal of Conflict Resolution*, 51(6):905–929, 2007.

[25] Federico Vazquez, Juan Carlos González-Avella, Víctor M. Eguíluz, and Maxi San Miguel. Time-scale competition leading to fragmentation and recombination transitions in the coevolution of network and states. *Phys. Rev. E*, 76:046120, Oct 2007.

[26] Demian Levis, Albert Diaz-Guilera, Ignacio Pagonabarraga, and Michele Starnini. Flocking-enhanced social contagion. *Phys. Rev. Research*, 2:032056, Sep 2020.

[27] Alain-Sol Sznitman. Topics in propagation of chaos. In Donald L. Burkholder, Etienne Pardoux, Alain-Sol Sznitman, and Paul-Louis Hennequin, editors, *Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989*, pages 165–251, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.

[28] Jürgen Gärtner. On the mckean-vlasov limit for interacting diffusions. *Mathematische Nachrichten*, 137(1):197–248, 1988.

[29] Nicolai V Krylov and Michael Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probability theory and related fields*, 131(2):154–196, 2005.

[30] David S Dean. Langevin equation for the density of a system of interacting Langevin processes. *Journal of Physics A: Mathematical and General*, 29(24):L613–L617, December 1996.

[31] Luzie Helfmann, Natasa Djurdjevac Conrad, Ana Djurdjevac, Stefanie Winkelmann, and Christof Schütte. From interacting agents to density-based modeling with stochastic pdes. *Communications in Applied Mathematics and Computational Science*, 16(1):1 – 32, 2021.

[32] G. Weisbuch, G. Deffuant, F. Amblard, and J.-P. Nadal. Interacting Agents and Continuous Opinions Dynamics. In M. Beckmann, H. P. Künzi, G. Fandel, W. Trockel, C. D. Aliprantis, A. Basile, A. Drexl, G. Feichtinger, W. Güth, K. Inderfurth, P. Korhonen, W. Kürsten, U. Schittko, P. Schönfeld, R. Selten, R. Steuer, F. Vega-Redondo, Robin Cowan, and Nicolas Jonard, editors, *Heterogenous Agents, Interactions and Economic Performance*, volume 521, pages 225–242. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003. Series Title: Lecture Notes in Economics and Mathematical Systems.

[33] Noah Friedkin and Eugene Johnsen. Social Influence Networks and Opinion Change. *Advances in Group Processes*, 16, January 1999.

[34] Federico Battiston, Giulia Cencetti, Iacopo Iacopini, Vito Latora, Maxime Lucas, Alice Patania, Jean-Gabriel Young, and Giovanni Petri. Networks beyond pairwise interactions: Structure and dynamics. *Networks beyond pairwise interactions: Structure and dynamics*, 874:1–92, August 2020.

[35] Federico Battiston, Enrico Amico, Alain Barrat, Ginestra Bianconi, Guilherme Ferraz de Arruda, Benedetta Franceschiello, Iacopo Iacopini, Sonia Kéfi, Vito Latora, Yamir Moreno, Micah M. Murray, Tiago P. Peixoto, Francesco Vaccarino, and Giovanni Petri. The physics of higher-order interactions in complex systems. *Nature Physics*, 17(10):1093–1098, October 2021.

[36] Leonie Neuhäuser, Michael T. Schaub, Andrew Mellor, and Renaud Lambiotte. Opinion Dynamics with Multi-body Interactions. In Samson Lasaulce, Panayotis Mertikopoulos, and Ariel Orda, editors, *Network Games, Control and Optimization*, pages 261–271, Cham, 2021. Springer International Publishing.

[37] Chu Wang, Qianxiao Li, Weinan E, and Bernard Chazelle. Noisy Hegselmann-Krause Systems: Phase Transition and the 2R-Conjecture. *Journal of Statistical Physics*, 166(5):1209–1225, March 2017.

[38] M Pineda, R Toral, and E Hernández-García. Noisy continuous-opinion dynamics. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(08):P08001, August 2009.

[39] Michael Mäs, Andreas Flache, and Dirk Helbing. Individualization as driving force of clustering phenomena in humans. *PLOS Computational Biology*, 6(10):1–8, 10 2010.

[40] Haiganoush K. Preisler, Alan A. Ager, Bruce K. Johnson, and John G. Kie. Modeling animal movements using stochastic differential equations. *Environmetrics*, 15(7):643–657, 2004.

[41] Yongzheng Sun and Wei Lin. A positive role of multiplicative noise on the emergence of flocking in a stochastic cucker-smale system. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(8):083118, 2015.

[42] Peter E Kloeden and Eckhard Platen. Higher-order implicit strong numerical schemes for stochastic differential equations. *Journal of statistical physics*, 66(1):283–314, 1992.

[43] Desmond J Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM review*, 43(3):525–546, 2001.

[44] Samuel Stern and Giacomo Livan. The impact of noise and topology on opinion dynamics in social networks. *Royal Society Open Science*, 8(4):201943, 2021.

[45] Unchitta Kan, Michelle Feng, and Mason A Porter. An adaptive bounded-confidence model of opinion dynamics on networks. *arXiv preprint arXiv:2112.05856*, 2021.

[46] Xicheng Zhang. Stochastic differential equations with sobolev diffusion and singular drift and applications. *The Annals of Applied Probability*, 26(5):2697–2732, 2016.

[47] Zimo Hao, Michael Röckner, and Xicheng Zhang. Strong convergence of propagation of chaos for mckean-vlasov sdes with singular interactions. *arXiv preprint arXiv:2204.07952*, 2022.

[48] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit and propagation of chaos for vlasov systems with bounded forces. *Journal of Functional Analysis*, 271(12):3588–3627, 2016.

[49] Gonçalo dos Reis, Stefan Engelhardt, and Greig Smith. Simulation of mckean–vlasov sdes with super-linear growth. *IMA Journal of Numerical Analysis*, 42(1):874–922, 2022.

[50] William RP Hammersley, David Šiška, and Łukasz Szpruch. Mckean–vlasov sdes under measure dependent lyapunov conditions. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 57, pages 1032–1057. Institut Henri Poincaré, 2021.

[51] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. Mean-field limit and quantitative estimates with singular attractive kernels. *arXiv preprint arXiv:2011.08022*, 2020.

[52] Daniel Lacker. On a strong form of propagation of chaos for mckean-vlasov equations. *Electronic Communications in Probability*, 23:1–11, 2018.

[53] R. M. Dudley. *Real Analysis and Probability*. Cambridge University Press, 2 edition, October 2002.

[54] Kyozi Kawasaki. Stochastic model of slow dynamics in supercooled liquids and dense colloidal suspensions. *Physica A: Statistical Mechanics and its Applications*, 208(1):35–64, 1994.

[55] Cédric Villani. *Optimal Transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.