IMA Journal of Applied Mathematics (2022) **87**, 80–110
https://doi.org/10.1093/imamat/hxab044
Advance Access publication on 13 November 2021

Noisy bounded confidence models for opinion dynamics: the effect of boundary conditions on phase transitions

B. D. Goddard*, B. Gooding and H. Short
School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh EH9 3FD, UK
*Corresponding author: b.goddard@ed.ac.uk

and

G. A. Pavliotis
Department of Mathematics, Imperial College London, London SW7 2AZ, UK

[Received on 22 March 2021; revised on 21 September 2021; accepted on 4 October 2021]

We study SDE and PDE models for opinion dynamics under bounded confidence, for a range of different boundary conditions, with and without the inclusion of a radical population. We perform exhaustive numerical studies with pseudo-spectral methods to determine the effects of the boundary conditions, suggesting that the no-flux case most faithfully reproduces the underlying mechanisms in the associated deterministic models of Hegselmann and Krause. We also compare the SDE and PDE models, and use tools from analysis to study phase transitions, including a systematic description of an appropriate order parameter.

Keywords: bounded confidence; opinion dynamics; no-flux boundary conditions; phase transitions.

2000 Math Subject Classification: 82C22; 82C26; 65M70.

1. Introduction and previous work

This work focuses on bounded confidence models for opinion dynamics under social influence (Castellano *et al.*, 2009; Proskurnikov & Tempo, 2017, 2018), which is part of the larger field of mathematical modelling in the social sciences (Castellano *et al.*, 2009; Turchin, 2016, 2018). In contrast to graph-based models in which communication occurs only between connected individuals, here communication is instead limited by the difference in their opinions, which is treated as a continuous variable. A typical example of a continuous opinion is an individual's political orientation, which is not restricted to a few discrete choices, such as extreme left or right, but rather can vary across a spectrum. The motivation for bounded confidence models comes from 'biased assimilation' (Lord *et al.*, 1979), which postulates that individuals are more strongly influenced by others with similar opinions to their own. A related consideration is that genuine discussion generally only occurs between individuals who already share some common ground, i.e., they already have sufficiently close opinions. This can be modelled by the coupling between individuals increasing as their opinions become more similar.

Here we consider models with a hard cut-off in opinion space; individuals with opinions differing by more than the confidence bound R do not influence each other. Such models were originally proposed as deterministic, discrete-time processes (Deffuant *et al.*, 2000; Hegselmann & Krause, 2002; Weisbuch *et al.*, 2002), which have also been extended to include noise or randomness (Pineda *et al.*, 2009, 2011, 2013; Su *et al.*, 2017), modelling uncertainty in observations or external influences. Related models include continuous time Ordinary Differential Equations (ODEs) (Blondel *et al.*, 2010; Yang *et al.*, 2014), Stochastic Differential Equations (SDEs) and (in the limit of many individuals) Partial Differential Equations (PDEs) (Kolarijani *et al.*, 2019; Wang *et al.*, 2017). See Castellano *et al.* (2009); Proskurnikov & Tempo (2017, 2018) for comprehensive reviews. In the below, due to the inclusion of noise in the SDE and (implicitly in) the PDE models, we refer to them as non-deterministic. More general formulations arise in mathematical biology as the Keller-Segel model (Keller & Segel, 1970) for slime mold, and other similar models (Mogilner & Edelstein-Keshet, 1999; Mogilner *et al.*, 2003; Topaz *et al.*, 2006). Since the individuals which interact changes according to the dynamics, such models are often described through 'co-evolutionary networks' (Proskurnikov & Tempo, 2018), e.g., the Vicsek model of phase transitions (Vicsek *et al.*, 1995), the Cucker-Smale model for flocking (Cucker & Smale, 2007a,b; Motsch & Tadmor, 2011) and robotics (Bullo *et al.*, 2009; Canuto *et al.*, 2012; Jadbabaie *et al.*, 2003). We also highlight recent mathematical work concerning the well-posedness and long-time behaviour of related mean-field models (Carrillo *et al.*, 2020; Chazelle *et al.*, 2017; Goddard *et al.*, 2020).

A typical question one asks about such models is how a uniform, or disordered, initial condition evolves under the interactions of individuals; typically the system is driven to a more ordered state (Castellano *et al.*, 2009; Kubo *et al.*, 2012; Yeomans, 1992). This has clear analogues with order–disorder phase transistions (Yeomans, 1992). Two popular models are due to Hegselmann and Krausse (HK) (Hegselmann & Krause, 2002) and Deffaunt and Weisbuch (DW) (Deffuant *et al.*, 2000; Weisbuch *et al.*, 2002). See Proskurnikov & Tempo (2018) for a recent review. Both are discrete in time and rely on the idea of repeated averaging under bounded confidence. In DW, agents interact in randomly chosen pairs who then either do or do not compromise (depending on the separation of their opinions). In HK, an individual moves to the average opinion of all agents within a distance R of themselves in opinion space. In such models, an initially homogeneous or uniform distribution of opinions is unstable. This is due to boundary effects, in which those with extreme opinions can only be influenced by those with less extreme opinions, and hence tend to move towards more moderate values, causing clusters. Three possible outcomes are: (i) consensus (a single cluster); (ii) polarization (two distinct clusters); (iii) fragmentation (more than two distinct clusters, or no distinct clusters). It has been noted that the number of clusters is typically $1/(2R)$, which is related to the $2R$-conjecture (Blondel *et al.*, 2007; Castellano *et al.*, 2009). This conjecture, which originally referred to the HK model, states that a uniform distribution of agents converges to clusters separated by distances of roughly $2R$. This was studied in Wang *et al.* (2017) for a PDE model with periodic boundary conditions; it would be interesting to investigate this further for the different boundary conditions proposed below. This is a topic of future work.

This convergence to a steady state is not normally observed in real-world systems (Ben-Naim, 2005; Carro *et al.*, 2013; Yildiz *et al.*, 2013). Two possible factors missing from these models are (i) influence of external agents or information, such as 'radicals', leaders or advertising; (ii) uncertainty in the dynamics. For (i), it is usually assumed that such influences are constant in time (Düring *et al.*, 2009; Hegselmann & Krause, 2015; Kolarijani *et al.*, 2019; Masuda, 2015; Wongkaew *et al.*, 2015; Yildiz *et al.*, 2013), or weakly susceptible to other opinions (Zhao *et al.*, 2016). Numerical simulations have demonstrated counterintuitive effects concerning the introduction of radials, such as increasing radical numbers decreasing the number of individuals sharing their opinion after long times. We

will demonstrate similar effects in our non-deterministic models; see Section 5. For (ii), it has been argued (Ben-Naim, 2005) that diffusion is an essential element of opinion dynamics, allowing the modelling of realistic political systems with disorder-order transitions and complex lifecycles. This has close links to the use of environmental noise in statistical physics models, which can lead to phase transitions (Ben-Naim, 2005; Czirók *et al.*, 1999; Grauwin & Jensen, 2012; Kolarijani *et al.*, 2019; Pineda *et al.*, 2009, 2011, 2013; Sznajd-Weron *et al.*, 2011; Wang *et al.*, 2017). Additionally, in many applications, it is not only the equilibrium that is interesting, but also the dynamical path to that equilibrium (Castellano *et al.*, 2009). For example, if the opinion measures political persuasion, then one would typically be interested in the distribution at a particular time (e.g., on an election date) than in the long-term equilibrium. We note here that existing approaches fail to correctly model the appropriate boundary conditions for such applications; we discuss this in detail later. In Sections 4 and 5, we systematically investigate both the effects of the noise strength and the dynamical paths to equilibrium.

Due to the non-linear, non-local nature of many opinion dynamics models, and the resulting challenges of analytical investigation, the use of careful and systematic computer simulations has been important in social dynamics for decades, see e.g., Hegselmann & Flache (1998); Hołyst *et al.* (2001); Stauffer (2001); Wang *et al.* (2017); Zhao *et al.* (2016). For N individuals, the agent-based, ODE and SDE models typically have computational costs that scale as N^2, which is the cost of determining the pairwise distances. An interesting regime, which is amenable to both mathematical analysis and decreased computational cost, is the mean-field limit where the number of individuals becomes very large, $N \to \infty$. Then the high-dimensional descriptions of individual opinions reduces to a 1+1 (space-time) dimensional non-local, non-linear PDE, for which there exist a range of accurate and efficient numerical approaches (Gomes *et al.*, 2020; Kolarijani *et al.*, 2019; Nold *et al.*, 2017; Wang *et al.*, 2017). Such PDE models have been referred to by a variety of names, including density-based (Lorenz, 2007), continuum (Blondel *et al.*, 2010; Hendrickx & Olshevsky, 2016), Eulerian (Canuto *et al.*, 2012; Mirtabatabaei *et al.*, 2014), hydrodynamic (Motsch & Tadmor, 2011, 2014), kinetic (Boudin & Salvarani, 2016; Brugna & Toscani, 2015; Düring *et al.*, 2009) or mean-field (Garnier *et al.*, 2017; Kolarijani *et al.*, 2019; Nordio *et al.*, 2017; Wang *et al.*, 2017). We note, in particular, that kinetic models provide an intermediate level of coarse-graining, between the discrete, 'particle', models and the continuum, 'hydrodynamics-type' PDEs. In kinetic models, interactions are modelled through instantaneous, local 'collisions', or interactions, between individuals. This contrasts with the 'soft' interactions included in the PDE models presented here.

Returning to the SDE and PDE models, there can be crucial differences between the two models: Firstly, order–disorder phase transitions are rigorously defined only in the thermodynamic limit as only then can there be non-uniqueness of invariant measures. However, such terms are still used for similar behaviour in the finite-N models (Toral & Tessone, 2007). Secondly, there are questions about how large the number of agents needs to be in order for the SDE and PDE models to be in good agreement. Many techniques used to study PDE models arise from statistical mechanics, but care must be taken when directly transferring them to the social sciences. This is due to the different notions of a 'large' number of individuals/particles; in molecular systems this could be of the order 10^{23}, whereas in social systems, it is more likely to be in the hundreds or thousands. In such cases, finite-size effects are likely to be important (Toral & Tessone, 2007); we demonstrate this in Supplementary Material Section SM3. However, these differences must be balanced by the relative computational complexities: as above, the computational cost of the SDE, agent-based model scales as N^2, whereas that of the PDE is independent of N, depending instead on the discretization.

The mean-field limit of a system of weakly interacting agents in the presence of noise is given by a non-linear, non-local PDE of Fokker-Planck type. For such PDEs (as well as the corresponding SDEs), it is necessary to specify boundary conditions. Previous work has been restricted to periodic domains, presumably for mathematical and computational ease, see e.g., Garnier *et al.* (2017); Kolarijani *et al.* (2019); Wang *et al.* (2017). However, care must be taken when comparing such implementations to the original deterministic models. For example, in the original models, if all initial opinions lie in an interval $[a, b]$, then this holds for all time (Proskurnikov & Tempo, 2018). To aid comparison, it would be natural to treat the non-deterministic models in the same way, requiring the individual opinions to lie in some prescribed interval, which, without loss of generality, may be chosen to be $[0, 1]$. One criticism of periodic boundary conditions is that they conflate the two extreme opinions at 0 and 1; this implies that two opposite extremes (e.g., on the political spectrum) are close. This is clearly not a realistic assumption. Whilst the even two-periodic choice in (Kolarijani *et al.*, 2019) overcomes this conflation, it does so at the expense or introducing a mirror system, which can strongly influence the dynamics and has (potentially) undesirable effects on quantities such as the order parameter; see Section 2.4. Here we introduce an additional choice of boundary conditions, namely no-flux. This can be regarded as a mixed (Robin) boundary condition for the mean-field Fokker-Planck-McKean-Vlasov PDE. As in the existing two cases, this preserves the mass conservation property of the model, but does not conflate extreme opinions nor require the introduction of an auxiliary system. *We propose that no-flux boundary conditions are much more compatible with the original, discrete HK models in terms of the mechanism of cluster formation from a uniform initial condition*. For example, as we will see in Section 4, large R (confidence bound) and small σ (noise) causes an initially uniform distribution to develop a cluster in the middle of the interval, rather than remaining in a uniform steady state, which is predicted when periodic boundary conditions are chosen. In particular, we note that previous results, such as the numerical experiments and resultant phase diagrams in Wang *et al.* (2017) and the rigorous analysis reported in Garnier *et al.* (2017) depend crucially on the assumption of periodic boundary conditions. Given that the choice of periodic boundary conditions is clearly not optimal from a modelling perspective, it is not obvious whether the numerical experiments reported in these papers are relevant to the study of opinion dynamics.

The challenge with no-flux boundary conditions is both analytical and numerical. The additional analytical challenges arise principally from the fact that, as already mentioned in Goddard *et al.* (2020), the no-flux boundary conditions for the McKean-Vlasov PDE are non-linear and non-local. This means, in particular, that the uniform distribution is no longer a stationary state. Furthermore, doing linear stability analysis, in the form of Garnier *et al.* (2017) [Sec. 4.1] is more involved, since the perturbation from the stationary state has to be chosen in such a way that it satisfies the boundary conditions and that it has zero mass. We note that, at least formally and in the absence of radical groups, the mean field PDE is a Wasserstein gradient flow for the free energy (Villani, 2003). We reiterate that care has to be taken when dealing with the nonlinear and nonlocal boundary conditions that we consider in this paper. See Goddard *et al.* (2020) for details. In terms of numerical methods for SDE models, reflecting boundary conditions are known to be challenging (Pilipenko, 2014; Skorokhod, 1961), and the corresponding PDE models can no longer use efficient Fourier methods. Here we implement efficient and robust Fourier and Chebyshev pseudo-spectral methods, based on Goddard *et al.* (2017); Nold *et al.* (2017). We emphasise the fact that, due to their greater realism and compatibility with the original deterministic models, we favour the use of no-flux boundary conditions in future studies; we include the periodic cases for comparison with the existing literature.

For completeness, we also briefly discuss the possible choice of zero Dirichlet boundary conditions. The key disadvantage in such as choice is that the system no longer preserves mass, i.e., the number

of agents can change over time. This may, however, be an appropriate boundary condition in some situations, e.g., if people with extreme political views stop interacting with others. Of course, if one considers extreme opinions to be rare, and the vast majority of the population has opinions localized in the centre of opinion space, then Dirichlet boundary conditions will give essentially identical results to any other choice. This could be achieved, e.g., in a low-noise system where the initial condition is centred towards the middle of the opinion space. However, when considering models with radicals who may aim to drive opinions to extremes, this is clearly not the case. An important example of such a system is a binary referendum, where the extremes of the system correspond to a strong likelihood of voting one way or the other. We do not consider such Dirichlet boundary conditions in this work, primarily because the lack of mass conservation prevents a meaningful comparison to the other choices. We mention that whether boundaries for mean-field SDEs modelling opinion formation are accessible or not is an interesting question. In particular, it is possible that, for appropriate choices of the interaction between agents, the dynamics never reaches the boundaries of opinion space. This is an interesting and intriguing question – in particular, the development of a Feller-type classification for mean field SDEs in a one dimensional bounded domain – that we plan to return to in future work.

Our main contributions are:

- a unification of existing and novel SDE and PDE models for bounded confidence opinion dynamics;
- a systematic numerical study of the transient and long-time dynamics of these models, under three different boundary conditions and a wide range of parameter regimes, with and without a 'radical' population;
- a careful discussion of an 'order parameter';
- the insight that the no-flux boundary conditions most faithfully reproduce the underlying mechanisms of the original deterministic models.

The remainder of the paper is organized as follows: In Section 2, we introduce the models, as well as the associated order parameter, and convergence to equilibrium. Section 3 describes the details of the numerical methods for both the PDE and SDE models. In Sections 4 and 5, we present the numerical experiments for systems without and with radicals, respectively. Section 6 contains our conclusions and a description of some open problems. The Supplementary Material contains more discussion on the order parameter, validation of the numerical methods against existing results from the literature, a thorough comparison of the SDE and PDE models, as well as some further examples.

2. Model

2.1 *Dynamics*

The original model of HK is discrete in time and space. It considers a set of N agents, with agent i having opinion x_i with $x = (x_1, \dots, x_N)$. Bounded confidence is introduced by defining a confidence level R and a set for each agent $I(i,x) = \{1 \le j \le N : |x_i - x_j| \le R\}$, i.e., the set of all individuals whose opinion is within R of that of individual i. At each time the opinion of individual i is updated through

$$x_i(t+1) = \left|I\big(i, x(t)\big)\right|^{-1} \sum_{j \in I(i,x(t))} x_j(t),$$

where here $|I|$ denotes the number of elements of I. In words, an individual's opinion at the next time step is given by the mean of the opinions of individuals within their confidence interval. It is clear to see here that if two groups are separated by a distance of R or more then they will form decoupled subsystems, which then never interact. Note that there exist alternative models in which the attraction increases with separation (Motsch & Tadmor, 2014).

In the SDE models (Kolarijani *et al.*, 2019; Wang *et al.*, 2017), which originate in statistical physics, the mean is taken not over the set I, but over all individuals, with zero weight on those outside I, i.e., $|I|$ is replaced by N in the normalization. This leads to a simpler mean-field PDE but it is also possible to retain the original normalization (Garnier *et al.*, 2016). The original choice of normalization is perhaps more physically relevant, especially in swarming/flocking models, in which the bounded confidence is based on physical, rather than opinion, distance (Cucker & Smale, 2007a,b). In such cases, it is plausible to assume that an individual is completely unaware of those outside its confidence bound. In opinion dynamics models, it is perhaps more reasonable to assume that an individual polls the opinion of all other individuals (as if in a completely connected network) and simply ignores the opinion of those individuals outside their confidence bound. Introducing the $1/N$ scaling results in dynamics, which are slowed down by the presence of individuals who do not interact, and are, in fact, unaware of each other (Motsch & Tadmor, 2011); other works have claimed that the results are insensitive to this choice (Canuto *et al.*, 2012; Garnier *et al.*, 2016).

For a system of N individuals with opinions x_i, and a confidence bound R, the dynamics that we consider in this paper, given a suitable initial condition, are described by (Wang *et al.*, 2017)

$$\mathrm{d}x_i = -\frac{1}{N}\sum_{j:|x_i-x_j|\le R}(x_i - x_j)\mathrm{d}t + \sigma\,\mathrm{d}W_t^{(i)}, \tag{2.1}$$

where $W_t^{(i)}$ are independent Wiener processes. As described in Section 1, the motivation for including noise comes from agents' 'free will', or uncertainty in measurement and communication. Taking the mean-field limit of $N \to \infty$ results in a Fokker-Planck PDE for the density of opinions $\rho(x,t)$ (Garnier *et al.*, 2017; Wang *et al.*, 2017):

$$\partial_t \rho(x,t) = \partial_x\left(\rho(x,t)\int (x-y)\rho(y,t)\mathbf{1}_{|x-y|\le R}(y)\,\mathrm{d}y\right) + \frac{\sigma^2}{2}\partial_{xx}\rho(x,t), \tag{2.2}$$

where ρ corresponds to the empirical measure $\rho^N(x,t) = N^{-1}\sum_j \delta_{x_j}(\mathrm{d}x)$ as $N \to \infty$.

2.2 *Boundary conditions*

It remains to discuss the boundary conditions (BC) imposed on Equations (2.1) and (2.2). As discussed above, we consider three separate cases: (i) periodic; (ii) no-flux; (iii) even two-periodic, all on $[0,1]$, representing opinion space. For **BC (i)** [Periodic] the natural interpretation is that $\rho(x,t)$ is extended periodically to the whole of $\mathbb{R}$ with the periodicity condition $\rho(x+1,t) = \rho(x,t)$, $\forall x \in \mathbb{R}$. We note that $\rho(\cdot,t)$ is not a probability density on $\mathbb{R}$, but does serve as a (normalized) probability density on [0,1]. If [0,1] represents the whole of opinion space then this is clearly not a good modelling assumption; in particular it conflates the two extreme opinions. For **BC (ii)** [No-Flux], it is helpful to rewrite (2.2) in

terms of the flux, j:

$$\partial_t \rho(x,t) = -\partial_x j(x,t), \qquad j(x,t) = -\rho(x,t)\int (x-y)\rho(y,t)\mathbf{1}_{|x-y|\leq R}(y)\,\mathrm{d}y - \frac{\sigma^2}{2}\partial_x \rho(x,t). \tag{2.3}$$

We then impose no-flux boundary conditions, i.e., $j(0,t) = j(1,t) = 0$. We emphasize that these boundary conditions are non-local and non-linear due to the nature of the flux in the PDE (2.2).

The mean field PDE (2.3) can be written as a gradient flow in the sense of Jordan-Kinderlehrer-Otto (Carrillo *et al.*, 2020)[Sec. 6.2]:

$$\partial_t \rho(x,t) = \partial_x \left(\rho \partial_x \frac{\delta \mathscr{F}}{\delta \rho}\right), \tag{2.4}$$

where the free energy for the Hegselmann-Krause model is

$$\mathscr{F}[\rho] = \frac{\sigma^2}{2}\int \rho \log(\rho)\,dx + \iint W(x-y)\rho(x)\rho(y)\,dxdy, \tag{2.5}$$

with

$$W(x) = -\left(\left(|x| - \frac{R}{2}\right)_-\right)^2 \tag{2.6}$$

where $(\cdot)_-$ denotes the negative part of a function. In Carrillo *et al.* (2020)[Sec. 6.2] it was shown that, for periodic boundary conditions and at sufficiently small R, an appropriately rescaled Hegselmann-Krause model exhibits a discontinuous phase transition. Similar results for the nonlinear, nonlocal no-flux boundary conditions that we consider in this paper will be presented elsewhere (Goddard *et al.*, 2021).

For **BC (iii)** [Even two-Periodic], we follow Kolarijani *et al.* (2019). Rather than considering $\rho(x,t)$ on $x \in [0,1]$, we consider its unique even, two-periodic extension which satisfies $\rho(-x,t) = \rho(x,t)$ and $\rho(x+2,t) = \rho(x,t)$, $\forall x \in \mathbb{R}$. Again, $\rho(\cdot,t)$ is a probability density on $[0,1]$. For clarity, in this case the sum over j in (2.1) is over the full, even two-periodic system with $2N$ individuals. *We will show that the no-flux boundary conditions most faithfully reproduce the behaviour of the original HK models*, e.g., the instability of a uniform initial condition, and the physical interpretation of the domain.

2.3 *Radicals*

As described in Hegselmann & Krause (2015); Kolarijani *et al.* (2019), a natural extension of bounded confidence models is to include $N_r > 0$ *radicals*, or extreme groups. These should be thought of as individuals with fixed opinions who, nevertheless, affect the opinions of those who interact with them. As discussed in Hegselmann & Krause (2015), such a formalism can also be used to model advertising, charismatic leaders and other external effects. In the language of statistical mechanics, radicals act as an external potential. Radicals (or stubborn agents) have also been included as a generalization of the voter model, see, e.g., Yildiz *et al.* (2011, 2013).

For the SDE, we retain the indexing of the 'normal' individuals as $1, \ldots, N$, and add in radicals indexed by $N+1, \ldots, N+N_r$. The dynamics are then governed by

$$\mathrm{d}x_i = -\frac{1}{N} \sum_{j:|x_i - x_j| \le R} (x_i - x_j)\mathrm{d}t + \sigma \mathrm{d}W_t^{(i)}, \quad i = 1, \ldots N$$

$$\mathrm{d}x_i = 0, \quad i = N+1, \ldots, N+N_r.$$

Note that the sum in the first equation now runs over $j = 1, \ldots, N+N_r$. The corresponding Fokker-Planck PDE is

$$\partial_t \rho(x,t) = -\partial_x j(x,t),$$

$$j(x,t) = -\rho(x,t) \int (x-y)\big(\rho(y,t) + M\rho_r(y)\big)\mathbf{1}_{|x-y|\le R}(y)\,\mathrm{d}y - \frac{\sigma^2}{2}\partial_x \rho(x,t),$$

where ρ_r determines the (fixed) distribution of the radicals. We find it convenient to fix ρ_r as a probability distribution (in the senses described above for the periodic boundary conditions) and scale the mass with a parameter M. For physical reasons, ρ_r should be non-negative (although it is interesting to consider negative/repulsive opinions), and also $M \ll 1$, otherwise the interpretation as ρ_r as the density of 'radicals' is lost. However, these restrictions are not intrinsic to the model itself. Note that, in the case of even two-periodic boundary conditions, $x \in [-1, 1]$, the full space, and the convolution is also taken over this whole domain.

2.4 *Order parameter*

To enable quantification of the resulting opinion densities, Wang *et al.* (2017) introduced the order parameter, which has both a discrete and continuum definition:

$$\tilde{Q}_d(t) = \frac{1}{N^2} \sum_{i,j=1}^{N} \mathbf{1}_{|x_i(t) - x_j(t)| \le R}, \qquad \tilde{Q}_c(t) = \iint \rho(x,t)\rho(y,t)\mathbf{1}_{|x-y|\le R}(y)\,\mathrm{d}x\mathrm{d}y.$$

In the sequel, we denote both quantities by $\tilde{Q}$, as their use is unambiguously defined by the data to which they refer. We will shortly explain the seemingly extraneous tilde notation.

The order parameter measures the order (or disorder) of the opinions. One interpretation is that it measures the proportion of pairs of individuals within a radius R of each other in opinion space. Naturally, disorder is described by a uniform distribution, whereas clustering (localized states) corresponds to consensus. For a uniform distribution $x_i = i/N$, for each i we find $\sum_j |x_i - x_j| = 2RN$, and so $\tilde{Q} = 2R$, which provides an R-dependent lower bound for $\tilde{Q}$; the same holds for $\rho(x,t) = 1$, $\forall x \in [0,1]$ in the continuum case. For a single cluster of individuals of width less than R, we find $\tilde{Q} = 1$, which is the maximum value of $\tilde{Q}$ for both the periodic and no-flux boundary conditions. For n equal clusters, of width less than R, and separated by at least R, we find that $\tilde{Q} = 1/n$, suggesting that the order parameter is essentially the inverse of the number of well-separated clusters. Note that there is not an injective mapping between densities and order parameters. The systematic derivation of order parameters for models of opinion dynamics will be studied in future work. We emphasize that the

identification of the order parameter for models of opinion dynamics is not a straightforward matter, since it is not clear what the 'physical' significance of such an order parameter is. This contrasts with other situations such as the magnetization in Ising-type models.

We now note that the order parameter for the even two-periodic case is somewhat different. We now have (at least) two choices: compute the order parameter for N individuals on $[0, 1]$, or for $2N$ individuals on $[-1, 1]$. In the first case, it is unclear how one should define the distance between two individuals on $[0, 1]$, since the dynamics are defined on $[-1, 1]$ with even/periodic boundary conditions. It is also desirable to use the same definition of I in both the dynamics and the order parameter; this is only possible when defining $\tilde{Q}$ using ρ on the whole domain $[-1, 1]$.

Consider a uniform distribution of $2N$ individuals on $[-1, 1]$ with periodic boundary conditions, which results in $\sum_{i,j=1}^{2N} \mathbf{1}_{|x_i - x_j| \leq R} = 2N^2 R$. Hence, to obtain the same value of $\tilde{Q}$ as in the periodic case, one requires a prefactor of $1/(2N^2)$. The same normalization gives the corresponding results for single and multiple clusters which are well-separated from each other and the domain boundaries. Motivated by this, we redefine the order parameters, removing tildes:

$$Q_d(t) = c_{\mathrm{BC}} \frac{1}{N^2} \sum_{i,j=1}^{N_{\mathrm{BC}}} \mathbf{1}_{|x_i(t) - x_j(t)| \leq R}, \quad Q_c(t) = c_{\mathrm{BC}} \iint \rho(x,t)\rho(y,t)\mathbf{1}_{|x-y|\leq R}(y)\,\mathrm{d}x\mathrm{d}y.$$

where $c_{\mathrm{BC}} = 1, N_{\mathrm{BC}} = N$ for periodic and no-flux boundary conditions and $c_{\mathrm{BC}} = 1/2, N_{\mathrm{BC}} = 2N$ for even two-periodic boundary conditions. Note that, as above, if R is small, or the opinion distribution is located away from the boundaries of the domain, then our definition is equivalent to that used previously (Kolarijani *et al.*, 2019). However, this leads to some non-standard results for Q in the even two-periodic case. Consider a single cluster of particles in $[1 - R/2, 1]$, and the corresponding mirror cluster in $[-1, -1 + R/2]$; this leads to a value of $Q = 2$. We use $Q > 1$ as a signature that the even two-periodic boundary conditions have had a significant effect on the dynamics. We give some examples of the behaviour of the order parameter in Supplementary Material Section SM1.

2.4.1 *Stationary states and convergence to equilibrium.* First, we remind the reader that stationary states for the HK dynamics are defined as stationary solutions of the mean field PDE, either in the absence or in the presence of radicals. When writing the mean field PDE as a conservation law, Equation (2.3), then the stationary Fokker-Planck equation becomes

$$-\partial_x j(x) = 0, \qquad j(x) = -\rho(x) \int (x - y)\rho(y)\mathbf{1}_{|x-y|\leq R}(y)\,\mathrm{d}y - \frac{\sigma^2}{2}\partial_x \rho(x). \tag{2.7}$$

In view of the no-flux boundary conditions $j(0) = j(1) = 0$, the stationary Fokker-Planck equation becomes

$$j(x) = 0.$$

Since the HK PDE admits a gradient flow formulation (2.4), it is possible to give a variational characterization of stationary states in terms of critical points of the free energy functional (2.5). The equivalence between the different characterizations of stationary states of the mean field PDE with periodic boundary conditions, as (a) stationary solutions of the Fokker-Planck equation, (b) critical points of the free energy functional, (c) global minimizers of the entropy dissipation functional and (d)

solutions of the Kirkwood-Monroe integral equation is presented in Carrillo *et al.* (2020)[Prop. 2.4]. Similar results for the case of no-flux boundary conditions will be presented elsewhere (Goddard *et al.*, 2021).

Away from the phase transition, we expect that we have exponentially fast convergence to equilibrium, measured either in relative entropy or in an appropriately weighted L^2 norm. It is useful to have a measure of how quickly the solution of (2.2) converges to equilibrium, i.e., of the exponent in the exponential estimate. We have found that a robust measure of being close to equilibrium is that the maximum value of $\partial_t \rho$ is lower than a given tolerance ε_{eq}. When this first happens defines an equilibrium time t_{eq} for the system. We note that this time clearly depends on the tolerance chosen, and so should be regarded as a measure of *relative* time to equilibrium for different parameter regimes, rather than a firm statement that we have reached equilibrium. Unless otherwise stated, we choose $\varepsilon_{\text{eq}} = 10^{-4}$.

3. Numerical methods

In this section, we give some details of the numerical methods employed for the SDE and PDE models, which were described in the previous section.

3.1 *Continuum model*

There are two principal methods that have been used to obtain accurate solutions of the continuum equations (2.2). The first, used by Wang *et al.* (2017), is a semi-implicit pseudo-spectral method, which is a standard approach used in fluids dynamics and statistical mechanics. Their method approximates the solution at a finite number of points in real space, with a corresponding Fourier transform. The derivatives are straightforward to approximate in Fourier space, by multiplication, whereas the convolutions can be approximated by inverse Fourier transformation to real space. For smooth solutions, such methods are exponentially accurate in the number of collocation points, and automatically satisfy the periodic boundary conditions and mass conservation. Time-stepping is performed with a semi-implicit update.

A related method is used in Kolarijani *et al.* (2019) where, rather than approximating the solution on a spatial grid, it is expanded in a basis of the corresponding Hilbert space, with the expansion truncated on a finite set. This leads to a system of ODEs for the coefficients of the basis functions, which then produce an approximation to the solution of the PDE. Once again, this is particularly suited to periodic geometries, as there is a clear basis set to use (Fourier or complex exponentials), and the periodic boundary conditions are applied through the use of periodic functions.

We note now that the no-flux case is more challenging in this regard. The boundary condition is now non-local, since the flux (see (2.3)) contains the interaction term. Our approach is based on the Matlab package 2DChebClass (Goddard *et al.*, 2017), which utilizes a Chebyshev pseudo-spectral method in which the convolution integrals are computed in real space. This has been shown to give highly accurate and efficient solutions for related non-local, non-linear PDEs in one and two dimensions. The main challenges over standard implementations (Boyd, 2001; Trefethen, 2000) are: (i) how to compute the convolutions with a finite support ($[x - R, x + R]$); (ii) how to enforce the boundary conditions.

3.1.1 *Computation of convolutions.* For (i), we note that the problem has already been solved in 2DChebClass (Goddard *et al.*, 2017), in particular for the computation of hard-rod free energies in (dynamical) density functional theory. In short, for a given computational point y_n, the interval of interest is $\mathscr{I}_{R,n} = [\max(0, y_n - R), \min(1, y_n + R)]$. For a pseudo-spectral method, given a vector of M spatial

collocation points $\boldsymbol{y} = \{y_m\}$, a function f is approximated by a vector of values at each of these points, which we denote by $\boldsymbol{f}[\boldsymbol{y}]$, such that $\boldsymbol{f}_m = f(y_m)$. Suppose we wish to evaluate the convolution

$$C_R(y) = \int_{\mathscr{I}_R} K(y - y')\rho(y')\mathrm{d}y'$$

for some explicit function K, an interval $\mathscr{I}_R = [\max(0, y' - R), \min(1, y' + R)]$, and a general ρ. This can be computed straightforwardly through matrix multiplication, as follows:

- Discretize the whole domain $[0, 1]$ with a vector of N Chebyshev collocation points, $\boldsymbol{y} = \{y_n\}$;
- Choose a discretization point y_i and determine the interval $\mathscr{I}_{R,i} = [\max(0, y_n - R), \min(1, y_n + R)]$;
- Discretise the interval I_n with $\hat{N}$ Chebyshev collocation points, $\hat{\boldsymbol{y}}_i = \{\hat{y}_{i,\hat{n}}\}$;
- Compute the corresponding integration weight (row) vector corresponding to $\hat{\boldsymbol{y}}_i$, denoted $\hat{\boldsymbol{w}}_i$, such that, $\hat{\boldsymbol{w}}_i \times \boldsymbol{f}[\hat{\boldsymbol{y}}_i] \approx \int_{\mathscr{I}_{R,i}} f(y)\mathrm{d}y$;
- Compute the $\hat{N} \times N$ (barycentric) interpolation matrix, B_i, from $\boldsymbol{y}$ to $\hat{\boldsymbol{y}}_i$, such that $B_i \times \boldsymbol{f}[\boldsymbol{y}] \approx \boldsymbol{f}[\hat{\boldsymbol{y}}_i]$.
- Evaluate the (explicit) kernel K at $y_i - \hat{\boldsymbol{y}}_i$ with values placed on the diagonal of a $\hat{N} \times \hat{N}$ matrix, with all other entries zero. We denote this matrix K_i;
- The i-th row of the $N \times N$ convolution matrix C_R is then given by $(C_{R,i}) = \hat{\boldsymbol{w}}_i \times K_i \times B_i$.

Then, for any length N vector $\boldsymbol{\rho}$ with $\boldsymbol{\rho}_n = \rho(y_n)$, we have $C \times \boldsymbol{\rho} \approx \int_{\mathscr{I}_R} K(y - y')\rho(y')\mathrm{d}y'$.

In words, for each collocation point, we first interpolate ρ from the full interval $[0, 1]$ onto the convolution domain $\mathscr{I}_R$, giving $\rho(y')$. We then multiply by the kernel evaluated at $y - y'$, giving $K(y - y')$. Taking the product of these two results and multiplying by $\hat{w}_i$ then results in an approximation to the convolution. The main advantage of this procedure is that the convolution matrix depends only on R and K, and thus needs only to be computed once for each such pair. It can then be reused at each timestep within a single computation and, indeed, for any computations involving the same convolution. Additionally, exactly the same procedure may be applied to the periodic domains, but with the simplification that the length of the convolution domain is the same for each collocation point.

3.1.2 *Enforcing boundary conditions.* It remains to determine how to accurately enforce the no-flux boundary conditions, i.e., point (ii) above. After discretizing space with a set $\{y_n\}$ of N collocation points, a PDE is reduced to a set of N (coupled) ODEs, describing the value of ρ at each of the y_n. In particular, there are $N - 2$ collocation points in the interior of the domain, and 2 collocation points on the boundary (here, the points at 0 and 1). Two standard approaches to tackling boundary conditions in pseudo-spectral methods are interpolant restriction and boundary bordering (Boyd, 2001; Trefethen, 2000).

For interpolant restriction, one would typically solve a problem with zero Dirichlet (or other simple) boundary conditions and then add on a particular solution that satisfies the desired boundary conditions. The challenge when applying this method lies primarily in determining a particular solution of the (non-local, non-linear) equations which satisfies the boundary conditions; this is far from straightforward in general, and also leads to a problem-specific numerical method.

Suppose the system of ODEs is written as $\mathrm{d}\rho/\mathrm{d}t = L\rho$, for some matrix L corresponding to a discretized linear operator. Then the boundary conditions may be imposed by deleting rows and columns of the matrix L and replacing them by rows and columns which impose the desired boundary conditions. Note that this is again a problem-specific approach, and the substitution effectively has to be performed 'by hand'. It is also still unclear how this could be performed straightforwardly for a non-local, non-linear problem, such as those studied here.

Instead, we use a similar, but more general approach of differential-algebraic equations. We replace the set of N ODEs with the $N-2$ ODEs on the interior, and two algebraic boundary conditions. This set of equations may then be solved with a standard numerical scheme, such as Matlab's ode15s. The key benefits here are that it is straightforward to implement for non-local, non-linear boundary conditions and one does not need to determine anything 'by hand'.

3.1.3 *Even two-periodic boundary conditions.* We now briefly describe how to enforce the even two-periodic boundary conditions. This is done by working on a periodic domain $[-1, 1]$, using an analogous discretization as in the standard periodic case. To enforce the evenness of the solution, we define it only on half of the domain, $[0, 1]$ and then, when operations such as differentiation, integration or interpolation need to be applied, we mirror it to the full domain $[-1, 1]$. Note that, for technical reasons, our numerical implementation actually stores ρ on $[-1, 0]$, but the overall idea is identical. This 'mirroring' operation automatically satisfies the periodic boundary conditions of the solution. However, we note that it can lead to numerical challenges since it does not enforce smoothness of the solution at 0 or ± 1. This can be overcome by choosing an initial condition which is smooth on the whole domain as, at least morally, the diffusive part of the PDE ensures that this smoothness is conserved. However, in our examples, since we define the initial condition as a periodic function on $[0, 1]$, such smoothness is not guaranteed.

3.1.4 *Computational details.* As described above, we use pseudo-spectral methods with Fourier (periodic, even two-periodic) or Chebyshev (no-flux) collocation points. In each case, we use 200 such points, distributed either on $[0, 1]$ (periodic, no-flux) or on $[-1, 1]$ (even two-periodic), and interpolate to 400 evenly spaced points for plotting. For the convolution integrals, we choose $\hat{N} = 100$. We note that these numbers may seem high for pseudo-spectral approaches, but here there are (at least) two motivating reasons. Firstly, we require highly accurate solutions over very long times (e.g., up to time 10^4 or 10^5) and, secondly, in many parameter regimes, the densities become strongly peaked with very high derivatives. We have found that these constraints lead to a requirement for a relatively large number of points *for certain parameter regimes*, notably those with small σ and large R. This is unexpected as such problems become highly advection-dominated. For ease of comparison, we use the same number of points for all computations. We note that a typical dynamical computation takes significantly less than a second on a standard laptop, whilst precomputing the convolution integral matrices (which needs to be done only once for each set of {function, boundary condition, R}) takes on the order of 20–30s for the periodic geometries, and 5s for the Chebyshev case.

For the resulting discretized system of ODEs, we use Matlab's ode15s routine, with relative and absolute tolerances of 10^{-6} and $\varepsilon_{\mathrm{eq}} = 10^{-4}$. We note that for the vast majority of the computations shown here it is possible to reduce the tolerances to around 10^{-10} or even 10^{-12}, with correspondingly smaller $\varepsilon_{\mathrm{eq}}$ but, again for consistency, we fix these values for all computations, unless otherwise stated.

3.1.5 *Validation against existing results.* We refer the reader to Section SM2 of the Supplementary Material for an extensive validation of our numerical codes against existing results from Wang *et al.* (2017) and Kolarijani *et al.* (2019).

3.2 *Discrete dynamics*

For the numerical solution of the SDE, we use the standard Euler-Maruyama method. We note that, since the noise in the model is purely additive, in our case this is equivalent to the Milstein method, and hence has both weak and strong orders of convergence equal to the time step, $\mathrm{d}t$. Unless otherwise stated, we average our results over 10 runs of 10^4 particles with $\mathrm{d}t = 10^{-2}$.

Since the method is standard, we only highlight some particular subtleties with the implementation of the boundary conditions. Here we consider a set of positions $\boldsymbol{x}(t) = \{x_1(t), \ldots, x_N(t), x_{N+1}, \ldots, x_{N+N_r}\}$ at time t, where the first N positions correspond to the normals, and the remaining N_r to the radicals. We define the (signed) separation between two positions x_i and x_j by $S(x_i, x_j)$; for the no-flux case this is simply $x_i - x_j$, whilst for the periodic cases it is the minimum separation of the two particles. For example, for a periodic interval of length 1, $S(0.8, 0.2)$ is –0.4, rather than 0.6. Note that the lengths of the periodic and even two-periodic intervals are one and two, respectively. We define the set of particles within the bounded confidence interval of particle i as $\mathscr{I}_i(\boldsymbol{x}) = \{j : |S(x_i, x_j)| \leq R\}$.

Given $\boldsymbol{x}(t)$, the proposed positions, ignoring the boundary conditions, at the next time step $t + \mathrm{d}t$ are given by

$$\tilde{x}_i(t + \mathrm{d}t) = x_i(t) - \frac{1}{N} \sum_{j \in \mathscr{I}_i(\hat{\boldsymbol{x}})} S(x_i, \hat{x}_j) + \sigma \sqrt{\mathrm{d}t} * w_i(t), \quad i = 1, \ldots, N$$

$$\tilde{x}_i(t + \mathrm{d}t) = x_i(t), \quad i = N + 1, \ldots, N + N_r,$$

where $w_i(t)$ is a random variable drawn from a normal distribution with mean zero and variance one. Here, for the periodic and no-flux boundary conditions, $\hat{\boldsymbol{x}} = \boldsymbol{x}$, whereas for the even two-periodic case, $\hat{\boldsymbol{x}} = \boldsymbol{x} \cup -\boldsymbol{x}$, i.e., the union of the positions and their reflections.

Given the proposed positions at time $t + \mathrm{d}t$, it is necessary to apply the boundary conditions. The periodic case is standard; after each time step we ensure that the positions lie in the interval $[0, 1]$ by taking $x_i(t+\mathrm{d}t) = \tilde{x}_i(t+\mathrm{d}t) \mod 1$. For the no-flux case, we use standard reflective boundary conditions (Pilipenko, 2014): If $\tilde{x}_i(t + \mathrm{d}t) \in [0, 1]$ then set $x_i(t + \mathrm{d}t) = \tilde{x}_i(t + \mathrm{d}t)$. If $\tilde{x}_i(t + \mathrm{d}t) \notin [0, 1]$ then propose a new $\tilde{x}_i(t + \mathrm{d}t)$ by reflecting at the closest boundary (i.e., 0 or 1); repeat until $\tilde{x}_i(t + \mathrm{d}t) \in [0, 1]$ and then set $x_i(t + \mathrm{d}t) = \tilde{x}_i(t + \mathrm{d}t)$. Note that, in practice, for sufficiently small $\mathrm{d}t$, it is highly unlikely to require more than one reflection at a given time step to ensure that $\tilde{x}_i$ lies in the interval. For the even two-periodic case, by the symmetry of the problem any individual moving out of the interval $[0, 1]$ has a corresponding 'mirror' individual that moves exactly the same distance into the interval across the same boundary. Thus, one way of implementing the even two-periodic boundary conditions is to apply the same procedure as for the no-flux case. Note that this *does not* result in the same dynamics due to the different choices of $\hat{\boldsymbol{x}}$.

Finally, we note that it is necessary to be able to sample from a given density of particles, both for the initial condition and for the radical population. For simple distributions, such as uniform or Gaussian, this can be achieved using standard pseudo-random number generators. However, to provide a general approach, e.g., to sample from the triangular radical distribution used in Kolarijani *et al.* (2019), we implemented a sampling algorithm using slice sampling, which is a Monte-Carlo approach capable of sampling from any given density. To compare such densities, and the results of the SDE dynamics, to

the PDE results, it is necessary to histogram the results and correctly normalise them; for a system with N particles, which is to be averaged over N_{runs} runs, and binned into a histogram with bin width $\mathrm{d}x$, with bin counts $\{b_i\}$, the correct normalization is $b_i/(NN_{\text{runs}}\mathrm{d}x)$.

3.3 *Comparison between the discrete and continuum models*

We direct the reader to Sections SM3 and SM4 of the Supplementary Material for a comparison between the SDE and PDE models. Of particular note is that a relatively large number of agents (of the order 10^3–10^4) are required to ensure good agreement between the two models. This is a significantly larger number of agents than used in previous studies, e.g., Kolarijani *et al.* (2019); Wang *et al.* (2017).

4. Numerical experiments: opinion dynamics in the absence of radicals

We begin our numerical experiments by studying noisy opinion dynamics models without radicals. For a given boundary condition, there are only three other properties of the model: the values of R and σ, and the initial condition.

4.1 *Uniform initial condition*

Our first system has a uniform initial condition, which is standard in many opinion dynamics models, see, e.g., Hegselmann & Krause (2015, 2002); Holzer & Khatri (2017); Kolarijani *et al.* (2019); Motsch & Tadmor (2014); Pluchino *et al.* (2006); Wang *et al.* (2017). This initial condition is a steady state for both the periodic and even two-periodic boundary conditions, which can be seen in 1, and easily shown using the stationary PDE. Here the different values of the order parameter are due only to the intrinsic dependence of Q on R. However, this is the first demonstration of the significant effect of the boundary conditions on the dynamics; the uniform distribution *is not* a steady state for the no-flux case, and the density tends to cluster in the centre of the interval. This is a result of the density being zero outside the domain $[0, 1]$, so there tends to be an inwards net force on the individuals. There is a clear trend as R and σ change: increasing R tends to enhance the cluster formation, whilst the opposite is true for increasing σ. This has a simple explanation: for larger R there is more interaction, and hence the individuals tend towards what is essentially the global mean, rather than the local mean with a small confidence interval; increasing σ adds more diffusion to the system, which tends to disperse clusters.

There is another striking feature of the middle plot in 1, namely that there seems to be a sudden switch in behaviour as R and/or σ are varied, with the long-time density being either strongly peaked or almost flat. This is particularly noticeable when σ is small. As such, in the middle panel of Figure 1, we zoom in to a transition region (left panel) and also consider smaller values of σ (right panel). The bottom panels of Figure 1 show the time to equilibrium, as defined in Section 2.4.1, on a log scale. The dynamics take appreciably longer to reach equilibrium in regions where the long-time dynamics is particularly sensitive to the choice of parameters. Note that in the small-σ case, it was necessary to increase the equilibration tolerance to $\varepsilon_{\text{eq}} = 5 \times 10^{-4}$; there is overlap between the two plots at $\sigma = 0.05$ to aid comparison.

In Figures 2 and 3, we show the dynamics of parameter pairs labelled in the middle and right subplots of Figure 1, respectively. We show snapshots of the density at various times (top) and the order parameter as a function of time (bottom), with coloured dots corresponding to the time–Q values for the snapshots in the upper panels. For the cases A–E in Figure 2, we see how the behaviour changes for fixed $\sigma = 0.05$ and increasing R. For small R, a shallow, almost uniform, cluster slowly develops. As R increases the dynamics become richer. At first, the density develops two peaks, with the order

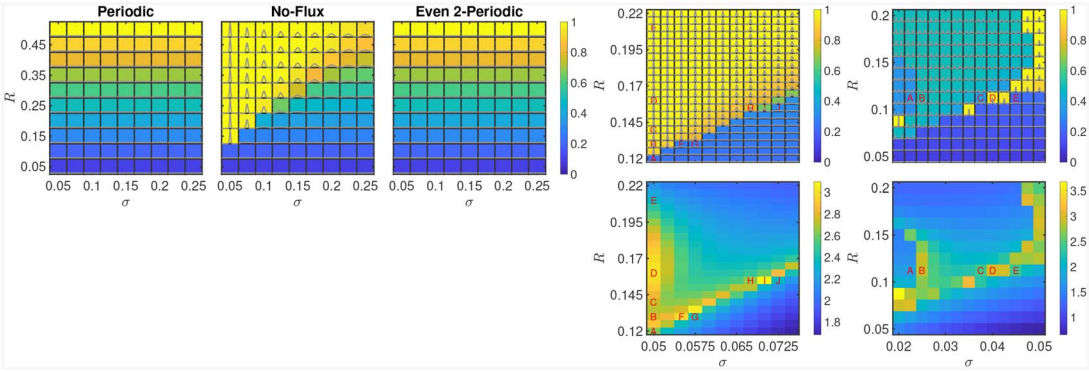

FIG. 1. Results for a uniform initial condition with varying R and σ. Colour denotes the order parameter at the final time (top) and equilibration time on a log scale (bottom); grey lines are the final densities. The middle and right plots are for no-flux boundary conditions and letters refer to snapshots in Figures 2 and 3.

parameter rising to a plateau. The two peaks then move together, before merging into a single cluster, indicated by a larger Q. In particular, for case E, where $R = 0.21$, the initial state has $Q \approx 0.4 \approx 2R$, as expected, it then rises to $Q \approx 0.5$, indicating the presence of two clusters, before ending at $Q \approx 1$, and a single cluster. Cases (B, F, G) and (H, I, J) demonstrate the effects of fixing R and increasing σ. In (B, F, G), we begin in the small-σ regime (B), where a single, steep cluster forms almost directly from the uniform state, As σ increases (F, G), diffusion dominates resulting in an approximately uniform distribution. Similar behaviour is observed for larger R (H, I, J).

Figure 3 shows the corresponding dynamics for the small-σ case, fixing $R = 0.125$ (see the right hand plot of Figure 1). Here we observe a much richer collection of possible long-time states and dynamics; to aid visualization, we plot the density as a function of time and space in the bottom panels. For small $\sigma = 0.0225$ (A), we observe a final state with three clusters, and final-order parameter approximately $1/3$. For slightly larger $\sigma = 0.025$ (B), the final state has two clusters, with order parameter around $1/2$, but the dynamics clearly pass through a transient state with three, non-equal clusters (see the snapshots in panel B at times 100 and 500). Increasing σ further next results in a direct transition to a two-cluster state (C), followed by direct formation of a single cluster (D) and eventually an essentially disordered/uniform long-time state (E). We note that there are similar transitions when fixing σ and varying R (not shown).

Here we find it informative to compare the right panel of Figure 1 with Figure 3 of the original HK paper (Hegselmann & Krause, 2002), which shows the long-time equilibria (for noiseless dynamics) as R is increased. For small R, the state is homogeneous, whilst increasing R results first in two clusters and then a single, central cluster, with rapid transitions as R increases. As stated in Hegselmann & Krause (2002), as R increases 'we step from fragmentation (plurality) over polarization (polarity) to consensus (conformity)'. This is a direct analogue of our results just described. In contrast, for the two periodic boundary conditions, the uniform initial condition is an equilibrium, and we see no such R-induced transitions. *This suggests that no-flux boundary conditions more faithfully reproduce the results, and underlying mechanisms, of the original models.*

We also compare to Figure 1 of Proskurnikov & Tempo (2018), which demonstrates the dependence on R (their d) of the final number of clusters and equilibration time. They note that both dependencies are non-monotonic, whereas it may be intuitively expected that increasing R causes a reduction in the number of clusters and a decrease in the equilibration time. Such effects are also visible when,

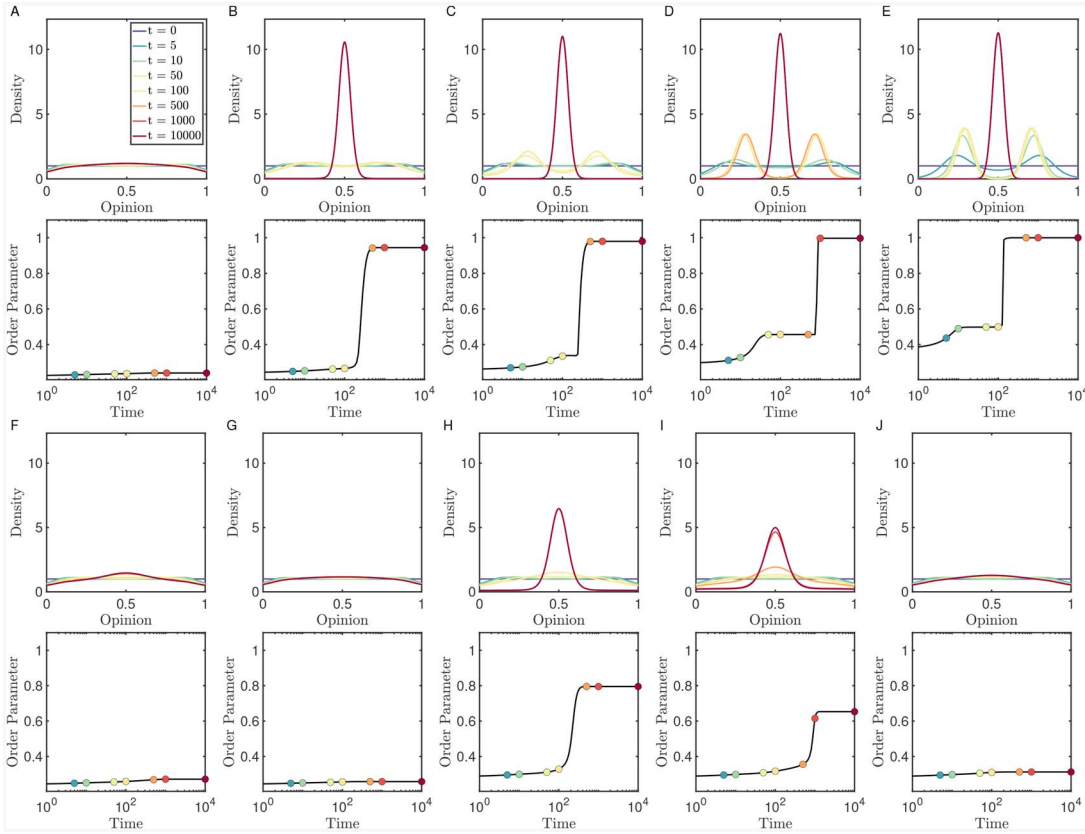

FIG. 2. Results for a uniform initial condition and no-flux boundary conditions. Letters correspond to labelling in the middle panel of Figure 1. In each case, we show snapshots of the densities at the indicated times (top) and the time evolution of the order parameter (bottom); coloured dots show the order parameter at the times of the corresponding snapshots in the top panel.

e.g., fixing $\sigma = 0.2$ and varying R in the right panel of Figure 1. This phenomenon of 'abnormally' slow convergence has also been demonstrated in Lorenz (2006), who described the resulting states as metastable. Similar sensitivities have also been observed in a noisy DW model (Carro *et al.*, 2013), which also noted the importance of the initial condition on determining the long-time dynamics; we will now investigate the further choices of initial condition. In Supplementary Material SM4, we show a comparison with the SDE for short times; the agreement is very good. The phenomenon of dynamical metastability for noisy opinion dynamics PDE models, also observed in Garnier *et al.* (2016), will be studied in future work.

4.2 *Single Gaussian initial condition*

In this section, we investigate the effects of a non-uniform initial condition with a single, relatively broad cluster/consensus. Following Wang *et al.* (2017), we choose

$$\rho_0(y) = Z^{-1} \exp\big(-C[d(y, y_0)]^2\big), \tag{4.1}$$

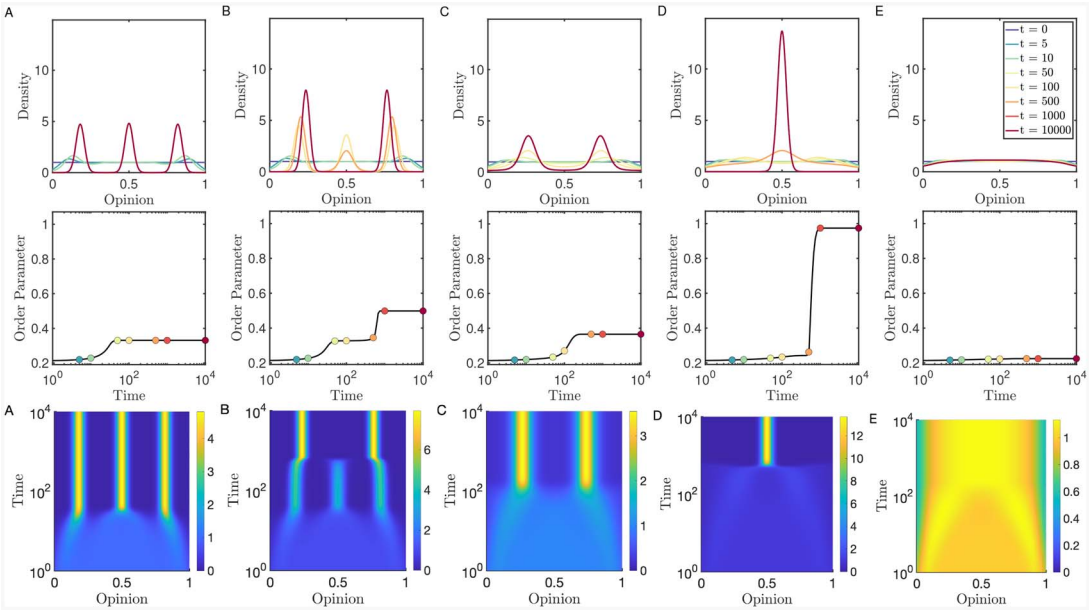

FIG. 3. As Figure 2, but for smaller values of σ; labels correspond to captions in the right panel of Figure 1. Additionally, we show time-space plots of the densities in the bottom row.

where $C = 20$ and Z is the normalization constant. Here $d(x, y)$ denotes the one-periodic distance between two points. In Wang *et al.* (2017), y_0 was chosen to be 0.5 but this is irrelevant beyond visualization in the periodic case. However, for the no-flux and even two-periodic cases, the choice of y_0 can result in qualitatively different dynamics. For example, if $y_0 = 0.5$ then the initial condition is symmetric and the periodic and even two-periodic cases are identical; for other y_0 this is not the case.

In the top panels of Figure 4, we display the final order parameters, densities and equilibration times for a range of values of R and σ, for all three boundary conditions, and y_0 equal to 0.3 (left), and 0.2 (right). We also show a zoom in parameter space for $y_0 = 0.3$ (bottom left), and results for an initial condition which is a Gaussian mixture (bottom right); see Section 4.3. White regions in the equilibration times denote simulations which have not converged.

We first discuss some general trends. With the exception of $y_0 = 0.2$ and even two-periodic boundary conditions (top right plot), the final result is either a single cluster, or an (almost) uniform state. This is to be expected as the only force which can break up the initial cluster is diffusion, which favours the uniform state. The formation of one large cluster on the left and one small cluster on the right for $y_0 = 0.2$, even two-periodic boundary conditions, small σ and large R is due to the periodicity of the initial condition. A small amount of mass is originally concentrated at the right end of the domain, near $y = 1$, and gets trapped due to the attraction to its periodic image near $y = -1$. Relatedly, we note the additional range of Q for the even two-periodic case, in particular for large R, where the density is significantly influenced by its periodic/even image. As expected, increasing σ also causes a trend to disordered states, whereas increasing R tends to increase the sharpness of the resulting cluster, with increased time-to-equilibrium in regions separating qualitatively different final densities.

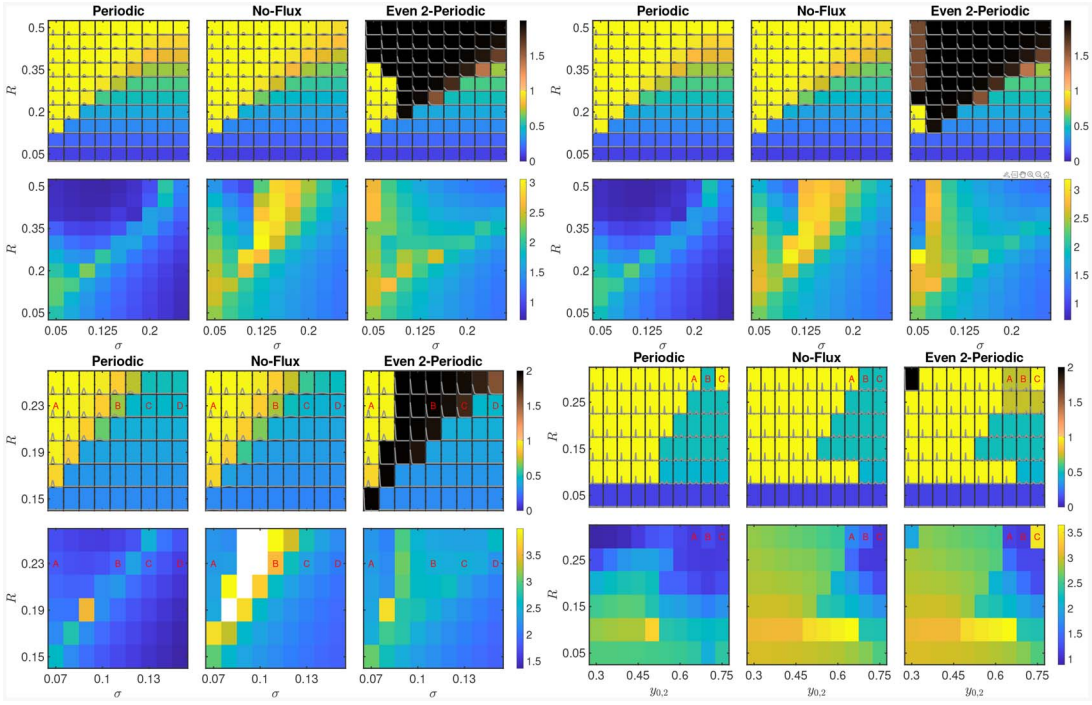

FIG. 4. As Figure 1 but for a Gaussian initial condition (4.1) with $y_0 = 0.3$ (top left, bottom left) and $y_0 = 0.2$ (top right), and an initial condition which is a linear combination of Gaussians (4.2) (bottom right) for different boundary conditions. Note the additional colour bar scale from 1 to 2 for the even two-periodic case. White denotes simulations which have not converged by the final time (10^4).

There are also some unexpected observations. A feature of the no-flux boundary conditions is the movement towards 0.5 of the final maximum opinion for fixed R, as σ increases. This is likely a consequence of the competition between diffusion and attraction; for smaller values of σ, the noise is not sufficiently strong to disperse the original cluster, whereas larger noise can cause the cluster to move. Examples of this dynamics can be seen in Figures 5; case B ($R = 0.23, \sigma = 0.11$) shows a clear drift in the mean opinion over time for the no-flux boundary conditions. Parameter sets B ($R = 0.23, \sigma = 0.11$) and C ($R = 0.23, \sigma = 0.13$) for the even two-periodic case demonstrate the significant effect that this choice of boundary condition has on the dynamics, with a strong cluster forming at the left hand edge. It is also interesting to note that the behaviour of the order parameter is not monotonic in a number of cases. We have found excellent agreement with the associated SDE agent-based model, at least for short times for which the computational cost of the SDEs is reasonable; see Supplementary Material Section SM4.

4.3 *Two Gaussians initial condition*

To demonstrate the effect of multi-modal initial conditions, we study a linear combination of two Gaussians, or a Gaussian mixture:

$$\rho_0(y) = Z^{-1}\Big[\exp\big(-C[d(y, y_{0,1})]^2\big) + \exp\big(-C[d(y, y_{0,2}))]^2\big)\Big], \tag{4.2}$$

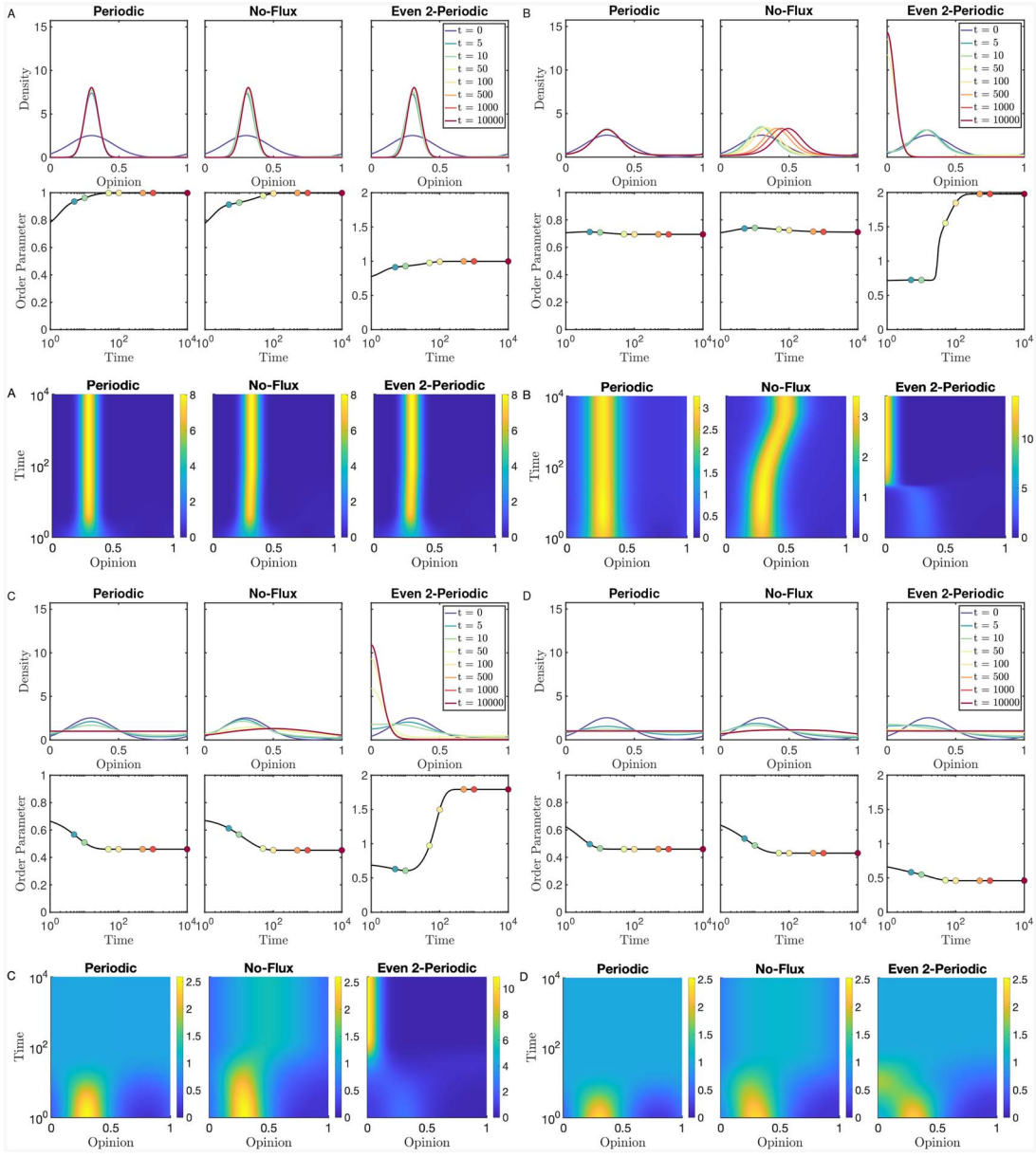

FIG. 5. As Figure 3 but for a Gaussian initial condition. Labels correspond to those in the bottom left panel of Figure 4, for A ($R = 0.23$, $\sigma = 0.07$), B ($R = 0.23$, $\sigma = 0.11$), C ($R = 0.23$, $\sigma = 0.13$) and D ($R = 0.23$, $\sigma = 0.15$).

where Z is the normalization constant. We note that there are many parameters in this setup, and presumably also a correspondingly large number of interesting transitions between regimes as the parameters are varied, but for conciseness, we fix $C = 80$, $y_{0,1} = 0.2$ and $\sigma = 0.03$. In the bottom right plot of Figure 4, we show the final-time density, corresponding order parameter and equilibration time as

R and $y_{0,2}$ are varied. We note a general trend, as in previous cases, that (with a few notable exceptions) increasing R for fixed $y_{0,2}$ increases the tendency for consensus to form, and also for consensus to be closer to the centre of the interval. For fixed R, increasing $y_{0,2}$ (i.e., separating the initial clusters) tends to favour the formation of two clusters, rather than a single consensus.

We focus on three particular pairs of parameters $(y_{0,2}, R)$, denoted A (0.65,0.3), B (0.7,0.3) and C (0.75,0.3) in Figures 4 and 6. In Figure 6, we plot the results up to $t = 10$, which captures much of the interesting dynamics and enables comparison to the SDE results for 10^4 particles. In all cases, the even two-periodic boundary condition result is qualitatively different; the long-time behaviour has two clusters, one close to zero and one close to the mean of the second initial cluster. The cause of this is the 'mirror' density, which attracts the initial clusters at 0.2 and –0.2 together. Note that the order parameter is close to 1, rather than the usual value of 1/2 denoting two clusters. Increasing $y_{0,2}$ in the other two boundary conditions results in a switch from a single cluster for $y_{0,2} = 0.65$ (A) to two clusters for $y_{0,2} = 0.70$ (B). The cause of this seems to be that increasing the separation of the clusters for fixed R causes the two clusters to sharpen, before they can begin to coalesce, and they are then too well-separated to be drawn together. This is related to the $2R$-conjecture regarding the separation of stable clusters (Wang *et al.*, 2017). Finally, increasing $y_{0,2}$ to 0.7 (C) demonstrates an additional effect of periodic boundary conditions. The long-term behaviour reverts to a single cluster, this time centred near 1, rather than towards the centre of this interval. The cause of this is the periodic nature of the domain – there are two 'distances' between the initial clusters, and the shorter of the two now crosses the 0–1 point. This causes the cluster to form to the right of the initial cluster at 0.7, rather than to the left. This is an issue for interpretation in terms of extreme opinions, and suggests that if this is an aim of the model then the no-flux boundary conditions are a more appropriate choice, both for stability and interpretability.

5. Numerical experiments: the effect of radicals

In this section, we will introduce a distribution of radicals, and investigate the sensitivity of the dynamics on this distribution. Motivated by the observation that the most interesting dynamics occur for small σ and a uniform initial distribution, we restrict to this regime here.

5.1 *Uniform initial condition and Gaussian radicals*

We first consider a single Gaussian radical distribution of the form

$$M\rho_r(y) = MZ^{-1} \exp\big(-C[d(y, y_0)]^2\big), \tag{5.1}$$

where Z is a normalization constant, and M determines the mass of the radical population. We now have five parameters in the system: R and σ, as before, and y_0, C and M, which describe the mean, width and mass of the radicals, respectively. To reduce the number of parameters, and to set the radical distribution to be relatively narrow, we fix $C = 800$.

In Figure 7, we fix two of the remaining parameters, and vary the other two: σ–R ($y_0 = 0.7$, $M = 0.1$), M–R ($\sigma = 0.02$, $y_0 = 0.7$), y_0–σ ($R = 0.1$, $M = 0.1$) and y_0–R ($\sigma = 0.02$, $M = 0.1$). The corresponding snapshots are given in Figures 8–10. The radical populations are shown in red (Fig. 7), or black (Figs 8 –10) although due to their small size and strongly peaked normal distributions, they are sometimes hard to distinguish.

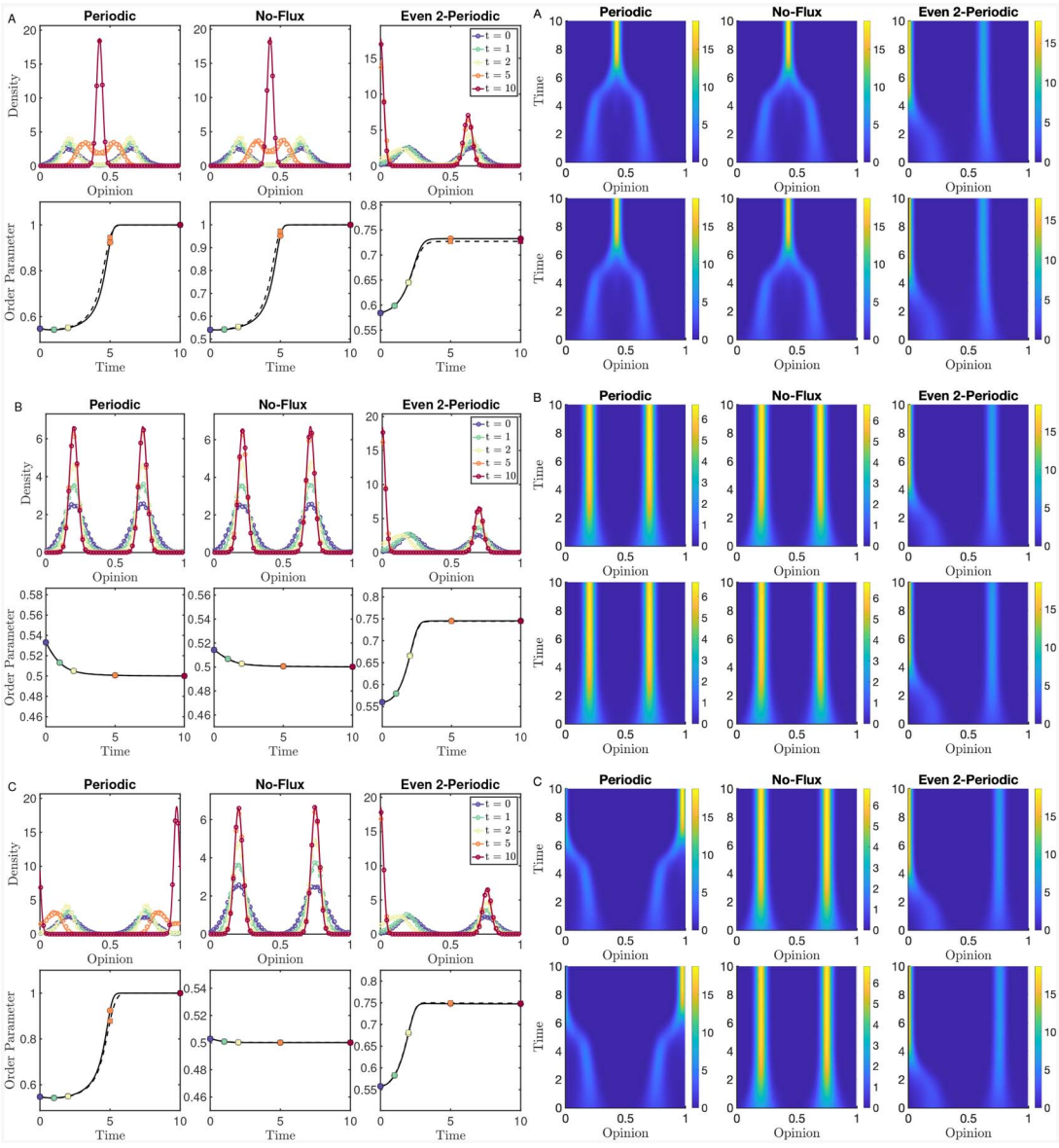

FIG. 6. As Figure 3 but for an initial condition which is a linear combination of Gaussians. Labels correspond to those in the bottom right panel of Figure 4, for A ($R = 0.3$, $y_{0,2} = 0.65$), B ($R = 0.3$, $y_{0,2} = 0.7$), and C ($R = 0.3$, $y_{0,2} = 0.75$). Additionally, we show the solutions of the corresponding SDEs with 10^4 particles. Densities shown by circles, order parameter by dashed lines and squares, and time-space plots in the bottom row of each panel.

5.1.1 *R–σ [Figures 7 (top left) and 8].* For small, fixed R, in all three boundary conditions we see a similar behaviour as in the case of no radicals; small σ results in a single cluster, which is now centred around the mean of the radical distribution, whilst increasing σ results in a uniform, or almost-

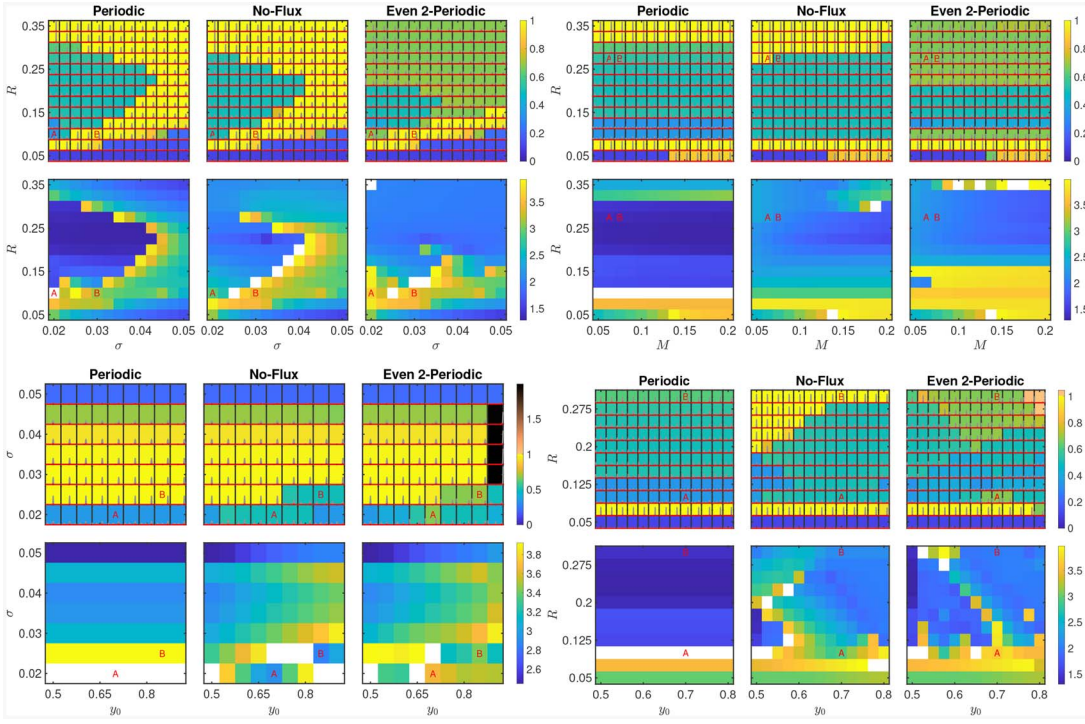

FIG. 7. As Figure 1 but for a uniform initial condition and a Gaussian radical distribution (5.1). Here we vary: (top left) the strength of the noise (σ) and the size of the confidence interval (R); (top right) the mass of the radicals (M) and the confidence bound (R); (bottom left) the mean position of the radical distribution (y_0) and the strength of the noise (σ); (bottom right) the mean position of the radical distribution (y_0) and the confidence bound (R). Radical distributions are shown in red.

uniform distribution. This is perhaps what one would expect intuitively. For larger R, the long-time dynamics is less intuitive. For the periodic and no-flux boundary conditions, fixing R and increasing σ causes a transition from multiple clusters (two or three) to a single one; one cluster is always centred around the radical distribution. For the even two-periodic case, there are three qualitatively different long-time distributions. For small sigma, two clusters emerge, as for the other boundary conditions, and for larger σ, there is once again a single cluster. However, for intermediate σ there is a different state with a cluster centred around the radical distribution, and another narrow cluster at zero; once again this additional state is due to the attractive nature of the mirror population. As before, we observe much longer equilibration times for regions of phase space on the border of different long-time distributions.

In Figure 8, we focus on two pairs of parameters with (R,σ) equal to (0.1,0.02) [A] and (0.1,0,03) [B]. For A, the boundary condition has a strong effect on the dynamics. The periodic case rapidly forms three clusters, corresponding to an order parameter of around $1/3$. As can be seen from the white region in Figure 7 this simulation has not reached our definition of equilibrium, so it is possible that this is not the final state. However, slightly increasing R appears to make the three-cluster state stable. In the no-flux case, there are initially four clusters, two of which are relatively weak, which eventually merge into two strong clusters, once centred around the radicals, and one at the other side of the interval. As with the other boundary conditions, this suggests that, with these parameter values, the radicals cause an

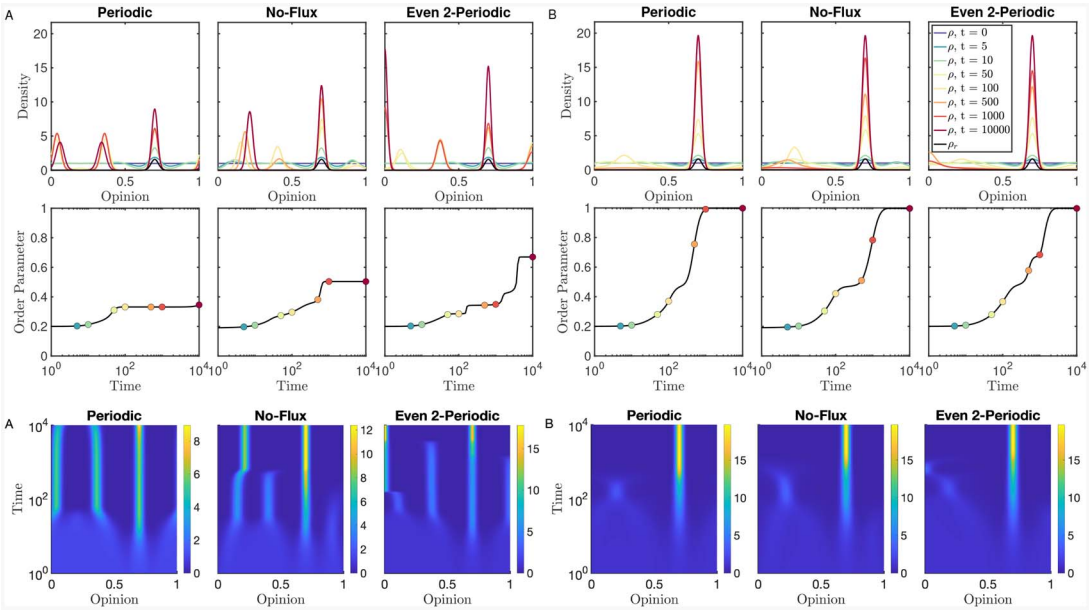

FIG. 8. As Figure 3 but for a uniform initial condition and a Gaussian radical distribution. Labels correspond to parameter values in the top left panel of 7, for A ($\sigma = 0.02$ and $R = 0.1$) and B ($\sigma = 0.03$ and $R = 0.1$).

initial division in the population, but do not have enough influence to cause a single consensus to form. Finally, in the even two-periodic case, the dynamics is even more complex; this is well demonstrated by the order parameter. The four clusters at times around time 10^2, look similar to those in the no-flux case, and once again the final state has two clusters. However, here the additional clusters persist for much longer, and the second cluster is near to zero, rather than centred at approximately 0.2.

For B, the final state for all three boundary conditions is relatively similar, with a single, strong cluster centred around the radicals. However, once again, the dynamical path to this distribution depends on the boundary conditions. In all three cases, at times up to around 100, a secondary cluster on the left of the interval is visible. In the periodic case, this then rapidly merges with the final cluster, whilst in the other two cases it is longer-lived, first moving away from the final cluster, before eventually merging into it. At time 500, the secondary cluster in the no-flux case is located near zero, as in case A; here we suggest that a larger value of σ increases the diffusion to a point at which the weaker clusters are dispersed and can be captured by the strong cluster centred around the radicals. The clusters, especially the large one, are also slightly wider with the increased value of σ, which may aid this coalescence.

5.1.2 *M–R [Figures 7 (top right) and 9 (left)].* From Figure 7, it is clear that the dependence on M is very weak in the two periodic cases, unless R is very small. For larger R, and no-flux boundary conditions, varying M can produce unintuitive results. Comparing A and B on the left of 9 we see that increasing M causes the long-time distribution to switch from a single cluster around the radical distribution (which is something one would expect), to a bimodal distribution, in which the increase in the number of radicals has caused polarization. A possible explanation for this is that a larger radical cluster attracts the nearby normal population more quickly, causing a split in the distribution after short

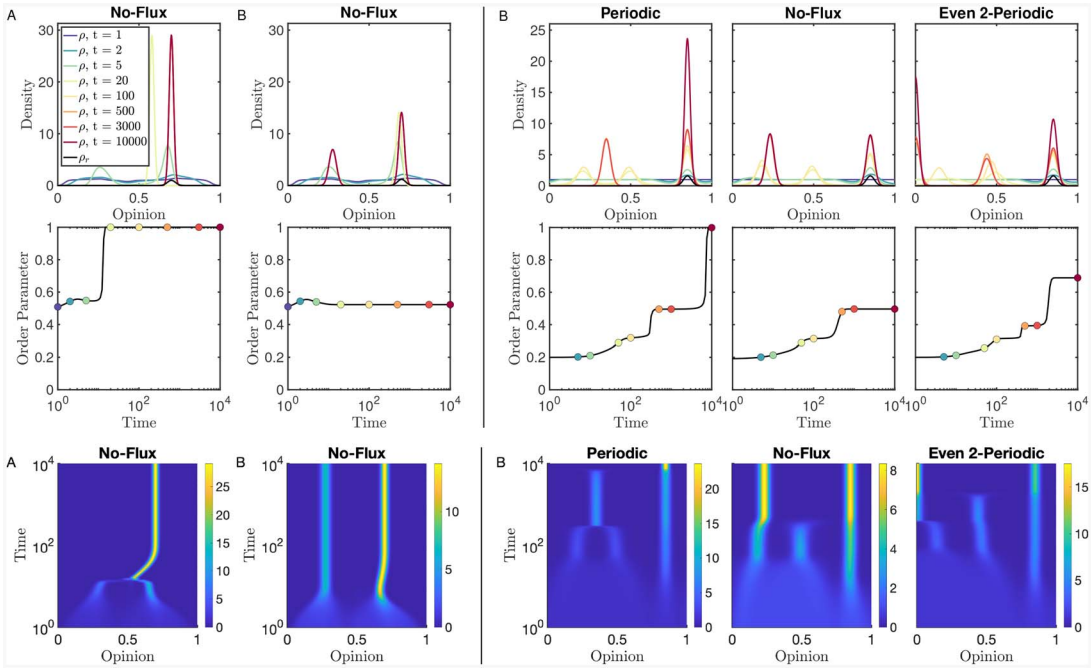

FIG. 9. As Figure 8 but for: (left) the labels in the top right panel of Figure 7 for A ($M = 0.0625$ and $R = 0.275$) and B ($M = 0.075$ and $R = 0.275$); (right) the labels in the bottom left panel of Figure 7 with B ($y_0 = 0.85$ and $\sigma = 0.025$).

times; this can be seen in both cases A and B. In case A, the two clusters move towards each other (with the one near the radicals actually moving away from the radical population), before forming a single cluster away from the radical distribution, which then migrates towards the radicals. In case B, it appears that the radicals are now sufficiently strong to prevent the initial cluster around them from moving towards the centre of the interval, leaving the two polarized clusters.

5.1.3 *y_0–σ [Figures 7 (bottom left) and 9 (right)].* As is clear from Figure 7, the results of the periodic case are independent of y_0, up to a shift in the opinion axis. In the other cases, the final order parameter and shape of the distribution are essentially independent of y_0 for all but small σ, or for large (or, by symmetry, small) y_0 when the no-flux and even two-periodic BCs become influential. As y_0 increases, there is a clear, corresponding shift in position of final-time cluster, and, as before, increasing σ leads to disorder. Case A corresponds to Case A in Figure 8. Case B (see right of Figure 9) corresponds to slightly larger σ and increased y_0, for which the dynamics is strongly-dependent on the choice of boundary conditions. The periodic case has a single cluster, whilst no-flux has two, which are almost the same size, and even two-periodic has two, but with one near zero and significantly taller.

5.1.4 *y_0–R [Figures 7 (bottom right) and 10].* In Figure 7, we see that the different regimes of the no-flux and even two-periodic cases are much richer, with multiple parameter regions with qualitatively different long-time dynamics. As could be expected, increasing R, and hence the interaction range, leads to bigger differences between the results for different boundary conditions. In Figure 10, we show some representative dynamics, demonstrating the richness and complexity as the parameters are varied.

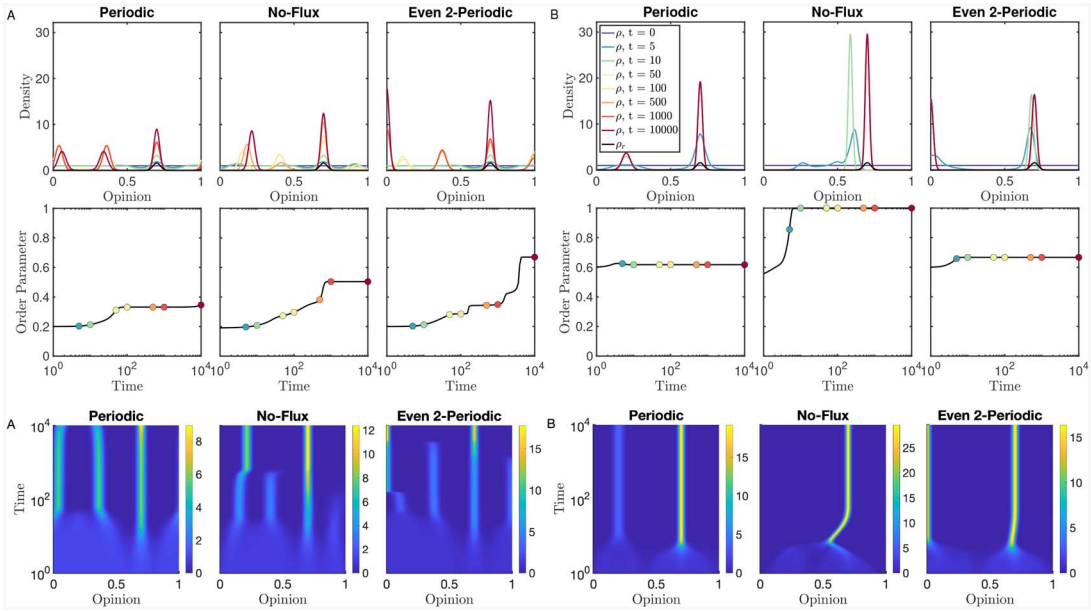

FIG. 10. As Figure 8 but for the labels in the bottom right panel of Figure 7, with A ($y_0 = 0.7$ and $R = 0.1$), and B ($y_0 = 0.7$ and $R = 0.3$).

5.2 *Uniform initial condition and double-Gaussian radicals*

As a second example, we consider a double-Gaussian distribution of radicals

$$M\rho_r(y) = MZ^{-1}\Big[\lambda \exp\big(-C[d(y, y_{0,1})]^2\big) + (1-\lambda)\exp\big(-C[d(y, y_{0,2})]^2\big)\Big], \tag{5.2}$$

where, once again, Z is the normalization constant for the term in square brackets, and M is the mass of the radical distribution. Here, $\lambda \in [0, 1]$ is a parameter describing the relative masses of the two Gaussians. The physical interpretation here is that we have two competing groups of radicals. To simplify exposition, whilst demonstrating interesting effects, we first fix $C = 800$ (to ensure a concentrated radical distribution) and $M = 0.1$. We consider two choices of λ, namely $\lambda = 0.5$, which corresponds to a completely even split of radicals, and $\lambda = 0.499$, which is a slight bias towards one of the radical opinions.

In Figure 11, we choose choose $y_{0,1} = 1 - y_{0,2} = 0.2$, and vary R and σ with $\lambda = 0.5$ (left) and $\lambda = 0.499$ (right). By symmetry, the periodic and even two-periodic cases are identical, for $\lambda = 0.5$. We note similar trends to before: increasing R favours the formation of a single cluster, whilst increasing σ tends to produce an almost uniform distribution. For the periodic and even two-periodic cases, for moderate R, increasing σ causes a transition from two clusters to a single one; we interpret this as the noise becoming large enough to overcome the attraction of the radical distribution. This does not occur in the no-flux case, at least for the range of parameters studied here. When a single cluster is formed, we see significant effects of the boundary conditions. For no-flux, the cluster is centred either in the middle of the interval ($\lambda = 0.5$), or on top of the dominant radical cluster ($\lambda = 0.499$). In contrast, the periodic and even two-periodic cases result in clusters centred around 0/1, caused by the

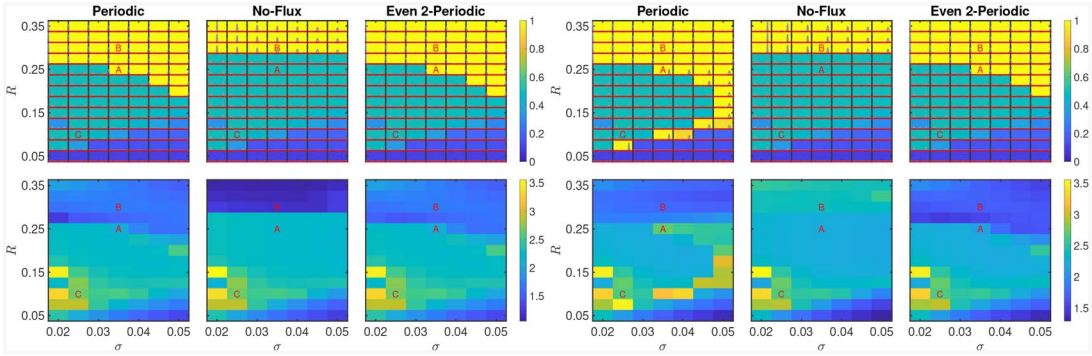

FIG. 11. As Figure 1 but for a uniform initial condition and a double-Gaussian radical distribution (5.2) with $\lambda = 0.5$ (left) and $\lambda = 0.499$ (right). Here we vary the strength of the noise (σ) and the confidence bound (R).

shortest distance between the radical clusters being across 0/1, rather than through 0.5. Physically, this means that two competing populations of radicals can result in either a moderate cluster (for no-flux), or two extreme clusters (for periodic and even two-periodic). This is a significant effect of the boundary conditions, and suggests issues when interpreting such results without giving careful consideration to the modelling choices. Similar sensitivities with respect to the choices of $y_{0,1}$ and $y_{0,2}$ are demonstrated in Supplementary Material Section SM5.

In cases A ($\sigma = 0.035$ and $R = 0.25$) and B ($\sigma = 0.035$ and $R = 0.3$), we see that strong dependence on the choice of λ. In the periodic case with parameters A, the short-time dynamics are similar for both values of λ, but the symmetry breaking with $\lambda = 0.499$ is very clear in the long-time dynamics. For the no-flux case with parameters B, two initial clusters merge into a single cluster at the centre of the interval, with the asymmetry becoming visible only at longer times when the cluster migrates to the right with $\lambda = 0.499$. We highlight that these are far from the only interesting and non-intuitive results from the model. For parameters C ($\sigma = 0.025$ and $R = 0.1$), we show only the $\lambda = 0.5$ periodic case, as the other cases are very similar. Here we interpret the two radical clusters as competing political viewpoints. Whilst the long-time behaviour in all cases is a pair of equal clusters centred around the radicals, our interest here lies in the intermediate dynamics, where there is a clear third peak centred around zero. As discussed in Boudin & Salvarani (2016) in the context of the Scottish independence referendum, this is a common feature of such situations in which the population mostly polarizes, but leaves a number of 'undecided' individuals, who only move to one of the popular opinions after long times. It would be interesting to study such situations further.

6. Conclusions and outlook

We have demonstrated the significant effect of the choice of boundary condition in SDE and PDE bounded confidence models, as well as the sensitivity of such models to small changes in various parameters, with and without the inclusion of radical distributions. This work clearly demonstrates that no-flux boundary conditions are the correct choice from a modelling perspective, and also in terms of bridging the gap between agent-based (Hegselmann & Krause, 2015, 2002) and PDE models. In particular, the no-flux choice most faithfully recreates the underlying dynamics of the original deterministic models Hegselmann & Krause (2015, 2002).

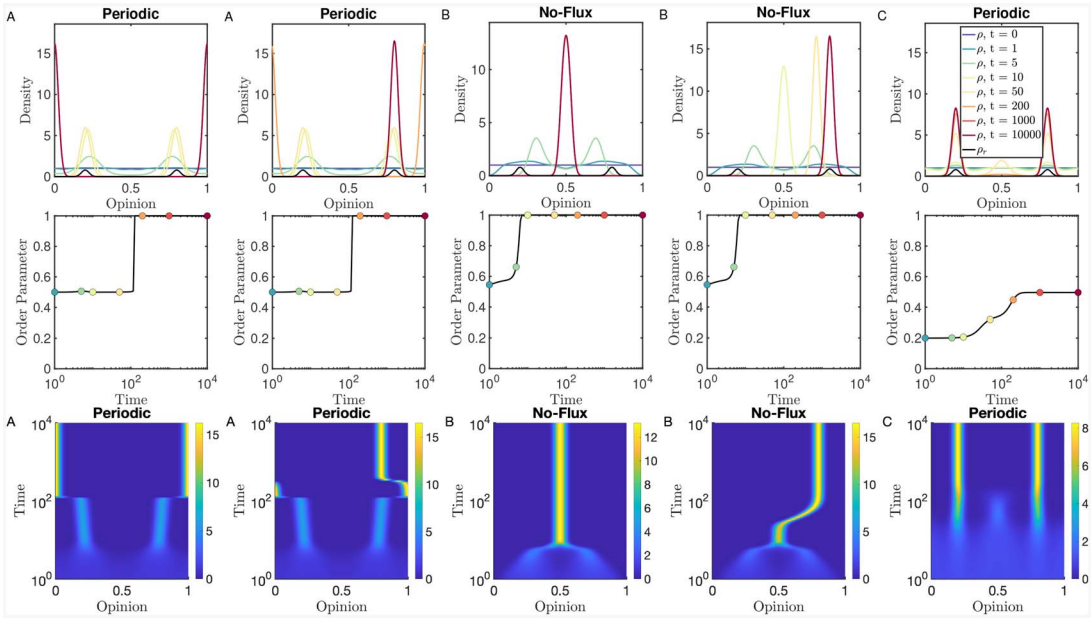

FIG. 12. As Figure 3, but with a uniform initial condition and double Gaussian radical distribution. Labels correspond to parameter values in Figure 11, with A ($\sigma = 0.035$ and $R = 0.25$). B ($\sigma = 0.035$ and $R = 0.3$). C ($\sigma = 0.025$ and $R = 0.1$).

There are many possible extensions, some of which have been studied in the literature for other models in opinion dynamics: confidence intervals which are asymmetric (Hegselmann & Krause, 2002), heterogeneous (Hendrickx & Olshevsky, 2016; Mirtabatabaei & Bullo, 2012; Zhao *et al.*, 2016) or time-dependent (Morarescu & Girard, 2010); influence that is negative (Proskurnikov & Tempo, 2018), or which increases with separation (Motsch & Tadmor, 2014); stubbornness or inertia (Proskurnikov & Tempo, 2018); persuasiveness or supportiveness (Hołyst *et al.*, 2001); multi-dimensional models (Bhattacharyya *et al.*, 2013; Etesami & Başar, 2015; Etesami *et al.*, 2013; Fortunato *et al.*, 2005; Martinsson, 2016; Nedić & Touri, 2012; Pluchino *et al.*, 2006); more general models for the noise, e.g., coloured, multiplicative or non-Gaussian (see Gomes *et al.* (2020)); slowly varying radical distributions; a rigorous, systematic derivation of the order parameter, perhaps using data-driven approaches. We note that most of these can be incorporated in both our models and numerical schemes in relatively straightforward ways, although there are open questions regarding SDEs with no-flux boundary conditions (Lions & Sznitman, 1984).

On the analytical side, there are interesting questions regarding rigorous analysis of the SDE model, as well as the rigorous derivation of the PDE model, For the no-flux case, the linear stability analysis of the stationary state(s) is a challenging problem, due to the non-linear and non-local boundary conditions. Furthermore, the study of fluctuations around the mean field limit, in particular close to the phase transition will provide us with detailed information about the behaviour of the finite-agent system. In the absence of radical groups and when the underlying network has fully connectivity (corresponding to the large R limit) and for when σ sufficiently small, a centred Gaussian is an approximate stationary state when we consider no-flux boundary conditions. This is similar to the case of periodic boundary conditions (Garnier *et al.*, 2017; Wang *et al.*, 2017), and follows from the fact that localized stationary

states do not feel the effect of the boundary conditions. The stability of such approximate localized states, corresponding to clusters, is a very interesting problem that we plan to address in future work.

In terms of applications, we plan to use real-world data, particularly from referenda, to study the use of such models in practical settings, with a focus on the transient dynamics rather than the long-term behaviour. There are also interesting questions regarding the use of such models when applied to advertising, or charismatic leaders (Hegselmann & Krause, 2015). Another important open question is how to determine the modelling parameters σ and R. This is related to inference for McKean SDEs, and we hope to address this topic in future work. We direct the interested reader to a study of estimating parameters in mean field models (Sharrock *et al.*, 2021), which could act as a starting point for such studies. By estimating modelling parameters, such as the noise, it should be possible to develop diagnostic tools for predicting consensus formation, or lack thereof.

Supplementary material

Supplementary material is available at *IMAMAT* online.

Funding

EPSRC (EP/P031587/1, EP/L024926/1 and EP/L020564/1 to G.A.P). This research was funded in part by JPMorgan Chase & Co. Any views or opinions expressed herein are solely those of the authors listed, and may differ from the views and opinions expressed by JPMorgan Chase & Co. or its affiliates. This material is not a product of the Research Department of J.P. Morgan Securities LLC. This material does not constitute a solicitation or offer in any jurisdiction.

AcknowledgEments

B.D.G. is grateful to Andrew Archer, Valerio Restocchi, and David Sibley for helpful discussions.

References

Ben-Naim, E. (2005) Opinion dynamics: rise and fall of political parties. *EPL*, **69**, 671.

Bhattacharyya, A., Braverman, M., Chazelle, B. & Nguyen, H. L. (2013) On the convergence of the Hegselmann-Krause system. *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science*, pp. 61–66.

Blondel, V. D., Hendrickx, J. M. & Tsitsiklis, J. N. (2007) On the 2R conjecture for multi-agent systems. *2007 European Control Conference (ECC)*. IEEE, pp. 874–881.

Blondel, V. D., Hendrickx, J. M. & Tsitsiklis, J. N. (2010) Continuous-time average-preserving opinion dynamics with opinion-dependent communications. *SIAM J. Control Optim.*, **48**, 5214–5240.

Boudin, L. & Salvarani, F. (2016) Opinion dynamics: kinetic modelling with mass media, application to the Scottish independence referendum. *Phys. A*, **444**, 448–457.

Boyd, J. P. (2001) *Chebyshev and Fourier Spectral Methods*. Courier Corporation.

Brugna, C. & Toscani, G. (2015) Kinetic models of opinion formation in the presence of personal conviction. *Phys. Rev. E (3)*, **92**, 052818.

Bullo, F., Cortes, J. & Martinez, S. (2009) *Distributed Control of Robotic Networks: A Mathematical Approach to Motion Coordination Algorithms*. Princeton University Press.

Canuto, C., Fagnani, F. & Tilli, P. (2012) An Eulerian approach to the analysis of Krause's consensus models. *SIAM J. Control Optim.*, **50**, 243–265.

Carrillo, J., Gvalani, R., Pavliotis, G. & Schlichting, A. (2020) Long-time behaviour and phase transitions for the McKean–Vlasov equation on the torus. *Arch. Ration. Mech. An.*, **235**, 635–690.

Carro, A., Toral, R. & San Miguel, M. (2013) The role of noise and initial conditions in the asymptotic solution of a bounded confidence, continuous-opinion model. *J. Statist. Phys.*, **151**, 131–149.

Castellano, C., Fortunato, S. & Loreto, V. (2009) Statistical physics of social dynamics. *Rev. Modern Phys.*, **81**, 591.

Chazelle, B., Jiu, Q., Li, Q. & Wang, C. (2017) Well-posedness of the limiting equation of a noisy consensus model in opinion dynamics. *J. Differer. Equations*, **263**, 365–397.

Cucker, F. & Smale, S. (2007a) Emergent behavior in flocks. *IEEE T. Automat. Contr.*, **52**, 852–862.

Cucker, F. & Smale, S. (2007b) On the mathematics of emergence. *Jap. J. Math*, **2**, 197–227.

Czirók, A., Barabási, A.-L. & Vicsek, T. (1999) Collective motion of self-propelled particles: kinetic phase transition in one dimension. *Phys. Rev. Lett.*, **82**, 209.

Deffuant, G., Neau, D., Amblard, F. & Weisbuch, G. (2000) Mixing beliefs among interacting agents. *Adv. Complex Syst.*, **3**, 87–98.

Düring, B., Markowich, P., Pietschmann, J.-F. & Wolfram, M.-T. (2009) Boltzmann and Fokker–Planck equations modelling opinion formation in the presence of strong leaders. *P. R. Soc. A*, **465**, 3687–3708.

Etesami, S. R. & Başar, T. (2015) Game-theoretic analysis of the Hegselmann-Krause model for opinion dynamics in finite dimensions. *IEEE T. Automat. Contr.*, **60**, 1886–1897.

Etesami, S. R., Başar, T., Nedić, A. & Touri, B. (2013) Termination time of multidimensional Hegselmann-Krause opinion dynamics. *2013 American Control Conference*. IEEE, pp. 1255–1260.

Fortunato, S., Latora, V., Pluchino, A. & Rapisarda, A. (2005) Vector opinion dynamics in a bounded confidence consensus model. *Int. J. Mod. Phys. C*, **16**, 1535–1551.

Garnier, J., Papanicolaou, G. & Yang, T.-W. (2016) Mean field model for collective motion bistability. *Discrete Contin. Dyn. Syst. - B*, **24**(2), 851–879.

Garnier, J., Papanicolaou, G. & Yang, T.-W. (2017) Consensus convergence with stochastic effects. *Vietnam J. Math.*, **45**, 51–75.

Goddard, B., Mills-Williams, R., Ottobre, M. & Pavliotis, G. (2021) Well-posedness of overdamped dynamic density functional theory. in preparation.

Goddard, B., Mills-Williams, R. & Pavliotis, G. (2020) Well-posedness and equilibrium behaviour of overdamped dynamic density functional theory. arXiv preprint arXiv:2002.11663.

Goddard, B. D., Nold, A. & Kalliadasis, S. (2017) *2DChebClass [Software]*. http://dx.doi.org/10.7488/ds/1991.

Gomes, S., Pavliotis, G. & Vaes, U. (2020) Mean field limits for interacting diffusions with colored noise: phase transitions and spectral numerical methods. *Multiscale Model. Simul.*, **18**, 1343–1370.

Grauwin, S. & Jensen, P. (2012) Opinion group formation and dynamics: structures that last from nonlasting entities. *Phys. Rev. E (3)*, **85**, 066113.

Hegselmann, R. & Flache, A. (1998) Understanding complex social dynamics: a plea for cellular automata based modelling. *JASSS*, **1**, 1.

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

Hegselmann, R., Krause, U. et al. (2002) Opinion dynamics and bounded confidence models, analysis, and simulation. *JASSS*, **5**, 1–33.

Hendrickx, J. M. & Olshevsky, A. (2016) On symmetric continuum opinion dynamics. *SIAM J. Control Optim.*, **54**, 2893–2918.

Hołyst, J. A., Kacperski, K. & Schweitzer, F. (2001) Social impact models of opinion dynamics. *Annual Reviews of Computational Physics IX*. World Scientific, pp. 253–273.

Holzer, M. & Khatri, R. (2017) Pattern formation, traveling fronts and consensus versus fragmentation in a model of opinion dynamics. *Phys. Lett. A*, **381**, 3197–3202.

Jadbabaie, A., Lin, J. & Morse, A. S. (2003) Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE T. Automat. Contr.*, **48**, 988–1001.

Keller, E. F. & Segel, L. A. (1970) Initiation of slime mold aggregation viewed as an instability. *J. Theor. Biol.*, **26**, 399–415.

Kolarijani, M., Proskurnikov, A. V. & Esfahani, P. M. (2019) Macroscopic noisy bounded confidence models with distributed radical opinions. arXiv preprint arXiv:1905.04057.

Kubo, R., Toda, M. & Hashitsume, N. (2012) *Statistical Physics II: Nonequilibrium Statistical Mechanics*, vol. **31**. Springer Science & Business Media.

Lions, P.-L. & Sznitman, A.-S. (1984) Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.*, **37**, 511–537.

Lord, C. G., Ross, L. & Lepper, M. R. (1979) Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *J. Pers. Soc. Psychol.*, **37**, 2098.

Lorenz, J. (2006) Consensus strikes back in the Hegselmann-Krause model of continuous opinion dynamics under bounded confidence. *JASSS*, **9**, 1.

Lorenz, J. (2007) Continuous opinion dynamics under bounded confidence: a survey. *Int. J. Mod. Phys. C*, **18**, 1819–1838.

Martinsson, A. (2016) An improved energy argument for the Hegselmann–Krause model. *J. Differ. Equ. Appl.*, **22**, 513–518.

Masuda, N. (2015) Opinion control in complex networks. *New J. Phys.*, **17**, 033031.

Mirtabatabaei, A. & Bullo, F. (2012) Opinion dynamics in heterogeneous networks: convergence conjectures and theorems. *SIAM J. Control Optim.*, **50**, 2763–2785.

Mirtabatabaei, A., Jia, P. & Bullo, F. (2014) Eulerian opinion dynamics with bounded confidence and exogenous inputs. *SIAM J. Appl. Dyn. Sys.*, **13**, 425–446.

Mogilner, A. & Edelstein-Keshet, L. (1999) A non-local model for a swarm. *J. Math. Biol.*, **38**, 534–570.

Mogilner, A., Edelstein-Keshet, L., Bent, L. & Spiros, A. (2003) Mutual interactions, potentials, and individual distance in a social aggregation. *J. Math. Biol.*, **47**, 353–389.

Morarescu, I.-C. & Girard, A. (2010) Opinion dynamics with decaying confidence: application to community detection in graphs. *IEEE T. Automat. Contr.*, **56**, 1862–1873.

Motsch, S. & Tadmor, E. (2011) A new model for self-organized dynamics and its flocking behavior. *J. Statist. Phys.*, **144**, 923.

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

Nedić, A. & Touri, B. (2012) Multi-dimensional hegselmann-krause dynamics. *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. Maui, HI: IEEE, pp. 68–73.

Nold, A., Goddard, B. D., Yatsyshin, P., Savva, N. & Kalliadasis, S. (2017) Pseudospectral methods for density functional theory in bounded and unbounded domains. *J. Comp. Phys.*, **334**, 639–664.

Nordio, A., Tarable, A., Chiasserini, C.-F. & Leonardi, E. (2017) Belief dynamics in social networks: a fluid-based analysis. *IEEE Trans. Netw. Sci. Eng.*, **5**, 276–287.

Pilipenko, A. (2014) *An Introduction to Stochastic Differential Equations with Reflection*, vol. **1**. Universitätsverlag Potsdam.

Pineda, M., Toral, R. & Hernandez-Garcia, E. (2009) Noisy continuous-opinion dynamics. *J. Statist. Phys.*, **2009**, P08001.

Pineda, M., Toral, R. & Hernández-García, E. (2011) Diffusing opinions in bounded confidence processes. *Eur. Phys. J. D*, **62**, 109–117.

Pineda, M., Toral, R. & Hernández-García, E. (2013) The noisy Hegselmann-Krause model for opinion dynamics. *Eur. Phys. J. B*, **86**, 490.

Pluchino, A., Latora, V. & Rapisarda, A. (2006) Compromise and synchronization in opinion dynamics. *Eur. Phys. J. B*, **50**, 169–176.

Proskurnikov, A. V. & Tempo, R. (2017) A tutorial on modeling and analysis of dynamic social networks. *Part I. Anu. Rev. Control*, **43**, 65–79.

Proskurnikov, A. V. & Tempo, R. (2018) A tutorial on modeling and analysis of dynamic social networks. Part II. *Anu. Rev. Control*, **45**, 166–190.

Sharrock, L., Kantas, N., Parpas, P. & Pavliotis, G. A. (2021) Parameter estimation for the McKean-Vlasov stochastic differential equation. arXiv preprint arXiv:2106.13751.

Skorokhod, A. V. (1961) Stochastic equations for diffusion processes in a bounded region. *Theor. Probab. Appl.*, **6**, 264–274.
Stauffer, D. (2001) Monte Carlo simulations of Sznajd models. *JASSS*, **5**, 1.
Su, W., Chen, G. & Hong, Y. (2017) Noise leads to quasi-consensus of Hegselmann–Krause opinion dynamics. *Automatica J. IFAC*, **85**, 448–454.
Sznajd-Weron, K., Tabiszewski, M. & Timpanaro, A. M. (2011) Phase transition in the Sznajd model with independence. *EPL*, **96**, 48002.
Topaz, C. M., Bertozzi, A. L. & Lewis, M. A. (2006) A nonlocal continuum model for biological aggregation. *B. Math. Biol.*, **68**, 1601.
Toral, R. & Tessone, C. J. (2007) Finite size effects in the dynamics of opinion formation. *Commun. Comput. Phys.*, **2**, 177–195.
Trefethen, L. N. (2000) *Spectral Methods in MATLAB*. SIAM.
Turchin, P. (2016) *Ages of Discord: A Structural-Demographic Analysis of American History*. Beresta Books.
Turchin, P. (2018) *Historical Dynamics: Why States Rise and Fall*, vol. 26. Princeton University Press.
Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. & Shochet, O. (1995) Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, **75**, 1226.
Villani, C. (2003) *Topics in Optimal Transportation*. Number 58. American Mathematical Soc.
Wang, C., Li, Q., Weinan, E. & Chazelle, B. (2017) Noisy Hegselmann-Krause systems: phase transition and the 2R-conjecture. *J. Statist. Phys.*, **166**, 1209–1225.
Weisbuch, G., Deffuant, G., Amblard, F. & Nadal, J.-P. (2002) Meet, discuss, and segregate!. *Complexity*, **7**, 55–63.
Wongkaew, S., Caponigro, M. & Borzi, A. (2015) On the control through leadership of the Hegselmann–Krause opinion formation model. *Math. Mod. Meth. Appl. S.*, **25**, 565–585.
Yang, Y., Dimarogonas, D. V. & Hu, X. (2014) Opinion consensus of modified Hegselmann–Krause models. *Automatica J. IFAC*, **50**, 622–627.
Yeomans, J. M. (1992) *Statistical Mechanics of Phase Transitions*. Clarendon Press.
Yildiz, E., Acemoglu, D., Ozdaglar, A. E., Saberi, A. & Scaglione, A. (2011) Discrete opinion dynamics with stubborn agents. *SSRN*, **1744113**, 1–39. http://dx.doi.org/10.2139/ssrn.1744113.
Yildiz, E., Ozdaglar, A., Acemoglu, D., Saberi, A. & Scaglione, A. (2013) Binary opinion dynamics with stubborn agents. *ACM Trans. Econ. Comput. (TEAC)*, **1**, 1–30.
Zhao, Y., Zhang, L., Tang, M. & Kou, G. (2016) Bounded confidence opinion dynamics with opinion leaders and environmental noises. *Comput. Oper. Res.*, **74**, 205–213.