Hydrodynamic models of preference formation in multi-agent societies

Lorenzo Pareschi[*]   Giuseppe Toscani[†]   Andrea Tosin[‡]   Mattia Zanella[§]

Abstract

In this paper, we discuss the passage to hydrodynamic equations for kinetic models of opinion formation. The considered kinetic models feature an opinion density depending on an additional microscopic variable, identified with the personal preference. This variable describes an opinion-driven polarisation process, leading finally to a choice among some possible options, as it happens e.g. in referendums or elections. Like in the kinetic theory of rarefied gases, the derivation of hydrodynamic equations is essentially based on the computation of the local equilibrium distribution of the opinions from the underlying kinetic model. Several numerical examples validate the resulting model, shedding light on the crucial role played by the distinction between opinion and preference formation on the choice processes in multi-agent societies.

Keywords: Opinion and preference formation, choice processes, kinetic modelling, hydrodynamic equations

Mathematics Subject Classification: 35L65, 35Q20, 35Q70, 35Q91, 82B21

1   Introduction

The mathematical modelling of opinion formation in multi-agent societies has enjoyed in recent years a growing attention [4, 5, 6, 29, 36, 46]. In particular, owing to their cooperative nature, the dynamics of opinion formation have been often dealt with resorting to methods typical of statistical mechanics [14, 16]. Among other approaches, kinetic theory served as a powerful basis to model fundamental interactions among the so-called agents [8, 9, 10, 17, 30, 47] and to provide a sound structure for related applications [1, 49]. In kinetic models, analogously to the kinetic theory of rarefied gases, the mechanism leading to the opinion variation is given by binary, i.e. pairwise, interactions among the agents. Then, depending on the parameters of such microscopic rules, the society develops a certain macroscopic equilibrium distribution [37, 40], which describes the formation of a relative consensus about certain opinions.

Two main aspects are usually taken into account in designing the elementary binary interactions. The first is the *compromise* [20, 52], namely the tendency of the individuals to reduce the distance between their respective opinions after the interaction. The second is the *self-thinking* [47], i.e. an erratic individual change of opinion inducing unpredictable deviations from the prescribed deterministic dynamics of the interactions.

Recently, many efforts have been devoted to include further details in the opinion formation models, so as to capture more and more realistic phenomena. The usual strategy consists in

---

[*]Department of Mathematics and Computer Sciences, University of Ferrara, Via Machiavelli 35, 44121 Ferrara, Italy (`lorenzo.pareschi@unife.it`)

[†]Department of Mathematics "F. Casorati", University of Pavia, Via Ferrata 1, 27100 Pavia, Italy (`giuseppe.toscani@unipv.it`)

[‡]Department of Mathematical Sciences "G. L. Lagrange", Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy (`andrea.tosin@polito.it`)

[§]Department of Mathematical Sciences "G. L. Lagrange", Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy (`mattia.zanella@polito.it`)

taking additional behavioural aspects into account, such as the *stubbornness* of the agents [34, 45], the emergence of *opinion leaders* [25, 27], the influence of *social networks* [2, 48], the *expertise* in decision making tasks [41], the *personal conviction* [3, 7, 11, 19]. Generally, the aim of such additional parameters is to model on one hand the resistance of the agents to change opinion and, on the other hand, the prominent role played by some individuals in attracting others towards their opinions. In all these contributions, the additional variables act as modifiers of the microscopic interactions. This means that they affect the process of opinion formation but are not affected in turn by the evolving opinions.

In the present paper we aim instead at modelling a parallel process to opinion formation, namely the formation of *preferences*. The preference, which is driven by, but need not coincide with, the opinion, is here understood as representative of the *choice* that an agent makes among some possible options, such as e.g. some candidates in an election or yes/no in a referendum [18]. As such, and unlike the opinion, the preference evolves towards necessarily *polarised* states, which reflect the available options.

To pursue this goal, we consider a novel class of inhomogeneous kinetic models for the joint distribution function $f(t, \xi, w)$, where ξ is the variable describing the preference and w the one describing the opinion. In particular, $f(t, \xi, w)d\xi dw$ is the proportion of agents who, at time $t \geq 0$, express a preference in the interval $[\xi, \xi + d\xi]$ and simultaneously an opinion in the interval $[w, w + dw]$. As far as the modelling of the interactions leading to the evolution of the opinion is concerned, we take advantage of the well consolidated background recalled before. Conversely, concerning the evolution of the preference, we assume transport-type dynamics of the form

$$\frac{d\xi}{dt} = (w - \alpha)\Phi(\xi).$$

Here, in analogy with the classical kinematics, w plays morally the role of the velocity, i.e. it drags the preference in time, however biased by a *perceived social opinion* α, which accounts for the predominant social feeling. Moreover, the zeros of the function Φ define the options where the preference may polarise.

These ingredients lead us to an inhomogeneous Boltzmann-type kinetic equation of the form

$$\partial_t f + (w - \alpha)\partial_\xi(\Phi(\xi)f) = Q(f, f),$$

where Q is a Boltzmann-type collision operator encoding the opinion formation interaction dynamics. From here, analogously to the classical kinetic theory of rarefied gases, we derive macroscopic equations for the density $\rho = \rho(t, \xi)$ and the mean opinion $m = m(t, \xi)$ of the agents with preference ξ at time t by means of a local equilibrium closure based on the identification of the local equilibrium distribution of the opinions – the equivalent of a local "Maxwellian". The precise type of hydrodynamic equations that we obtain in this way depends on whether the mean opinion of the agents is or is not conserved in time by the microscopic dynamics of opinion formation. For instance, if it is conserved we get the following system of *conservation laws*:

$$\begin{cases} \partial_t \rho + \partial_\xi \left(\Phi(\xi)\rho(m - \alpha)\right) = 0, \\ \partial_t(\rho m) + \partial_\xi \left(\Phi(\xi)(M_{2,\infty} - \alpha\rho m)\right) = 0, \end{cases}$$

where $M_{2,\infty}$ denotes the energy of the local Maxwellian. For special classes of local equilibrium distributions, that we compute from the collisional kinetic equation by an asymptotic procedure reminiscent of the *grazing collision limit* of the classical kinetic theory [47, 50], we can express such an energy analytically in terms of the hydrodynamic parameters ρ, m. We recover therefore a self-consistent macroscopic model, which we show to be *hyperbolic* for all the physically admissible values of (ρ, m) and able to reproduce the preference polarisations, viz. the choices, discussed above.

In more detail, the paper is organised as follows. In Section 2 we give preliminary microscopic insights into the joint process of opinion and preference formation, stressing in particular the role of the perceived social opinion α. In Section 3 we move to an aggregate analysis of the opinion

formation by means of Boltzmann-type kinetic models, studying in particular their steady states which, as set forth above, pave the way for the identification of the local equilibrium distributions needed in the passage to the hydrodynamic equations. In Section 4 we derive various types of macroscopic models of preference formation out of the aforementioned inhomogeneous Boltzmann-type equation and we link them precisely to key features of the microscopic interactions among the agents. In Section 5 we present several numerical tests, both at the kinetic and at the macroscopic scales, which exemplify the distinction between the preference formation and the opinion formation processes and show how much such a distinction enhances the interpretation of the social dynamics. Finally, in Section 6 we discuss further developments and research prospects.

2 A microscopic look at the opinion-preference interplay

The first mathematical models of consensus formation in opinion dynamics were proposed in the form of systems of ordinary differential equations (ODEs) describing the behaviour of a finite number of agents. After the pioneering works [21, 28], which introduced simple agent-based models to understand the effects of the influence among connected individuals, many research efforts have been devoted to the construction of sophisticated differential models of opinion formation. Besides those dealing with simple consensus dynamics, in recent years new models of more realistic social phenomena have been proposed to capture additional aspects. Without intending to review all the literature, we give here some references on certain classes of models for finite systems: the celebrated Hegselmann-Krause model [32], which considers bounded-confidence-type interactions to stress the impact of homophily in learning processes, see also [15, 35]; models incorporating leader-follower effects [38]; models of social interactions on realistic networks [51]; models of opinion control [12].

An aspect which, to our knowledge, has been so far basically disregarded in mathematical models of opinion dynamics, in spite of its realism, is the distinction between the *opinion* in the strict sense of the individuals about single issues and their overall *preference*, which, in some cases, is actually mainly responsible for their choices. For instance, in case of referendums or political elections, the preference of a voter can be identified with his/her *voting intention*, which may not always coincide with his/her opinion on every topic debated during the election campaign. Obviously, the preference evolves in time with the opinion, in such a way that if a certain opinion persists for a sufficiently long time it can affect the preference considerably. For example, a voter with a voting intention biased towards right (left, respectively) parties, who however finds him/herself frequently in agreement with the positions taken by left (right, respectively) parties on key topics of the election campaign, may end up with a final vote opposite to his/her original intention.

As a foreword to the subsequent contents of the paper, in this section we present some preliminary considerations about the effect of the interplay between opinion and preference, taking advantage of a deterministic microscopic model for a finite number of agents. Let us consider then a system composed by N agents with microscopic state given by a pair $(\xi_i, w_i) \in [-1, 1]^2$, where w_i is the opinion of the ith agent, ξ_i is his/her preference and $i = 1, \ldots, N$. Sticking to a standard custom in the literature of models of opinion dynamics, we describe mathematically the opinion of an agent as a bounded scalar variable w_i conventionally taken in the interval $[-1, 1]$. In particular, we understand the values $w_i = \pm 1$ as the two extreme opinions, while $w_i = 0$ as the maximum of indecisiveness. Since, from the physical point of view, the preference is commensurable with an opinion, we adopt the same mathematical conventions also for the variable ξ_i.

In order to highlight the different roles of the opinion and the preference variables, we rely on the analogy with the classical kinetic theory of rarefied gases. There, a molecule moving on a line is characterised by its position $x \in \mathbb{R}$ and its velocity $v \in \mathbb{R}$. In the absence of external forces, the velocity remains constant, whereas the position varies according to the kinematic law

$$\frac{dx}{dt} = v. \tag{1}$$

3

In practice, a particle with positive velocity will move rightwards, while a particle with negative velocity will move leftwards. In first approximation, it seems natural to assume that the opinion plays the role of the velocity and the preference that of the position. Indeed, at least in the case in which an agent has to end up with one of the two preferences ± 1, like e.g. in a referendum, one can assume that a large part of the agents with positive opinion will move their preferences rightwards, while agents with negative opinion will move their preferences leftwards. Clearly, one cannot resort to a law like (1), which allows the position to increase or decrease indefinitely in time: a correction is required in order to maintain the preference variable in the allowed interval $[-1, 1]$.

A primary example of the analogy just discussed is provided below, where the time evolution of the opinions and the preferences of the agents is modelled by the ODE system

$$
\begin{cases}
\dfrac{d\xi_i}{dt} = (w_i - \alpha)\Phi(\xi_i) & (2) \\[2ex]
\dfrac{dw_i}{dt} = \dfrac{1}{N}\sum_{j=1}^{N}P(w_i,\, w_j)(w_j - w_i) & (3)
\end{cases}
$$

for $i = 1, \ldots, N$, supplemented with initial conditions $(\xi_i(0),\, w_i(0)) = (\xi_{0,i},\, w_{0,i}) \in [-1,\, 1]^2$. The second equation describes standard alignment dynamics among the opinions of the agents, i.e. *consensus*, driven by the interaction/compromise function $0 \leq P(\cdot,\, \cdot) \leq 1$, see e.g., [40, 47]. The first equation describes instead the evolution of the preference of the ith agent based on the signed distance between its true opinion w_i and a reference opinion $\alpha \in [-1,\, 1]$ perceived in the society, which we will refer to as the *perceived social opinion*. The function $\Phi : [-1,\, 1] \to [0,\, 1]$ has to be primarily chosen so as to guarantee that $\xi_i(t) \in [-1,\, 1]$ for all $t > 0$. However, as we will see in a moment, this function will be also useful to take into account meaningful *polarisations* of the preferences.

In model (2)-(3), the coupling between opinion and preference is actually one-directional, indeed the evolution of ξ_i depends on that of the w_i's but not vice versa. In particular, the system (3) for the w_i's can be solved *a priori*, before analysing the dynamics (2) of the ξ_i's. Let us consider, in particular, the case of interactions with bounded confidence, which are described by taking

$$P(w_i, w_j) = \chi(|w_j - w_i| \leq \Delta), \tag{4}$$

where χ denotes the characteristic function and $\Delta \in [0,\, 2]$ is a given confidence threshold, above which agents do not interact because their opinions are too far away from each other. If $\Delta = 0$ then only agents with the very same opinion interact, whereas if $\Delta = 2$ then we speak of all-to-all interactions, considering that $|w_j - w_i| \leq 2$ for all $w_i,\, w_j \in [-1,\, 1]$. The latter case is actually equivalent to choosing $P \equiv 1$.

Depending on the value of Δ, one can observe a loss of global consensus. Asymptotically, the opinions may form several clusters, whose number is dictated by Δ and by the initial conditions $w_{0,1}, \ldots, w_{0,N}$, see Figure 1. However, since the function P given in (4) is symmetric, i.e. $P(w_i,\, w_j) = P(w_j,\, w_i)$ for all $i,\, j = 1, \ldots, N$, the mean opinion $\frac{1}{N}\sum_{i=1}^{N} w_i$ is conserved in time and for all $t > 0$ coincides, in particular, with the mean opinion at $t = 0$.

Now we give some insights into the preference dynamics modelled by (2), at least under special forms of the function Φ. In order to avoid that the preference ξ_i leaves the interval $[-1,\, 1]$, a very natural condition is $\Phi(\pm 1) = 0$. This implies that the constant functions $\xi_i(t) = -1$ and $\xi_i(t) = 1$ are indeed stationary solutions to (2) and may therefore represent attractive or repulsive equilibria of the system, depending on the sign of $w_i - \alpha$.

For instance, we may choose

$$\Phi(\xi) = 1 - |\xi|.$$

Taking for granted from (3) that $w_i(t) \leq 1$ for all $t > 0$, if we integrate (2) starting from an initial condition $\xi_{0,i} \in [0,\, 1]$ then for all times $t > 0$ in which ξ_i remains non-negative we find

$$\xi_i(t) \leq 1 - (1 - \xi_{0,i})e^{-(1-\alpha)t} \leq 1.$$

4

(a) $\Delta = 1$	(b) $\Delta = 0.4$	(c) $\Delta = 0.2$

Figure 1: Solution of (3) with $N = 50$ agents and P given by (4) for decreasing values of the confidence threshold Δ. The initial opinions $w_{0,i}$ have been sampled uniformly in $[-1, 1]$. The ODE system has been integrated numerically via a standard fourth order Runge-Kutta method.

Likewise, taking for granted from (3) that $w_i(t) \geq -1$ for all $t > 0$, if we start from an initial condition $\xi_{0,i} \in [-1, 0]$ then for all times $t > 0$ in which ξ_i remains non-positive we deduce

$$\xi_i(t) \geq -1 + (1 + \xi_{0,i})e^{-(1+\alpha)t} \geq -1.$$

This argument, applied to the various time intervals in which ξ_i has constant sign, shows indeed that $\xi_i(t) \in [-1, 1]$ for all $t \geq 0$. Nevertheless, we cannot solve (2) exactly, because from (3) we cannot calculate exactly the function $t \mapsto w_i(t)$. On the other hand, we can get a useful idea at least of the large time dynamics of (2) by fixing w_i to its asymptotic value, say $w_{\infty,i} \in [-1, 1]$, and considering the equation

$$\frac{d\xi_i}{dt} = (w_{\infty,i} - \alpha)(1 - |\xi_i|),$$

whose solution reads

$$\xi_i(t) = \begin{cases} -1 + (1 + \xi_{0,i})e^{(w_{\infty,i} - \alpha)t} & \text{if } \xi_i(t) \leq 0 \\ 1 - (1 - \xi_{0,i})e^{-(w_{\infty,i} - \alpha)t} & \text{if } \xi_i(t) \geq 0. \end{cases}$$

From here we easily deduce that:

- if $\xi_{0,i} < 0$ and $w_{\infty,i} < \alpha$ then $\xi_i \to -1$ for $t \to +\infty$;

- if $\xi_{0,i} > 0$ and $w_{\infty,i} > \alpha$ then $\xi_i \to 1$ for $t \to +\infty$.

In both cases, the final preference confirms and consolidates the initial one, because $w_{\infty,i} - \alpha$ has the same sign as $\xi_{0,i}$. Conversely,

- if $\xi_{0,i} < 0$ but $w_{\infty,i} > \alpha$ then $\xi_i \to 1$ for $t \to +\infty$;

- if $\xi_{0,i} > 0$ but $w_{\infty,i} < \alpha$ then $\xi_i \to -1$ for $t \to +\infty$.

In these cases, the final preference reverses the initial one, because $w_{\infty,i} - \alpha$ has opposite sign with respect to $\xi_{0,i}$.

Another possible choice of the function Φ is:

$$\Phi(\xi) = |\xi| \left(1 - \xi^2\right), \tag{5}$$

which vanishes also at $\xi = 0$. Thus we are led to consider the equation

$$\frac{d\xi_i}{dt} = (w_{\infty,i} - \alpha) |\xi_i| \left(1 - \xi_i^2\right),$$

whose solution reads

$$\xi_i(t) = \frac{\xi_{0,i}}{\sqrt{\xi_{0,i}^2 + \left(1 - \xi_{0,i}^2\right) e^{-2 \operatorname{sgn} \xi_{0,i}(w_{\infty,i} - \alpha)t}}}.$$

Now the asymptotic trend of the preference can be summarised as follows.

(a) $\Delta = 1$, $\alpha = -0.3$　　　　　　　　(b) $\Delta = 1$, $\alpha = 0.3$

Figure 2: The curves $t \mapsto \xi_i(t)$ generated by the coupled system (2)-(3) with $N = 50$ agents and P, Φ like in (4), (5) with $\Delta = 1$ and $\alpha = \pm 0.3$. The initial values $w_{0,i}$, $\xi_{0,i}$ have been sampled uniformly in the interval $[-1,\ 1]$. The ODE system has been integrated numerically via a standard fourth order Runge-Kutta method.

- For $w_{\infty,i} < \alpha$:
  - if $\xi_{0,i} < 0$ then $\xi_i \to -1$ for $t \to +\infty$;
  - if $\xi_{0,i} > 0$ then $\xi_i \to 0$ for $t \to +\infty$.

- For $w_{\infty,i} > \alpha$:
  - if $\xi_{0,i} < 0$ then $\xi_i \to 0$ for $t \to +\infty$;
  - if $\xi_{0,i} > 0$ then $\xi_i \to 1$ for $t \to +\infty$.

We observe that the large time behaviour of the preference is again a polarisation in poles coinciding with the zeroes of the function Φ. Unlike the previous case, however, the presence of an intermediate pole at $\xi = 0$ prevents a complete reversal of the initial preference when the latter has opposite sign with respect to $w_{\infty,i} - \alpha$. In such a situation, the agents simply become indecisive, their preference tending indeed to zero.

In order to illustrate the actual coupled dynamics of (2)-(3), we solve numerically the coupled system of equations with $N = 50$ agents and with the functions P, Φ given in (4), (5), respectively.

In Figure 2 we present the curves $t \mapsto \xi_i(t)$ in the case $\Delta = 1$ and for $\alpha = \pm 0.3$. Since the agents reach a global consensus around the centrist opinion $w = 0$, cf. Figure 1(a), with a leftward-biased perceived social opinion $\alpha < 0$ we observe the polarisation of the preferences either towards the indecisiveness $\xi = 0$, if the initial preference was in turn leftward-biased, i.e. $\xi_{0,i} < 0$, or in $\xi = 1$, if the initial preference was rightward-biased, i.e. $\xi_{0,i} > 0$, cf. Figure 2(a). Conversely, with a rightward-biased perceived social opinion $\alpha > 0$ we observe indecisiveness if $\xi_{0,i} > 0$ and consolidation in $\xi = -1$ if $\xi_{0,i} < 0$, cf. Figure 2(b).

Such rather simple dynamics may become more complex under the formation of multiple opinion clusters. To exemplify this case, we consider now $\Delta = 0.4$, like in Figure 1(b), and again the two cases $\alpha = \pm 0.3$, cf. Figure 3(a, b) along with also $\alpha = \pm 0.6$, cf. Figure 3(c, d). In this case, simultaneous polarisations in $\xi = \pm 1$ can also be observed, depending on the distribution of the pairs $(\xi_{0,i},\ w_{0,i})$ at the initial time.

6

(a) $\Delta = 0.4$, $\alpha = -0.3$ (b) $\Delta = 0.4$, $\alpha = 0.3$

(c) $\Delta = 0.4$, $\alpha = -0.6$ (d) $\Delta = 0.4$, $\alpha = 0.6$

Figure 3: The same as in Figure 2 but with $\Delta = 0.4$, cf. Figure 1(b), and $\alpha = \pm 0.3$ (top row), $\alpha = \pm 0.6$ (bottom row).

3 Aggregate analysis of opinion dynamics

The discussion set forth in the previous section shows that it is in general quite hard to analyse exactly the interplay between opinions and preferences from a strictly microscopic point of view. Due to the severe dependence of the microscopic system on the particular initial state and trajectory of each agent, the main difficulty is, as usual, to grasp the essential facts able to explain the big picture, namely to depict the collective behaviour. For this reason, from this section we move to a more aggregate analysis, which, starting from a description of opinion dynamics by methods of statistical physics and kinetic theory, will finally lead us to macroscopic equations for the preference dynamics written in terms of hydrodynamic parameters such as the density of the agents and their mean opinion.

3.1 Microscopic binary interactions

In order to approach the opinion dynamics (3) from the point of view of kinetic theory, we need to set up a consistent scheme of binary, i.e. pairwise, interactions among the agents. To this purpose, inspired by [13], we consider (3) for just two agents, say i, j, and we discretise the differential equation with the forward Euler formula during a small time step $0 < \gamma < 1$. Setting

$$w := w_i(t), \quad w_* := w_j(t), \quad w' := w_i(t + \gamma), \quad w'_* := w_j(t + \gamma)$$

we obtain the binary rules

$$
\begin{aligned}
w' &= w + \gamma P(w, w_*)(w_* - w) + D(w)\eta, \\
w'_* &= w_* + \gamma P(w_*, w)(w - w_*) + D(w_*)\eta,
\end{aligned}
\tag{6}
$$

where we have also added a random contribution, given by a centred random variable η, modelling stochastic fluctuations induced by the self-thinking of the agents. Here, $D(\cdot) \geq 0$ is an opinion-dependent diffusion coefficient modulating the amplitude of the stochastic fluctuations, that is the variance of η.

In general, the binary interactions (6) are such that

$$
\langle w' + w'_* \rangle = w + w_* + \gamma \left(P(w, w_*) - P(w_*, w) \right) (w_* - w),
\tag{7}
$$

where $\langle \cdot \rangle$ denotes the expectation with respect to the distribution of η. Hence the mean opinion is in general not conserved on average in a single binary interaction unless P is symmetric, i.e. $P(w, w_*) = P(w_*, w)$ for all $w, w_* \in [-1, 1]$. Furthermore, at leading order for γ small enough we have

$$
\begin{aligned}
\langle (w')^2 + (w'_*)^2 \rangle = {}& w^2 + w_*^2 + 2\gamma \left(wP(w, w_*) - w_*P(w_*, w) \right) (w_* - w) \\
& + \left(D^2(w) + D^2(w_*) \right) \sigma^2 + o(\gamma),
\end{aligned}
\tag{8}
$$

where $\sigma^2 > 0$ denotes the variance of η. Therefore, in general, also the energy is not conserved on average in a single binary interaction, not even for a symmetric function P.

Equations (7), (8) show that a particularly interesting case is when P is constant, for then from (7) we deduce that the mean opinion is conserved in each binary interaction, while from (8) we see that, at least in the absence of stochastic fluctuations (i.e. formally for $\sigma^2 = 0$), the average energy is dissipated:

$$
\langle (w')^2 + (w'_*)^2 \rangle = w^2 + w_*^2 - 2\gamma(w_* - w)^2 + o(\gamma) \leq w^2 + w_*^2 + o(\gamma).
$$

In order to be physically admissible, the interaction rules (6) have to be such that $|w'|, |w'_*| \leq 1$ for $|w|, |w_*| \leq 1$. Observing that

$$
\begin{aligned}
|w'| &= |(1 - \gamma P(w, w_*))w + \gamma P(w, w_*)w_* + D(w)\eta| \\
&\leq (1 - \gamma P(w, w_*))|w| + \gamma P(w, w_*) + D(w)|\eta|,
\end{aligned}
$$

where we have used the fact that $|w_*| \leq 1$, we see that a sufficient condition for $|w'| \leq 1$ is

$$
D(w)|\eta| \leq (1 - \gamma P(w, w_*))(1 - |w|),
$$

which is satisfied if there exists a constant $c > 0$ such that

$$
\begin{cases}
|\eta| \leq c(1 - \gamma P(w, w_*)) \\
cD(w) \leq 1 - |w|,
\end{cases}
\qquad \forall\, w, w_* \in [-1, 1].
\tag{9}
$$

Considering that $P(w, w_*) \leq 1$ by assumption, the first condition can be further enforced by requiring $|\eta| \leq c(1 - \gamma)$, which implies that η has to be chosen as a compactly supported random variable. The second condition forces instead $D(\pm 1) = 0$. Taking inspiration from [47], possible choices are: $D(w) = 1 - |w|$ and $c = 1$, which produces $|\eta| \leq 1 - \gamma$; or $D(w) = 1 - w^2$ and $c = \frac{1}{2}$, which yields $|\eta| \leq \frac{1}{2}(1 - \gamma)$. Another less obvious option is

$$
D(w) = \sqrt{\left(1 - (1 + \gamma^s)w^2\right)_+} \quad \text{and} \quad c = \frac{\gamma^{s/2}}{\sqrt{1 + \gamma^s}}, \quad s > 0,
\tag{10}
$$

where $(\cdot)_+ := \max\{0, \cdot\}$ denotes the positive part, which produces $|\eta| \leq \frac{\gamma^{s/2}(1 - \gamma)}{\sqrt{1 + \gamma^s}}$. This function D converges uniformly to $\sqrt{1 - w^2}$ in $[-1, 1]$ when $\gamma \to 0^+$. Notice, however, that such a uniform limit does not comply with (10) regardless of choice of $c > 0$, because of the infinite derivative at $w = \pm 1$.

Exactly the same considerations hold true for the second interaction rule in (6).

3.2 Kinetic description and steady states

Introducing the distribution function $f = f(t, w) : \mathbb{R}_+ \times [-1, 1] \to \mathbb{R}_+$, such that $f(t, w)dw$ is the fraction of agents with opinion in $[w, w + dw]$ at time t, the binary rules (6) can be encoded in a Boltzmann-type kinetic equation, which, in weak form, writes:

$$\frac{d}{dt} \int_{-1}^{1} \varphi(w) f(t, w) \, dw$$
$$= \frac{1}{2} \int_{-1}^{1} \int_{-1}^{1} \langle \varphi(w') + \varphi(w'_*) - \varphi(w) - \varphi(w_*) \rangle f(t, w) f(t, w_*) \, dw \, dw_*, \quad (11)$$

where $\varphi : [-1, 1] \to \mathbb{R}$ is an arbitrary test function, i.e. any observable quantity depending on the microscopic state of the agents. Choosing $\varphi(w) = 1$, we obtain that the integral of f with respect to w is constant in time, i.e. that the total number of agents is conserved. This also implies that, up to normalisation at the initial time, f can be thought of as a probability density for every $t > 0$. Choosing instead $\varphi(w) = w$ we discover

$$\frac{d}{dt} \int_{-1}^{1} w f(t, w) \, dw = \frac{\gamma}{2} \int_{-1}^{1} \int_{-1}^{1} (P(w, w_*) - P(w_*, w))(w_* - w) f(t, w) f(t, w_*) \, dw \, dw_*, \quad (12)$$

therefore the mean opinion $M_1 := \int_{-1}^{1} w f(t, w) \, dw$ is either conserved in time, if P is symmetric so that the right-hand side of the previous equation vanishes, or not conserved, if P is non-symmetric. This difference has important consequences on the steady distributions of (11), which in turn will impact considerably on the equations describing the formation of the preferences. Therefore, in what follows we investigate it in some detail.

3.2.1 Symmetric P

The prototype of a symmetric P is the constant function $P \equiv 1$. In this case, from (11) we can recover an explicit expression of the asymptotic distribution function at least in the so-called *quasi-invariant regime*, i.e. the one in which the variation of the opinion in each binary interaction is small. To describe such a regime, we scale the parameters γ, σ^2 in (6) as

$$\gamma \to \epsilon \gamma, \qquad \sigma^2 \to \epsilon \sigma^2, \quad (13)$$

where $\epsilon > 0$ is an arbitrarily small scaling coefficient. Parallelly, in order to study the large time behaviour of the system, we introduce the new time scale $\tau := \epsilon t$ and we scale the distribution function as $g(\tau, w) := f(\frac{\tau}{\epsilon}, w)$. In this way, it is clear that, at every fixed $\tau > 0$ and in the limit $\epsilon \to 0^+$, g describes the large time trend of f. Since $\partial_\tau g = \frac{1}{\epsilon} \partial_t f$, substituting in (11) and using the symmetry of the interactions (6) with $P \equiv 1$ we see that the equation satisfied by g is

$$\frac{d}{d\tau} \int_{-1}^{1} \varphi(w) g(\tau, w) \, dw = \frac{1}{\epsilon} \int_{-1}^{1} \int_{-1}^{1} \langle \varphi(w') - \varphi(w) \rangle g(\tau, w) g(\tau, w_*) \, dw \, dw_*. \quad (14)$$

Now, because of the scaling (13), if φ is sufficiently smooth then the difference $\langle \varphi(w') - \varphi(w) \rangle$ is small and can be expanded about w to give:

$$\langle \varphi(w) - \varphi(w') \rangle = \varphi'(w) \langle w' - w \rangle + \frac{1}{2} \varphi''(w) \langle (w' - w)^2 \rangle + \frac{1}{6} \varphi'''(\bar{w}) \langle (w' - w)^3 \rangle,$$

where $\min\{w, w'\} < \bar{w} < \max\{w, w'\}$. Plugging into (14) this produces

$$\frac{d}{d\tau} \int_{-1}^{1} \varphi(w) g(\tau, w) \, dw = \gamma \int_{-1}^{1} \varphi'(w)(m - w) g(\tau, w) \, dw$$
$$+ \frac{\sigma^2}{2} \int_{-1}^{1} \varphi''(w) D^2(w) g(\tau, w) \, dw + R_\varphi(g, g),$$

9

Figure 4: Asymptotic opinion distribution (16) with mean $m = 0.25$ and four different values of the parameter λ.

where we have denoted by $m \in [-1,\, 1]$ the constant mean opinion and where $R_\varphi(g,\, g)$ is a reminder such that $|R_\varphi(g,\, g)| = O(\sqrt{\epsilon})$ under the assumption that η has finite third order moment, i.e. $\langle |\eta|^3 \rangle < +\infty$, cf. [47] for details. Hence for $\epsilon \to 0^+$ it results $R_\varphi(g,\, g) \to 0$ and we get

$$\frac{d}{d\tau} \int_{-1}^{1} \varphi(w) g(\tau,\, w)\, dw = \gamma \int_{-1}^{1} \varphi'(w)(m-w)g(\tau,\, w)\, dw + \frac{\sigma^2}{2} \int_{-1}^{1} \varphi''(w) D^2(w) g(\tau,\, w)\, dw.$$

Integrating back by parts the terms on the right-hand side and assuming $\varphi(\pm 1) = \varphi'(\pm 1) = 0$, due to the arbitrariness of φ this can be recognised as a weak form of the *Fokker-Planck equation*

$$\partial_\tau g = \frac{\sigma^2}{2} \partial_w^2 \left(D^2(w) g \right) + \gamma \partial_w((w-m)g). \tag{15}$$

Fixing[1] $D(w) = \sqrt{1-w^2}$, the unique asymptotic ($\tau \to +\infty$) solution with unitary mass, say $g_\infty(w)$, to (15) reads

$$g_\infty(w) = \frac{(1+w)^{\frac{1+m}{\lambda}-1}(1-w)^{\frac{1-m}{\lambda}-1}}{2^{\frac{2}{\lambda}-1} \mathrm{B}\left(\frac{1+m}{\lambda},\, \frac{1-m}{\lambda}\right)}, \qquad \lambda := \frac{\sigma^2}{\gamma}, \tag{16}$$

where $\mathrm{B}(\cdot,\, \cdot)$ denotes the Beta function. Notice that such a g_∞ is a Beta probability density function on the interval $[-1,\, 1]$. Using the known formulas for the moments of Beta random variables, we easily check that its mean is indeed m and we compute its energy as

$$M_{2,\infty} := \int_{-1}^{1} w^2 g_\infty(w)\, dw = \frac{2m^2 + \lambda}{2 + \lambda}. \tag{17}$$

In Figure 4 we illustrate some typical trends of the distribution function (16) with positive mean, $m = 0.25$ in this example. We observe that, depending on the value of λ, such a distribution may depict a transition from a strong consensus around the mean ($\lambda = 0.1$) to a milder consensus ($\lambda = 0.4$) and further to a radicalisation in the extreme opinion $w = 1$ ($\lambda = 1$) up to the appearance of a double radicalisation in the two opposite extreme opinions $w = \pm 1$ ($\lambda = 2$).

3.2.2   Non-symmetric P

A natural prototype of a non-symmetric function P is a linear perturbation of a constant P depending on only one of the two variables w, w_*. More specifically, we consider

$$P(w,\, w_*) = P(w_*) = pw_* + q, \tag{18}$$

---

[1]In view of the scaling (13), as $\epsilon \to 0^+$ the function (10) converges uniformly to $\sqrt{1-w^2}$, which can therefore be chosen as diffusion coefficient in the Fokker-Planck equation (15) *after* performing the quasi-invariant limit.

where p, $q \in \mathbb{R}$ have to be chosen in such a way that $pw_* + q \in [0, 1]$ for all $w_* \in [-1, 1]$. This is obtained if

$$0 \leq q \leq 1, \qquad |p| \leq \min\{q, \, 1 - q\}.$$

With respect to model (6), such a function P describes a situation in which agents with opinion $w_* > 0$ are more persuasive than agents with opinion $w_* < 0$ if $p > 0$ and vice versa if $p < 0$.

Using (18) in (12) we obtain that the evolution of the mean opinion $M_1 = M_1(t)$ is ruled by

$$\frac{dM_1}{dt} = \frac{p\gamma}{2} \int_{-1}^{1} \int_{-1}^{1} (w_* - w)^2 f(t, \, w) f(t, \, w_*) \, dw \, dw_*,$$

whence we see that the sign of the time derivative $\frac{dM_1}{dt}$ coincides with that of p. Thus, if $p > 0$ the mean opinion is non-decreasing, while if $p < 0$ the mean opinion is non-increasing. Continuing the previous calculation, we further find:

$$\frac{dM_1}{dt} = \frac{p\gamma}{2} \left(M_2 - M_1^2 \right),$$

which indicates that at the steady state ($t \to +\infty$) it results invariably $M_{2,\infty} = M_{1,\infty}^2$. This implies that the asymptotic distribution has zero variance, thus that it is necessarily a Dirac delta centred in the asymptotic mean opinion, i.e. $f_\infty(w) = \delta(w - M_{1,\infty})$. Plugging this into (11) we discover

$$\langle \varphi(M_{1,\infty} + D(M_{1,\infty})\eta) \rangle - \varphi(M_{1,\infty}) = 0,$$

which has to hold for every test function φ. As a consequence, we deduce $D(M_{1,\infty}) = 0$, whence $M_{1,\infty} = \pm 1$ if the only zeroes of the diffusion coefficient are $w = \pm 1$ like in the examples considered in Section 3.1.

In conclusion, with the non-symmetric function P given by (18) we fully characterise the asymptotic distribution function as:

- $f_\infty(w) = \delta(w + 1)$ if $p < 0$, the mean opinion decreasing from its initial value to $M_{1,\infty} = -1$;

- $f_\infty(w) = \delta(w - 1)$ if $p > 0$, the mean opinion increasing from its initial value to $M_{1,\infty} = 1$.

The considerations above can be generalised to the following function P:

$$P(w, \, w_*) = rw + pw_* + q, \tag{19}$$

where $p \neq r$, so that P is non-symmetric, and where the coefficients p, q, $r \in \mathbb{R}$ have to be chosen in such a way that $rw + pw_* + q \in [0, 1]$ for all $(w, \, w_*) \in [-1, 1]^2$. Repeating the previous calculations, we conclude that:

- $f_\infty(w) = \delta(w + 1)$ if $p - r < 0$; in this case, the mean opinion decreases from its initial value to $M_{1,\infty} = -1$;

- $f_\infty(w) = \delta(w - 1)$ if $p - r > 0$; in this case, the mean opinion increases from its initial value to $M_{1,\infty} = 1$.

From the modelling point of view, we may interpret the difference $p - r$ as a balance between the persuasion ability of the agents, expressed by p, and their tendency to be persuaded, expressed by r. Notice indeed that for $p = 0$ and $r \neq 0$ we obtain the mirror case of (18), in which agents with opinion $w > 0$ are more inclined to change their opinion than agents with opinion $w < 0$ if $r > 0$ and vice versa if $r < 0$.

The discussion above clearly shows that an arbitrarily small perturbation of a constant P, by destroying the conservation of the mean opinion, may drag the system towards asymptotic configurations much less variegated than (16) independently of the parameters γ, σ^2 of the interactions.

4 Macroscopic description of preference formation

According to model (2)-(3), the opinions of the agents evolve through mutual interactions independent of the preferences; on the other hand, the preference of each agent is transported in time by his/her opinion. This suggests that a proper way to account for the interplay between opinion and preference in an aggregate manner is by means of an inhomogeneous Boltzmann-type kinetic equation, whose transport term describes the evolution of the preference and whose "collisional" term accounts simultaneously for the changes in the opinions.

4.1 Inhomogeneous Boltzmann-type description and hydrodynamics

A Boltzmann-type description of the opinion dynamics in the form of binary interactions (6) coupled to the transport of the preference (2) is obtained by introducing the kinetic distribution function

$$f = f(t, \xi, w) : \mathbb{R}_+ \times [-1, 1] \times [-1, 1] \to \mathbb{R}_+,$$

such that $f(t, \xi, w)d\xi\, dw$ is the proportion of agents that at time t have a preference in $[\xi, \xi + d\xi]$ and an opinion in $[w, w+dw]$. The distribution function f satisfies the following weak Boltzmann-type equation:

$$\partial_t \int_{-1}^1 \varphi(w) f(t, \xi, w)\, dw + \partial_\xi \left(\Phi(\xi) \int_{-1}^1 (w - \alpha)\varphi(w) f(t, \xi, w)\, dw \right)$$
$$= \frac{1}{2} \int_{-1}^1 \int_{-1}^1 \langle \varphi(w') + \varphi(w'_*) - \varphi(w) - \varphi(w_*) \rangle f(t, \xi, w) f(t, \xi, w_*)\, dw\, dw_*, \quad (20)$$

where the transport term (second term on the left-hand side) has been written taking into account that, according to (2), the transport velocity of the preference ξ is $(w - \alpha)\Phi(\xi)$ and where w', w'_* on the right-hand side are given by (6).

From the distribution function f, by integration with respect to the opinion w, we can compute macroscopic quantities in the space of the preferences, such as the *density* of the agents with preference ξ at time t:

$$\rho(t, \xi) := \int_{-1}^1 f(t, \xi, w)\, dw$$

and the *mean opinion* of the agents with preference ξ at time t:

$$m(t, \xi) := \frac{1}{\rho(t, \xi)} \int_{-1}^1 w f(t, \xi, w)\, dw.$$

The interest in (20) is that it allows one to obtain evolution equations directly for the quantities ρ, m, provided one is able to characterise the large time statistical trends of the opinions, like in Section 3. The underlying key idea is to consider a so-called *hydrodynamic regime*, in which the opinions reach a local equilibrium much more quickly than the preferences, pretty much in the spirit of the microscopic investigations performed in Section 2.

Let $0 < \delta \ll 1$ be a small parameter, which we use to define a macroscopic time scale $\tau := \delta t$, i.e. the time scale of the evolution of the preferences, which then turns out to be much larger, viz. slower, than the characteristic one of the binary interactions among the agents. If we want that, on this new scale, the preference dynamics remain the same, from (2) we see that we need to scale simultaneously the transport speed of the preference by letting $\Phi(\xi) \to \delta\Phi(\xi)$.

Let $g(\tau, \xi, w) := f(\frac{\tau}{\delta}, \xi, w)$, whence $\partial_\tau g = \frac{1}{\delta}\partial_t f$. Plugging into (20) we find that g satisfies the equation

$$\partial_\tau \int_{-1}^1 \varphi(w) g(\tau, \xi, w)\, dw + \partial_\xi \left(\Phi(\xi) \int_{-1}^1 (w - \alpha)\varphi(w) g(\tau, \xi, w)\, dw \right)$$

$$= \frac{1}{2\delta} \int_{-1}^{1} \int_{-1}^{1} \langle \varphi(w') + \varphi(w'_*) - \varphi(w) - \varphi(w_*) \rangle g(\tau, \xi, w) g(\tau, \xi, w_*) \, dw \, dw_*. \quad (21)$$

Basically, the aforesaid scaling produces the coefficient $1/\delta$ in front of the interaction term, hence δ is analogous to the Knudsen number in classical fluid dynamics. Since we are assuming that δ is small, a hydrodynamic regime is justified and, in particular, it can be described by a splitting of (21), cf. [26], totally analogous to the one often adopted in the numerical solution of the inhomogeneous Boltzmann equation, see e.g. [22, 23, 39]. One first solves the fast interactions:

$$\partial_\tau \int_{-1}^{1} \varphi(w) g(\tau, \xi, w) \, dw$$
$$= \frac{1}{2\delta} \int_{-1}^{1} \int_{-1}^{1} \langle \varphi(w') + \varphi(w'_*) - \varphi(w) - \varphi(w_*) \rangle g(\tau, \xi, w) g(\tau, \xi, w_*) \, dw \, dw_*, \quad (22)$$

which, owing to the high frequency $1/\delta$, reach quickly an equilibrium described by a local (in ξ and τ) asymptotic distribution function playing morally the role of a local Maxwellian. Notice indeed that (22) is actually an equation on the time scale of the microscopic interactions, because τ can be scaled back to t using the factor $1/\delta$. Next, one transports such a local equilibrium distribution according to the remaining terms of (21) on the slower hydrodynamic scale:

$$\partial_\tau \int_{-1}^{1} \varphi(w) g(\tau, \xi, w) \, dw + \partial_\xi \left(\Phi(\xi) \int_{-1}^{1} (w - \alpha) \varphi(w) g(\tau, \xi, w) \, dw \right) = 0. \quad (23)$$

Due to (22), and taking the definition of ρ into account, the local "Maxwellian" can be given the form $g(\tau, \xi, w) = \rho(\tau, \xi) g_\infty(w)$, where g_∞ is one of the asymptotic opinion distribution functions found in Section 3. This is the distribution transported by (23), hence we finally obtain

$$\partial_\tau \left(\rho \int_{-1}^{1} \varphi(w) g_\infty(w) \, dw \right) + \partial_\xi \left(\Phi(\xi) \rho \int_{-1}^{1} (w - \alpha) \varphi(w) g_\infty(w) \, dw \right) = 0 \quad (24)$$

and we can use the knowledge of g_∞ to compute explicitly the remaining integral terms.

4.2 First order hydrodynamic models

Let us consider at first the case of the non-symmetric functions P (18), (19) discussed in Section 3.2.2. The asymptotic opinion distribution is either $g_\infty(w) = \delta(w + 1)$ or $g_\infty(w) = \delta(w - 1)$, depending on the asymmetry of P. Plugging into (24) along with the choice $\varphi(w) = 1$ we find therefore either

$$\partial_\tau \rho - (1 + \alpha) \partial_\xi (\Phi(\xi) \rho) = 0 \quad (25)$$

or

$$\partial_\tau \rho + (1 - \alpha) \partial_\xi (\Phi(\xi) \rho) = 0. \quad (26)$$

In both cases, we get a self-consistent equation for the sole density ρ and we speak thus of *first order* hydrodynamic model.

Unlike typical conservation laws, in (25) and (26) the flux does not only depend on the variable ξ through the conserved quantity ρ but also explicitly through the function Φ. An analogous characteristic is found, for instance, in conservation-law-based macroscopic models of vehicular traffic featuring different flux functions in different roads, see [31].

We observe that both (25) and (26) admit the family of stationary distributional solutions

$$\rho_\infty(\xi) = \sum_{k=1}^{M} \rho_k \delta(\xi - \xi_k), \qquad \rho_k \geq 0, \quad (27)$$

where the ξ_k's are the zeroes of the function Φ. This indicates that models (25) and (26) reproduce the asymptotic polarisation of the agents in the preference poles individuated by the points where

Φ vanishes. The coefficients ρ_k represent the masses concentrating in each pole. Furthermore, (25) describes invariably a leftward transport of ρ in the space of the preferences, because $-(1+\alpha) < 0$ for all $\alpha \in (-1, 1]$ (if $\alpha = -1$ the density is simply not transported). Conversely, (26) describes invariably a rightward transport of ρ, since $1 - \alpha > 0$ for all $\alpha \in [-1, 1)$ (now the density is not transported if $\alpha = 1$).

4.3   Second order hydrodynamic model

We now consider the symmetric case $P \equiv 1$ discussed in Section 3.2.1, which produces the asymptotic opinion distribution g_∞ given by (16). Notice that this distribution is parametrised by the (local) mean opinion $m = m(\tau, \xi)$, because the latter is conserved by the opinion dynamics. This implies that, if we plug such a g_∞ into (24) together with the choice $\varphi(w) = 1$, we do not get a self-consistent equation for the density ρ. In fact, we find:

$$\partial_\tau \rho + \partial_\xi \left(\Phi(\xi)\rho(m - \alpha)\right) = 0,$$

with both hydrodynamic parameters ρ, m unknown. In order to close the macroscopic equations, we need a further equation relating ρ and m, which we can obtain from (24) with $\varphi(w) = w$ and recalling also (17):

$$\partial_\tau (\rho m) + \partial_\xi \left(\Phi(\xi)\rho \left(\frac{2m^2 + \lambda}{2 + \lambda} - \alpha m \right) \right) = 0.$$

On the whole, we get the *second order* (i.e., composed of a self-consistent pair of equations) hydrodynamic model

$$\begin{cases} \partial_\tau \rho + \partial_\xi \left(\Phi(\xi)\rho(m - \alpha)\right) = 0 \\ \partial_\tau (\rho m) + \partial_\xi \left(\Phi(\xi)\rho \left(\dfrac{2m^2 + \lambda}{2 + \lambda} - \alpha m \right) \right) = 0, \end{cases} \tag{28}$$

where the parameter $\lambda = \sigma^2/\gamma$, which here enters the game through the energy of the stationary opinion distribution (16), is reminiscent of the self-thinking (diffusion) of the agents.

Also in this case, (27) is a family of admissible stationary distributional solutions. Hence model (28) can in turn reproduce the asymptotic polarisation of the preferences already observed in the microscopic model.

It is useful to ascertain under which conditions system (28) is hyperbolic in the natural state space $\{(\rho, m) \in \mathbb{R}_+ \times [-1, 1]\}$. To this purpose, we rewrite it in the quasilinear matrix form

$$\partial_\tau U + \Phi(\xi)A(U)\partial_\xi U + \Phi'(\xi)F(U) = 0,$$

where $U := (\rho, m)^T$, $A(U)$ is the matrix

$$A(U) := \begin{pmatrix} m - \alpha & \rho \\ \frac{\lambda(1-m^2)}{\rho(2+\lambda)} & \frac{(2-\lambda)m}{2+\lambda} - \alpha \end{pmatrix} \tag{29}$$

and $F(U)$ denotes lower order terms, which is not important to write explicitly. Since Φ is real-valued, system (28) is hyperbolic if both eigenvalues of $A(U)$ are real. To check this, we compute the discriminant $\Delta(U)$ of the characteristic polynomial of $A(U)$:

$$\Delta(U) := \operatorname{tr}^2 A(U) - 4 \det A(U) = \frac{4\lambda}{\lambda + 2} \left(1 - \frac{2m^2}{\lambda + 2} \right).$$

Since $m \in [-1, 1]$, thus $m^2 \in [0, 1]$, and $\lambda \geq 0$, we easily see that $\Delta(U)$ is always non-negative. Therefore, we conclude:

Proposition 4.1. *System* (28) *is hyperbolic in the whole state space* $\{(\rho, m) \in \mathbb{R}_+ \times [-1, 1]\}$ *for every choice of the parameters* $\alpha \in [-1, 1]$, $\lambda \geq 0$ *and for every function* $\Phi : [-1, 1] \to [0, 1]$.

4.4 General first and second order hydrodynamic models

If the asymptotic opinion distribution g_∞ is not known analytically, like e.g. in the significant case (4), the hydrodynamic models can still be written from (24), although only in a semi-analytical form.

Assume that the microscopic dynamics (6) do not conserve the mean opinion. Then the sole conserved quantity is the mass of the agents and from (24) with $\varphi(w) = 1$ we obtain the first order model

$$\partial_\tau \rho + (M_{1,\infty} - \alpha)\partial_\xi(\Phi(\xi)\rho) = 0$$

in the unknown $\rho = \rho(\tau, \xi)$, where $M_{1,\infty} := \int_{-1}^{1} w g_\infty(w)\,dw$ is the asymptotic mean opinion. The latter may be computed e.g. from (12), by means of an appropriate numerical approach.

Conversely, if the microscopic dynamics (6) conserve the mean opinion then g_∞ is parametrised by m and from (24) with $\varphi(w) = 1$, w we obtain the second order model

$$\begin{cases} \partial_\tau \rho + \partial_\xi\left(\Phi(\xi)\rho(m - \alpha)\right) = 0 \\ \partial_\tau(\rho m) + \partial_\xi\left(\Phi(\xi)\rho\left(M_{2,\infty}(m) - \alpha m\right)\right) = 0 \end{cases} \tag{30}$$

in the unknowns $\rho = \rho(\tau, \xi)$, $m = m(\tau, \xi)$. Here, $M_{2,\infty}(m) := \int_{-1}^{1} w^2 g_\infty(w)\,dw$ is the energy of the asymptotic opinion distribution, expressed as a function of the conserved quantity m.

The precise calculation of $M_{2,\infty}$ requires, in general, an accurate numerical reconstruction of g_∞. The latter is a stationary solution to the Fokker-Planck equation

$$\partial_\tau g = \frac{\sigma^2}{2}\partial_w^2\left(D^2(w)g\right) + \gamma\partial_w(\mathcal{B}[g]g) \tag{31}$$

with

$$\mathcal{B}[g](\tau, w) := \int_{-1}^{1} P(w, w_*)(w - w_*)g(\tau, w_*)\,dw_*, \tag{32}$$

which is obtained in the quasi-invariant regime starting from the binary interactions (6) with a symmetric but not necessarily constant compromise function P. In particular, the following implicit representation of g_∞ can be given:

$$g_\infty(w) = \frac{C}{D^2(w)}\exp\left(-\frac{2}{\lambda}\int\frac{\mathcal{B}[g_\infty](w)}{D^2(w)}\,dw\right), \tag{33}$$

where $C > 0$ is a normalisation constant and the integral on the right-hand side denotes any antiderivative of the function $w \mapsto \mathcal{B}[g_\infty](w)/D^2(w)$. For instance, if P is the function (19) with

$$r = p, \qquad 0 \le q \le 1, \qquad |p| \le \frac{1}{2}\min\{q,\,1 - q\},$$

so that P is symmetric and $P(w, w_*) \in [0,\,1]$ for all $(w, w_*) \in [-1,\,1]^2$, then from (33) we find the semi-explicit expression

$$g_\infty(w) = Ce^{\frac{2p}{\lambda}w}(1 + w)^{\frac{q(1+m) - p(1 - M_{2,\infty})}{\lambda} - 1}(1 - w)^{\frac{q(1-m) + p(1 - M_{2,\infty})}{\lambda} - 1},$$

which brings the calculation of $M_{2,\infty}$ back to the numerical solution of the non-linear system of equations

$$\begin{cases} \int_{-1}^{1} g_\infty(w)\,dw = 1 \\ \int_{-1}^{1} w^2 g_\infty(w)\,dw = M_{2,\infty} \end{cases}$$

parametrised by m.

In general, however, we observe that the type of dependence of $M_{2,\infty}$ on m valid for $P \equiv 1$, cf. (17), is somewhat paradigmatic. In fact, let us consider (14) in the quasi-invariant limit $\epsilon \to 0^+$

for binary interactions (6) with a symmetric P. Fixing $D(w) = \sqrt{1 - w^2}$, we obtain the following equation for M_2:

$$\frac{dM_2}{d\tau} = 2\gamma \int_{-1}^{1} \int_{-1}^{1} w(w_* - w) P(w, w_*) g(\tau, w) g(\tau, w_*) \, dw \, dw_* + \sigma^2 (1 - M_2).$$

Set $a := \inf_{w, w_* \in [-1, 1]} P(w, w_*)$, $0 \leq a \leq 1$. Then

$$-(2 + \lambda) M_2 + 2am^2 + \lambda \leq \frac{1}{\gamma} \cdot \frac{dM_2}{d\tau} \leq -(2a + \lambda) M_2 + 2m^2 + \lambda,$$

which produces asymptotically

$$\frac{2am^2 + \lambda}{2 + \lambda} \leq M_{2,\infty} \leq \frac{2m^2 + \lambda}{2a + \lambda}.$$

This suggests that a perhaps rough but possibly useful approximation of $M_{2,\infty}$, to be used in (30), is the average of these lower and upper bounds, i.e.:

$$M_{2,\infty} \approx \left(\frac{a}{\lambda + 2} + \frac{1}{2a + \lambda} \right) m^2 + \frac{\lambda(1 + a + \lambda)}{(\lambda + 2)(2a + \lambda)},$$

which for $a = 0$, like e.g. in case (4), yields $M_{2,\infty} \approx \frac{m^2}{\lambda} + \frac{\lambda + 1}{\lambda + 2}$.

5 Numerical tests

In this section we exemplify, by means of several numerical tests, the main features of the formation of preferences at the kinetic and hydrodynamic scales as described by the models presented in the previous sections.

The numerical approach is essential, in particular, to investigate the cases in which the compromise function P does not allow for an explicit computation of the asymptotic opinion distribution g_∞. Therefore, first we will briefly review *Structure Preserving* (SP) numerical methods, which are able to capture the large time solution to possibly non-local Fokker-Planck equations with non-constant diffusion, such as those introduced in Sections 3.2.1 and 4.4, see [42, 43]. Next, we will compare the large time distributions so computed with those obtained from the numerical solution of the original Boltzmann-type equation (14) in the quasi-invariant limit ($\epsilon \ll 1$) by means of classical *Monte Carlo* (MC) methods for kinetic equations [23, 40]. After validating in this way the accuracy of the numerical solver for the sole opinion dynamics (homogeneous kinetic model), we will investigate the inhomogeneous kinetic model (20) as well as the hydrodynamic models derived therefrom.

5.1 MC and SP methods for the homogeneous kinetic equation (14)

We begin we rewriting the Boltzmann-type equation (14) in strong form:

$$\partial_\tau g = \frac{1}{\epsilon} \left(Q^+(g, g) - g \right), \tag{34}$$

where Q^+ is the gain part of the kinetic collision operator:

$$Q^+(g, g)(\tau, w) := \left\langle \int_{-1}^{1} \frac{1}{{}'J} g(\tau, {}'w) g(\tau, {}'w_*) \, dw_* \right\rangle.$$

Here, $({}'w, {}'w_*)$ are the pre-interaction opinions generating the post-interaction opinions (w, w_*) according to the binary interaction rule (6) and ${}'J$ is the Jacobian of the transformation from the former to the latter.

To compute the solution of (34), we adopt a direct MC scheme based on the Nanbu algorithm for Maxwellian molecules [40]. We introduce a uniform time grid $\tau^n := n\Delta\tau$ with fixed step $\Delta\tau > 0$ and we denote $g^n(w) := g(\tau^n, w)$. A forward discretisation of (34) on such a mesh reads then

$$g^{n+1} = \left(1 - \frac{\Delta\tau}{\epsilon}\right) g^n + \frac{\Delta\tau}{\epsilon} Q^+(g^n, g^n). \tag{35}$$

From (34), owing to mass conservation, we see that $\int_{-1}^1 Q^+(g, g)(\tau, w)\, dw = 1$ for all $\tau > 0$ if $\int_{-1}^1 g(0, w)\, dw = 1$, therefore $Q^+(g, g)(\tau, \cdot)$ can be regarded as a probability density function at all times. From (35), under the restriction $\Delta\tau \leq \epsilon$, we obtain therefore that g^{n+1} is a convex combination of two probability density functions and is therefore in turn a probability density function. The probabilistic interpretation of (35) is clear: with probability $\frac{\Delta\tau}{\epsilon}$ any two particles interact during the time step $\Delta\tau$; with complementary probability $1 - \frac{\Delta\tau}{\epsilon}$ they do not. This is the basis on which to ground an MC-type numerical method for the approximate solution of (34).

However, it is in general numerically demanding to obtain from (35) an accurate reconstruction of the asymptotic distribution g_∞. To obviate this difficulty, one can take advantage of the fact that, for ϵ sufficiently small, the large time trend of (34) is well approximated by the Fokker-Planck equation (31). In [43], an SP numerical scheme has been specifically designed to capture the large time behaviour of the solution to (31) with arbitrary accuracy and no restriction on the w-mesh size. Moreover, in the transient regime that scheme is second order accurate, preserves the non-negativity of the solution and is entropic for specific problems with gradient flow structure. See also [24, 33, 42] for further applications.

To derive SP schemes for the Fokker-Planck equation (31), we rewrite the latter in flux form:

$$\partial_\tau g = \mathcal{F}[g], \tag{36}$$

where the flux is

$$\mathcal{F}[g](\tau, w) := \mathcal{C}[g](\tau, w)g(\tau, w) + \frac{\sigma^2}{2} D^2(w)\partial_w g(\tau, w)$$

and

$$\mathcal{C}[g](\tau, w) := \gamma \int_{-1}^1 P(w, w_*)(w - w_*)g(\tau, w_*)dw_* + \frac{\sigma^2}{2}(D^2)'(w).$$

Next, we introduce a uniform grid $\{w_i\}_{i=1}^N \subset [-1, 1]$ such that $w_{i+1} - w_i = \Delta w > 0$, we denote by $g_i(\tau)$ an approximation of the grid value $g(\tau, w_i)$ and we consider the conservative discretisation of (36)

$$\frac{dg_i}{d\tau} = \frac{\mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2}}{\Delta w}, \tag{37}$$

where $\mathcal{F}_{i\pm1/2}$ is an approximation of $\mathcal{F}$ at $w_{i\pm1/2} := w_i \pm \frac{\Delta w}{2}$. In particular, we choose a numerical flux of the form

$$\mathcal{F}_{i+1/2} := \tilde{\mathcal{C}}_{i+1/2}\tilde{g}_{i+1/2} + \frac{\sigma^2}{2} D_{i+1/2}^2 \frac{g_{i+1} - g_i}{\Delta w},$$

with $\tilde{g}_{i+1/2}$ defined as a convex combination of g_i, g_{i+1}:

$$\tilde{g}_{i+1/2} := \left(1 - \delta_{i+1/2}\right) g_{i+1} + \delta_{i+1/2}g_i.$$

The coefficient $\delta_{i+1/2} \in [0, 1]$ has to be properly chosen. Setting in particular

$$\tilde{\mathcal{C}}_{i+1/2} := \frac{\sigma^2 D_{i+1/2}^2}{\Delta w} \left(\frac{\gamma}{\sigma^2} \int_{w_i}^{w_{i+1}} \frac{\mathcal{B}[g](\tau, w)}{D^2(w)}\, dw + \log\frac{D_{i+1}}{D_i}\right), \tag{38}$$

where $\mathcal{B}[g]$ is given by (32), we obtain explicitly

$$\delta_{i+1/2} := \frac{1}{\lambda_{i+1/2}} + \frac{1}{1 - \exp(\lambda_{i+1/2})} \quad \text{with} \quad \lambda_{i+1/2} := \frac{2\Delta w\tilde{\mathcal{C}}_{i+1/2}}{\sigma^2 D_{i+1/2}^2}.$$

The order of this scheme for large times coincides with that of the quadrature formula employed for computing the integral contained in (38). In particular, if a standard Gaussian quadrature rule is used then spectral accuracy is achieved in the w variable. In the transient regime, instead, the scheme is always second order accurate.

5.1.1 Comparison of the numerical solutions for large times

We now compare the large time numerical solution of the Boltzmann-type equation (34), obtained by means of the MC scheme with the following specifications:

- 10^5 particles;
- quasi-invariant regime approximated by taking either $\epsilon = 10^{-1}$ or $\epsilon = 10^{-2}$,

with the numerical solution of the Fokker-Planck equation (31), obtained by means of the SP scheme with the following specifications:

- $N = 81$ grid points for the mesh $\{w_i\}_{i=1}^N \subset [-1, 1]$, yielding a mesh step $\Delta w = 2.5 \cdot 10^{-2}$;
- fourth order Runge-Kutta method for the time integration of (37);
- Gaussian quadrature rule, with 10 quadrature points in each cell $[w_i, w_{i+1}]$, for the approximation of the integral in (38).

We use the symmetric bounded-confidence-type compromise function P given by (4) with several choices of the confidence threshold $\Delta \in [0, 2]$. At the initial time $\tau = 0$, we prescribe the uniform distribution in $[-1, 1]$, i.e.

$$g(0, w) = \frac{1}{2}\chi(w \in [-1, 1]).$$

In Figure 5 we fix $\Delta = 1$ (left panels) and $\Delta = 0.4$ (right panels) and we take $\lambda = \sigma^2/\gamma = 5 \cdot 10^{-3}$. We observe that, as expected, the smaller ϵ the more the MC solution coincides with the SP solution of the Fokker-Planck equation for either value of the confidence threshold Δ. Furthermore, the asymptotic profiles compare qualitatively well with those obtained with the deterministic microscopic model (3), cf. Figure 1(a, b), in terms of number and location of the opinion clusters.

In Figure 6 we repeat the same comparisons between the MC and SP numerical solutions but with $\Delta = 0.2$. For $\lambda = 10^{-4}$ (left panels) we recover both a transient behaviour and an asymptotic trend of the solution fully consistent with those already observed with the deterministic microscopic model (3). In particular, four opinion clusters emerge in the long run. Interestingly, for a slightly larger parameter $\lambda = 10^{-3}$, indicating a higher relevance of the self-thinking (stochastic fluctuation) in the behaviour of the individuals, two opinion clusters merge, thereby giving rise to just three clusters in the long run. This aggregate phenomenon can only be observed if some microscopic randomness is duly taken into account in the model.

5.2 Inhomogeneous kinetic equation (21)

We now pass to the inhomogeneous kinetic model, in which the formation of the preferences is driven by an interplay with the opinion dynamics studied before.

We start by outlining the procedure by which we solve the inhomogeneous Boltzmann-type equation (21). Since the Knudsen-like number δ is assumed to be small, at each time step we adopt the very same splitting procedure already discussed in Section 4.1. Therefore, upon introducing a time discretisation $\tau^n := n\Delta\tau$, with $\Delta\tau > 0$ constant, we proceed as follows.

Interaction step. At time $\tau = \tau^n$, we solve the interactions towards the equilibrium during half a time step:

$$\begin{cases} \partial_\tau G(\tau, \xi, w) = \dfrac{1}{\delta}Q(G, G)(\tau, \xi, w), & \tau \in (\tau^n, \tau^{n+1/2}] \\ G(\tau^n, \xi, w) = g(\tau^n, \xi, w) \end{cases} \quad (39)$$

Figure 5: **Top row**: contours of the distribution function g computed numerically for $\tau \in (0, T]$, $T = 50$, from the Fokker-Planck equation (31) with the SP scheme. **Bottom row**: comparison of the numerical approximations at $\tau = T$ of the large time distribution g_∞ obtained with the previous SP scheme and with the MC scheme for the Boltzmann-type equation (34) with two decreasing values of the parameter ϵ simulating the quasi-invariant regime. In both rows, the confidence thresholds are $\Delta = 1$ (left) and $\Delta = 0.4$ (right).

for all $\xi = \xi_i$ belonging to a suitable mesh $\{\xi_i\}_i \subset [-1, 1]$. In this step, we take advantage of the MC scheme introduced in Section 5.1, which has proved to give asymptotic solutions comparable to those of the more accurate SP scheme, provided the parameter δ is sufficiently small. In particular, we use a sample of 10^6 particles and we fix $\delta = 10^{-2}$.

In (39), Q denotes the collision operator that appears on the right-hand side of (21) once this equation has been written in strong form.

Transport step. Next, we take the asymptotic distribution obtained in the interaction step as the input of a pure transport towards the next time step τ^{n+1}:

$$
\begin{cases}
\partial_\tau g(\tau, \xi, w) + (w - \alpha)\partial_\xi \left(\Phi(\xi)g(\tau, \xi, w)\right) = 0, \quad \tau \in (\tau^{n+1/2}, \tau^{n+1}] \\
g(\tau^{n+1/2}, \xi, w) = G(\tau^{n+1/2}, \xi, w).
\end{cases}
$$

In the tests of this section, unless otherwise specified, we prescribe the uniform distribution in the variables ξ, w as initial datum:

$$
g(0, \xi, w) := \frac{1}{4}\chi((\xi, w) \in [-1, 1]^2) \tag{40}
$$

we fix $\lambda = 10^{-3}$ and we take the function Φ given in (5).

19

Figure 6: The same as Figure 5 but with $\Delta = 0.2$. In the left panels we use $\lambda = 10^{-4}$, in the right panels $\lambda = 10^{-3}$. In the latter case, the asymptotic distribution features only three opinion clusters, well reproduced by both the SP Fokker-Planck solution and the MC Boltzmann solution (especially with $\epsilon = 10^{-2}$), because the two central clusters merge during the transient due to a higher relevance of the self-thinking (diffusion) with respect to the tendency to compromise (transport).

(a) $\tau = 1$ (b) $\tau = 3$ (c) $\tau = 5$

Figure 7: Contours of the inhomogeneous kinetic distribution $g(\tau, \xi, w)$ at different times with $\Delta = 1$ and $\alpha = -0.3$.

5.2.1 Symmetric P

First, we consider symmetric interactions described again by the bounded confidence compromise function (4) with $\Delta = 1$. In Figures 7, 8 we show the evolution of the inhomogeneous kinetic model

(a) $\tau = 1$ (b) $\tau = 3$ (c) $\tau = 5$

Figure 8: The same as Figure 7 but with $\alpha = 0.3$.

(a) $\tau = 1$ (b) $\tau = 3$ (c) $\tau = 5$

Figure 9: The same as Figure 7 but with $\Delta = 0.4$ and $\alpha = 0.3$.

(a) w-marginal at $T = 10$ (b) ξ-marginal at $T = 10$

Figure 10: Marginal distributions of the opinions (a) and of the preferences (b) for the numerical test of Figure 9.

for two different choices of the perceived social opinion, $\alpha = \pm 0.3$ respectively. We clearly observe that while the opinions distribute around the conserved mean opinion $m = 0$, as expected, the preferences polarise in two possible ways. For $\alpha = -0.3$, cf. Figure 7, polarisations emerge in $\xi = 0$ and $\xi = 1$. Specifically, individuals with an initial preference in $[-1, 0]$ tend to polarise in $\xi = 0$, whereas individuals with an initial preference in $(0, 1]$ tend to polarize in $\xi = 1$. For $\alpha = 0.3$, cf. Figure 8, the mirror trends emerge. These polarisation patterns of the preferences are very much consistent with those observed in Section 2 with the deterministic microscopic model (2)-(3), cf. Figure 2.

Next, we consider the same symmetric compromise function P as before but now we fix $\Delta = 0.4$.

In Figure 9 we depict the evolution of the inhomogeneous kinetic model for $\alpha = 0.3$. As far as the opinion dynamics are concerned, we recognise that individuals tend to cluster in two well distinct positions, see Figure 10(a), directly comparable with the emerging clusters shown in Figure 5(right) and also, up to diffusion, in Figure 1(b). Nevertheless, the social detail is now higher, because we clearly distinguish that individuals with the same asymptotic opinion may actually polarise in different preferences. More specifically, the opinion cluster near $w = -0.5$ is formed by individuals with preferences polarised in either $\xi = -1$ or $\xi = 0$, while the opinion cluster near $w = 0.5$ is formed by individuals with preference polarised in either $\xi = 0$ or $\xi = 1$, see Figure 9(c). Remarkably, three polarisations of the preference emerge on the whole in the long run, see Figure 10(b), because the opinions do not reach a global consensus.

Also these polarisation patterns of the preferences are consistent with those discussed in Section 2, indeed the deterministic microscopic model can account in principle for three preference poles. The fact that Figure 3(b) shows asymptotically only two of them depends essentially on the choice of the initial conditions, which in a particle model hardly allow one to observe the representative average trend in a single realisation.

The case $\alpha = -0.3$ is qualitatively analogous to the one just discussed, therefore we do not report it in detail.

5.2.2   Non-symmetric P

Finally, we investigate the effect of a non-symmetric compromise function P. As already discussed in Section 3.2.2, we recall that the asymmetry of P can be understood as a systematic bias of the individuals, who for some reason are more prone to change opinion in a specific direction. In this numerical example, we remain in the class of the bounded confidence models and, taking inspiration from [32], we consider

$$P(w, w_*) = \chi(-\Delta_L \leq w_* - w \leq \Delta_R), \tag{41}$$

where Δ_L, $\Delta_R \in [0, 2]$ are two confidence thresholds.

In order to understand the effect of function (41), we observe that if $w \leq w_*$ then interactions are allowed provided $|w_* - w| = w_* - w \leq \Delta_R$. Otherwise, if $w \geq w_*$ then interactions are allowed provided $|w_* - w| = w - w_* \leq \Delta_L$. Thus, if e.g. $\Delta_R > \Delta_L$ then an individual with opinion w is more incline to interact with other individuals with opinion $w_* \geq w$. The converse holds if instead $\Delta_R < \Delta_L$.

Remark 5.1. If $\Delta_L = \Delta_R$ then (41) actually reduces to (4) with $\Delta = \Delta_R$.

We choose $\Delta_L = 0.3$ and $\Delta_R = 0.7$, meaning that individuals compromise preferentially with other individuals with an opinion located on the right of their own. Moreover, we consider the perceived social opinion $\alpha = 0.3$. In Figure 11 we show the evolution of the inhomogeneous kinetic model starting from the uniform distribution (40).

We observe that initially the mean opinion is neutral at any preference, indeed

$$\int_{-1}^{1} w g(0, \xi, w) \, dw = 0, \quad \forall \xi \in [-1, 1].$$

Nevertheless, due to the non-symmetric interactions, the mean opinion is not conserved in time, cf. Figure 11(d). In particular, owing to the bias induced by $\Delta_R > \Delta_L$, the opinions tend to shift on the whole rightwards, cf. Figure 11(e), while the preferences polarise in the three poles $\xi = -1, 0, 1$, cf. Figure 11(f). Again, we notice that the joint picture preference-opinion is a lot more informative than the sole opinion dynamics, because it allows us to observe e.g. that two clusters with nearly the same asymptotic opinion about $w \approx 0.5$ actually include individuals expressing strongly different preferences ($\xi = 0, 1$), cf. Figures 11(b, c).

(a) $\tau = 1$ (b) $\tau = 5$ (c) $\tau = 10$

(d) Mean opinion trend   (e) w-marginal at $T = 10$ (f) ξ-marginal at $T = 10$

Figure 11: **Top row**: Contours of the inhomogeneous kinetic distribution $g(\tau, \xi, w)$ at different times with the non-symmetric compromise function (41) featuring $\Delta_L = 0.3$, $\Delta_R = 0.7$. **Bottom row**: Time trend of the mean opinion (the symmetric case is plotted for duly comparison) and marginal distributions of opinions and preferences at time $T = 10$.

5.3 Hydrodynamic model

Now we test the hydrodynamic model of preference formation derived in Section 4. In particular, since the dynamics predicted by the first order models of Section 4.2 are quite well understood analytically, we focus on the second order model presented in Section 4.3, cf. (28).

To discretise the system of conservation laws (28), we introduce a uniform mesh in the preference domain $[-1, 1]$ made of 300 grid points. Furthermore, we choose a time step $\Delta\tau > 0$ such that the following CFL condition is met at each computational time:

$$\frac{\Delta\tau}{\Delta\xi} \max_{\xi \in [-1, 1]} \left[\Phi(\xi) \max\{|\mu_1(\tau, \xi)|, |\mu_2(\tau, \xi)|\}\right] \leq 1,$$

where μ_1, μ_2 are the eigenvalues of the matrix (29). Then we use a WENO reconstruction in the variable ξ with a Godunov-type numerical flux, coupled with a third order Runge-Kutta integration in time. See [44] for a detailed description of the numerical scheme.

In all the numerical tests of this section we consider the following initial condition:

$$\rho(0, \xi) = \frac{1}{2}\chi(\xi \in [-1, 1]), \quad m(0, \xi) = 0, \tag{42}$$

which represents a uniform distribution of the density over the whole range of preferences with a null mean opinion, denoting initial indecisiveness, at all preferences.

In the first test, whose results are displayed in Figure 12 at time $\tau = 3$, we fix the perceived social opinion $\alpha = 0$ and we consider three increasing values of the parameter λ, in particular $\lambda = 10^{-2}, 10^{-1}, 1$, denoting a progressively stronger influence of the self-thinking. The numerical solution clearly shows that, due to $\alpha = 0$, the symmetry of the initial density is preserved, cf. Figure 12(a). In other words, the system is not driven spontaneously towards a specific preference, hence three symmetric polarisations emerge in $\xi = -1, 0, 1$. The inspection of the

Figure 12: **Top row**: density (a) and mean opinion (b) at time $\tau = 3$ computed from the second order hydrodynamic model (28) with $\alpha = 0$ and increasing values of λ, starting from the initial condition (42). **Bottom row**: time evolution of the mass of agents concentrating about the three preference poles $\xi = -1, 0, 1$ for (c) small and (d) large λ.

trend of the mean opinion, cf. Figure 12(b), reveals that individuals ending in one of the two choices $\xi = \pm 1$ tend to develop opinions in agreement with their preference, while individuals ending in the neighbourhood of the choice $\xi = 0$ tend to develop opinions markedly opposite to their preference. Such shapes of the solution are more and more evident for increasing λ. The role of the parameter λ is further stressed by Figures 12(c), (d), which show the time trend of the mass of agents concentrating in suitable neighbourhoods of the preference poles $\xi = -1, 0, 1$, specifically the intervals $[-1, -\frac{3}{4}]$, $[-\frac{1}{8}, \frac{1}{8}]$, $[\frac{3}{4}, 1]$. For small λ, cf. Figure 12(c), there is a perfect equilibrium among the concentrations in the three poles. For relatively larger λ, cf. Figure 12(d), denoting an increased relevance of the self-thinking, the individuals who concentrate about $\xi = 0$ raise considerably with respect to those who concentrate instead in $\xi = \pm 1$. On the other hand, the latter remain symmetric.

In the second test, cf. Figure 13, we consider the same situation described before but we set the perceived social opinion to $\alpha = 0.3$, which induces a bias in the transport of the preference. The system exhibits again three polarisations, but now with an overall trend towards $\xi = -1, 0$ and a residual polarisation in $\xi = 1$, cf. Figure 13(a). The reason is that now the perceived social opinion is higher than the initial mean opinion of the individuals, therefore initially the dominant drift is leftwards. During time, mean opinions higher than α emerge, cf. Figure 13(b), which give

(a) Density   (b) Mean opinion

(c) $\lambda = 10^{-2}$ (d) $\lambda = 1$

Figure 13: The same as Figure 12 but with $\alpha = 0.3$.

rise to the polarisation in $\xi = 1$ and to a further contribution to the polarisation in $\xi = 0$. The latter subtracts mass to the polarisation in $\xi = -1$, especially for large λ, cf. Figures 13(c), (d).

In order to validate the macroscopic model, in Figure 14 we compare the hydrodynamic quantities ρ, m with the numerical marginals $\int_{-1}^{1} g(\tau, \xi, w) \, dw$, $\int_{-1}^{1} wg(\tau, \xi, w) \, dw$ computed from the solution g to the inhomogeneous Boltzmann model (21). In particular, we solve the kinetic equation via the numerical procedure outlined in Section 5.2 with 50 grid points in the ξ-mesh. Starting from the initial condition (40) for the Boltzmann model, to which there corresponds the initial condition (42) at the hydrodynamic level, we observe that the macroscopic model describes consistently the time evolution of the density of agents and of their mean opinion, as expected.

Finally, in Figure 15 we show that the mirror behaviour is observed with $\alpha = -0.3$, with an overall trend towards $\xi = 0$, 1 and a residual polarisation in $\xi = -1$.

6 Summary and outlook

In this paper we have proposed a development of classical consensus-based opinion formation models by including an additional variable, that we have called preference, which is transported in time by the evolving opinions of the agents. Although commensurable with an opinion, the preference does not simply replicate the opinion dynamics. It is rather the expression of a final choice, which is often necessarily much sharper, i.e. more polarised, than the opinion, like for

Figure 14: Comparison of the numerical solution obtained in Figure 13 in the case $\lambda = 10^{-2}$ with the numerical solution of the Boltzmann-type equation (21) in the quasi-invariant scaling with $\delta = 10^{-2}$ at two successive times: $\tau = 1$ (**top row**) and $\tau = 3$ (**bottom row**).

(a) Density

(b) Mean opinion

Figure 15: The same as Figure 12 but with $\alpha = -0.3$.

instance in case of polls, referendums, elections.

At the agent-based level, the inspiration to model the interplay between opinion and preference is brought from the classical kinematic relation between position and velocity in mechanics: roughly speaking, the time variation of the preference ξ is dictated by the sign of the instantaneous opinion w. However, here we introduce two main differences: (i) a polarisation mechanism of the preference on some values denoting the choices available to the agents; (ii) a bias linked to the

general feeling perceived in the society, that we identify with a perceived social opinion α. Hence, ultimately, the time variation of ξ depends on the signed distance $w - \alpha$. The meaning is clear: an agent will tend to move his/her preference, hence to orient his/her choice, according to the relative collocation in the society that s/he perceives for him/herself. For instance, if there are three possible choices, viz. preference poles, say $\xi = -1, 0, 1$, and the perceived social opinion is $\alpha > 0$, an agent with preference $\xi > 0$ and opinion $w > \alpha$ will move his/her preference towards $\xi = 1$, because $w - \alpha > 0$; conversely, if the opinion is $w < \alpha$, the agent will move his/her preference towards $\xi = 0$, because $w - \alpha < 0$. Such a distinction between opinion and preference is crucial to explain how the choice processes, although originating from the opinion dynamics, produce finally outcomes different from those of the opinion formation processes.

Taking advantage of the methods of statistical physics, in particular of the kinetic theory, we have given an aggregate description of the interplay between opinion and preference by means of an inhomogeneous Boltzmann-type equation, in which w plays morally the role of the "velocity" and ξ that of the "position". Specifically, the collision term of this equation accounts for the opinion formation due to binary interactions among the agents, while the transport term accounts for the opinion-driven preference formation. From here, we have finally derived macroscopic hydrodynamic models of preference formation with the technique of the local equilibrium closure, thanks to the possibility to identify precisely, from the kinetic model, the local equilibrium distribution of the opinions.

The analytical investigation of our models, and the related numerical results further extending the scope of the qualitative analysis, have shown the soundness and consistency of our hierarchical approach across all the considered scales. Moreover, they have highlighted the importance of the parameter α in shaping the collective distribution of the preferences. In this work, we have considered α simply as a given constant. Future research may instead address a variable perceived social opinion, which changes in time driven by various social forces. For instance, the control of α could be the goal of antithetical communication strategies commonly seen in contemporary political scenarios. Majority parties might try to weaken the opposition parties by creating the possibly exaggerated perception of a social opinion strongly oriented in their favour. At the same time, opposition parties might try to exaggerate the social bias towards the majority parties, in order to guard the electorate against the risks of too extremist positions. On the other hand, truly extremist parties might prefer to falsely soften the perceived social opinion, in order to avoid scaring the electorate and losing consensus. All these competitive strategies amount to controlling α with the aim of optimising a certain cost functional. This implies controlling the flux of the macroscopic models of preference formation proposed in this paper, taking however into account that the control strategy depends necessarily on the instantaneous opinion distribution, because α is itself an opinion. From the mathematical point of view, this requires to control the inhomogeneous Boltzmann-type kinetic equation and then to study the passage from the controlled kinetic model to hydrodynamic equations. Suitable techniques need to be explored, in order to make such a passage feasible from both the analytical and the numerical point of view.

Acknowledgements

References

[1] G. Albi, L. Pareschi, G. Toscani, and M. Zanella. Recent advances in opinion modeling: control and social influence. In N. Bellomo, P. Degond, and E. Tadmor, editors, *Active Particles Volume 1, Theory, Methods, and Applications*, Modeling and Simulation in Science, Engineering and Technology. Birkhäuser, 2016.

[2] G. Albi, L. Pareschi, and M. Zanella. Opinion dynamics over complex networks: kinetic modelling and numerical methods. *Kinet. Relat. Models*, 10(1):1–32, 2017.

[3] C. Anteneodo and N. Crokidakis. Symmetry breaking by heating in a continuous opinion model. *Phys. Rev. E*, 95(4), 2017.

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

[5] E. Ben-Naim, P. L. Krapivsky, and S. Redner. Bifurcations and patterns in compromise processes. *Phys. D*, 183(3):190–204, 2003.

[6] E. Ben-Naim, P. L. Krapivsky, F. Vazquez, and S. Redner. Unity and discord in opinion dynamics. *Phys. A*, 330(1):99–106, 2003.

[7] S. Biswas. Mean-field solutions of kinetic-exchange opinion models. *Phys. Rev. E*, 84:056106, 2011.

[8] L. Boudin, A. Mercier, and F. Salvarani. Conciliatory and contradictory dynamics in opinion formation. *Phys. A*, 391(22):5672–5684, 2012.

[9] L. Boudin and F. Salvarani. A kinetic approach to the study of opinion formation. *ESAIM Math. Model. Numer. Anal.*, 43(3):507–522, 2009.

[10] L. Boudin and F. Salvarani. The quasi-invariant limit for a kinetic model of sociological collective behavior. *Kinet. Relat. Models*, 2(3):433–449, 2009.

[11] C. Brugna and G. Toscani. Kinetic models of opinion formation in the presence of personal conviction. *Phys. Rev. E*, 92(5):052818/1–9, 2015.

[12] C. Canuto, F. Fagnani, and P. Tilli. An Eulerian approach to the analysis of Krause's consensus models. *SIAM J. Control Optim.*, 50(1):243–265, 2012.

[13] J. A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani. Asymptotic flocking dynamics for the kinetic Cucker-Smale model. *SIAM J. Math. Anal.*, 42(1):218–236, 2010.

[14] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Rev. Mod. Phys.*, 81:591–646, 2009.

[15] F. Ceragioli and P. Frasca. Continuous and discontinuous opinion dynamics with bounded confidence. *Nonlinear Anal. Real World Appl.*, 13(3):1239–1251, 2012.

[16] A. Chatterjee. Socio-economic inequalities: A statistical physics perspective. In F. Abergel, H. Aoyama, B. Chakrabarti, A. Chakraborti, and A. Ghosh, editors, *Econophysics and Data Driven Modelling of Market Dynamics*, New Economic Windows, pages 287–324. Springer, 2015.

[17] V. Comincioli, L. Della Croce, and G. Toscani. A Boltzmann-like equation for choice formation. *Kinet. Relat. Models*, 2(1):135–149, 2009.

[18] E. Cristiani and A. Tosin. Reducing complexity of multiagent systems with symmetry breaking: an application to opinion dynamics with polls. *Multiscale Model. Simul.*, 16(1):528–549, 2018.

[19] N. Crokidakis. Role of noise and agents' convictions on opinion spreading in a three-state voter-like model. *J. Stat. Mech. Theory Exp.*, 2013:P07008, 2013.

[20] G. Deffuant, F. Amblard, G. Weisbuch, and T. Faure. How can extremism prevail? A study on the relative agreement interaction model. *JASSS*, 5(4), 2002.

[21] M. H. DeGroot. Reaching a consensus. *J. Amer. Statist. Assoc.*, 69(345):118–121, 1974.

[22] G. Dimarco, R. Loubère, J. Narski, and T. Rey. An efficient numerical method for solving the Boltzmann equation in multidimensions. *J. Comput. Phys.*, 353:46–81, 2018.

[23] G. Dimarco and L. Pareschi. Numerical methods for kinetic equations. *Acta Numer.*, 23:369–520, 2014.

[24] G. Dimarco, L. Pareschi, and M. Zanella. Uncertainty quantification for kinetic models in socio-economic and life sciences. In L. Pareschi and S. Jin, editors, *Uncertainty quantification for hyperbolic and kinetic equations*, volume 14 of *SEMAI SIMAI Springer Series*. Springer International Publishing, 2018.

[25] B. Düring, P. Markowich, J.-F. Pietschmann, and M.-T. Wolfram. Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders. *Proc. R. Soc. A*, 465(2112):3687–3708, 2009.

[26] B. Düring and G. Toscani. Hydrodynamics from kinetic models of conservative economies. *Phys. A*, 384(2):493–506, 2007.

[27] B. Düring and M.-T. Wolfram. Opinion dynamics: inhomogeneous Boltzmann-type equations modelling opinion leadership and political segregation. *Proc. R. Soc. A*, 471(2182):20150345/1–21, 2015.

[28] J. R. P. French Jr. A formal theory of social power. *Psychol. Rev.*, 63(3):181–194, 1956.

[29] S. Galam. Rational group decision making: A random field Ising model at $T = 0$. *Phys. A*, 238(1):66–80, 1997.

[30] S. Galam. Heterogeneous beliefs, segregation, and extremism in the making of public opinions. *Phys. Rev. E*, 71:046123, 2005.

[31] M. Garavello, R. Natalini, B. Piccoli, and A. Terracina. Conservation laws with discontinuous flux. *Netw. Heterog. Media*, 2(1):159–179, 2007.

[32] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence: Models, analysis, and simulation. *J. Artif. Soc. Soc. Simulat.*, 5(3):1–33, 2002.

[33] M. Herty, A. Tosin, G. Visconti, and M. Zanella. Hybrid stochastic kinetic description of two-dimensional traffic dynamics. *SIAM J. Appl. Math.*, 78(5):2737–2762, 2018.

[34] M. Lallouache, A. S. Chakrabarti, A. Chakraborti, and B. K. Chakrabarti. Opinion formation in kinetic exchange models: Spontaneous symmetry-breaking transition. *Phys. Rev. E*, 82:056112, 2010.

[35] J. Lorenz. Continuous opinion dynamics under bounded confidence: a survey. *Internat. J. Modern Phys. C*, 18(12):1819–1838, 2007.

[36] A. C. R. Martins and S. Galam. Building up of individual inflexibility in opinion dynamics. *Phys. Rev. E*, 87:042807, 2013.

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

[38] W. Ni and D. Cheng. Leader-following consensus of multi-agent systems under fixed and switching topologies. *Systems Control Lett.*, 59(3-4):209–217, 2010.

[39] L. Pareschi and G. Russo. Time relaxed Monte Carlo methods for the Boltzmann equation. *SIAM J. Sci. Comput.*, 23(4):1253–1273, 2001.

[40] L. Pareschi and G. Toscani. *Interacting Multiagent Systems: Kinetic equations and Monte Carlo methods*. Oxford University Press, 2013.

[41] L. Pareschi, P. Vellucci, and M. Zanella. Kinetic models of collective decision-making in the presence of equality bias. *Phys. A*, 467:201–217, 2017.

[42] L. Pareschi and M. Zanella. Structure preserving schemes for mean-field equations of collective behavior. In Westdickenberg M. Klingenberg C., editor, *Theory, Numerics and Applications of Hyperbolic Problems II. HYP 2016*, volume 237 of *Springer Proceedings in Mathematics & Statistics*, pages 405–421. Springer, Cham, 2018.

[43] L. Pareschi and M. Zanella. Structure preserving schemes for nonlinear Fokker-Planck equations and applications. *J. Sci. Comput.*, 74(3):1575–1600, 2018.

[44] C.-W. Shu. High order weighted essentially nonoscillatory schemes for convection dominated problems. *SIAM Rev.*, 51(1):82–126, 2009.

[45] L. Stella, F. Bagagiolo, D. Bauso, and G. Como. Opinion dynamics and stubbornness through mean-field games. In *52nd IEEE Conference on Decision and Control*, pages 2519–2524, Florence, Italy, December 2013.

[46] K. Sznajd-Weron and J. Sznajd. Opinion evolution in closed community. *Internat. J. Modern Phys. C*, 11(6):1157–1165, 2000.

[47] G. Toscani. Kinetic models of opinion formation. *Commun. Math. Sci.*, 4(3):481–496, 2006.

[48] G. Toscani, A. Tosin, and M. Zanella. Opinion modeling on social media and marketing aspects. *Phys. Rev. E*, 98(2):022315/1–15, 2018.

[49] A. Tosin and M. Zanella. Boltzmann-type models with uncertain binary interactions. *Comm. Math. Sci.*, 16(4):962–984, 2018.

[50] C. Villani. On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Ration. Mech. Anal.*, 143(3):273–307, 1998.

[51] D. J. Watts and P. S. Dodds. Influentials, networks, and public opinion formation. *J. Consum. Res.*, 34(4):441–458, 2007.

[52] W. Weidlich. *Sociodynamics: a Systematic Approach to Mathematical Modelling in the Social Sciences*. Harwood Academic Publishers, Amsterdam, 2000.