Opinion modeling on social media and marketing aspects

Giuseppe Toscani* Andrea Tosin† Mattia Zanella‡

Abstract

We introduce and discuss kinetic models of opinion formation on social networks in which the distribution function depends on both the opinion and the connectivity of the agents. The opinion formation model is subsequently coupled with a kinetic model describing the spreading of popularity of a product on the web through a social network. Numerical experiments on the underlying kinetic models show a good qualitative agreement with some measured trends of hashtags on social media websites and illustrate how companies can take advantage of the network structure to obtain at best the advertisement of their products.

Keywords: Kinetic modeling; opinion formation; social networks; marketing through social networks.
Mathematics Subject Classification: 35Q20, 35Q84, 35Q91, 82B21, 91D30

1 Introduction

Social media are nowadays an important transmission vehicle of information. In the last twenty years we are witnessing an explosion of Internet-based messages transmitted through these media. They have become a major factor in influencing various aspects of consumer behavior including awareness, information acquisition, opinions, attitudes, purchase behavior, to name a few [21, 24, 28, 32, 33, 34]. Therefore, more and more companies started to make use of social media in promotional efforts. On the other hand, maybe in reason of the recent development of these commercial strategies, even though social media is magnifying the impact that consumer-to-consumer conversations have in the marketplace, consolidated methods for increasing the impact of products through social networks have not yet been articulated.

Kinetic models of opinion formation in a multi-agent system of individuals characterized also by further parameters have been addressed in various papers. Among others, the modeling of the first part of this paper has points of contact with a recent work by Düring et al. [18], cf. also [17]. There, the opinion variable is coupled with a further parameter denoting the assertiveness of the agents (with high assertiveness corresponding to leadership). Also the dynamics of opinion formation in a society with a marked presence of zealot or stubborn individuals, i.e., agents who tend to maintain their strong opinions after interacting with other agents, have been dealt with in a number of papers. The effect of conviction was studied mainly in discrete models of opinion dynamics related to consensus formation, voting dynamics, game theory models, and diffusion of innovations among other applications [11, 16, 22, 25, 39, 41]. In these works it is shown, mainly through simulations, how agents with a certain stubbornness can affect the process of consensus formation, especially as far as the kind of expected equilibria that could arise due to their influence and the time needed to convergence are concerned. Last, to stress the importance of understanding opinion dynamics in the modern societies, we quote the recent attempts to investigate the process

*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)

of opinion formation in presence of uncertain interactions between agents [36] and to act on it by means of suitable control strategies [2, 3].

In recent years, the empirical study of social networks and of their role in the formation and spreading of opinions gained a lot of momentum thanks to the extraordinary large amount of data coming from online platforms. In particular, large-scale aggregate statistical descriptions showed patterns and periodic structures in the connectivity of real world networks [14, 23, 27]. In this article, resorting to the powerful methodology of statistical physics [1, 30], we will try to explore some of the aforementioned issues about the popularity dynamics on social media by introducing mathematical models of kinetic type able to follow the marketing of products on social networks. Recent researches on marketing aspects of social media [24] revealed that social media are considered an important component of the promotional strategy, and therefore are incorporated as an integral part of the marketing strategy. Mangold and Faulds [24] describe social media as a *hybrid element of promotion mix* for its dual marketing functions. First, social media can be utilized as a traditional integrated marketing communications tool (e.g., direct marketing), where companies control the content, timing, and frequency of information being shared with consumers. Secondly, social media enables consumers to communicate with each other within their social networks, which creates a further attracting effect for companies. This hybrid marketing tool brings a new challenge to marketers, because they need to learn how to effectively spread information on a product over the largest audience of their target consumers. While discussing about possible strategies, Kaplan and Haenlein [21] suggested to deeply investigate social processes to better understand the social networks behavior. Accordingly, it seems appropriate in a first step to study opinion formation on social networks.

Unlike classical modeling of opinion formation in a multi-agent society [5, 10, 12, 13, 15, 35], where agents are considered indistinguishable, here we will take into account a further parameter, the network range, or *connectivity*, of the agents, which can be reasonably measured in terms of followers. Highly connected agents are identified as *influencers*, because, thanks to their large number of followers, their opinions can reach and influence many other users of the social media [19]. As a matter of fact, however, influencers obey the same mechanism of opinion sharing as any other agent. In particular, they are not necessarily recognized as leaders and typically do not operate in a coordinated manner. The key point is simply the assumption that the opinion of the agents with a high connectivity, i.e., a large number of followers, results more convincing than that of the agents with a low number of followers. As a main difference with [4, 9], where the authors consider dynamic networks, here we structure the population of the agents by means of a fixed background network with realistic connectivity distribution.

Starting from these assumptions, in Section 2 we will introduce a kinetic equation modeling opinion formation in presence of the connectivity parameter. Then, in Section 3 we will couple to such kinetic model a further kinetic equation describing the time evolution of the popularity of a certain product. Finally, in Section 4 we will perform numerical experiments based on the previous model equations, trying to ascertain the better strategy that a company can find to promote at best its products.

The analysis performed in the present work enlightens that the role of an underlying social network is of paramount importance in the formation of opinion patterns among the agents, such as e.g., consensus/dissent. Moreover, it clearly demonstrates that highly connected agents are fundamental to increase in time the popularity of a product, thereby allowing it to be sold at best. The mechanism described by the coupling of the opinion formation model with the popularity spreading model is such that it gives, in some cases, apparently counterintuitive but finally realistic answers. Indeed, it can happen that the natural choice to spread out the message on a large number of agents with low connectivity, while giving an immediate increase in popularity, does not actually guarantee the persistence of the popularity in time. On the other hand, by addressing few people with high connectivity the initial decay of the popularity turns out to be only a local effect that connectivity will subsequently remedy.

2 Kinetic modeling of opinion dynamics over networks

2.1 Microscopic binary model

In the socio-physics community, a customary procedure for modeling the formation of opinions in a population of agents consists in representing the opinion of an individual, with respect to a certain subject, by a real number. This number can vary in some discrete set or in a fixed interval, say $[-1, 1]$, where ± 1 denote the extremal opinions. Individual changes of opinion are assumed to be a result of random binary interactions between pairs of agents. Specifically, the pre-interaction opinions w and w_* of two agents will turn into the new post-interaction opinions w' and w'_* as a consequence of the discussion opened by the two agents, of the influence of external factors such as media or propaganda, and of the spontaneous self-thinking [30, 35].

As briefly discussed in the introduction, modeling opinion dynamics over a social network naturally requires to couple the opinion variable with further parameters able to characterize the *range* of the agents in the social network. Consequently, we assume that the microscopic state of each agent is given by the pair (w, c), where $w \in [-1, 1]$ is the opinion of the agent and $c \in \mathbb{R}_+$ is his/her *connectivity* in the social network, represented e.g., by the number of followers.

The connectivity parameter c is a measure of the credibility of an agent. The higher c, the higher the credibility conferred to that agent by the users of the social network and, consequently, the higher his/her influence on the other agents. Agents with high credibility are called *influencers*: they are able to influence the opinion of other agents while being in turn virtually unaffected by the latter. Unlike [4], in the present work we take the statistical distribution of the connectivity of the agents constant in time. This corresponds to assuming that the connectivity distribution possibly evolves over a time scale by far much slower than that of the opinion changes, so that it can be considered an almost stationary background.

Similarly to [35], individual changes of opinion result from random binary interactions between the agents, where now, in addition to the classical rules of change, the connectivity (viz. credibility) enters to modify the post-interaction opinions. In a microscopic binary interaction between two agents with states (w, c) and (w_*, c_*) the opinion variables update now according to

$$\begin{cases} w' = w + \gamma\kappa(c, c_*)(w_* - w) + D_{\mathrm{op}}(w, c)\eta \\ w'_* = w_* + \gamma\kappa(c_*, c)(w - w_*) + D_{\mathrm{op}}(w_*, c_*)\eta_*. \end{cases} \tag{1}$$

In (1), $\gamma > 0$ is a proportionality parameter. The function κ, expressing the rate of relaxation of either opinion toward that of the other agent (the compromise), is here depending on the connectivities of the agents. From the previous discussion about the role of connectivity, it is natural to require that κ satisfies the assumptions:

$$\kappa(c, c_*) \to 1 \quad \text{for} \quad \frac{c}{c_*} \to 0, \qquad \kappa(c, c_*) \to 0 \quad \text{for} \quad \frac{c}{c_*} \to +\infty,$$

which heuristically mean that the post-interaction opinion w' is greatly influenced by w_* if $c_* \gg c$ while it is virtually unaffected by w_* if, conversely, $c_* \ll c$. Possible examples of functions with these characteristics are

$$\begin{aligned} \kappa(c, c_*) &= \frac{c_*}{c + c_*} = \frac{1}{1 + c/c_*}, \\ \kappa(c, c_*) &= e^{-c/c_*}\left(1 - e^{-c_*/c}\right). \end{aligned} \tag{2}$$

Note that both functions in (2) satisfy the further property $0 \le \kappa(c, c_*) \le 1$ for all $c, c_* \in \mathbb{R}_+$.

The terms $D_{\mathrm{op}}(w, c)\eta$ and $D_{\mathrm{op}}(w_*, c_*)\eta_*$ measure the rate of change of opinion due to the self-thinking of the individuals, namely the possibility that agents change randomly (and independently of each other) their opinion. Following [35], we assume that η and η_* are independent and identically distributed random variables with zero mean and variance $\sigma^2 > 0$, while the coefficient $D_{\mathrm{op}}(w, c) = D_{\mathrm{op}}(|w|, c)$ is nonnegative for all $(w, c) \in [-1, 1] \times \mathbb{R}_+$, non-increasing in $|w|$ and vanishing for $|w| \to 1$.

2.2 Analysis of the microscopic interactions

It is clear that, in order to have a physically acceptable model, the post-interaction opinions cannot cross the extremal values ± 1. For this reason, it is important to verify that the binary interactions (1) preserve the bounds, i.e., that w', $w'_* \in [-1, 1]$ if w, $w_* \in [-1, 1]$. Since $|w_*| \leq 1$, we deduce from (1) the bound

$$\begin{aligned} |w'| &= |(1 - \gamma\kappa(c, c_*))w + \gamma\kappa(c, c_*)w_* + D_{\mathrm{op}}(w, c)\eta| \\ &\leq |(1 - \gamma\kappa(c, c_*)w| + \gamma\kappa(c, c_*) + D_{\mathrm{op}}(w, c)\,|\eta|\,. \end{aligned}$$

If we further assume $0 \leq \gamma$, $\kappa(c, c_*) \leq 1$,

$$|w'| \leq (1 - \gamma\kappa(c, c_*))\,|w| + \gamma\kappa(c, c_*) + D_{\mathrm{op}}(w, c)\,|\eta|\,.$$

Thus $|w'| \leq 1$ if $D_{\mathrm{op}}(w, c)\,|\eta| \leq (1-\gamma\kappa(c, c_*))(1-|w|)$, which is satisfied if there exists a constant $\alpha > 0$ such that

$$\begin{cases} |\eta| \leq \alpha(1 - \gamma\kappa(c, c_*)) \\ \alpha D_{\mathrm{op}}(w, c) \leq 1 - |w|\,. \end{cases}$$

The first condition can be enforced by requiring $|\eta| \leq \alpha(1-\gamma)$, which gives a bound on the random variable η independent of the connectivity distribution in the network. The second condition implies $D_{\mathrm{op}}(\pm 1, c) = 0$ for all $c \in \mathbb{R}_+$, which characterizes the stubbornness of the agents with extremal opinions.

Next, denoting by $\langle\cdot\rangle$ the average with respect to the distributions of the random variables η, η_*, we observe that

$$\langle w' + w'_* \rangle = w + w_* + \gamma(\kappa(c, c_*) - \kappa(c_*, c))(w_* - w).$$

Hence the mean opinion remains unchanged in a binary interaction only if the two agents have the same connectivity ($c = c_*$), or if the function κ is symmetric, so that $\kappa(c, c_*) = \kappa(c_*, c)$.

Likewise, for γ small we compute

$$\begin{aligned} \left\langle (w')^2 + (w'_*)^2 \right\rangle = &-2\gamma(\kappa(c_*, c)w_* - \kappa(c, c_*)w)(w_* - w) \\ &+ \sigma^2(D^2_{\mathrm{op}}(w, c) + D^2_{\mathrm{op}}(w_*, c_*)) + o(\gamma). \end{aligned}$$

We remark that, unlike the case of classical opinion dynamics (without social network) [35], the second moment (the energy) is not necessarily dissipated in a binary interaction, not even in the absence of self-thinking (i.e., for $\sigma^2 = 0$), unless again κ is symmetric in its arguments.

2.3 Boltzmann-type description

Let us denote by $p = p(t, w, c)$ the proportion of agents in the population with opinion w and connectivity c at time $t \geq 0$. It is then possible to describe the time evolution of $p = p(t, w, c)$ by resorting to a Boltzmann-type equation, whose collision part reflects the dynamics of opinion changes because of the interactions among the agents. Under the binary rules (1), the Boltzmann-type equation reads

$$\begin{aligned} \partial_t p &= Q_{\mathrm{op}}(p, p) \\ &= \left\langle \int_{\mathbb{R}_+} \int_{-1}^{1} \left(\frac{1}{{}'J_{\mathrm{op}}} p(t, {}'w, c)p(t, {}'w_*, c_*) - p(t, w, c)p(t, w_*, c_*) \right) dw_*\, dc_* \right\rangle, \end{aligned} \tag{3}$$

where Q_{op} is the collisional operator that takes into account opinion variations due to interactions, and the average $\langle\cdot\rangle$ is taken with respect to the distribution of the random variables η, η_*. In (3), ${}'w$, ${}'w_*$ denote the pre-interaction opinions which generate the post-interaction opinions w, w_* and

$'J_{\mathrm{op}}$ is the Jacobian of the transformation (1) as a function of the variables $'w$, $'w_*$. The weak form of this equation, which avoids the explicit computation of the Jacobian, writes

$$\begin{aligned}&\frac{d}{dt}\int_{\mathbb{R}_+}\int_{-1}^{1}\phi(w,\,c)p(t,\,w,\,c)\,dw\,dc\\&\quad=\left\langle\int_{\mathbb{R}_+}\int_{-1}^{1}\int_{\mathbb{R}_+}\int_{-1}^{1}(\phi(w',\,c)-\phi(w,\,c))p(t,\,w,\,c)p(t,\,w_*,\,c_*)\,dw\,dc\,dw_*\,dc_*\right\rangle,\end{aligned}\tag{4}$$

where $\phi : [-1,\,1]\times\mathbb{R}_+\to\mathbb{R}$ is a test function, namely any observable quantity which can be expressed as a function of the microscopic state $(w,\,c)$, and w' is the post-interaction opinion directly given by (1).

In order to simplify the kinetic description delivered by (4), and to make it more amenable to mathematical analysis, we introduce the following argument. When picking randomly an agent of the system, we may assume that his/her opinion w and connectivity c are two independent variables. In other words, we assume that a generic microscopic state $(w,\,c)$ can be built by sampling independently the opinion w from the statistical distribution of the opinions and the connectivity c from the statistical distribution of the connectivity. In fact at the aggregate level there might be no *a priori* reason to believe that a certain opinion is expressed only by individuals with a certain connectivity nor that a given connectivity level is possessed only by individuals expressing a certain opinion. Hence we postulate

$$p(t,\,w,\,c)=f(t,\,w)g(c),\tag{5}$$

where $f(t,\,w)$ is the probability density function of the opinion at time t and $g(c)$ is the probability density function of the connectivity. We anticipate that we will deal with a more general case later in Section 2.5.

Plugging (5) into (4) and choosing the test function of the form $\phi(w,\,c)=\varphi(w)\psi(c)$ gives

$$\begin{aligned}&\left(\int_{\mathbb{R}_+}\psi(c)g(c)\,dc\right)\frac{d}{dt}\int_{-1}^{1}\varphi(w)f(t,\,w)\,dw\\&\quad=\left\langle\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\psi(c)\left(\int_{-1}^{1}\int_{-1}^{1}(\varphi(w')-\varphi(w))f(t,\,w)f(t,\,w_*)\,dw\,dw_*\right)g(c)g(c_*)\,dc\,dc_*\right\rangle.\end{aligned}$$

In particular, for $\psi(c)\equiv 1$ we end up with

$$\begin{aligned}&\frac{d}{dt}\int_{-1}^{1}\varphi(w)f(t,\,w)\,dw\\&\quad=\left\langle\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\left(\int_{-1}^{1}\int_{-1}^{1}(\varphi(w')-\varphi(w))f(t,\,w)f(t,\,w_*)\,dw\,dw_*\right)g(c)g(c_*)\,dc\,dc_*\right\rangle.\end{aligned}\tag{6}$$

Now we can take advantage of (6) to investigate the aggregate trend of the mean opinion and of the energy of the agent's system.

Let $m(t):=\int_{-1}^{1}wf(t,\,w)\,dw$ be the mean opinion at time t. Choosing $\varphi(w)=w$ in (6) we get

$$\frac{dm}{dt}=\gamma\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\kappa(c,\,c_*)\left(\int_{-1}^{1}\int_{-1}^{1}(w_*-w)f(t,\,w)f(t,\,w_*)\,dw\,dw_*\right)g(c)g(c_*)\,dc\,dc_*=0,$$

therefore m is conserved in time in spite of the fact that, as seen in Section 2.2, at the level of a single binary interaction it is not.

In order to study the large time trend of the opinion energy, which will provide insights into the convergence of the system toward equilibria, it is convenient to apply to (6) the *quasi-invariant opinion limit* introduced in [35], namely an asymptotic procedure reminiscent of the grazing collision limit of the classical kinetic theory [37]. In practice, we consider the regime of weak but

frequent binary interactions, which amounts to taking $\gamma,\, \sigma^2 \to 0^+$ in (1) while simultaneously scaling the time as $\tau := \gamma t$. Introducing the scaled distribution function $\tilde{f}(\tau,\, w) := f(\tau/\gamma,\, w)$, we easily obtain from (6)

$$\frac{d}{d\tau}\int_{-1}^{1}\varphi(w)\tilde{f}(\tau,\, w)\,dw = \frac{1}{\gamma}\left\langle \int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\left(\int_{-1}^{1}\int_{-1}^{1}(\varphi(w')-\varphi(w))\tilde{f}(\tau,\, w)\tilde{f}(\tau,\, w_*)\,dw\,dw_*\right) g(c)g(c_*)\,dc\,dc_*\right\rangle. \tag{7}$$

Notice that for γ small $t = \tau/\gamma$ is large. Therefore, for every fixed $\tau > 0$ the limit $\gamma \to 0^+$ describes the large time trend of $f(t,\, w)$ and of its statistical moments. In parallel, since for $\tau \to +\infty$ it results $t \to +\infty$ as well, the asymptotic behavior of $\tilde{f}(\tau,\, w)$ and of its statistical moments approximates well that of $f(t,\, w)$ and of the corresponding statistical moments. If we define the opinion energy as $E(\tau) := \int_{-1}^{1} w^2\tilde{f}(\tau,\, w)\,dw$ and choose $\varphi(w) = w^2$ in (7) we get, in the quasi-invariant opinion limit $\gamma,\, \sigma^2 \to 0^+$,

$$\frac{dE}{d\tau} = \mathcal{C}(m^2 - E) + \lim_{\gamma,\,\sigma^2\to 0^+}\frac{\sigma^2}{\gamma}\int_{\mathbb{R}_+}\int_{-1}^{1} D_{\mathrm{op}}^2(w,\, c)\tilde{f}(\tau,\, w)g(c)\,dw\,dc \tag{8}$$

where

$$\mathcal{C} := 2\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\kappa(c,\, c_*)g(c)g(c_*)\,dc\,dc_*. \tag{9}$$

If $\sigma^2/\gamma \to 0$ then from (8) we deduce

$$E(\tau) = (E_0 - m^2)e^{-\mathcal{C}\tau} + m^2,$$

where $E_0 := E(0)$. Hence the energy converges exponentially fast to the asymptotic value $E^\infty = m^2$. The coefficient $\mathcal{C}$, which depends on the statistical properties of the connectivity of the social network, gives the speed of convergence of E to E^∞. In other words, $1/\mathcal{C}$ is proportional to the half-life of the exponential decay of the energy. Moreover, it is straightforward to see that

$$W_2(\tilde{f}(\tau,\, \cdot),\, \delta_m) = \sqrt{E(\tau) - m^2},$$

where W_2 at the left-hand side denotes the 2-Wasserstein distance in the space of the probability measures, cf. [6], and δ_m is the Dirac delta centered at the mean opinion m. Hence we further get

$$W_2(\tilde{f}(\tau,\, \cdot),\, \delta_m) = \sqrt{E_0 - m^2}\, e^{-\frac{\mathcal{C}}{2}\tau},$$

indicating that $\tilde{f}(\tau,\, w)$ converges asymptotically to δ_m (*consensus*) with an exponential speed determined by the background network through $\mathcal{C}/2$. Remarkably, in this case the energy is globally dissipated although it is not necessarily so in each binary interaction (cf. Section 2.2).

If instead $\sigma^2/\gamma \to \lambda > 0$ then from (8) we have

$$\frac{dE}{d\tau} = \mathcal{C}(m^2 - E) + \lambda\int_{\mathbb{R}_+}\int_{-1}^{1} D_{\mathrm{op}}^2(w,\, c)\tilde{f}(\tau,\, w)g(c)\,dw\,dc.$$

Choosing for instance

$$D_{\mathrm{op}}(w,\, c) = \beta(c)\sqrt{1 - w^2} \tag{10}$$

with $\beta(c) \geq 0$ for all $c \in \mathbb{R}_+$, and setting

$$\mathcal{B} := \int_{\mathbb{R}_+}\beta^2(c)g(c)\,dc, \tag{11}$$

the equation of the energy becomes

$$\frac{dE}{d\tau} = \mathcal{C}(m^2 - E) + \lambda\mathcal{B}(1 - E),$$

whose solution reads

$$E(\tau) = E_0 e^{-(\mathcal{C}+\lambda\mathcal{B})\tau} + \frac{\mathcal{C}m^2 + \lambda\mathcal{B}}{\mathcal{C}+\lambda\mathcal{B}}\left(1 - e^{-(\mathcal{C}+\lambda\mathcal{B})\tau}\right).$$

The asymptotic value is now $E^\infty := \frac{\mathcal{C}m^2+\lambda\mathcal{B}}{\mathcal{C}+\lambda\mathcal{B}}$ and the exponential speed at which it is approached depends on the background network through $\mathcal{C} + \lambda\mathcal{B}$ or, in other words, the half-life of the exponential decay is proportional to $1/(\mathcal{C} + \lambda\mathcal{B})$. Since λ, $\mathcal{B} \geq 0$, this speed is in general higher than the one of the case without self-thinking.

For a more accurate characterization of the asymptotic distribution reached in this case we refer the reader to the next Section 2.4.

Remark 2.1. Let us define

$$\|\kappa\|_\infty := \operatorname*{ess\,sup}_{(c,\,c_*)\in\mathbb{R}_+\times\mathbb{R}_+} \kappa(c,\,c_*).$$

Then, definition (9) implies that $\mathcal{C} \leq 2\,\|\kappa\|_\infty$. In particular, if $\kappa(c,\,c_*) \leq 1$ for all $c,\,c_* \in \mathbb{R}_+$ then $\mathcal{C} \leq 2$. Let us observe that, in absence of the social network, the binary interaction rules are often of the form (1) with $\kappa \equiv 1$ (cf. [35]), which obviously gives $\mathcal{C} = 2$. Here "without social network" means that in an opinion exchange there is no hierarchy among the agents based on their connectivity, viz. their credibility. In this case, convergence of the system toward the steady state is the fastest possible one. In other words, the presence of the network slows down this process.

2.4 Fokker-Planck asymptotic analysis with self-thinking

In order to gain more detailed insights into the asymptotic opinion distribution reached by the system in the balanced interactive-diffusive regime $\sigma^2/\gamma \to \lambda > 0$ we perform the quasi-invariant opinion limit in (7) with a generic test function φ, which now we require to be smooth enough, say $\varphi \in C^3(-1,\,1)$, and such that $\varphi(\pm1) = \varphi'(\pm1) = 0$.

Since γ, σ^2 are taken small, from (1) we have that $w' - w$ is small as well and we can expand it in Taylor series to obtain

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

with $\bar{w} \in (\min\{w,\,w'\},\,\max\{w,\,w'\})$. Plugging this expansion into (7) and developing the computations yields

$$\begin{aligned}
&\frac{d}{dt}\int_{-1}^{1}\varphi(w)\tilde{f}(\tau,\,w)\,dw \\
&\quad= \frac{\mathcal{C}}{2}\int_{-1}^{1}\int_{-1}^{1}\varphi'(w)(w_* - w)\tilde{f}(\tau,\,w)\tilde{f}(\tau,\,w_*)\,dw\,dw_* \\
&\qquad+ \frac{\gamma}{2}\left(\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\kappa^2(c,\,c_*)g(c)g(c_*)\,dc\,dc_*\right)\int_{-1}^{1}\int_{-1}^{1}\varphi''(w)(w_* - w)^2\tilde{f}(\tau,\,w)\tilde{f}(\tau,\,w_*)\,dw\,dw_* \\
&\qquad+ \frac{\sigma^2}{2\gamma}\int_{\mathbb{R}_+}\int_{-1}^{1}\varphi''(w)D_{\mathrm{op}}^2(w,\,c)\tilde{f}(\tau,\,w)g(c)\,dw\,dc \\
&\qquad+ R(\varphi),
\end{aligned}$$

where the coefficient $\mathcal{C}$ has been defined in (9) while the remainder is

$$R(\varphi) := \frac{1}{6\gamma}\left\langle\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\int_{-1}^{1}\int_{-1}^{1}\varphi'''(\bar{w})\left(\gamma\kappa(w_* - w) + D_{\mathrm{op}}\eta\right)^3\tilde{f}(w)\tilde{f}(w_*)g(c)g(c_*)\,dw\,dw_*\,dc\,dc_*\right\rangle.$$

If $\langle|\eta|^3\rangle < +\infty$ then $\langle|\eta|^3\rangle \sim (\sigma^2)^{3/2}$; due to the further boundedness of φ''', κ, D_{op}, it results $|R(\varphi)| \sim \gamma + \sqrt{\sigma^2}$, whence $R(\varphi) \to 0$ for $\gamma,\,\sigma^2 \to 0^+$. On the whole, in the quasi-invariant opinion

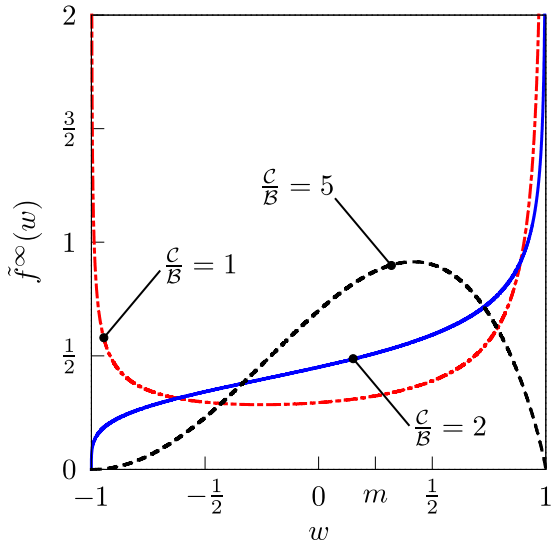


Figure 1: The asymptotic distribution (13) with $m = \frac{1}{4}$, $\lambda = 1$ and different values of $\frac{\mathcal{C}}{\mathcal{B}}$ representative of the three cases discussed in the text.

limit the previous equation becomes

$$\frac{d}{dt}\int_{-1}^{1}\varphi(w)\tilde{f}(\tau,\,w)\,dw = \frac{\mathcal{C}}{2}\int_{-1}^{1}\int_{-1}^{1}\varphi'(w)(w_*-w)\tilde{f}(\tau,\,w)\tilde{f}(\tau,\,w_*)\,dw\,dw_* + \frac{\lambda}{2}\int_{\mathbb{R}_+}\int_{-1}^{1}\varphi''(w)D_{\mathrm{op}}^2(w,\,c)\tilde{f}(\tau,\,w)g(c)\,dw\,dc,$$

which, integrating back by parts at the right-hand side and using the boundary conditions on φ, can be recognized as a weak form of the Fokker-Planck equation

$$\partial_t\tilde{f} = \frac{\lambda}{2}\partial_w^2(\mathcal{D}_{\mathrm{op}}(w)\tilde{f}) + \frac{\mathcal{C}}{2}\partial_w((w-m)\tilde{f}), \tag{12}$$

where we have defined

$$\mathcal{D}_{\mathrm{op}}(w) := \int_{\mathbb{R}_+} D_{\mathrm{op}}^2(w,\,c)g(c)\,dc.$$

The asymptotic solution $\tilde{f}^{\infty}(w) := \lim_{\tau\to+\infty}\tilde{f}(\tau,\,w)$ to this equation, which is obtained by equating the right-hand side of (12) to zero, reads

$$\tilde{f}^{\infty}(w) = \frac{K}{\mathcal{D}_{\mathrm{op}}(w)}\exp\left(\frac{\mathcal{C}}{\lambda}\int\frac{m-w}{\mathcal{D}_{\mathrm{op}}(w)}\,dw\right),$$

where $K > 0$ is a normalization constant to be chosen in such a way that $\int_{-1}^{1}\tilde{f}^{\infty}(w)\,dw = 1$. With $D_{\mathrm{op}}(w,\,c)$ given in particular by (10) we get $\mathcal{D}_{\mathrm{op}}(w) = \mathcal{B}(1-w^2)$, where the coefficient $\mathcal{B}$ has been defined in (11), and explicitly

$$\tilde{f}^{\infty}(w) = \frac{(1+w)^{\frac{\mathcal{C}(1+m)}{2\lambda\mathcal{B}}-1}(1-w)^{\frac{\mathcal{C}(1-m)}{2\lambda\mathcal{B}}-1}}{2^{\frac{\mathcal{C}}{\lambda\mathcal{B}}-1}\,\mathrm{B}\left(\frac{\mathcal{C}(1+m)}{2\lambda\mathcal{B}},\,\frac{\mathcal{C}(1-m)}{2\lambda\mathcal{B}}\right)}, \tag{13}$$

where $\mathrm{B}(x,\,y) := \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$ is the beta function.

The function in (13) is a beta probability density function on the interval $[-1,\,1]$ parametrized by the constants $m \in [-1,\,1]$, $\lambda > 0$, $\frac{\mathcal{C}}{\mathcal{B}} > 0$. Its mean is m and its variance is $\frac{1-m^2}{\frac{\mathcal{C}}{\lambda\mathcal{B}}+1}$. It is

interesting to discuss the different trends of $\tilde{f}^\infty$ depending on the parameter $\frac{\mathcal{C}}{\mathcal{B}}$, which summarizes the statistical connectivity properties of the social network. By inspecting the exponents of the expression (13) we conclude in particular that:

(i) if $\frac{\mathcal{C}}{\mathcal{B}} \geq \max\left\{\frac{2\lambda}{1+m}, \frac{2\lambda}{1-m}\right\}$ then opinions mostly distribute around the mean m. This can be interpreted as a tendency to a mild consensus (cf. the dashed black curve in Figure 1);

(ii) if $\min\left\{\frac{2\lambda}{1-m}, \frac{2\lambda}{1+m}\right\} \leq \frac{\mathcal{C}}{\mathcal{B}} < \max\left\{\frac{2\lambda}{1-m}, \frac{2\lambda}{1+m}\right\}$ then opinions tend to concentrate in $w = -1$ if $m < 0$ or in $w = 1$ if $m > 0$, which indicates a tendency to a radicalized consensus (cf. the solid blue curve in Figure 1);

(iii) if $\frac{\mathcal{C}}{\mathcal{B}} < \min\left\{\frac{2\lambda}{1-m}, \frac{2\lambda}{1+m}\right\}$ then opinions tend to radicalize at the two extreme values $w = \pm 1$ with a substantial splitting of the population between them (cf. the dot-dashed red curve in Figure 1).

We remark that by replacing $\sqrt{1-w^2}$ in (10) with a few different functions of w, such as $1-w^2$ or $1-|w|$, different families of asymptotic distributions $\tilde{f}^\infty$ can be obtained analytically, cf. [35], which are still parametrized by the coefficient $\frac{\mathcal{C}}{\mathcal{B}}$. Then a similar analysis of the influence of the social network on the steady opinion distribution can be performed also in those cases. Here we stick to the choice (10) because it gives rise to an explicit closed form of relevant statistical moments of $\tilde{f}^\infty$.

2.5 A more general case

As mentioned in Section 2.3, in this Section we will investigate the more general case in which we do not make the assumption of statistical independence of the variables w and c, leading to (5). To this purpose, and to avoid inessential difficulties, we slightly modify the interaction rules (1) by assuming that the function κ would depend only on the connectivity of the individual who is changing opinion, hence:

$$\begin{cases} w' = w + \gamma\kappa(c)(w_* - w) + D_{\mathrm{op}}(w,\, c)\eta \\ w'_* = w_* + \gamma\kappa(c_*)(w - w_*) + D_{\mathrm{op}}(w_*,\, c_*)\eta_*. \end{cases} \tag{14}$$

This simplification, which maintains most of the properties of the post-interaction opinions, allows us to explore this case with enough analytical detail. To this extent, we have to require that κ be nonnegative and bounded. Moreover

$$\lim_{c\to 0^+} \kappa(c) = 1, \qquad \lim_{c\to+\infty} \kappa(c) = 0.$$

This translates the idea that users of the social network with low connectivity, viz. with scarce credibility, may be more influenceable by possibly different opinions while those with high connectivity may tend to stick to their guns (sometimes they may need to do that in order to maintain the consensus).

The Boltzmann equation (4) with interaction rules (14) and without assumption (5) is such that the mean opinion $m(t) := \int_{\mathbb{R}_+}\int_{-1}^1 wp(t,\, w,\, c)\, dw\, dc$ is in general not conserved. In fact, letting $\phi(w,\, c) = w$ in (4) yields:

$$\frac{dm}{dt} = \gamma\left(\int_{\mathbb{R}_+}\int_{-1}^1 \kappa(c)p(t,\, w,\, c)\, dw\, dc\right)m - \gamma\int_{\mathbb{R}_+}\int_{-1}^1 \kappa(c)wp(t,\, w,\, c)\, dw\, dc. \tag{15}$$

In particular, denoting by $p^\infty(w,\, c) := \lim_{t\to+\infty} p(t,\, w,\, c)$ the steady distribution, it results that the stationary value m^∞ of the mean opinion is formally given by

$$m^\infty = \frac{\int_{\mathbb{R}_+}\int_{-1}^1 w\,\kappa(c)\,p^\infty(w,\, c)\, dw\, dc}{\int_{\mathbb{R}_+}\int_{-1}^1 \kappa(c)\,p^\infty(w,\, c)\, dw\, dc}.$$

We can make this formula more expressive by writing $p(t,\, w,\, c) = f_c(t,\, w)g(c)$, where $f_c(t,\, w)$ is the *conditional distribution of w given c*, whereas $g(c)$ is the same as in (5). Notice that $f_c(t,\, w)$ is in general different from $f(t,\, w)$ in (5), equality holding if and only if w and c are taken independent. Asymptotically $p^\infty(w,\, c) = f_c^\infty(w)g(c)$, which implies

$$m^\infty = \frac{\int_{\mathbb{R}_+} \kappa(c) \left(\int_{-1}^1 w f_c^\infty(w)\, dw \right) g(c)\, dc}{\int_{\mathbb{R}_+} \kappa(c) g(c)\, dc} = \frac{\int_{\mathbb{R}_+} \kappa(c) m_c^\infty g(c)\, dc}{\int_{\mathbb{R}_+} \kappa(c) g(c)\, dc}, \tag{16}$$

$m_c^\infty := \int_{-1}^1 w f_c^\infty(w)\, dw$ being the mean of the conditional distribution f_c^∞.

By applying the quasi-invariant opinion limit to (4) with interaction rules (14), and proceeding like in Section 2.4, one obtains the Fokker-Planck equation for the scaled distribution function $\tilde{p}(\tau,\, w,\, c) := p(\tau/\gamma,\, w,\, c)$ in the large time scale $\tau = \gamma t$, which is now given by

$$\partial_\tau \tilde{p} = \frac{\lambda}{2} \partial_w^2 (D_{\mathrm{op}}^2(w,\, c)\tilde{p}) + \partial_w \big(\kappa(c)(w - m(\tau))\tilde{p} \big). \tag{17}$$

In equation (17) the constant $\lambda > 0$ is, as before, the limit of the ratio σ^2/γ when $\gamma,\, \sigma^2 \to 0^+$. Notice that, in spite of the fact that $\tilde{p}$ depends on both w and c, this equation involves only w-derivatives (besides the time τ) because the distribution of the connectivity is constant by assumption. Setting $\tilde{p}(\tau,\, w,\, c) = \tilde{f}_c(\tau,\, w)g(c)$, with $\tilde{f}_c(\tau,\, w) = f_c(\tau/\gamma,\, w)$, (17) turns out to be actually an equation for $\tilde{f}_c$:

$$\partial_\tau \tilde{f}_c = \frac{\lambda}{2} \partial_w^2 (D_{\mathrm{op}}^2(w,\, c)\tilde{f}_c) + \partial_w \big(\kappa(c)(w - m(\tau))\tilde{f}_c \big),$$

whose stationary solution $\tilde{f}_c^\infty(w) := \lim_{\tau \to +\infty} \tilde{f}_c(\tau,\, w)$ satisfies the first-order differential equation

$$\frac{\lambda}{2} \partial_w^2 \big(D_{\mathrm{op}}^2(w,\, c)\tilde{f}_c^\infty \big) + \partial_w \big(\kappa(c)(w - m^\infty)\tilde{f}_c^\infty \big) = 0.$$

Although m^∞ is in general unknown, we can solve this equation by regarding it as a constant which parametrizes the solution $\tilde{f}_c^\infty$. We have:

$$\tilde{f}_c^\infty(w) = \frac{K}{D_{\mathrm{op}}^2(w,\, c)} \exp\left(\frac{2\kappa(c)}{\lambda} \int \frac{m^\infty - w}{D_{\mathrm{op}}^2(w,\, c)}\, dw \right),$$

where the two constants $K > 0$, $m^\infty \in [-1,\, 1]$ are determined by imposing

$$\begin{cases} \displaystyle\int_{-1}^1 \tilde{f}_c^\infty(w)\, dw = 1 & (\tilde{f}_c^\infty \text{ is a probability density w.r.t. } w \text{ for all } c \in \mathbb{R}_+) \\ \displaystyle\int_{\mathbb{R}_+} \int_{-1}^1 \tilde{f}_c^\infty(w) g(c)\, dw\, dc = m^\infty & (m^\infty \text{ is the mean of } \tilde{p}^\infty(w,\, c) = \tilde{f}_c^\infty(w)g(c)). \end{cases} \tag{18}$$

The second condition may be equivalently replaced by (16).

Choosing the diffusion coefficient as in (10) we obtain

$$\tilde{f}_c^\infty(w) = \frac{(1+w)^{\frac{\kappa(c)(1+m^\infty)}{\lambda\beta^2(c)} - 1} (1-w)^{\frac{\kappa(c)(1-m^\infty)}{\lambda\beta^2(c)} - 1}}{2^{\frac{2\kappa(c)}{\lambda\beta^2(c)} - 1}\, \mathrm{B}\left(\frac{\kappa(c)(1+m^\infty)}{\lambda\beta^2(c)},\, \frac{\kappa(c)(1-m^\infty)}{\lambda\beta^2(c)} \right)}, \tag{19}$$

which shows that, for every $c \in \mathbb{R}_+$, the steady distribution is a beta probability density function on the interval $[-1,\, 1]$. Here again B denotes the beta function. Interestingly, the mean of (19) coincides precisely with m^∞, thus it is in particular independent of c. Hence in this case conditions (18) are automatically satisfied for every choice of $m^\infty \in [-1,\, 1]$. This result can be fruitfully commented.

(i) In general, studying the time-asymptotic problem may not be enough to identify univocally the stationary distributions. For instance, when the steady state is given as in (19) the value of m^∞ needs to be determined from the time-evolutionary problem (15), and this requires the knowledge of $p(t, w, c)$ from (4) for all $t > 0$. However, for different choices of the diffusion coefficient $D_{\mathrm{op}}(w, c)$, such as e.g., those inspired by [35], conditions (18) may give rise to a non-linear equation for m^∞ which may admit solutions (though not necessarily unique). This method is similar to the one proposed in [38] for recovering the stationary fundamental diagrams of vehicular traffic from a kinetic approach.

(ii) If the initial distribution $p_0(w, c) := p(0, w, c) = f_{c,0}(w)g(c)$, with $f_{c,0}(w) := f_c(0, w)$, is chosen in such a way that the mean $m_c(0) := \int_{-1}^{1} w f_{c,0}(w)\, dw$ is independent of c, then, under the ansatz that also $m_c(t) := \int_{-1}^{1} w f_c(t, w)\, dw$ is independent of c for every $t > 0$, it follows that $m_c(t)$ coincides actually with the w-mean $m(t)$ of p and further, from (15), that this value is conserved in time. Hence, similarly to (13), $m^\infty = m(0)$ in (19), with the significant difference that, unlike $\tilde{f}^\infty$, here $\tilde{f}_c^\infty$ is not the w-marginal distribution of p^∞ or, in other words, that it is not the asymptotic opinion distribution. The latter is instead given by

$$\tilde{f}^\infty(w) = \int_{\mathbb{R}_+} \tilde{f}_c^\infty(w) g(c)\, dc.$$

3 Spreading of the popularity of a product

In this section we consider the problem of the *spreading of the popularity* of a certain product induced by the opinion dynamics described by the model of Section 2. The "product" may be any piece of information reaching the users of the social network: news, videos, advertisements, and the like, that individuals possibly repost to their followers depending on how much their current opinion is aligned with it.

3.1 Microscopic model

We quantify the popularity of a product by means of a variable $v \in \mathbb{R}_+$, whose evolution depends on the interaction with the opinion w and the connectivity c of the agents that the product reaches. In particular, we propose the following microscopic update rule:

$$v' = (1 - \mu)v + P(w, c) + D_{\mathrm{pop}}(v)\xi, \tag{20}$$

where $\mu \in (0, 1)$ is the natural decay rate of the popularity of a product which is not reposted, $P : [-1, 1] \times \mathbb{R}_+ \to \mathbb{R}_+$ is a function expressing the increase in popularity due to reposting, and finally ξ is a random variable with zero mean and finite variance $\varsigma^2 > 0$ modeling a stochastic fluctuation of the popularity with popularity-dependent strength $D_{\mathrm{pop}}(v) \geq 0$. A possible choice for P is

$$P(w, c) = \nu c \mathbb{1}_{[0,\,\Delta]}(|w - \hat{w}|), \tag{21}$$

where $\nu, \Delta > 0$ are parameters, $\mathbb{1}_{[0,\,\Delta]}$ is the characteristic function of the interval $[0, \Delta]$, and $\hat{w} \in [-1, 1]$ is the *target opinion*, i.e., the opinion that the product is mainly addressed at. Hence, according to model (21), the increase in the popularity depends on whether an individual decides to repost the product, which happens if his/her opinion is closer than the threshold Δ to the target opinion. In such a case, the increase in popularity is proportional to the connectivity of the individual, viz. to the number of followers reached when reposting the product.

In order to guarantee from (20) that $v' \geq 0$ we observe that, since $P(w, c) \geq 0$, it is enough to require $D_{\mathrm{pop}}(v)\xi \geq (\mu - 1)v$, which is satisfied if $\xi \geq \mu - 1$ and $D_{\mathrm{pop}}(v) \leq v$. Hence the stochastic fluctuation can take negative values provided the diffusion coefficient is not larger than v. The latter characteristic induces in particular $D_{\mathrm{pop}}(0) = 0$. The most straightforward choice for D_{pop}

is $D_{\rm pop}(v) = v$, which implies a larger and larger diffusion for increasing popularity. Other options are:

$$D_{\rm pop}(v) = \min\{v,\, V_0\}, \qquad D_{\rm pop}(v) = \min\left\{v,\, \frac{V_0^2}{v}\right\},$$

where $V_0 > 0$. The first function expresses a saturation of the diffusion coefficient for high popularity ($v > V_0$). The second function expresses instead a decay of the popularity fluctuations for highly popular products (again $v > V_0$).

3.2 Boltzmann-type description

We now implement a Boltzmann-type kinetic description of the microscopic dynamics (20), coupled with either (1) or (14), by introducing the distribution function $h = h(t,\, v) : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ such that $h(t,\, v)\, dv$ is the fraction of products with popularity in $[v,\, v + dv]$ at time t. If $p(t,\, w,\, c)$ is the distribution of the pair $(w,\, c)$ at time t as introduced in Section 2.3 we have:

$$\begin{cases} \partial_t p = Q_{\rm op}(p,\, p) \\ \partial_t h = Q_{\rm pop}(h,\, p), \end{cases} \tag{22}$$

where the collisional operator $Q_{\rm op}$ has been defined in (3) whereas $Q_{\rm pop}$ writes as follows:

$$Q_{\rm pop}(h,\, p)(t,\, v) = \left\langle \int_{\mathbb{R}_+} \int_{-1}^{1} \left(\frac{1}{{}'J_{\rm pop}} h(t,\, {}'v) - h(t,\, v) \right) p(t,\, w,\, c)\, dw\, dc \right\rangle.$$

Here ${}'v$ is the pre-interaction value of the popularity which generates the post-interaction value v according to the transformation (20) while ${}'J_{\rm pop}$ is the Jacobian of such a transformation as a function of the variable ${}'v$. In weak form, the second equation in (22) reads

$$\frac{d}{dt} \int_{\mathbb{R}_+} \varphi(v) h(t,\, v)\, dv = \left\langle \int_{\mathbb{R}_+} \int_{-1}^{1} \int_{\mathbb{R}_+} \left(\varphi(v') - \varphi(v)\right) h(t,\, v) p(t,\, w,\, c)\, dv\, dw\, dc \right\rangle, \tag{23}$$

where $\langle \cdot \rangle$ denotes here the average with respect to the distribution of the random variable ξ in (20) and $\varphi : \mathbb{R}_+ \to \mathbb{R}$ is a test function, i.e., any observable quantity of the popularity v.

If we assume in particular that opinion dynamics are a much quicker process than the spreading of the popularity then in (23) we can replace p with the asymptotic distribution $p^\infty(w,\, c)$ to get

$$\frac{d}{dt} \int_{\mathbb{R}_+} \varphi(v) h(t,\, v)\, dv = \left\langle \int_{\mathbb{R}_+} \int_{-1}^{1} \int_{\mathbb{R}_+} \left(\varphi(v') - \varphi(v)\right) h(t,\, v) p^\infty(w,\, c)\, dv\, dw\, dc \right\rangle. \tag{24}$$

Choosing $\varphi(v) = 1$ shows that $\int_{\mathbb{R}_+} h(t,\, v)\, dv$ is constant in time, hence if we choose an initial distribution $h_0(v) := h(0,\, v)$ which satisfies the normalization condition $\int_{\mathbb{R}_+} h_0(v)\, dv = 1$ then $h(t,\, v)$ will be a probability density function for all $t > 0$.

Moreover, denoting by $M(t) := \int_{\mathbb{R}_+} v h(t,\, v)\, dv$ the average of the popularity and letting $\varphi(v) = v$ in (24) yields

$$\frac{dM}{dt} = -\mu M + \mathcal{P}, \qquad \mathcal{P} := \int_{\mathbb{R}_+} \int_{-1}^{1} P(w,\, c) p^\infty(w,\, c)\, dw\, dc,$$

that is

$$M(t) = \left(M_0 - \frac{\mathcal{P}}{\mu} \right) e^{-\mu t} + \frac{\mathcal{P}}{\mu}$$

with $M_0 := M(0)$. As a result, the mean popularity reaches asymptotically the value $M^\infty = \frac{\mathcal{P}}{\mu}$, which depends on the statistical properties of the opinion distribution and of the social network through $\mathcal{P}$.

Likewise, denoting by $\mathcal{E}(t) := \int_{\mathbb{R}_+} v^2 h(t,\, v)\, dv$ the energy of the distribution h, we can study the asymptotic behavior of $\mathcal{E}$ by taking advantage of the quasi-invariant limit procedure illustrated in Sections 2.3, 2.4. Precisely, let $0 < \epsilon \ll 1$ be a small parameter and scale $\mu = \mu_0\epsilon$ in (20) and $\nu = \nu_0\epsilon$ in (21) with $\mu_0,\, \nu_0 > 0$. Moreover, let $P(w,\, c) = \epsilon P_0(w,\, c)$ with $P_0(w,\, c) = \nu_0 c \mathbb{1}_{[0,\,\Delta]}(|w - \hat{w}|)$. In the large time scale $\tau := \epsilon t$ the scaled distribution function $\tilde{h}(\tau,\, v) = h(\tau/\epsilon,\, v)$ satisfies the equation

$$\frac{d}{d\tau}\int_{\mathbb{R}_+} \varphi(v)\tilde{h}(\tau,\, v)\, dv = \frac{1}{\epsilon}\left\langle \int_{\mathbb{R}_+}\int_{-1}^{1}\int_{\mathbb{R}_+} (\varphi(v') - \varphi(v))\, \tilde{h}(\tau,\, v) p^{\infty}(w,\, c)\, dv\, dw\, dc \right\rangle. \tag{25}$$

Choosing $\varphi(v) = v^2$ and taking the quasi-invariant interaction limit $\epsilon,\, \varsigma^2 \to 0^+$ (recall that $\varsigma^2 = \langle \xi^2 \rangle$) we obtain

$$\frac{d\mathcal{E}}{d\tau} = -2\mu_0\mathcal{E} + 2\mathcal{P}_0 M + \lim_{\epsilon,\, \varsigma^2 \to 0^+} \frac{\varsigma^2}{\epsilon}\int_{\mathbb{R}_+} D^2_{\mathrm{pop}}(v)\tilde{h}(\tau,\, v)\, dv,$$

where we have denoted $\mathcal{P}_0 := \int_{\mathbb{R}_+}\int_{-1}^{1} P_0(w,\, c) p^{\infty}(w,\, c)\, dw\, dc$.

If $\varsigma^2/\epsilon \to 0^+$ then this equation reduces to

$$\frac{d\mathcal{E}}{d\tau} = -2\mu_0\mathcal{E} + 2\mathcal{P}_0 M,$$

which implies for the energy the asymptotic value $\mathcal{E}^{\infty} = \mathcal{P}_0 M^{\infty}/\mu_0$. Noticing that with the ϵ-scaling introduced above it results $M^{\infty} = \mathcal{P}_0/\mu_0$, we finally obtain $\mathcal{E}^{\infty} = (\mathcal{P}_0/\mu_0)^2 = (M^{\infty})^2$. This indicates that the asymptotic popularity distribution $\tilde{h}^{\infty}(v) := \lim_{\tau \to +\infty} \tilde{h}(\tau,\, v)$ has zero variance, hence $\tilde{h}^{\infty}(v) = \delta_{\mathcal{P}_0/\mu_0}(v)$.

If conversely $\varsigma^2/\epsilon \to \zeta > 0$ and if we choose in particular $D_{\mathrm{pop}}(v) = v$ then the previous equation gives

$$\frac{d\mathcal{E}}{d\tau} = (\zeta - 2\mu_0)\mathcal{E} + 2\mathcal{P}_0 M, \tag{26}$$

which implies $\mathcal{E}^{\infty} = \frac{2\mathcal{P}_0^2}{\mu_0(2\mu_0 - \zeta)}$. From the constraint $\mathcal{E}^{\infty} \geq 0$ we deduce that this value is admissible only if $\zeta < 2\mu_0$. Instead if $\zeta \geq 2\mu_0$, i.e., for a too strong diffusion in the limit, (26) indicates that $\mathcal{E}^{\infty} \to +\infty$, hence that $\tilde{h}^{\infty}$ has infinite variance. In any case, for $\zeta > 0$ the asymptotic popularity distribution is no longer a Dirac delta, i.e., it does not collapse onto the asymptotic mean value M^{∞}. We defer to the next Section 3.3 a more detailed analysis of the large time behavior in this case.

3.3 Fokker-Planck asymptotic analysis with popularity diffusion

We can apply to (25) the same quasi-invariant limit procedure discussed in Section 2.4, taking $\epsilon,\, \varsigma^2 \to 0^+$ and assuming $\varsigma^2/\epsilon \to \zeta > 0$. This describes an asymptotic regime in which the effects of natural decay plus reposting and diffusion balance. The resulting Fokker-Planck equation for the distribution $\tilde{h}(\tau,\, v)$ is

$$\partial_\tau \tilde{h} = \frac{\zeta}{2}\partial_v^2(D^2_{\mathrm{pop}}(v)\tilde{h}) + \partial_v((\mu_0 v - \mathcal{P}_0)\tilde{h}),$$

which admits the following stationary solution:

$$\tilde{h}^{\infty}(v) = \frac{K}{D^2_{\mathrm{pop}}(v)} \exp\left(\frac{2}{\zeta}\int \frac{\mathcal{P}_0 - \mu_0 v}{D^2_{\mathrm{pop}}(v)}\, dv\right)$$

where $K > 0$ is a normalization constant such that $\int_{\mathbb{R}_+} \tilde{h}^{\infty}(v)\, dv = 1$.

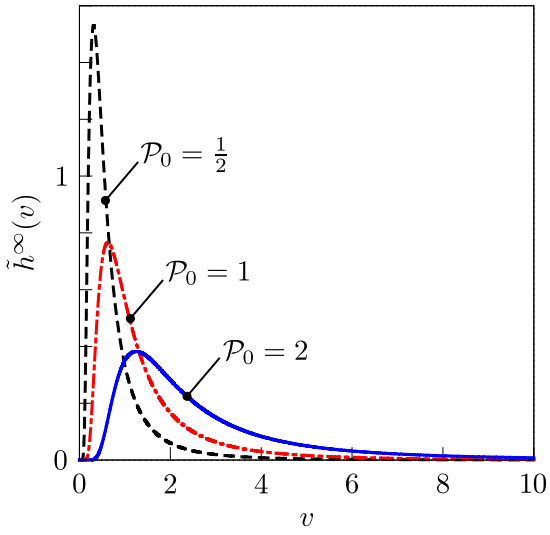


Figure 2: The asymptotic distribution (27) with $\mu_0 = \frac{3}{5}$, $\zeta = 1$ and different values of $\mathcal{P}_0$.

For $D_{\text{pop}}(v) = v$ we obtain in particular

$$\tilde{h}^\infty(v) = \frac{(2\mathcal{P}_0/\zeta)^{1+\frac{2\mu_0}{\zeta}}}{\Gamma\left(1+\frac{2\mu_0}{\zeta}\right)} \cdot \frac{e^{-\frac{2\mathcal{P}_0}{\zeta v}}}{v^{2\left(1+\frac{\mu_0}{\zeta}\right)}}, \tag{27}$$

where $\Gamma(z) := \int_0^{+\infty} t^{z-1}e^{-t}\,dt$ is the gamma function. Notice that (27) is a fat-tailed inverse gamma distribution, indeed $\tilde{h}^\infty(v) \sim v^{-2\left(1+\frac{\mu_0}{\zeta}\right)}$ when $v \to +\infty$. Precisely, this distribution exhibits a *Pareto tail*, cf. [20], which indicates that products reaching very high popularity levels may be rare in general but not that improbable. The mean of the distribution (27) is $\mathcal{P}_0/\mu_0$, consistently with the result found in Section 3.2. Notice also that $v^2\tilde{h}^\infty(v) \sim v^{-\frac{2\mu_0}{\zeta}}$ for $v \to +\infty$, which confirms that the energy and the variance of the distribution (27) are finite only if $\zeta < 2\mu_0$.

Figure 2 shows different profiles of the distribution (27) for fixed μ_0, ζ and increasing values of the parameter $\mathcal{P}_0$, which accounts for the amount of popularity pumped into the system by the users of the social network through reposting.

For $D_{\text{pop}}(v) = \min\{v, V_0\}$, $V_0 > 0$, the stationary distribution reads instead

$$\tilde{h}^\infty(v) = \begin{cases} K\dfrac{e^{-\frac{2\mathcal{P}_0}{\zeta v}}}{v^{2\left(1+\frac{\mu_0}{\zeta}\right)}} & \text{if } v \le V_0 \\[2ex] \dfrac{Ke^{\frac{\mu_0}{\zeta}\left(1-\frac{4\mathcal{P}_0}{\mu_0 V_0}\right)}}{V_0^{2\left(1+\frac{\mu_0}{\zeta}\right)}} e^{\frac{2}{\zeta V_0^2}\left(\mathcal{P}_0-\frac{\mu_0}{2}v\right)v} & \text{if } v > V_0. \end{cases} \tag{28}$$

Notice that now $\tilde{h}^\infty(v) \sim e^{-\frac{\mu_0}{\zeta V_0^2}v^2}$ when $v \to +\infty$, thus fat tails disappear immediately if the diffusion coefficient is assumed to saturate at high popularity.

Finally, for $D_{\text{pop}}(v) = \min\left\{v, \frac{V_0^2}{v}\right\}$, $V_0 > 0$, the stationary distribution becomes

$$\tilde{h}^\infty(v) = \begin{cases} K\dfrac{e^{-\frac{2\mathcal{P}_0}{\zeta v}}}{v^{2\left(1+\frac{\mu_0}{\zeta}\right)}} & \text{if } v \le V_0 \\[2ex] \dfrac{Ke^{\frac{\mu_0}{2\zeta}\left(1-\frac{16\mathcal{P}_0}{\mu_0 V_0}\right)}}{V_0^{2\left(2+\frac{\mu_0}{\zeta}\right)}} v^2 e^{\frac{2}{3\zeta V_0^4}\left(\mathcal{P}_0-\frac{3\mu_0}{4}v\right)v^3} & \text{if } v > V_0, \end{cases} \tag{29}$$

which has an even thinner tail than (28), indeed $\tilde{h}^\infty(v) \sim v^2 e^{-\frac{\mu_0}{2\zeta V_0^4} v^4}$ for $v \to +\infty$.

The distribution of popularity $\tilde{h}^\infty$ supported the most by empirical evidence is (27). For instance, in [32] the author finds that the statistical distribution of the popularity of scientific papers, measured in terms of the received citations, features a power-law-type tail with Pareto exponent near to 3. In [34] the authors find that the popularity of the movies in the United States, measured in terms of their box office gross income, exhibits completely analogous statistical properties.

The Pareto exponent of the distribution (27) can be computed out of the probability that the popularity be greater than a given threshold $v > 0$:

$$\frac{C e^{-\frac{2\mathcal{P}_0}{\zeta v}}}{v^{1+\frac{2\mu_0}{\zeta}}} \leq \int_v^{+\infty} \tilde{h}^\infty(u)\, du \leq \frac{C}{v^{1+\frac{2\mu_0}{\zeta}}},$$

where $C > 0$ is a constant. Hence $\int_v^{+\infty} \tilde{h}^\infty(u)\, du \sim v^{-\left(1+\frac{2\mu_0}{\zeta}\right)}$ when $v \to +\infty$, giving the Pareto exponent $1 + \frac{2\mu_0}{\zeta}$. Owing to the constraint $\zeta < 2\mu_0$, this exponent is invariably greater than 2 and is near 3 if μ_0/ζ is near 1, i.e., if the natural decay rate of popularity is of the same order as the variance of the stochastic fluctuations.

4 Numerical experiments

This section is devoted to a numerical investigation of the kinetic models (4), (23) of the opinion dynamics over a background social network and of the resulting spreading of the popularity of products. The numerical approximation of the Boltzmann equations is done by means of Monte Carlo methods in the Fokker-Planck scaling, see [29, 30, 31]. In the following we use in particular samples of $N = 10^5$ particles.

4.1 Test 1: Opinion dynamics with independent w, c

In this first test we consider the opinion dynamics model (1)-(3) under the independence assumption (5) between w and c, which leads to equation (6) for the evolution of the marginal opinion distribution $f(t, w)$. We remark that in this case the asymptotic profile of f can be analytically computed in the quasi-invariant opinion regime, cf. Section 2.4 and in particular (13). Therefore this test serves as a benchmark for successive numerical experiments.

We consider the following connectivity distributions for $c \in \mathbb{R}_+$:

$$g_1(c) = \frac{e^{-\frac{1}{c}}}{c^3}, \qquad g_2(c) = e^{-c}, \tag{30}$$

which are, respectively, a power law of degree 3 (for c large) and an exponential distribution. Notice that both distributions imply a unitary mean connectivity, indeed

$$\int_{\mathbb{R}_+} c g_1(c)\, dc = \int_{\mathbb{R}_+} c g_2(c)\, dc = 1.$$

The choice of g_1 in (30) is motivated by the experimental literature, see e.g., [7, 14, 27] among others, according to which the connectivity distribution of many networks resulting from social interactions, such as e.g., contact/follower networks in online platforms, scientific collaboration networks, economic networks, is of power-law type. This feature is commonly assumed to be generated by the tendency of the individuals to create links with highly connected vertices of the network rather than with poorly connected ones. Conversely the choice of g_2 in (30) is motivated by the fact that many social networks feature the so-called *small-world* structure, i.e., they are highly clustered with small characteristic path lengths [8, 40].

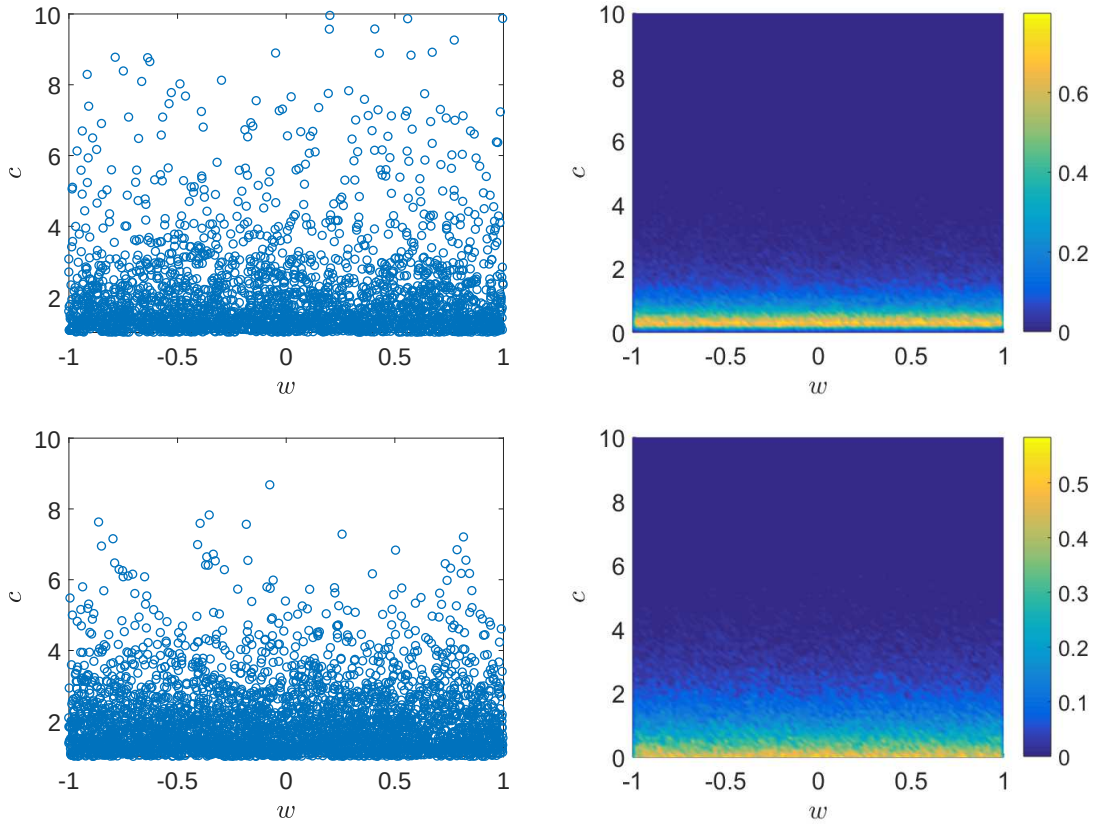


Figure 3: **Test 1, Section 4.1**. Initial distributions (32) (top row) and (33) (bottom row). Left column: 10^3 particles sampled from the prescribed initial distributions. Right column: continuous numerical approximations of such distributions.

For both networks modeled by g_1 and g_2 we compute numerically the large time trend of the solution f to the Boltzmann equation (6) using in the binary interaction rules (1) the functions κ given by (2) and D_{op} given by (10) with in particular

$$\beta(c) = \frac{1}{\frac{1}{10} + c}. \tag{31}$$

This function models a local diffusion which becomes weaker and weaker as the connectivity increases, meaning that people with higher credibility tend to be less prone to erratic changes of opinions due to self-thinking.

As initial opinion distribution we take

$$f_0(w) := f(0, w) = \frac{1}{2}\mathbb{1}_{[-1,\,1]}(w),$$

i.e., the uniform distribution in the interval $[-1, 1]$ with mean $m = 0$, which, as discussed in Section 2.3, is conserved in time. The initial bivariate distribution function $p_0(w, c) := p(0, w, c)$ of the kinetic model is then either

$$p_0(w, c) = \frac{1}{2}\mathbb{1}_{[-1,\,1]}(w)\frac{e^{-\frac{1}{c}}}{c^3}, \tag{32}$$

if the network connectivity is described by the distribution g_1 in (30), or

$$p_0(w, c) = \frac{1}{2}\mathbb{1}_{[-1,\,1]}(w)e^{-c}, \tag{33}$$

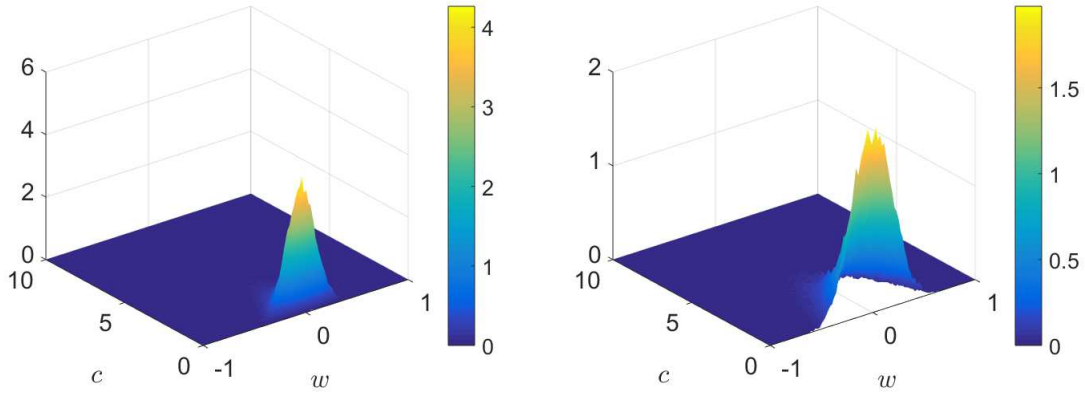


Figure 4: **Test 1, Section 4.1**. Large time bivariate solution $p^\infty(w, c)$ of the Boltzmann-type model (3)-(4) with power law (left) and exponential (right) connectivity distribution, cf. (30), and with the independence assumption (5) of the variables w, c. The initial conditions are the distributions (32), (33), respectively, displayed in Figure 3.

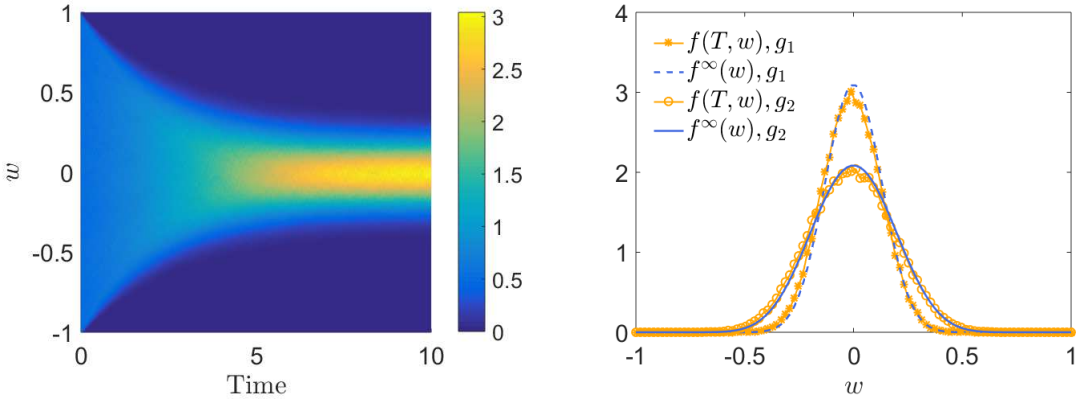


Figure 5: **Test 1, Section 4.1**. Left: transient trend of the marginal distribution $f(t, w)$ in the time interval $[0, 10]$ with power law connectivity distribution. Right: large time opinion distribution (at the final computational time $T = 10$) in the two cases of connectivity distribution (g_1: power law; g_2: exponential). The blue curves are the analytical steady distribution plotted from (13).

if it is described by the distribution g_2 in (30). See Figure 3.

Figure 4 shows an approximation, for a sufficiently large time, of the asymptotic bivariate distribution $p^\infty(w, c) = f^\infty(w)g_{1,2}(c)$, where $f^\infty(w)$ is obtained by solving numerically the scaled Boltzmann equation (7) with $\gamma = 10^{-3}$. Not surprisingly, it turns out to be close to the asymptotic distribution (13) computed in the quasi-invariant opinion regime from the Fokker-Planck equation (12), as it is further clearly shown by Figure 5. The latter displays the transient behavior of $f(t, w)$ in the time interval $[0, 10]$ and the asymptotic profile of $f^\infty(w)$, which agrees perfectly with (13) computed using the values of $\mathcal{C}$, $\mathcal{B}$ evaluated from (9), (11) by numerical integration of the functions κ, β, $g_{1,2}$ of the present test and moreover with $\lambda = 3 \cdot 10^{-3}$.

It is interesting to observe, from the right panel of Figure 5, that, thanks to its heavier tail, the power law distribution g_1 induces a greater opinion consensus around the conserved mean $m = 0$ than that produced by the exponential distribution g_2. This is somehow reminiscent of the dichotomy between clustering and consensus dynamics analyzed in [26].

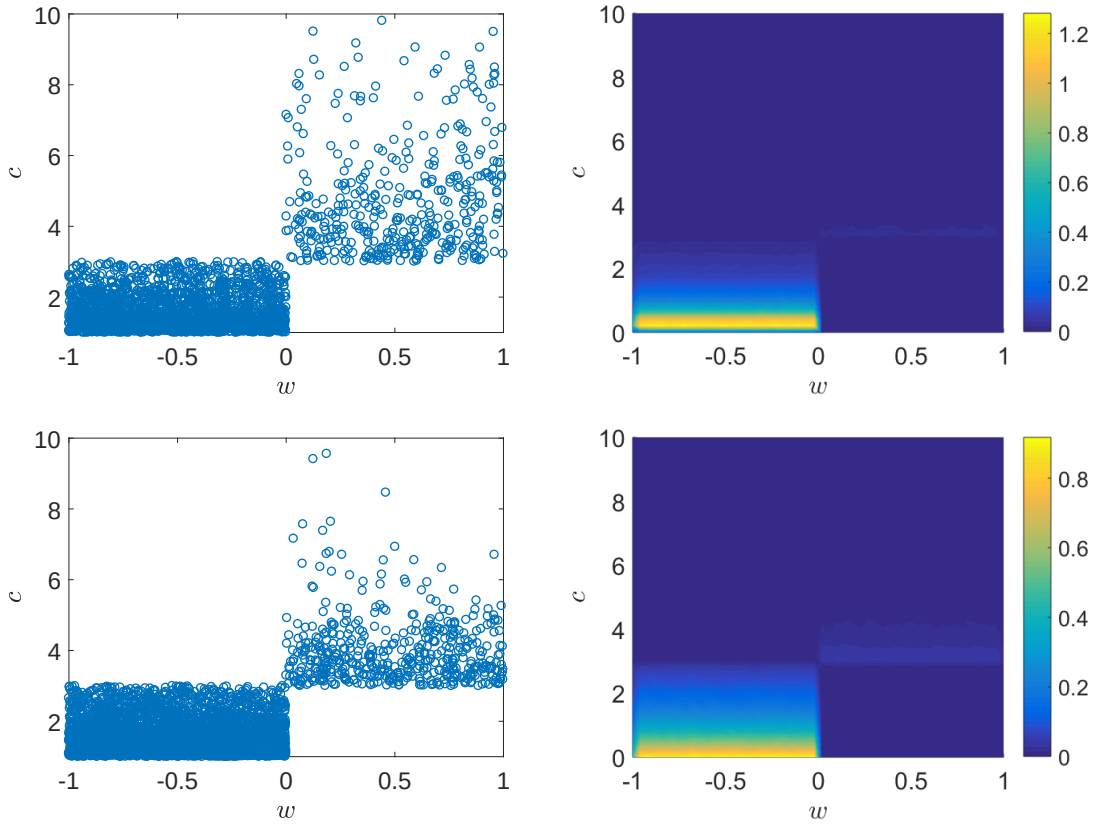


Figure 6: **Test 2, Section 4.2**. Initial distributions (34) with $c_0 = 3$ and connectivity distribution given by either the power law g_1 in (30) (top row) or the exponential law g_2 in (30) (bottom row). Left column: 10^3 particles sampled from the prescribed initial distributions. Right column: continuous numerical approximations of such distributions.

4.2 Test 2: Opinion dynamics with dependent w, c

In this second test we consider the Boltzmann-type model (1)-(3) for the bivariate kinetic distribution $p(t, w, c)$ *without* the independence assumption (5) between the variables w and c. The reference theoretical discussion is the one set forth in Section 2.5. In particular, we choose as initial condition a distribution $p_0(w, c) := p(0, w, c)$ which does not write as the product of its marginals, precisely:

$$p_0(w, c) = \left[\mathbb{1}_{[-1,\,0]\times[0,\,c_0)}(w, c) + \mathbb{1}_{(0,\,1]\times[c_0,\,+\infty)}(w, c)\right] g_{1,2}(c), \tag{34}$$

where $g_{1,2}(c)$ is either connectivity distribution given in (30). The function

$$f_{c,0}(w) = \mathbb{1}_{[-1,\,0]\times[0,\,c_0)}(w, c) + \mathbb{1}_{(0,\,1]\times[c_0,\,+\infty)}(w, c),$$

with $c_0 \in \mathbb{R}_+$, is the initial conditional distribution of w given c. In practice, such an $f_{c,0}$ says that individuals with a connectivity lower than c_0 have initially an opinion uniformly distributed in the interval $[-1, 0]$ whereas individuals with a connectivity greater than or equal to c_0 have initially an opinion uniformly distributed in the interval $[0, 1]$. This models polarized opinions around the mean values $\mp\frac{1}{2}$ of $f_{c,0}$ (for $c < c_0$ and $c \geq c_0$, respectively) depending on the connectivity distribution.

Figure 6 shows either distribution (34) with $c_0 = 3$ for the two different choices (30) of the distribution g. It is worth pointing out that with this threshold c_0 the initial fraction of individuals

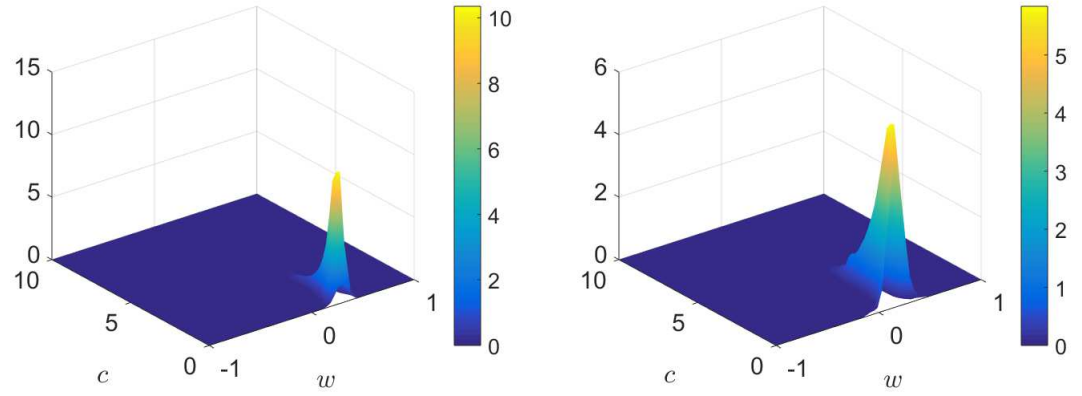


Figure 7: **Test 2, Section 4.2**. Large time bivariate solution $p^\infty(w, c)$ of the Boltzmann-type model (3)-(4) with power law (left) and exponential (right) connectivity distribution, cf. (30) and without the independence assumption (5) of the variables w, c. The initial conditions are the distributions (34) displayed in Figure 6.

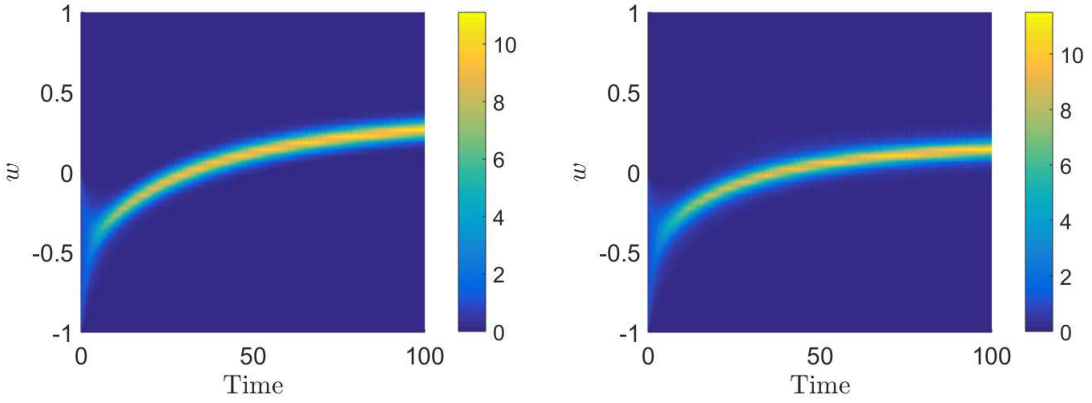


Figure 8: **Test 2, Section 4.2**. Time evolution of the marginal opinion density $f(t, w)$ in the time interval $[0, 100]$ for power law (left) and exponential (right) connectivity distribution.

with opinion $w \leq 0$, namely $\int_{-1}^{0} \int_{\mathbb{R}_+} p_0(w, c)\, dc\, dw$, is invariably larger than that of individuals with opinion $w > 0$, namely $\int_{0}^{1} \int_{\mathbb{R}_+} p_0(w, c)\, dc\, dw$, for both choices of the distribution g.

We compute the large time trend of the system choosing the connectivity-based interaction function κ (2) and the local diffusion coefficient D_{op} (10) with β given by (31). Notice in particular that the choice (2) of κ makes this numerical experiment depart from the analysis performed in Section 2.5, where the function κ was supposed to depend only on the connectivity of one individual of the interacting pair. This gives us the opportunity to explore a scenario which was not possible to cover analytically in detail.

Figure 7 shows the large time solution of the Boltzmann-type model (1)-(3) computed under the time scaling $\tau = \gamma t$ with $\gamma = 10^{-3}$ (cf. the quasi-invariant opinion limit described in Section 2.4). The numerical solution is shown at the computational time $T = 100$, which, owing to the above scaling, yields a good approximation of the asymptotic profile. It is interesting to observe that, in contrast to the test considered in Section 4.1, here the statistical dependence of the variables w, c tends to drive the kinetic distribution p toward the region of the state space where $w \geq 0$. This happens in spite of the fact that initially such a region contains the lower fraction of individuals, who are however the most influential ones of the social network, namely those with an over-

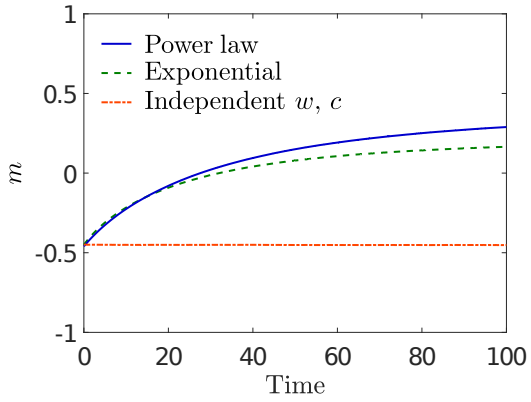


Figure 9: **Test 2, Section 4.2**. Time evolution of the mean opinion in the time interval [0, 100] with the two connectivity distributions given in (30). The evolution of the mean opinion in the case of independent w, c, cf. Section (2.3), is also reported for duly comparison.

threshold connectivity c. This trend is even more evident in Figure 8, which shows the time evolution of the marginal opinion distribution $f(t, w)$, and in Figure 9, which shows the time evolution of the mean opinion $m(t) = \int_{-1}^{1} wf(t, w)\, dw$. Unlike the case addressed analytically and numerically in Sections 2.3, 4.1, here m is no longer constant in time (as expected) and, in particular, is attracted toward a nonnegative asymptotic value (with small variance, cf. Figure 8) dictated by the initial distribution of the influencers, i.e., the individuals with $c \geq c_0$.

It is also interesting to observe from Figure 9 that with the power law connectivity distribution g_1 the mean opinion m tends asymptotically to a value closer to the initial mean opinion of the influencers, namely $\frac{1}{2}$, than with the exponential connectivity distribution g_2. This is consistent with the fact that the heavier tail of g_1 makes the presence of influencers, i.e., individuals with a very high connectivity, much more probable.

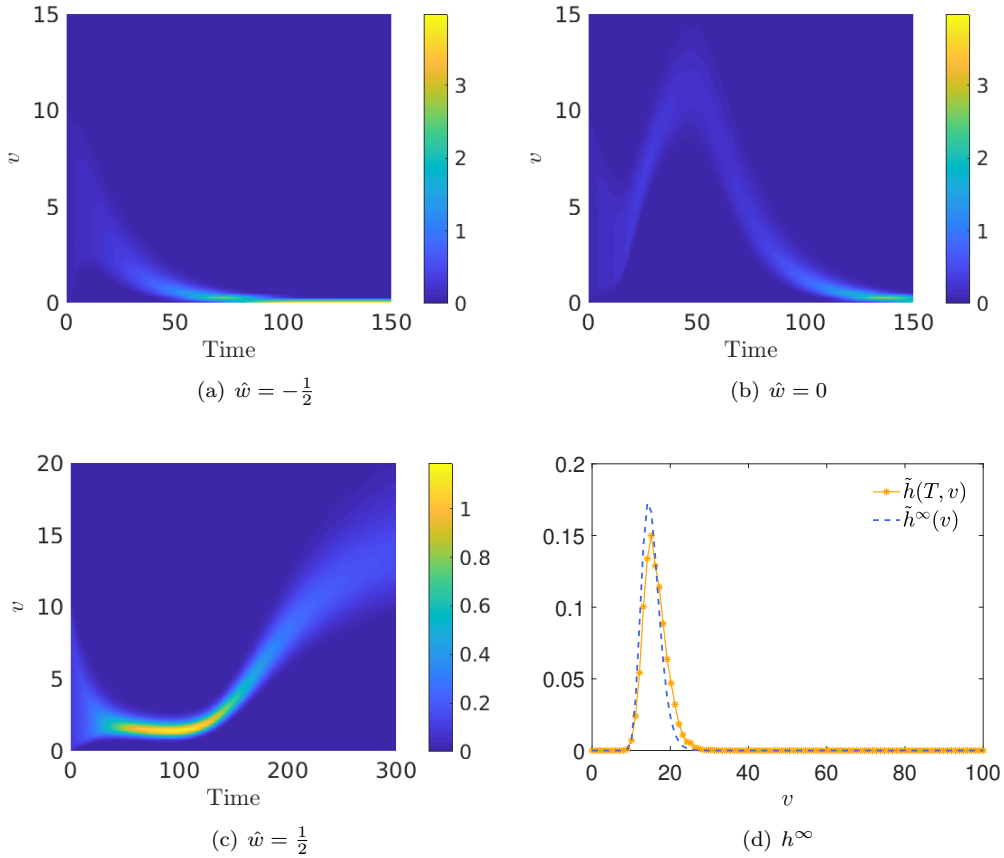
4.3 Test 3: Spreading of the popularity

In this test we consider model (22) for the spreading of the popularity of a product in connection with the opinion dynamics investigated so far. In particular, in contrast to the analysis performed in Section 3, here we solve the dynamically coupled model, i.e., we do not assume that opinion dynamics are necessarily a much quicker process than the spreading of the popularity. This allows us to account for a richer time trend of the relevant statistical quantities than the simple exponential decay found in Section 3.2.

We adopt the power law connectivity distribution g_1 given in (30) and, as far as the opinion distribution is concerned, we take as initial condition the corresponding probability density function (34) with $c_0 = 3$, see also the top row of Figure 6. This means, in particular, that we consider the case of dependent variables w, c. In parallel, we assume that initially the popularity v is uniformly distributed in the interval [0, 10], hence

$$h_0(v) := h(0, v) = \frac{1}{10}\mathbb{1}_{[0,\,10]}(v).$$

We fix the other parameters of the microscopic interactions, cf. (20)-(21), as follows: $\epsilon = 10^{-3}$ (the scaling parameter of the quasi-invariant regime, cf. Sections 3.2, 3.3), $\mu = 10^{-5}$, whence $\mu_0 = \mu/\epsilon = 10^{-2}$, $\nu = 10^{-3}$, whence $\nu_0 = \nu/\epsilon = 1$, $\Delta = 0.15$, $\varsigma^2 = 10^{-5}$, whence $\zeta = \varsigma^2/\epsilon = 10^{-2}$. Finally, we choose the diffusion coefficient $D_{\text{pop}}(v) = v$. Notice that this choice of the microscopic parameters implies in particular $\mu_0/\zeta = 1$, which is the condition found in Section 3.3 for obtaining an asymptotic distribution of the popularity with Pareto exponent 3, the value supported by empirical observations.



(a) $\hat{w} = -\frac{1}{2}$ (b) $\hat{w} = 0$

(c) $\hat{w} = \frac{1}{2}$ (d) h^∞

Figure 10: **Test 3, Section 4.3**. (a)–(c) Time evolution of the popularity distribution $h(t, v)$ coupled to the opinion dynamics examined in Test 2, Section 4.2 for the three different choices of the target opinion $\hat{w} = -\frac{1}{2}$, 0, $\frac{1}{2}$. (d) Analytical and numerical asymptotic distributions of the popularity computed respectively from (27) and by numerical integration of model (22) up to the computational time $T = 700$.

The initial condition $p_0(w, c)$ in (34) with the fat-tailed connectivity distribution $g_1(c)$, cf. (30), implies that initially the mean opinion of the population is

$$m_0 := \int_{\mathbb{R}_+} \int_{-1}^{1} w p_0(w, c)\, dw\, dc = \frac{1}{2} - \Gamma\left(2, \frac{1}{3}\right) \approx -0.45,$$

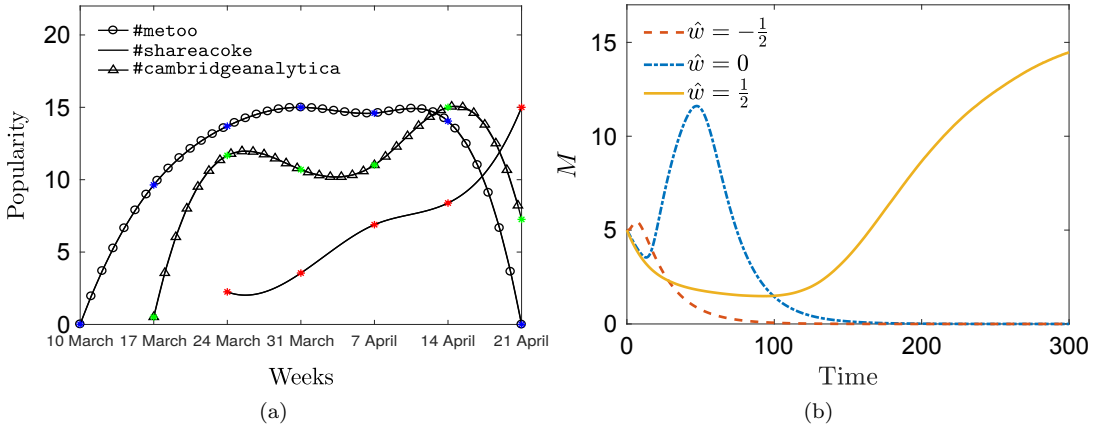
where $\Gamma(s, x) := \int_x^{+\infty} t^{s-1} e^{-t}\, dt$ is the incomplete gamma function. By looking only at this aggregate characteristic of the society one may argue that a good strategy for enhancing the popularity of a product is to target an opinion close to m_0. For instance, one may take $\hat{w} = -\frac{1}{2}$ in (21). Nevertheless, our mathematical model shows that this is by no means a good choice, in fact in such a case the popularity decays rapidly to zero, cf. Figure 10a.

Another possibly common strategy may be to target the neutral opinion $\hat{w} = 0$ in the hope that this may render the product more generically suited to most individuals. Our model clearly shows that also this choice turns out to be ineffective, because, after a temporary rise, the popularity of the product invariably vanishes in the long run, cf. Figure 10b.

The drawback of both strategies is that they choose the target opinion $\hat{w}$ by basically ignoring statistical facts related to the connectivity distribution in the population. Considering that most individuals are actually not influencers, in fact the fraction of agents with low connectivity ($c <$

Table 1: Total number of social interactions involving three hashtags in the most popular social networks (Facebook, Twitter, Instagram, Google+, and Youtube) in the period March-April 2018. Source: `brand24.com`.

Week	`#metoo`	`#shareacoke`	`#cambridgeanalytica`
10 March	17	–	–
17 March	10521	–	468
24 March	14954	194	10662
31 March	16375	306	9768
7 April	15936	594	10055
14 April	15337	723	13497
21 April	0	1293	6626



(a) (b)

Figure 11: **Test 3, Section 4.3**. (a) Third order spline interpolation of the popularity time trend of the three hashtags reported in Table 1. For ease of representation data have been normalized with respect to their maxima. Colored markers are the non-interpolated data. (b) Time evolution of the mean popularity M in the three scenarios (a)–(c) considered in Figure 10.

$c_0 = 3$) is

$$\int_0^3 \int_{-1}^1 p_0(w,\, c)\, dw\, dc = \int_0^3 g_1(c)\, dc = \Gamma\left(2,\, \frac{1}{3}\right) \approx 95\%,$$

one understands that the value m_0 is mainly determined by people who will not be able to dictate the opinion trends. Owing to this, the strategy of targeting the initial mean opinion of the influencers ($c \geq c_0$), namely $\hat{w} = \frac{1}{2}$, turns out to be successful in spite of the very small fraction of influencers in the social network (about 5%). In fact, as Figure 10c shows, the popularity suffers from an initial short decrease due to the fact that the target opinion is somehow elitist at the beginning; then it goes through a non-zero plateau during the influencer-driven phase of reorganization of the opinions; and finally it steadily rises when the influencers manage to drag the opinions of the population in their area. For this last case we also compare in Figure 10d the asymptotic distributions of the popularity computed analytically from (27) and by numerical integration of model (22) up to the computational time $T = 700$. Remarkably function (27), in spite of having been obtained under the simplification of steady opinion dynamics, cf. Section 3, turns out to be a good approximation of the actual popularity distribution computed from model (22) in which popularity trends are dynamically coupled with opinion formation.

Among the possible measures of popularity in social networks, the number of social interactions involving hashtags (such as e.g., posts on online social platforms, their shares and likes) is adopted

by several media-marketing websites. As an example, in Table 1 we report the week data collected by the reference webapp `brand24.com` in the period March-April 2018 concerning three hashtags which have characterized recent social media campaigns:

(i) `#metoo` about the movement against sexual harassment;

(ii) `#shareacoke` about the advertisement of a popular soft drink;

(iii) `#cambridgeanalytica` about the improper use of personal data of social network users.

In Figure 11a we plot the time trends of the popularity of these hashtags obtained by means of a third order spline interpolation of the data in Table 1. We point out that data in the plot have been normalized with respect to their maxima for an easier visualization. In Figure 11b we show instead the time evolution of the mean popularity in the scenarios (a)–(c) illustrated in Figure 10. The comparison shows that the popularity trends simulated by our model capture qualitatively well the observable characteristics of representative real trends, thereby confirming that our model may explain the basic microscopic mechanisms of popularity spreading in connection with opinion sharing on social networks.

5 Conclusions

In this paper we introduced mathematical models of kinetic type able to follow the marketing of products by using social networks. Our analysis is twofold. First, we considered the problem of opinion formation in a social network in which the change of opinion also depends on the connectivity of the agents. Specifically, in agreement with the literature, we assumed that agents with low connectivity are more prone to be influenced by well connected agents (the so-called influencers). Then, we investigated the spreading of the popularity of a selected product (such as e.g., news, video, advertisement) by coupling it to the opinion dynamics taking place over the social network. In particular, we assumed that the product launched onto the social network targets a given opinion and can be possibly reposted by the agents that it reaches depending on how much their current opinion is aligned with the targeted one. By means of analytical and numerical results we were able to recover the typical fat-tailed statistical distribution of the popularity well acknowledged by empirical observations. In addition to this, our results pointed out the importance of the agents with high connectivity, and of their opinions, in the possible success of the product under consideration. More in general, they highlighted clearly the role of the social network in promoting different forms of consensus, and consequently different trends of popularity, depending on the connectivity distribution of the agents. Finally, we showed the potential effectiveness of our modeling approach by means of a qualitative comparison of the simulated popularity trends with real trends of a few media campaigns.

Further extensions of the proposed model may include, for instance, a variable target opinion which follows the ongoing opinion dynamics over the social network as well as a dynamic network in which connections among the agents can possibly change in time, see e.g. [4]. Moreover the notion of popularity, here considered especially in connection with product marketing, may be thought of more in general as a measure of the permeation of a given message into the population of agents. Therefore the approach developed in this paper can constitute a formal basis for the investigation of effective communication strategies in awareness campaigns concerning important contemporary social issues such as e.g., homeland security, national health or other social programs.

Acknowledgements

This work has been written within the activities of GNFM (Gruppo Nazionale per la Fisica Matematica) and GNCS (Gruppo Nazionale per il Calcolo Scientifico) groups of INdAM (National Institute of High Mathematics). G. T. acknowledges support from the MIUR project "Optimal mass transportation, geometrical and functional inequalities with applications" and from IMATI

institute of the National Council for Research (Pavia, Italy). M. Z. acknowledges support from "Compagnia di San Paolo" (Torino, Italy).

References

[1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(1):1–47, 2002.

[2] G. Albi, L. Pareschi, and M. Zanella. Boltzmann-type control of opinion consensus through leaders. *Phil. Trans. R. Soc. A*, 372(2028):20140138/1–18, 2014.

[3] G. Albi, L. Pareschi, and M. Zanella. On the optimal control of opinion dynamics on evolving networks. In L. Bociu, J. A Désidéri, and A. Habbal, editors, *System Modeling and Optimization. CSMO 2015*, volume 494 of *IFIP Advances in Information and Communication Technology*. Springer, Cham, 2016.

[4] 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.

[5] G. Aletti, G. Naldi, and G. Toscani. First-order continuous models of opinion formation. *SIAM J. Appl. Math.*, 67(3):837–853, 2007.

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

[7] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[8] A.-L. Barabási, R. Albert, and H. Jeong. Mean-field theory for scale-free random networks. *Phys. A*, 272(1-2):73–187, 1999.

[9] J. Barré, J. A. Carrillo, P. Degond, D. Peurichard, and E. Zatorska. Particle interactions mediated by dynamical networks: assessment of macroscopic descriptions. *J. Nonlinear Sci.*, 28(1):235–268, 2018.

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

[11] S. Biswas. Mean-field solutions of kinetic-exchange opinion models. *Phys. Rev. E*, 84(5):056106/1–9, 2011.

[12] C. M. Bordogna and E. V. Albano. Dynamic behavior of a social model for opinion formation. *Phys. Rev. E*, 76(6):061125/1–6, 2007.

[13] 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.

[14] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Rev.*, 51(4):661–703, 2009.

[15] 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.

[16] N. Crokidakis. Role of noise and agents' convictions on opinion spreading in a three-state voter-like model. *J. Stat. Mech. Theory Exp.*, 07:P07008/1–14, 2013.

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

[18] 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.

[19] K. Freberg, K. Graham, K. McGaughey, and L. A. Freberg. Who are the social media influencers? A study of public perceptions. *Public Relat. Rev.*, 37(1):90–92, 2011.

[20] S. Gualandi and G. Toscani. Pareto tails in socio-economic phenomena: a kinetic description. *Economics*, 2017(111):1–18, 2017.

[21] A. M. Kaplan and M. Haenlein. Users of the world, unite! The challenges and opportunities of social media. *Bus. Horizons*, 53(1):59–68, 2010.

[22] A. Lipowski, D. Lipowska, and A. L. Ferreira. Agreement dynamics on directed random graphs. *J. Stat. Mech. Theory Exp.*, 2017:063408/1–13, 2017.

[23] Y.-Y. Liu and A.-L. Barabási. Control principles of complex networks. *Rev. Mod. Phys.*, 88(3):035006, 2016.

[24] G. W. Mangold and D. J. Faulds. Social media: The new hybrid element of the promotion mix. *Bus. Horizons*, 52(4):357–365, 2009.

[25] M. Mobilia, A. Petersen, and S. Redner. On the role of zealotry in the voter model. *J. Stat. Mech. Theory Exp.*, 2007:P08029/1–17, 2007.

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

[27] M. E. J. Newman, D. J. Watts, and S. H. Strogatz. Random graph models of social networks. *Proc. Natl. Acad. Sci. USA*, 99(suppl 1):2566–2572, 2002.

[28] R. K. Pan and S. Sinha. The statistical laws of popularity: universal properties of the box-office dynamics of motion pictures. *New J. Phys.*, 12:115004/1–23, 2010.

[29] L. Pareschi and G. Russo. An introduction to Monte Carlo method for the Boltzmann equation. *ESAIM: Proc.*, 10:35–75, 2001.

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

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

[32] S. Redner. How popular is your paper? An empirical study of the citation distribution. *Eur. Phys. J. B*, 4:131–134, 1998.

[33] S. Sinha and R. K. Pan. Blockbusters, bombs and sleepers: the income distribution of movies. In A. Chatterjee, S. Yarlagadda, and B. K. Chakrabarti, editors, *Econophysics of Wealth Distributions*, New Economic Windows, pages 43–47. Springer, 2005.

[34] S. Sinha and S. Raghavendra. Hollywood blockbusters and long-tailed distributions. *Eur. Phys. J. B*, 42(2):293–296, 2004.

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

[36] A. Tosin and M. Zanella. Boltzmann-type models with uncertain binary interactions. *Commun. Math. Sci.*, 2018. To appear (preprint: arXiv:1709.02353).

[37] 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.

[38] G. Visconti, M. Herty, G. Puppo, and A. Tosin. Multivalued fundamental diagrams of traffic flow in the kinetic Fokker-Planck limit. *Multiscale Model. Simul.*, 15(3):1267–1293, 2017.

[39] A. Waagen, G. Verma, K. Chan, A. Swami, and R. D'Souza. Effect of zealotry in high-dimensional opinion dynamics models. *Phys. Rev. E*, 91(2):022811/1–10, 2014.

[40] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.

[41] J. Xie, S. Sreenivasan, G. Korniss, W. Zhang, Lim C., and B. K. Szymanski. Social consensus through the influence of committed minorities. *Phys. Rev. E*, 84(1):011130/1–8, 2011.