Consensus in the Hegselmann–Krause Model

Nicolas Lanchier[1] · Hsin-Lun Li[1]

Abstract
This paper is concerned with the probability of consensus in a multivariate, socially structured version of the Hegselmann–Krause model for the dynamics of opinions. Individuals are located on the vertices of a finite connected graph representing a social network, and are characterized by their opinion, with the set of opinions Δ being a general bounded convex subset of a finite dimensional normed vector space. Having a confidence threshold τ, two individuals are said to be compatible if the distance (induced by the norm) between their opinions does not exceed the threshold τ. Each vertex x updates its opinion at rate the number of its compatible neighbors on the social network, which results in the opinion at x to be replaced by a convex combination of the opinion at x and the nearby opinions: α times the opinion at x plus $(1-\alpha)$ times the average opinion of its compatible neighbors. The main objective is to derive a lower bound for the probability of consensus when the opinions are initially independent and identically distributed with values in the opinion set Δ.

Keywords Interacting particle systems · Hegselmann–Krause model · Opinion dynamics · Martingale · Martingale convergence theorem · Optional stopping theorem · Confidence threshold · Consensus

AMS 2000 Subject Classifications: Primary 60K35

1 Introduction

As illustrated by the popular Axelrod model [1], opinion and cultural dynamics are driven by two key components: social influence, the tendency of individuals to become more similar when they interact, and homophily, the tendency to interact more frequently with individuals

✉ Nicolas Lanchier
   nicolas.lanchier@asu.edu

   Hsin-Lun Li
   hsinlunl@asu.edu

[1] School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA

who are more similar [4]. Another important component in the context of social dynamics is the topology of the social network whose connections indicate who interacts with whom, an aspect that is well captured by stochastic models known as interacting particle systems [27, 29].

The voter model The simplest example of interacting particle system modeling the dynamics of opinions is the voter model introduced independently in [5, 15]. Individuals located on the vertices of a connected graph (interpreted as a social network) are characterized by one of two competing opinions, and update their opinion independently at rate one by mimicking one of their neighbors chosen uniformly at random. In particular, this model accounts for the topology of the social network as well as social influence, but not homophily. Most of the results about the voter model rely on a duality relationship with a system of coalescing random walks. When the individuals are located on the d-dimensional integer lattice, the system clusters, and so a local consensus is reached, in dimensions $d \leq 2$, whereas the process converges to a nontrivial stationary distribution, and so the two opinions coexist at equilibrium, in higher dimensions [15]. The asymptotic behavior of the cluster size for the one- and the two-dimensional voter models is studied in [3, 7]. In higher dimensions, even though the two opinions coexist at equilibrium, as time evolves, spatial correlations build up due to the presence of local interactions. The spatial correlations in the infinite time limit are studied in [2, 31]. On finite connected graphs, the voter model always reaches a consensus, and the time to consensus on the large d-dimensional torus is studied in [6].

Threshold voter models Natural variants of the voter model that also describe the dynamics of opinions on social networks are the threshold voter models. In this case, individuals switch their opinion at rate one if and only if they disagree with at least θ of their neighbors. Looking again at the process on the d-dimensional integer lattice, when the threshold is equal to one, coexistence occurs except in one dimension [28]. While increasing the threshold θ, the process exhibits two phase transitions, from coexistence to clustering, and then from clustering to fixation where both opinions again coexist, but coexistence is now due to the system becoming frozen locally as opposed to convergence to a unique nontrivial stationary distribution. Fixation occurs in particular when the individuals switch their opinion if and only if they disagree with a strict majority of their neighbors [9], a particular case referred to as the majority vote model.

The majority rule model Another process related to the majority vote model that includes the simultaneous updates of multiple individuals is the majority rule model [10]. In the original version of this model, the population is finite, individuals are again characterized by one of two competing opinions, and a single update consists in choosing a random number of individuals uniformly (referred to as a discussion group), which results in all the individuals in the group adopting the majority opinion within the group before the interaction. In case of a tie, all the individuals adopt a preferred opinion representing the status quo. A spatial version of the majority rule model where individuals are located on the integer lattice is introduced in [19]. The discussion groups consist of all the $n \times \cdots \times n$ hypercubes and are chosen to be updated independently at rate one. In this case, the behavior depends on the parity of n: the system clusters with the density of each opinion being preserved when n is odd whereas the status quo wins when n is even.

The Axelrod model The classical and threshold voter models and the majority rule model account for social influence but not homophily, the tendency to interact more frequently with individuals who are more similar. The most popular interacting particle system that includes both social influence and homophily is the Axelrod model [1]. Individuals are now characterized by one of q possible opinions about F different issues, which results in q^F possible vectors of opinions representing cultural profiles. Homophily is modeled by

assuming that neighbors interact at a rate equal to the fraction of opinions or cultural features they have in common while social influence is modeled by matching one of the opinions the two neighbors do not share, if any. Depending on the values of F and q, the system either clusters so that a local consensus is reached or fixates in a fragmented configuration where disagreements persist. Intuitively, increasing the value of F promotes consensus while increasing the value of q promotes discordance. For the one-dimensional model where the individuals are located on the integers, consensus is studied in [16, 23] while discordance is studied in [20]. See also [26] for a study of the two-dimensional process.

Constrained voter models Homophily in the Axelrod model is incorporated by assuming that the rate at which individuals interact is proportional to the amount of agreement between neighbors. In many other models, this component is modeled instead through the inclusion of a so-called confidence threshold: neighbors either interact at a constant rate or do not interact at all depending on whether their level of disagreement is within a reasonable threshold or not. The simplest such model is the constrained voter model [30] where individuals are either leftist, centrist or rightist, and mimic a randomly chosen neighbor unless one of the individuals is leftist and the other one rightist. This model has been generalized in [22] where the set of opinions is represented by the vertices of a general finite connected graph and where the level of disagreement is measured using the geodesic distance on the opinion graph: neighbors interact at rate one, which results in one neighbor mimicking the other neighbor if and only if the geodesic distance (or graph distance) between their opinions does not exceed some confidence threshold τ. It is proved in [22] that the one-dimensional process clusters when the confidence threshold exceeds the radius of the opinion graph while sufficient conditions for fixation in a fragmented configuration are also given. A lower bound for the probability of consensus for the process where individuals are located on a finite connected graph is derived in [12]. The constrained voter model [30] corresponds to the particular case where the opinion graph consists of the path with three vertices and two edges and where $\tau = 1$ while the particular case where the opinion graph consists of the hypercube with 2^n vertices is referred to as the vectorial Deffuant model introduced in [8] and studied in [21].

The Deffuant model In another (more popular) version of the Deffuant model [8], the opinion space is uncountable and consists of the unit interval. Edges become active independently at rate one and the two individuals connected by that edge interact if and only if the distance between their opinions does not exceed some confidence threshold τ, which results in the two neighbors' opinions getting closer to each other. If the opinions are initially independent and uniform, it is conjectured that the system on infinite graphs converges to a consensus when $\tau > 1/2$ whereas disagreements persist when $\tau < 1/2$. This conjecture was first established in [17] for the one-dimensional process while [11] gave an alternative simpler proof. Multivariate versions of the model where the opinion space consists of a bounded convex subset of a normed vector space are also studied in [14] for the one-dimensional system and in [18] for the process on general finite connected graphs.

The Hegselmann–Krause model The Hegselmann–Krause (HK) model [13] is a general interacting particle system. The opinion space again consists of the unit interval but, in contrast with the Deffuant model which evolves via pairwise interactions, the new opinion at a vertex x is now replaced by a weighted average of all the opinions in the population that are within some confidence threshold τ of the opinion at x. General conditions under which certain initial configurations and certain sets of parameters can lead to asymptotic stability and/or consensus are discussed in [24, 25]. In the version of the HK-model considered in [4], the matrix of the weights between individuals is chosen to be the adjacency matrix of a graph (the weight is one if the two individuals are connected by an edge and zero otherwise) thus inducing an opinion dynamics on a social network: Individuals are located on the vertices

of a graph representing a social network, and each update at vertex x results in the new opinion at x to be the average of the opinions of the neighbors that are compatible with x (at opinion distance less than τ). The objective of this paper is to derive a lower bound for the probability of consensus for a multivariate version of this model.

2 Model Description and Results

To define the model rigorously, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a finite connected graph representing a social network: each vertex is occupied by an individual and any two individuals are connected by an edge if and only if they are *socially* connected or acquainted. Individuals are characterized by their opinion, and we assume that the opinions are initially independent and distributed across a bounded convex subset $\Delta \subset \mathbb{R}^n$, thus extending the opinion space $[0, 1]$ of the original HK-model. To measure the disagreements between two neighbors, meaning from now on two individuals who are acquainted, and define the dynamics of the process, we also assume that $\mathbb{R}^n$ is equipped with an arbitrary norm $\| \cdot \|$. The state of the process at time t is a configuration of opinions

$$\xi_t : \mathcal{V} \to \Delta \quad \text{where} \quad \xi_t(x) = \text{opinion at vertex } x \text{ at time } t.$$

Having a confidence threshold $\tau > 0$, two neighbors are said to be compatible at time t if the distance (induced by the norm) between their opinions does not exceed this threshold, and we denote by $N_t(x)$ the set of neighbors of x that are compatible with x, i.e.,

$$N_t(x) = \{y \in \mathcal{V} : (x, y) \in \mathcal{E} \text{ and } \|\xi_t(x) - \xi_t(x)\| \leq \tau\}. \tag{1}$$

The process we study evolves as follows: Individuals update their opinion independently at a rate equal to the number of their compatible neighbors, the new opinion being equal to the average opinion of their compatible neighbors. More precisely, letting $\mathcal{F}_t$ be the natural filtration of the process, we assume that

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} P\left(\xi_{t+\epsilon}(x) = \frac{1}{|N_t(x)|} \sum_{y \in N_t(x)} \xi_t(y) \,\Big|\, \mathcal{F}_t \right) = |N_t(x)| \tag{2}$$

when $N_t(x) \neq \varnothing$ whereas the transition rate is zero when x has no compatible neighbors. These transitions indicate that the individuals are open-minded in the sense that the new opinion only depends on the opinions of the compatible neighbors. Following [24, 25], we can include more realistically a degree of stubbornness by assuming that the new opinion is given by

$$\alpha \, \xi_t(x) + \frac{1 - \alpha}{|N_t(x)|} \sum_{y \in N_t(x)} \xi_t(y) \quad \text{where} \quad 0 \leq \alpha < 1.$$

The parameter α represents the degree of stubbornness of the individuals. To simplify the algebra, we assume from now on that $\alpha = 0$, which corresponds to model (2), but we point out that all our proofs and our main result easily extend to the model with stubbornness. Indeed, when $0 < \alpha < 1$, each update results in an individual's opinion moving closer to the average opinion of its compatible neighbors, and only this aspect of the dynamics is used throughout our proofs.

The main objective is to study the probability of consensus, i.e., the probability that all the opinions converge to the same limit, when starting from a configuration where the opinions are independent, identically distributed with values in the opinion space Δ. Because

the opinion space is uncountable, it is also natural to assume that the initial distribution is continuous. Under this assumption, replacing the large inequality in (1) by a strict inequality, i.e., assuming that neighbors at exactly opinion distance τ from each other are incompatible, does not change the behavior of the process. Indeed, for any given opinion $c \in \Delta$, because the initial opinions are independent continuous random variables, the probability that an individual has initially or at any time exactly opinion c is equal to zero. It follows that the probability of two neighbors being exactly at opinion distance τ (or any given positive real number) of each other is equal to zero so our model and the model with a strict inequality are equal with probability one. There is however a positive probability that two neighbors have the exact same opinion (opinion distance zero), which occurs when there is an update at a vertex that has exactly one compatible neighbor. To state our result, let

$$\mathscr{C} = \left\{ \lim_{t \to \infty} \max_{x,y \in \mathscr{V}} \| \xi_t(x) - \xi_t(y) \| = 0 \right\}$$

be the consensus event. We also define

$$\mathbf{r} = \inf\{r > 0 : \Delta \subset B(c, r) \text{ for some } c \in \Delta\}$$
$$= \inf\{r > 0 : \sup_{a \in \Delta} \|a - c\| \leq r \text{ for some } c \in \Delta\},$$

which can be viewed as the radius of the opinion space, and

$$\mathbf{c} \in \Delta \quad \text{such that} \quad \Delta \subset B(\mathbf{c}, \mathbf{r}),$$

which can be viewed as the center of the opinion space. Throughout this paper, we also assume that the opinions across the social network are initially independent and identically distributed, and we let X be the Δ-valued random variable describing the initial opinions:

$$P(\xi_0(x) \in B) = P(X \in B) \quad \text{for all} \quad x \in \mathscr{V} \text{ and Borel set } B \subset \Delta.$$

Then, we have the following universal lower bound that holds uniformly over all possible finite connected graphs (social networks).

Theorem 1 (probability of consensus) *For all $\tau > \mathbf{r}$, $P(\mathscr{C}) \geq 1 - E\|X - \mathbf{c}\|/(\tau - \mathbf{r})$.*

This result is identical to [18,Theorem 2.1] which applies to the Deffuant model. However, because the HK dynamics is significantly different from the Deffuant dynamics, our proof is quite different from the proof in [18].

To begin with, we prove that, for all $c \in \mathbb{R}^n$, the process $X_t(c)$ that keeps track of the total disagreement between a virtual external agent with opinion c and the vertices of the graph is a supermartingale (Lemma 1). Looking at 2^n of these processes for values of c that surround the opinion space and applying the martingale convergence theorem, we deduce that the opinion model converges almost surely to a random configuration ξ_∞ (Lemma 2). The evolution rules of the HK-model imply that, in the configuration ξ_∞, any two neighbors on the social network either completely agree or disagree by more than the confidence threshold τ (Lemma 3). Using this characterization, we deduce that, for all $\epsilon > 0$, the first time T_ϵ at which any two neighbors disagree by at most ϵ or at least τ is almost surely finite (Lemma 4). Note that the assumption $\tau > \mathbf{r}$ means that, regardless of their opinion, all the individuals are compatible with the center of the opinion space. In particular, a stubborn individual with that opinion would lead the system to a consensus. Using this idea, we prove that the event $\mathscr{A}$ that the opinion at a vertex is at distance less than $\tau - \mathbf{r}$ from the center $\mathbf{c}$ of the opinion space at time T_ϵ for some ϵ small always drives the process to a global consensus (Lemma 5). Applying the optional stopping theorem to the bounded supermartingale $X_t(\mathbf{c})$ stopped at time T_ϵ gives a

lower bound for the probability of $\mathscr{A}$ (Lemma 6) which, together with Lemma 5, implies the theorem.

Numerical example Before going into the details of the proof, we give a numerical example to show how our result applies to the original HK-model. Assume that the opinion space is

$$\Delta = B(a, r) = \{c \in \mathbb{R}^n : \|a - c\| < r\},$$

the ball with center a and radius r, and that the opinions are initially independent and uniformly distributed: $X = \text{Uniform}(\Delta)$. Then, $\mathbf{c} = a$ and $\mathbf{r} = r$. In addition, for all $s \leq r$,

$$P(\|X - a\| \leq s) = P(X \in B(a, s)) = \frac{\lambda(B(0, s))}{\lambda(B(0, r))} = \frac{s^n \, \lambda(B(0, 1))}{r^n \, \lambda(B(0, 1))} = \left(\frac{s}{r}\right)^n$$

where $\lambda(B)$ refers to the Lebesgue measure of set $B \subset \mathbb{R}^n$. It follows that

$$E \, \|X - a\| = \int_0^\infty P(\|X - a\| > s) \, ds = \int_0^r \left(1 - \left(\frac{s}{r}\right)^n\right) ds = \frac{nr}{n + 1}.$$

Therefore, the theorem implies that

$$P(\mathscr{C}) \geq 1 - \frac{E \, \|X - \mathbf{c}\|}{\tau - \mathbf{r}} = 1 - \frac{E \, \|X - a\|}{\tau - r} = 1 - \frac{nr}{(n + 1)(\tau - r)} \quad \text{when} \quad \Delta = B(a, r).$$

In particular, for the original HK-model ($n = 1$ and $a = r = 1/2$),

$$P(\mathscr{C}) \geq \frac{4\tau - 3}{4\tau - 2} > 0 \quad \text{for all} \quad \tau > 3/4 \quad \text{when} \quad \Delta = [0, 1].$$

3 Underlying Supermartingales

In this section, we prove that the processes that keep track of the total disagreement between a fixed point c and the model's opinions, i.e.,

$$X_t(c) = \sum_{x \in \mathscr{V}} \|\xi_t(x) - c\| \quad \text{for all} \quad c \in \mathbb{R}^n,$$

are supermartingales. Applying the martingale convergence theorem to these supermartingales will show almost sure convergence of the HK-model to a limiting configuration ξ_∞. This result will also be used later in combination with the optional stopping theorem to derive the lower bound for the probability of consensus in the theorem.

Lemma 1 *The processes $X_t(c)$ are supermartingales for the filtration $\mathscr{F}_t$.*

Proof Let $\mathscr{V}_t$ be the set of vertices with at least one compatible neighbor, i.e., at least one neighbor at opinion distance at most τ of the vertex. To simplify the notation, we let

$$\bar{\xi}_t(x) = \frac{1}{|N_t(x)|} \sum_{y \in N_t(x)} \xi_t(y) \quad \text{for all} \quad x \in \mathscr{V}_t.$$

By the absolute homogeneity of the norm and the triangle inequality, for all $x \in \mathscr{V}_t$,

$$|N_t(x)| \cdot \|\bar{\xi}_t(x) - c\| = |N_t(x)| \cdot \left\| \frac{1}{|N_t(x)|} \sum_{y \in N_t(x)} \xi_t(y) - c \right\|$$

$$= \left\| \sum_{y \in N_t(x)} \xi_t(y) - |N_t(x)| \cdot c \right\| = \left\| \sum_{y \in N_t(x)} (\xi_t(y) - c) \right\| \leq \sum_{y \in N_t(x)} \|\xi_t(y) - c\|.$$

In particular, recalling the transition rates of the HK-model, we get

$$\lim_{\epsilon \to 0} E\left(\frac{X_{t+\epsilon}(c) - X_t(c)}{\epsilon} \,\Big|\, \mathscr{F}_t\right) = \sum_{x \in \mathscr{V}_t} |N_t(x)| \cdot \left(\|\bar{\xi}_t(x) - c\| - \|\xi_t(x) - c\|\right)$$

$$= \sum_{x \in \mathscr{V}_t} \left(\left(\sum_{y \in N_t(x)} \|\xi_t(y) - c\|\right) - |N_t(x)| \cdot \|\xi_t(x) - c\|\right).$$

To conclude, observe that

$$|\{x \in \mathscr{V}_t : y \in N_t(x)\}| = |\{x : (x, y) \in \mathscr{E} \text{ and } \|\xi_t(x) - \xi_t(y)\| \le \tau\}|$$
$$= |\{x : (y, x) \in \mathscr{E} \text{ and } \|\xi_t(y) - \xi_t(x)\| \le \tau\}| = |N_t(y)|.$$

In other words, the number of vertices x such that y is in the neighborhood of x and compatible with x at time t is equal to the number of vertices that are in the neighborhood of y and compatible with y at that time. It follows that

$$\lim_{\epsilon \to 0} E\left(\frac{X_{t+\epsilon}(c) - X_t(c)}{\epsilon} \,\Big|\, \mathscr{F}_t\right) \le \sum_{x \in \mathscr{V}_t} \left(\sum_{y \in N_t(x)} \|\xi_t(y) - c\|\right)$$

$$- \sum_{x \in \mathscr{V}_t} |N_t(x)| \cdot \|\xi_t(x) - c\|$$

$$= \sum_{y \in \mathscr{V}_t} |N_t(y)| \cdot \|\xi_t(y) - c\|$$

$$- \sum_{x \in \mathscr{V}_t} |N_t(x)| \cdot \|\xi_t(x) - c\| = 0,$$

which proves that $X_t(c)$ is a supermartingale. $\qquad\qquad\square$

Using Lemma 1, we now prove almost sure convergence of the HK-model.

Lemma 2 *There is a random configuration $\xi_\infty : \mathscr{V} \to \Delta$ such that*

$$\lim_{t \to \infty} \xi_t(x) = \xi_\infty(x) \quad \text{almost surely for all } x \in \mathscr{V}.$$

Proof Let $r > 0$ be sufficiently large that $\Delta \subset \Lambda = [-r, r]^n$ (see Figure 1). Assuming that the configuration ξ_t does not converge, there exists $\epsilon > 0$ such that

$$\max_{x \in \mathscr{V}} \|\xi_t(x) - \xi_{t-}(x)\|_1 = \max_{x \in \mathscr{V}} \sum_{i=1}^n |\langle \xi_t(x) - \xi_{t-}(x), e_i\rangle| > \epsilon \quad \text{at arbitrary large times}$$

(3)

where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ denotes the ith unit vector in $\mathbb{R}^n$ and where $\xi_{t-}(x)$ refers to the opinion at vertex x just before time t. By the choice of r and by the equivalence of the norms in finite dimension, there exists $m_1 > 0$ such that, for at least one of the corners of the n-dimensional cube Λ, say corner c,

$$|X_t(c) - X_{t-}(c)| \ge m_1 \max_{x \in \mathscr{V}} |\|\xi_t(x) - c\|_1 - \|\xi_{t-}(x) - c\|_1|$$
$$= m_1 \max_{x \in \mathscr{V}} \|\xi_t(x) - \xi_{t-}(x)\|_1 > \epsilon m_1 > 0.$$

In particular, for at least one corner c,

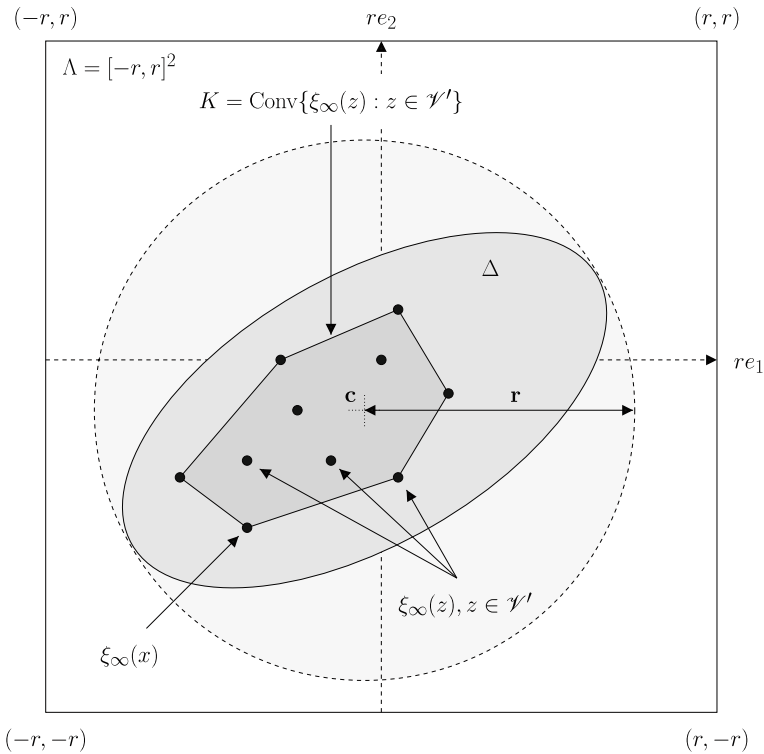$$|X_t(c) - X_{t-}(c)| > \epsilon m_1 \quad \text{at arbitrary large times,} \qquad (4)$$

Fig. 1 Picture related to the proof of Lemmas 2 and 3

showing that $X_t(c)$ does not converge. Now, using again the equivalence of the norms, there exists $m_2 < \infty$ such that the process is bounded by

$$|X_t(c)| \leq \max_{a \in \Delta} \|a - c\| \operatorname{card}(\mathcal{V}) \leq m_2 \max_{a,b \in \Lambda} \|a - b\|_1 \operatorname{card}(\mathcal{V}) = 2nrm_2 \operatorname{card}(\mathcal{V}) < \infty.$$

Because the process $X_t(c)$ is also a supermartingale according to Lemma 1, it follows from the martingale convergence theorem that it converges almost surely. In particular, the events in (3) and (4) and the divergence of ξ_t occur with probability zero, which proves the lemma. $\qquad \square$

To study the limiting configuration ξ_∞, we now let $\mathcal{G}_\infty = (\mathcal{V}, \mathcal{E}_\infty)$ be the random subgraph of the social network $\mathcal{G}$ induced by the edges that are compatible in the limit:

$$(x, y) \in \mathcal{E}_\infty \quad \text{if and only if} \quad (x, y) \in \mathcal{E} \quad \text{and} \quad \|\xi_\infty(x) - \xi_\infty(y)\| \leq \tau. \tag{5}$$

The next lemma shows that, in the limiting configuration, individuals on the same connected component of the graph $\mathcal{G}_\infty$ share the same opinion.

Lemma 3 *Configuration ξ_∞ is constant on each of the connected components of $\mathcal{G}_\infty$.*

Proof Let $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ be a connected component of the graph $\mathcal{G}_\infty$, and observe that, because ξ_∞ is an absorbing state for the HK-model, it follows from the evolution rules that

$$N_\infty(x) = \lim_{t \to \infty} N_t(x) = \varnothing \quad \text{or} \quad \xi_\infty(x) = \frac{1}{\operatorname{card}(N_\infty(x))} \sum_{y \in N_\infty(x)} \xi_\infty(y)$$

for all $x \in \mathcal{V}'$. When $\mathcal{V}'$ reduces to one vertex, ξ_∞ is trivially constant on $\mathcal{V}'$. Otherwise, because the subgraph $\mathcal{G}'$ is connected and all its edges are compatible in the limit, we have

$$N_\infty(x) \neq \varnothing \quad \text{so} \quad \xi_\infty(x) = \frac{1}{\text{card}(N_\infty(x))} \sum_{y \in N_\infty(x)} \xi_\infty(y) \quad \text{for all} \quad x \in \mathcal{V}'. \tag{6}$$

In this case, let K be the convex envelope of the limiting opinions in $\mathcal{V}'$,

$$K = \text{Conv}\{\xi_\infty(z) : z \in \mathcal{V}'\} \subset \Delta,$$

and let $\xi_\infty(x) \in K$, $x \in \mathcal{V}'$, be an extreme point of K, i.e., a point which is not an inner point of segments of the convex envelope (see Fig. 1). Now, assume by contradiction that

$$\xi_\infty(y) \neq \xi_\infty(x) \quad \text{for some} \quad y \in N_\infty(x). \tag{7}$$

Then, it follows from (6) that

$$\xi_\infty(x) \in \text{Int}(K') \quad \text{where} \quad K' = \text{Conv}\{\xi_\infty(y) : y \in N_\infty(x)\}. \tag{8}$$

Here, $\text{Int}(K')$ refers to the interior of the set K'. Because

$$N_\infty(x) \subset \mathcal{V}' \quad \text{and so} \quad K' \subset K,$$

we deduce from (8) that $\xi_\infty(x) \in \text{Int}(K)$, contradicting the fact that $\xi_\infty(x)$ is an extreme point of the convex set K. This shows that the statement in (7) is false therefore all the neighbors of x on the connected component must share the same opinion as x in the limit. This also implies that the neighbors' opinions are (identical) extreme points of K so, by the same reasoning, the neighbors of the neighbors of x must share the same opinion as x, and using a simple induction, we deduce that the convex envelope reduces to one point: $K = \{\xi_\infty(x)\}$, which proves the lemma. □

Note that the proof of the lemma can be simplified when dealing with the original HK-model, where $\Lambda = (0, 1)$ and $\|\cdot\|$ refers to the Euclidean norm, using that, according to (6), the restriction of the limiting configuration to each of the connected components of $\mathcal{G}_\infty$ is harmonic. Indeed, in this case, let $x \in \mathcal{V}'$ be a vertex where ξ_∞ reaches its maximum:

$$\xi_\infty(x) = \max\{\xi_\infty(z) : z \in \mathcal{V}'\}.$$

Then, because $\xi_\infty(x) \geq \xi_\infty(y)$ for all $y \in N_\infty(x) \subset \mathcal{V}'$ and

$$\frac{1}{\text{card}(N_\infty(x))} \sum_{y \in N_\infty(x)} (\xi_\infty(x) - \xi_\infty(y)) = \xi_\infty(x) - \frac{1}{\text{card}(N_\infty(x))} \sum_{y \in N_\infty(x)} \xi_\infty(y) = 0,$$

we must have $\xi_\infty(x) = \xi_\infty(y)$ for all $y \in N_\infty(x)$.

4 Optional Stopping and Consensus

As previously mentioned, the next step to prove the theorem is to apply the optional stopping theorem to the supermartingale $X_t(\mathbf{c})$. In order to apply the optional stopping theorem, we also define the collection of stopping times

$$T_\epsilon = \inf \{t : \|\xi_t(x) - \xi_t(y)\| \notin [\epsilon, \tau] \text{ for all } (x, y) \in \mathcal{E}\} \quad \text{for all} \quad \epsilon \in (0, \tau).$$

The next lemma shows that these stopping times are almost surely finite, which is one of the required assumptions to apply the optional stopping theorem.

Lemma 4 *For all $\epsilon \in (0, \tau)$, time T_ϵ is almost surely finite.*

Proof On the event that $T_\epsilon = \infty$, there exists $(x, y) \in \mathscr{E}$ such that

$$\|\xi_t(x) - \xi_t(y)\| \in [\epsilon, \tau] \quad \text{at arbitrary large times.} \tag{9}$$

This, together with Lemma 2, implies that

$$\|\xi_\infty(x) - \xi_\infty(y)\| = \lim_{t \to \infty} \|\xi_t(x) - \xi_t(y)\| \in [\epsilon, \tau]. \tag{10}$$

Observing also that, by the definition of $\mathscr{E}_\infty$ in (5),

$$(x, y) \in \mathscr{E} \text{ and } \|\xi_\infty(x) - \xi_\infty(y)\| \leq \tau \implies (x, y) \in \mathscr{E}_\infty$$
$$\|\xi_\infty(x) - \xi_\infty(y)\| \geq \epsilon \implies \xi_\infty(x) \neq \xi_\infty(y),$$

we deduce that the limiting configuration ξ_∞ is not constant on the connected components of the random graph $\mathscr{G}_\infty$ which, according to Lemma 3, is an event with probability zero. In conclusion, the events in (9) and (10) and the event that $T_\epsilon = \infty$ all have probability zero. □

Next, we prove that if, at time T_ϵ for ϵ small, the opinion of at least one vertex is close enough to the center **c** of the opinion space, an event that we call $\mathscr{A}$, then all the opinions across the network should be close to that opinion. This, together with a convexity argument and Lemma 3, implies consensus. In particular, the probability of consensus is larger than the probability of $\mathscr{A}$, which will be used later to deduce the theorem.

Lemma 5 *For all $\epsilon' \in (0, \tau/2)$ and $\tau > \mathbf{r} + \epsilon'$, there is $\epsilon > 0$ such that*

$$\mathscr{A} = \mathscr{A}(\epsilon, \epsilon') = \left\{ \|\xi_{T_\epsilon}(x) - \mathbf{c}\| < \tau - \mathbf{r} - \epsilon' \text{ for some } x \in \mathscr{V} \right\} \subset \mathscr{C}.$$

Proof First, we observe that, on the event $\mathscr{A}$, for all $y \in \mathscr{V}$,

$$\|\xi_{T_\epsilon}(x) - \xi_{T_\epsilon}(y)\| \leq \|\xi_{T_\epsilon}(x) - \mathbf{c}\| + \|\xi_{T_\epsilon}(y) - \mathbf{c}\| < \tau - \mathbf{r} - \epsilon' + \mathbf{r} = \tau - \epsilon'. \tag{11}$$

Now, we prove by induction that, on the event $\mathscr{A}$,

$$\|\xi_{T_\epsilon}(x) - \xi_{T_\epsilon}(y)\| \leq d(x, y)\epsilon \quad \text{for all} \quad y \in \mathscr{V} \quad \text{when} \quad \epsilon = \epsilon'/\text{card}(\mathscr{V}). \tag{12}$$

Here, $d(x, y)$ refers to the graph distance between vertex x and vertex y.
Base case Assume that $d(x, y) = 0$. Then,

$$y = x \quad \text{and} \quad \|\xi_{T_\epsilon}(x) - \xi_{T_\epsilon}(y)\| = 0 = 0 \times \epsilon.$$

Induction step: Let $y \neq x$ and $d = d(x, y)$ and assume that (12) holds for all vertices at graph distance $d - 1$ from vertex x. Letting z be a neighbor of y at distance $d - 1$ from x,

$$\|\xi_{T_\epsilon}(x) - \xi_{T_\epsilon}(z)\| \leq (d - 1)\epsilon. \tag{13}$$

Because the longest self-avoiding path on the social network contains less than $\text{card}(\mathscr{V})$ edges, we must have $d < \text{card}(\mathscr{V})$. Assuming by contradiction that $\|\xi_{T_\epsilon}(y) - \xi_{T_\epsilon}(z)\| > \epsilon$. By definition of the stopping time T_ϵ, the two neighbors y and z must be incompatible at that time:

$$\|\xi_{T_\epsilon}(y) - \xi_{T_\epsilon}(z)\| > \tau. \tag{14}$$

Using the triangle inequality, then (11) and (14), we deduce that

$$\|\xi_{T_\epsilon}(x) - \xi_{T_\epsilon}(z)\| \geq \|\xi_{T_\epsilon}(z) - \xi_{T_\epsilon}(y)\| - \|\xi_{T_\epsilon}(y) - \xi_{T_\epsilon}(x)\|$$
$$\geq \tau - (\tau - \epsilon') = \epsilon' = \epsilon \, \text{card}(\mathscr{V}) > (d - 1)\epsilon,$$

which contradicts (13), therefore $\|\xi_{T_\epsilon}(y) - \xi_{T_\epsilon}(z)\| \leq \epsilon$ and

$$\|\xi_{T_\epsilon}(x) - \xi_{T_\epsilon}(y)\| \leq \|\xi_{T_\epsilon}(x) - \xi_{T_\epsilon}(z)\| + \|\xi_{T_\epsilon}(y) - \xi_{T_\epsilon}(z)\| \leq (d-1)\epsilon + \epsilon = d\epsilon.$$

This completes the proof of (12).

Recalling that $d(x, y) < \mathrm{card}(\mathcal{V})$ for all $y \in \mathcal{V}$, we deduce that, on the event $\mathscr{A}$,

$$\|\xi_{T_\epsilon}(x) - \xi_{T_\epsilon}(y)\| < \epsilon\,\mathrm{card}(\mathcal{V}) = \epsilon' \quad \text{and so} \quad \xi_{T_\epsilon}(y) \in B(\xi_{T_\epsilon}(x), \epsilon') \quad \text{for all} \quad y \in \mathcal{V}.$$

Since at each update of the process the new opinion is contained in the convex envelope of the other opinions, we deduce that, on the event $\mathscr{A}$,

$$\xi_\infty(y) \in \mathrm{Conv}\{\xi_{T_\epsilon}(z) : z \in \mathcal{V}\} \subset B(\xi_{T_\epsilon}(x), \epsilon') \quad \text{and so} \quad \|\xi_\infty(y) - \xi_\infty(z)\| < 2\epsilon' < \tau$$

for all $y, z \in \mathcal{V}$. In conclusion, $\mathscr{G}_\infty$ is connected and $\mathscr{C}$ occurs by Lemma 3. $\qquad\square$

The last step to prove the theorem is to apply the optional stopping theorem to the supermartingale $X_t(\mathbf{c})$ stopped at time T_ϵ to get a lower bound for the probability of the sub-event $\mathscr{A}$.

Lemma 6 *The event* $\mathscr{A} = \mathscr{A}(\epsilon, \epsilon')$ *defined in Lemma 5 has probability*

$$P(\mathscr{A}) \geq 1 - E\,\|X - \mathbf{c}\|/(\tau - \mathbf{r} - \epsilon').$$

Proof Observing that, on the complement of $\mathscr{A}$,

$$\|\xi_{T_\epsilon}(x) - \mathbf{c}\| \geq \tau - \mathbf{r} - \epsilon' \quad \text{for all} \quad x \in \mathcal{V},$$

we deduce that the conditional expectation of $X_{T_\epsilon}(\mathbf{c})$ is bounded from below by

$$E\left(X_{T_\epsilon}(\mathbf{c}) \,|\, \mathscr{A}^c\right) = E\left(\left.\sum_{x \in \mathcal{V}} \|\xi_{T_\epsilon}(x) - \mathbf{c}\| \,\right|\, \mathscr{A}^c\right) \geq (\tau - \mathbf{r} - \epsilon') \cdot \mathrm{card}(\mathcal{V}).$$

In particular, conditioning on the partition $\{\mathscr{A}, \mathscr{A}^c\}$, we get

$$\begin{aligned} E(X_{T_\epsilon}(\mathbf{c})) &= E(X_{T_\epsilon}(\mathbf{c}) \,|\, \mathscr{A})P(\mathscr{A}) + E(X_{T_\epsilon}(\mathbf{c}) \,|\, \mathscr{A}^c)P(\mathscr{A}^c) \\ &\geq E(X_{T_\epsilon}(\mathbf{c}) \,|\, \mathscr{A}^c)P(\mathscr{A}^c) \geq (\tau - \mathbf{r} - \epsilon')\,\mathrm{card}(\mathcal{V})P(\mathscr{A}^c). \end{aligned} \tag{15}$$

Because the process $X_t(\mathbf{c})$ is a (bounded) supermartingale according to Lemma 1 and because T_ϵ is an almost surely finite stopping time according to Lemma 4, we may apply the optional stopping theorem. This, together with the previous inequality (15), gives

$$P(\mathscr{A}) = 1 - P(\mathscr{A}^c) \geq 1 - \frac{E(X_{T_\epsilon}(\mathbf{c}))}{(\tau - \mathbf{r} - \epsilon')\,\mathrm{card}(\mathcal{V})} \geq 1 - \frac{E(X_0(\mathbf{c}))}{(\tau - \mathbf{r} - \epsilon')\,\mathrm{card}(\mathcal{V})}. \tag{16}$$

Recalling that the opinions across the social network $\mathscr{G}$ are initially independent and identically distributed with values in the opinion space Δ, we also have

$$E(X_0(\mathbf{c})) = E\left(\sum_{x \in \mathcal{V}} \|\xi_0(x) - \mathbf{c}\|\right) = \sum_{x \in \mathcal{V}} E\,\|\xi_0(x) - \mathbf{c}\| = E\,\|X - \mathbf{c}\| \cdot \mathrm{card}(\mathcal{V}). \tag{17}$$

The result follows from (16) and (17). $\qquad\square$

To deduce the theorem, let $\tau > \mathbf{r}$. Then, for all $\epsilon' \in (0, \tau - \mathbf{r})$, we have $\tau > \mathbf{r} + \epsilon'$. Therefore, according to Lemmas 5 and 6, there exists $\epsilon > 0$ such that

$$P(\mathscr{C}) \geq P(\mathscr{A}(\epsilon, \epsilon')) \geq 1 - E\,\|X - \mathbf{c}\|/(\tau - \mathbf{r} - \epsilon').$$

Because this holds for all $\epsilon' > 0$ arbitrary small,

$$P(\mathscr{C}) \geq 1 - \lim_{\epsilon' \to 0} E \|X - \mathbf{c}\|/(\tau - \mathbf{r} - \epsilon') = 1 - E \|X - \mathbf{c}\|/(\tau - \mathbf{r}).$$

This completes the proof of Theorem 1.

References

1. Axelrod, R.: The dissemination of culture: a model with local convergence and global polarization. J. Conflict Resolut. **41**(2), 203–226 (1997)
2. Bramson, M., Griffeath, D.: Renormalizing the 3-dimensional voter model. Ann. Probab. **7**(3), 418–432 (1979)
3. Bramson, M., Griffeath, D.: Clustering and dispersion rates for some interacting particle systems on Z. Ann. Probab. **8**(2), 183–213 (1980)
4. Castellano, C., Fortunato, S., Loreto, V.: Statistical physics of social dynamics. Rev. Mod. Phys. **81**(2), 591 (2009)
5. Clifford, P., Sudbury, A.: A model for spatial conflict. Biometrika **60**, 581–588 (1973)
6. Cox, J.T.: Coalescing random walks and voter model consensus times on the torus in Z^d. Ann. Probab. **17**(4), 1333–1366 (1989)
7. Cox, J.T., Griffeath, D.: Diffusive clustering in the two-dimensional voter model. Ann. Probab. **14**(2), 347–370 (1986)
8. Deffuant, G., Neau, D., Amblard, F., Weisbuch, G.: Mixing beliefs among interacting agents. Adv. Complex Syst. **3**(01n04), 87–98 (2000)
9. Durrett, R., Steif, J.E.: Fixation results for threshold voter systems. Ann. Probab. **21**(1), 232–247 (1993)
10. Galam, S.: Minority opinion spreading in random geometry. Eur. Phys. J. B **25**(4), 403–406 (2002)
11. Häggström, O.: A pairwise averaging procedure with application to consensus formation in the Deffuant model. Acta Appl. Math. **119**, 185–201 (2012)
12. Hardin, M., Lanchier, N.: Probability of consensus in spatial opinion models with confidence threshold. ArXiv:1912.06746
13. Hegselmann, R., Krause, U.: Opinion dynamics and bounded confidence models, analysis, and simulation. J. Artif. Soc. Soc. Simul. **5**(3), 3–33 (2002)
14. Hirscher, T.: The Deffuant model on $\mathbb{Z}$ with higher-dimensional opinion spaces. ALEA Lat. Am. J. Probab. Math. Stat. **11**(1), 409–444 (2014)
15. Holley, R.A., Liggett, T.M.: Ergodic theorems for weakly interacting infinite systems and the voter model. Ann. Probab. **3**(4), 643–663 (1975)
16. Lanchier, N.: The Axelrod model for the dissemination of culture revisited. Ann. Appl. Probab. **22**(2), 860–880 (2012)
17. Lanchier, N.: The critical value of the Deffuant model equals one half. ALEA Lat. Am. J. Probab. Math. Stat. **9**(2), 383–402 (2012)
18. Lanchier, N., Li, H.-L.: Probability of consensus in the multivariate Deffuant model on finite connected graphs. Electron. Commun. Probab. **25**, 12 (2020)
19. Lanchier, N., Neufer, J.: Stochastic dynamics on hypergraphs and the spatial majority rule model. J. Stat. Phys. **151**(1–2), 21–45 (2013)
20. Lanchier, N., Scarlatos, S.: Fixation in the one-dimensional Axelrod model. Ann. Appl. Probab. **23**(6), 2538–2559 (2013)
21. Lanchier, N., Scarlatos, S.: Clustering and coexistence in the one-dimensional vectorial Deffuant model. ALEA Lat. Am. J. Probab. Math. Stat. **11**(1), 541–564 (2014)
22. Lanchier, N., Scarlatos, S.: Limiting behavior for a general class of voter models with confidence threshold. ALEA Lat. Am. J. Probab. Math. Stat. **14**(1), 63–92 (2017)
23. Lanchier, N., Schweinsberg, J.: Consensus in the two-state Axelrod model. Stoch. Process. Appl. **122**(11), 3701–3717 (2012)
24. Li, H.-L.: Mixed Hegselmann-Krause dynamics. ArXiv:2010.03050

25. Li, H.-L.: Mixed Hegselmann-Krause dynamics – nondeterministic case. ArXiv:2105.00577
26. Li, J.: Axelrod's model in two dimensions. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)–Duke University (2014)
27. Liggett, T.M.: Interacting Particle Systems. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 276. Springer-Verlag, New York (1985)
28. Liggett, T.M.: Coexistence in threshold voter models. Ann. Probab. **22**(2), 764–802 (1994)
29. Liggett, T.M.: Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Grundlehren der Mathematischen Wissenschaften, vol. 324. Springer-Verlag, Berlin (1999)
30. Vazquez, F., Krapivsky, P.L., Redner, S.: Constrained opinion dynamics: freezing and slow evolution. J. Phys. A **36**(3), L61–L68 (2003)
31. Zähle, I.: Renormalization of the voter model in equilibrium. Ann. Probab. **29**(3), 1262–1302 (2001)