Electron. Commun. Probab. **25** (2020), article no. 79, 1–12.
https://doi.org/10.1214/20-ECP359
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Probability of consensus in the multivariate Deffuant model on finite connected graphs*

Nicolas Lanchier[†] Hsin-Lun Li[‡]

Abstract

The Deffuant model is a spatial stochastic model for the dynamics of opinions in which individuals are located on a connected graph representing a social network and characterized by a number in the unit interval representing their opinion. The system evolves according to the following averaging procedure: at each time step, two neighbors are randomly chosen and interact if and only if the distance between their opinions does not exceed a certain confidence threshold, with each interaction resulting in the neighbors' opinions getting closer to each other. Most of the analytical results established so far about this model assume that the individuals are located on the integers. In contrast, we study the more realistic case where the social network can be any finite connected graph. In addition, we extend the opinion space to any bounded convex subset of a normed vector space where the norm is used to measure the level of disagreement or distance between the opinions. Our main result gives a lower bound for the probability of consensus. Our proof leads to a universal lower bound that depends on the confidence threshold, the opinion space (convex subset and norm) and the initial distribution, but not on the size or the topology of the social network.

Keywords: Interacting particle systems; Deffuant model; opinion dynamics; martingale; optional stopping theorem; confidence threshold; consensus.
MSC2020 subject classifications: 60K35.
Submitted to ECP on May 27, 2020, final version accepted on November 3, 2020.

1 Introduction

This paper is concerned with opinion dynamics on connected graphs. The first and most popular stochastic model in this topic is the voter model, introduced in [5, 15]. The main mechanism in the voter model is social influence, the tendency of individuals to become more similar when they interact. More precisely, individuals located on the vertex set of a connected graph (traditionally the d-dimensional integer lattice) are characterized by one of two competing opinions, and update their opinion at rate one by simply mimicking one of their neighbors chosen uniformly at random. Using a duality relationship between the voter model and a system of coalescing random walks, it can be

*Nicolas Lanchier was partially supported by NSF grant CNS-2000792.
[†]School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA.
E-mail: nicolas.lanchier@asu.edu
[‡]School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA.
E-mail: hsinlunl@asu.edu

proved that the process on the infinite square lattice clusters in one and two dimensions whereas opinions coexist at equilibrium in higher dimensions [15]. While mathematicians studied analytically various aspects of the model such as the asymptotics for the cluster size in one and two dimensions [3, 7], the spatial correlations at equilibrium in higher dimensions [2], and the occupation time of the process [6], social scientists and statistical physicists developed and studied numerically more realistic models of opinion dynamics. We refer to [18, 25] for reviews of the main results about the voter model, and to [4] for a review of more recent stochastic models of opinion dynamics introduced by applied scientists.

Apart from social influence, an important component of opinion dynamics is homophily, the tendency to interact more frequently with individuals who are more similar. The most popular spatial model that includes social influence and homophily is probably the Axelrod model [1] where individuals are now characterized by a vector of cultural features, and interact with their neighbors at a rate proportional to the number of features they share (homophily), which results in the two neighbors having one more feature in common (social influence). For a mathematical treatment of the Axelrod model, we refer to [16, 19, 20, 23, 24]. Other spatial stochastic models of opinion dynamics include homophily in the form of a confidence threshold: individuals interact with their neighbors on the graph if and only if the level of disagreement between the two individuals before the interaction does not exceed a certain threshold. The simplest such model is the constrained voter model [27], the voter model with three opinions (leftist, centrist and rightist) where leftists and rightists do not interact. Extensions of this model where the opinion space takes the form of a finite connected graph and the level of disagreement is measured using the geodesic distance on this graph were introduced and studied analytically in [22, 26]. The Deffuant model [8] and the Hegselmann-Krause model [12] are two other important spatial stochastic models that include social influence and homophily in the form of a confidence threshold.

In the original version of the Deffuant model [8], individuals are located on a general finite connected graph representing a social network and characterized by opinions that are initially chosen independently and uniformly at random in the unit interval. At each time step, an edge is chosen at random and the two neighbors connected by this edge interact if and only if the distance between their opinions before the interaction does not exceed a confidence threshold τ (homophily), which results in the two neighbors' opinions getting closer to each other after the interaction (social influence). Because [8] is purely based on numerical simulations, the authors only considered specific graphs: the complete graph and the two-dimensional torus. Their simulations on large graphs suggest the following conjecture for the infinite system obtained by assuming that pairs of neighbors are now chosen in continuous time at rate one: the process exhibits a phase transition at the critical threshold one-half in that a consensus is reached when $\tau > 1/2$ whereas disagreements persist in the long run when $\tau < 1/2$. This conjecture was first established for the process on the integers in [17] using probabilistic and geometric techniques while a slightly stronger result was proved shortly after in [10] using a different approach. The existence of a phase transition along with lower and upper bounds for the critical threshold were also proved for variants of the model: a multivariate version where the opinion space is a (subset of a) finite-dimensional vector space and certain metrics are used to quantify the disagreement between individuals [9, 13, 14], and a discrete version called the vectorial Deffuant model also introduced in [8] where the opinion space is the hypercube and the disagreement between individuals is quantified using the Hamming distance [21].

In this paper, we study a version of the model where both the opinion space and the social network are fairly general. The opinion space is a bounded convex subset of a finite-dimensional normed vector space (where the norm is used to measure the disagreements). Under the averaging procedure [8], convexity is a necessary assumption because future opinions must be on the segment connecting past opinions, but we also point out that an extension of the model has been introduced in [9] where the opinion space is a general path-connected set, the opinion distance is measured by the length of some geodesics connecting the opinions and each update displaces the opinions along these geodesics. More importantly, while all the previous analytical results assume that the inviduals are located on the integers, with the notable exception of [11] where the process is also studied on the d-dimensional lattice and even the infinite bond percolation cluster, we assume more realistically that the individuals are located on a general finite connected graph, meaning any possible real-world social networks. But unlike [8] that relies on simulations for specific graphs, our results apply to all possible finite connected graphs. Due to the finiteness of the graph, the existence of a phase transition at a specific critical threshold no longer holds, and we instead derive a general lower bound for the probability of consensus. While our bound depends on the choice of the opinion space, it is uniform in all possible choices of the social network.

2 Model description and main results

The two key components of the model studied in this paper are the social network on which the individuals are located and the opinion space. To define these two components,

- we let $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ be a finite connected graph and
- we let $\Delta \subset \mathbb{R}^n$ be a bounded convex subset and $\| \cdot \|$ be a norm on $\mathbb{R}^n$.

The multivariate Deffuant model is a discrete-time Markov chain whose state at time t is a configuration of opinions on the graph:

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

Following all the previous works in this topic, we assume that the process starts from a constant product measure, meaning that the initial opinions $\xi_0(x)$, $x \in \mathscr{V}$, are independent and identically distributed, and we let X be the random variable with distribution

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

The evolution rules are based on two parameters: the confidence threshold $\tau > 0$ and the convergence parameter $\mu \in (0, 1/2]$. At each time step, an edge is chosen uniformly at random, which results in a potential update of the system at the two vertices connected by this edge. More precisely, assuming that edge $(x, y) \in \mathscr{E}$ is selected at time t, we let

$$\begin{aligned} \xi_t(x) &= \xi_{t-1}(x) + \mu \, (\xi_{t-1}(y) - \xi_{t-1}(x)) \, \mathbf{1}\{\|\xi_{t-1}(x) - \xi_{t-1}(y)\| \leq \tau\} \\ \xi_t(y) &= \xi_{t-1}(y) + \mu \, (\xi_{t-1}(x) - \xi_{t-1}(y)) \, \mathbf{1}\{\|\xi_{t-1}(x) - \xi_{t-1}(y)\| \leq \tau\} \end{aligned}$$

while the opinions at the other vertices remain unchanged. In words, the two neighbors that are selected interact if and only if their opinion distance or level of disagreement before the interaction does not exceed the confidence threshold τ, which results in a partial averaging of their opinions by a factor μ, called the convergence parameter.

Our main result gives a lower bound for the probability of consensus that applies to any finite connected graph, any opinion space (convex set and norm), and any initial distribution with value in the opinion space. To state this result, we let

$$\mathbf{r} = \inf\{r > 0 : \Delta \subset B(c, r) \text{ for some } c \in \Delta\} \quad \text{where} \quad B(c, r) = \{a \in \mathbb{R}^n : \|a - c\| \leq r\}$$

and fix $\mathbf{c} \in \Delta$ such that $\Delta \subset B(\mathbf{c}, \mathbf{r})$. By definition of $\mathbf{r}$, which we call the radius of the opinion space, and because the opinion space is convex, such a point $\mathbf{c}$ indeed exists.

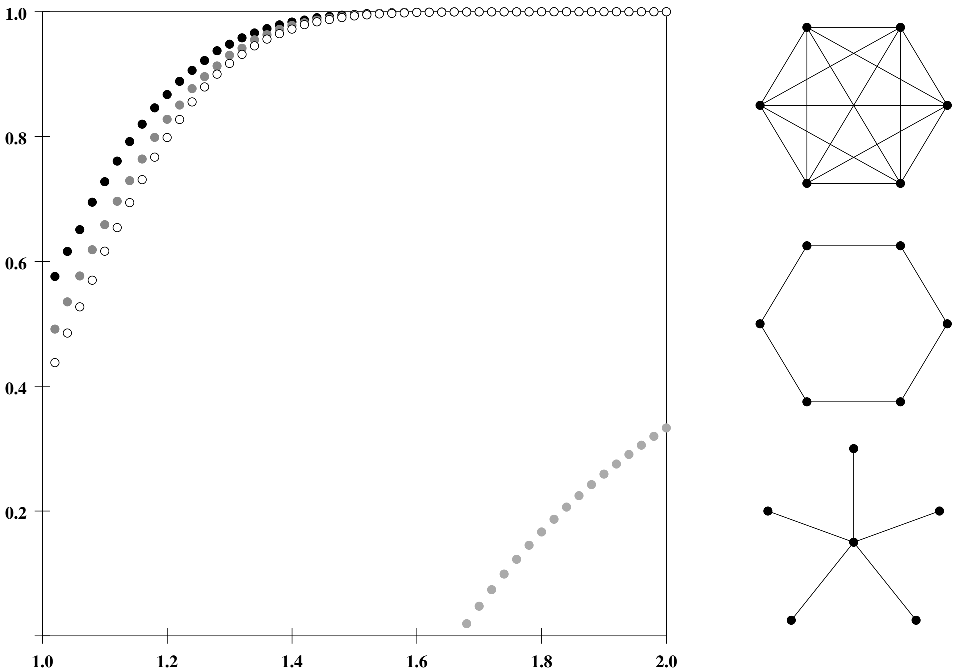

Figure 1: Simulation results for the probability of consensus as a function of τ for the process on the complete graph (black dots), the ring (grey dots) and the star (white dots) with six vertices depicted on the right. The convergence parameter $\mu = 1/4$ while the opinion space is the unit Euclidean ball equipped with the Euclidean norm in two dimensions. Each simulation point is obtained from the average of 100,000 realizations of the process with confidence threshold ranging from the radius to the diameter of the opinion space. The grey dots at the bottom right show the universal lower bound (valid for all finite connected graphs) derived from Theorem 2.1.

Theorem 2.1 (Probability of consensus)**.** *For all $\tau > \mathbf{r}$,*

$$P(\mathscr{C}) \geq 1 - \frac{E\,\|X - \mathbf{c}\|}{\tau - \mathbf{r}} \quad \text{where} \quad \mathscr{C} = \left\{ \lim_{t \to \infty} \max_{x,y \in \mathscr{V}} \|\xi_t(x) - \xi_t(y)\| = 0 \right\}.$$

Recall that the simulations in [8] suggest that, when the individuals are located on an infinite connected graph and the initial opinions are chosen uniformly at random in the unit interval, the process exhibits a phase transition from coexistence to consensus at the critical threshold $\tau = 1/2$, a result that was proved rigorously in some particular cases. In view of this conjecture, it is reasonable to believe that, again for infinite connected graphs and uniformly distributed initial opinions but a more general opinion space, there is now a phase transition at the critical value $\tau =$ the radius of the opinion space. In particular, we conjecture that, for infinite connected graphs, we have almost sure consensus under the assumption of our theorem: $\tau > \mathbf{r}$. The reason why the probability of consensus when $\tau > \mathbf{r}$ is strictly less than one when switching to finite graphs is simply due to the presence of strong random fluctuations on finite graphs. The law of large numbers used in [10] no longer operates. For the same reasons, while the condition $\tau < \mathbf{r}$ should lead to coexistence for infinite graphs, the probability of consensus on finite graphs is strictly positive. Indeed, it follows from the argument of convexity in the proof of Lemma 3.5 below that, when all the opinions are initially in the same ball with radius $\tau/2$, which occurs with probability at least the Lebesgue measure of this ball over the Lebesgue measure of the opinion space raised at the power the number of vertices, consensus occurs. In contrast with the lower bound in our theorem, the lower bound above strongly depends on the size of the graph.

The key to proving the theorem is to study a collection of auxiliary processes (see (3.1) below) that keep track of the cumulative disagreement between a fixed opinion $c \in \Delta$ and the opinions at each of the vertices at time t. Using a triangle-type inequality (Lemma 3.1), we first prove that all these auxiliary processes are almost surely nonincreasing, meaning that, for all c, the averaging procedure can only decrease the overall level of disagreement between an observer with fixed opinion c and the population (Lemma 3.2). Almost sure monotonicity implies two important results:

1. The opinion model converges almost surely to a (random) limiting configuration.

   In addition, due to the evolution rules, each limiting configuration is characterized by a partition of the graph into connected components such that all the individuals in the same component share the same opinion and the distance between opinions in two adjacent components exceeds the confidence threshold τ (Lemma 3.5).

2. All the auxiliary processes are bounded supermartingales.

   In particular, one may apply the optional stopping theorem to these supermartingales and a certain stopping time (Lemma 4.1) to obtain a lower bound for the probability that the random partition above consists of only one set, meaning that all the individuals in the limiting configuration share the same opinion and consensus occurs.

Our proof leads to a lower bound that depends on the confidence threshold, the opinion space and the initial distribution, but not on the size and/or the topology of the social network. In particular, our lower bound is universal in the sense that it is uniform over all possible choices of the network, but we point out that, as shown in Figure 1, the (exact) probability of consensus should depend on the network. Indeed, our simulations suggest for instance that the complete graph promotes consensus more than the star graph.

The rest of the paper is devoted to proofs. In the next section, we show convergence almost surely to a (random) limiting configuration in which neighbors either share the same opinion or disagree too much to interact. Then, we use the optional stopping theorem for supermartingales to derive the universal lower bound for the probability of consensus.

3 Limiting configurations

The objective of this section is to prove that, regardless of the initial configuration, the process converges almost surely to a limiting configuration in which any two neighbors either share the same opinion or disagree too much to interact, i.e.,

(P1) $\lim_{t \to \infty} \xi_t(x) = \xi_\infty(x)$ exists for all $x \in \mathscr{V}$

(P2) $\|\xi_\infty(x) - \xi_\infty(y)\| \notin (0, \tau]$ for all edges $(x, y) \in \mathscr{E}$.

From now on, we let $(X_t(c))$ be the process defined by

$$X_t(c) = \sum_{x \in \mathscr{V}} \|\xi_t(x) - c\| \quad \text{for all} \quad c \in \mathbb{R}^n. \tag{3.1}$$

That is, the process keeps track of the cumulative disagreement between a fixed opinion c possibly outside Δ and the opinions at each of the vertices. In particular, this collection of processes is somewhat reminiscent of the concept of energy in [10] in the sense that

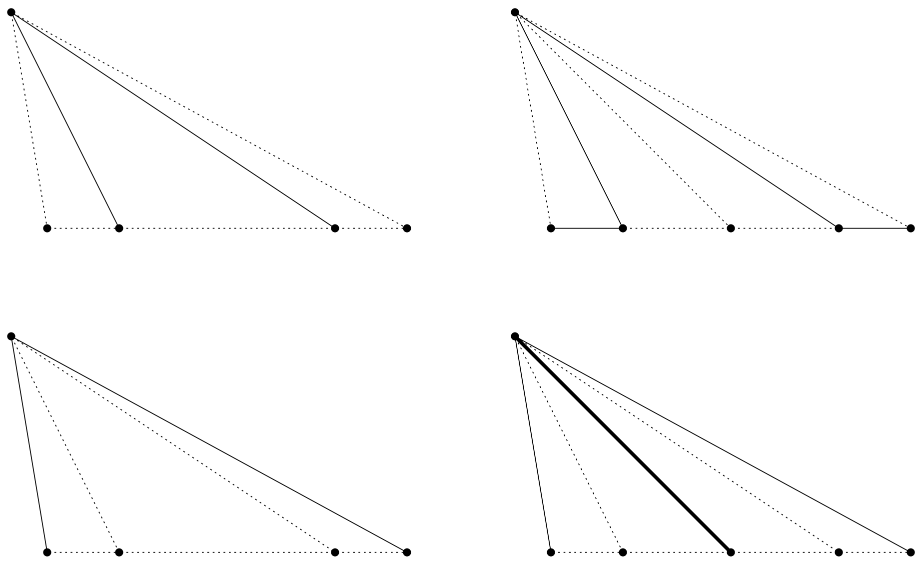

Figure 2: Illustration of Lemma 3.1. The lemma simply states that the sum of the norms of the vectors in solid lines is larger for the pictures at the bottom than for the pictures at the top, where the median in thick line in the bottom right picture is counted twice.

both can be viewed as measures of the overall disorder in the process that is expected to decrease under the influence of the averaging procedure. To shorten the notation, we let

$$\phi : \Delta \times \Delta \to \Delta \quad \text{defined as} \quad \phi(a,b) = (1-\mu)a + \mu b = a + \mu(b-a).$$

In particular, whenever a vertex x that has opinion a interacts with a vertex y that has a compatible opinion $b \in B(a,\tau)$, the opinion at x becomes $\phi(a,b)$ and the opinion at y becomes $\phi(b,a)$. Although the details are somewhat more complicated, the basic idea to prove the two properties above is to show that the processes $(X_t(c))$ converge almost surely. To begin with, we prove the following lemma which is illustrated in Figure 2 and gives two variants of the triangle inequality.

Lemma 3.1 (Triangle inequalities)**.** *For all* $a, b \in \Delta$ *and all* $c \in \mathbb{R}^n$*,*

$$\begin{aligned} \|\phi(a,b) - c\| + \|\phi(b,a) - c\| &\leq \|a - c\| + \|b - c\| \\ \|\phi(a,b) - c\| + \|\phi(b,a) - c\| &\leq \|a - c\| + \|b - c\| - 2\|\phi(a,b) - a\| + \|a + b - 2c\|. \end{aligned}$$

Proof. Using the triangle inequality and absolute homogeneity, we get

$$\begin{aligned} \|\phi(a,b) - c\| + \|\phi(b,a) - c\| &= \|(1-\mu)a + \mu b - c\| + \|(1-\mu)b + \mu a - c\| \\ &= \|(1-\mu)(a-c) + \mu(b-c)\| + \|(1-\mu)(b-c) + \mu(a-c)\| \\ &\leq \|(1-\mu)(a-c)\| + \|\mu(b-c)\| + \|(1-\mu)(b-c)\| + \|\mu(a-c)\| = \|a-c\| + \|b-c\| \end{aligned}$$

which proves the first inequality. Now, because $0 < \mu \leq 1/2$, the opinions

$$a, \quad \phi(a,b), \quad c_0 = (a+b)/2, \quad \phi(b,a), \quad b$$

all lie on the segment line $[a,b]$ in this specific order going from point a to point b, therefore using again the triangle inequality and absolute homogeneity, we obtain

$$\begin{aligned} \|\phi(a,b) - c\| + \|\phi(b,a) - c\| &\leq \|\phi(a,b) - c_0\| + \|c_0 - c\| + \|\phi(b,a) - c_0\| + \|c_0 - c\| \\ &= \|\phi(a,b) - \phi(b,a)\| + 2\|c_0 - c\| = \|a-b\| - \|\phi(a,b) - a\| - \|\phi(b,a) - b\| + \|2(c_0 - c)\| \\ &\leq \|a-c\| + \|b-c\| - 2\|\phi(a,b) - a\| + \|a+b-2c\| \end{aligned}$$

which proves the second inequality. This completes the proof. □

In the next lemma, we use the first inequality in Lemma 3.1 to prove that, for all $c \in \Delta$, the processes $(X_t(c))$ are almost surely nonincreasing.

Lemma 3.2 (Monotonicity)**.** *For all* $c \in \Delta$*,*

$$0 \leq X_t(c) \leq X_s(c) \leq 2\mathbf{r} \cdot \mathrm{card}(\mathscr{V}) \quad \textit{for all} \quad s \leq t.$$

Proof. At each update of the processes, say at time s,

$$\xi_s(x) = \phi(\xi_{s-1}(x), \xi_{s-1}(y)) \quad \text{and} \quad \xi_s(y) = \phi(\xi_{s-1}(y), \xi_{s-1}(x)) \quad \text{for some} \quad (x,y) \in \mathscr{E}.$$

In particular, applying Lemma 3.1 with $a = \xi_{s-1}(x)$ and $b = \xi_{s-1}(y)$, we get

$$\begin{aligned} X_s(c) - X_{s-1}(c) \ &= \ \|\xi_s(x) - c\| + \|\xi_s(y) - c\| - \|\xi_{s-1}(x) - c\| - \|\xi_{s-1}(y) - c\| \\ &= \ \|\phi(a,b) - c\| + \|\phi(b,a) - c\| - \|a - c\| - \|b - c\| \leq 0. \end{aligned}$$

In addition, because $c \in \Delta$ and $\Delta \subset B(\mathbf{c}, \mathbf{r})$, we have

$$\begin{aligned} 0 \leq X_t(c) \ &= \ \sum_{x \in \mathscr{V}} \|\xi_t(x) - c\| \\ &\leq \ \sum_{x \in \mathscr{V}} \Big(\|\xi_t(x) - \mathbf{c}\| + \|c - \mathbf{c}\| \Big) \leq \sum_{x \in \mathscr{V}} 2\mathbf{r} = 2\mathbf{r} \cdot \mathrm{card}(\mathscr{V}) < \infty. \end{aligned}$$

This completes the proof. □

Note that Lemma 3.2 implies that the processes $(X_t(c))$ are bounded supermartingales, which will be used later with the optional stopping theorem to derive our universal lower bound for the probability of consensus. By the martingale convergence theorem, each of these processes converges almost surely to a finite random variable, which suggests almost sure convergence of the interacting particle system. The main difficulty to prove this result is that whenever two vertices with compatible opinions a and b interact, the process $(X_t(c))$ does not "see the update" when a, b, c are aligned in this order. For some norms, the lack of alignment is not even a sufficient condition for the process to see the change of opinions so it is not clear how to deduce convergence of the system. To prove this result, we now use Lemma 3.2 and the second inequality in Lemma 3.1 to show that the process keeps slowing down in the sense that the jumps at each vertex get smaller and smaller.

Lemma 3.3 (Slow-down)**.** *For all* $\epsilon > 0$*, there is* $S = S(\epsilon)$ *almost surely finite such that*

$$\|\xi_s(x) - \xi_{s-1}(x)\| < \epsilon \quad \textit{for all} \quad s \geq S \textit{ and } x \in \mathscr{V}.$$

Proof. Assume by contradiction that, for some $\epsilon > 0$ and $x \in \mathscr{V}$, the opinion at x jumps by more than ϵ infinitely often with positive probability, and let (s_i) be the times of these updates. In particular, we have

$$\|\xi_{s_i}(x) - \xi_{s_i - 1}(x)\| \geq \epsilon \quad \text{for all} \quad i > 0.$$

Letting $y_i \in \mathscr{V}$ be the vertex that interacts with x at time s_i, setting

$$a_i = \xi_{s_i - 1}(x), \quad b_i = \xi_{s_i - 1}(y) \quad \text{and} \quad c_i = (a_i + b_i)/2,$$

and applying the second inequality in Lemma 3.1 with $a = a_i$ and $b = b_i$, we get

$$\begin{aligned} X_{s_i}(c) - X_{s_i - 1}(c) &= \|\xi_{s_i}(x) - c\| + \|\xi_{s_i}(y) - c\| - \|\xi_{s_i - 1}(x) - c\| - \|\xi_{s_i - 1}(y) - c\| \\ &= \|\phi(a_i, b_i) - c\| + \|\phi(b_i, a_i) - c\| - \|a_i - c\| - \|b_i - c\| \\ &\leq -2\|\phi(a_i, b_i) - a_i\| + \|a_i + b_i - 2c\| = -2\|\xi_{s_i}(x) - \xi_{s_i - 1}(x)\| + 2\|c_i - c\| \\ &\leq -2\epsilon + 2\|c_i - c\| \leq -\epsilon \end{aligned} \tag{3.2}$$

for all $c \in B(c_i, \epsilon/2)$. Now, observe that there exists $\epsilon' > 0$ such that

$$B(c, \epsilon/2) \cap \Delta(\epsilon') \neq \varnothing \quad \text{for all} \quad c \in \Delta \quad \text{where} \quad \Delta(\epsilon') = \Delta \cap (\epsilon' \mathbb{Z})^n \tag{3.3}$$

and where $(\epsilon' \mathbb{Z})^n$ is a grid with mesh size ϵ' in n dimensions. It follows from the Pythagorean theorem that the Euclidean distance from c to the grid is bounded by

$$\sqrt{(\epsilon'/2)^2 + \cdots + (\epsilon'/2)^2} = \sqrt{n(\epsilon'/2)^2} = \sqrt{n}\,(\epsilon'/2)$$

which implies that (3.3) holds for $\epsilon' < \epsilon/\sqrt{n}$. This and the equivalence of the norms in finite dimensions imply that, for each norm, there indeed exists $\epsilon' > 0$ such that (3.3) holds. In addition, because the opinion space Δ is bounded, and again the dimension is finite,

$$\text{card}(\Delta(\epsilon')) < \infty \quad \text{for all} \quad \epsilon' > 0. \tag{3.4}$$

Combining (3.3) and (3.4), we deduce that

$$\Delta'(\epsilon') = \{c \in \Delta(\epsilon') : \text{card}\{i : c \in B(c_i, \epsilon/2)\} = \infty\} \neq \varnothing.$$

In particular, there exists

$$c' \in \Delta'(\epsilon') \quad \text{such that} \quad I = \{i \in \mathbb{N} : c \in B(c_i, \epsilon/2)\} \text{ is infinite.}$$

This, together with (3.2) and Lemma 3.2, implies that

$$\lim_{t \to \infty} X_t(c') \leq X_0(c') + \sum_{i \in I} (X_{s_i}(c') - X_{s_i - 1}(c')) = X_0(c') + \sum_{i \in I} (-\epsilon) = -\infty,$$

which contradicts the fact that $(X_t(c'))$ is positive. □

The next lemma shows that the jumps getting smaller and smaller implies that, for large times, neighbors must either be incompatible or have almost the same opinion.

Lemma 3.4 (Clustering). *For all $0 < \epsilon < \tau$, there is $T = T(\epsilon)$ almost surely finite such that*

$$\|\xi_s(x) - \xi_s(y)\| \notin [\epsilon, \tau] \quad \textit{for all} \quad s \geq T \textit{ and } (x, y) \in \mathscr{E}.$$

Proof. Assume by contradiction that there exist $\epsilon > 0$ and $(x, y) \in \mathscr{E}$ such that the opinion distance along the edge belongs to $[\epsilon, \tau]$ infinitely often, meaning that

$$\xi_{s_i}(x) - \xi_{s_i}(y) \in [\epsilon, \tau] \quad \text{for an increasing sequence} \quad (s_i) \subset \mathbb{N}.$$

Letting A_i be the event that edge (x, y) is selected at time $s_i + 1$, because the edge selected at each time step is chosen uniformly at random and independently of everything else,

$$\sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{\infty} \frac{1}{\text{card}(\mathscr{E})} = \infty \quad \text{and} \quad \text{the events } (A_i) \text{ are independent.}$$

In particular, it follows from the second Borel-Cantelli lemma that

$$P\Big(\limsup_{i \to \infty} A_i\Big) = P(\text{card}\{i \geq 1 : A_i \text{ occurs}\} = \infty) = 1. \tag{3.5}$$

In addition, on the event A_i,

$$\|\xi_{s_i+1}(x) - \xi_{s_i}(x)\| = \|\phi(\xi_{s_i}(x), \xi_{s_i}(y)) - \xi_{s_i}(x)\| = \|\mu(\xi_{s_i}(x) - \xi_{s_i}(y))\| > \mu\epsilon. \tag{3.6}$$

Combining (3.5) and (3.6), we deduce that, with probability one, the opinion at x jumps by more than $\mu\epsilon$ infinitely often, which contradicts Lemma 3.3. □

To deduce almost sure convergence of the particle system from the previous lemma, the last step is to prove that neighbors who almost totally agree cannot randomly oscillate together, which follows from an argument of convexity. The proof of the next lemma shows in fact a little bit more: there is a partition of the graph into connected components such that all the opinions in the same component are eventually trapped in a fixed ball with arbitrarily small radius while opinions in two adjacent components are incompatible, which implies in particular properties (P1) and (P2).

Lemma 3.5 (Convergence)**.** *Properties (P1) and (P2) hold.*

Proof. Let $N = \mathrm{card}(\mathscr{V})$ and $0 < \epsilon < \tau/N$. According to Lemma 3.4, there exists a random but almost surely finite time T such that

$$\|\xi_s(x) - \xi_s(y)\| \notin [\epsilon/N, \tau] \quad \text{for all} \quad s \geq T \text{ and } (x,y) \in \mathscr{E}$$

and we write $x \leftrightarrow y$ if there exist $x_0 = x, x_1, \ldots, x_j = y$ all distinct such that

$$(x_i, x_{i+1}) \in \mathscr{E} \text{ and } \|\xi_s(x_i) - \xi_s(x_{i+1})\| < \epsilon/N \quad \text{for all} \quad 0 \leq i < j \text{ and } s \geq T.$$

In particular, by the triangle inequality,

$$\|\xi_T(x) - \xi_T(y)\| \leq \sum_{i=0}^{j-1} \|\xi_T(x_i) - \xi_T(x_{i+1})\| < \frac{j\epsilon}{N} \leq \epsilon. \tag{3.7}$$

The relationship $\leftrightarrow$ defines an equivalence relationship so it induces a partition of the vertex set into equivalence classes $\mathscr{V}_1, \mathscr{V}_2, \ldots, \mathscr{V}_k$ that correspond to connected components of the graph. In addition, by (3.7) and the definition of $\leftrightarrow$, there exist $c_1, c_2, \ldots, c_k \in \Delta$ such that

(a) for all $i = 1, 2, \ldots, k$, we have $\xi_T(x) \in B(c_i, \epsilon)$ for all $x \in \mathscr{V}_i$ and
(b) whenever $\mathscr{V}_i$ and $\mathscr{V}_j$ are connected by $(x,y) \in \mathscr{E}$, we have $\|\xi_T(x) - \xi_T(y)\| > \tau$.

Assume that properties (a) and (b) hold from time T to time $s > T$ and that edge (x,y) is selected at time $s+1$. Then, either $x \leftrightarrow y$, say $x, y \in \mathscr{V}_i$, in which case

$$\begin{aligned}[\xi_{s+1}(x), \xi_{s+1}(y)] &= [(1-\mu)\,\xi_s(x) + \mu\,\xi_s(y), (1-\mu)\,\xi_s(y) + \mu\,\xi_s(x)] \\ &\subset [\xi_s(x), \xi_s(y)] \subset B(c_i, \epsilon)\end{aligned}$$

by convexity of $B(c_i, \epsilon)$, or edge (x,y) connects two different classes in which case

$$\|\xi_s(x) - \xi_s(y)\| > \tau \quad \text{therefore} \quad \xi_{s+1}(x) = \xi_s(x) \text{ and } \xi_{s+1}(y) = \xi_s(y).$$

In either case, properties (a) and (b) remain true after the interaction. Because $\epsilon > 0$ can be chosen arbitrarily small, this proves that properties (P1) and (P2) hold. □

4 Stopping time and consensus event

This section is devoted to the proof of Theorem 2.1. As mentioned after Lemma 3.2, the processes $(X_t(c))$ are bounded supermartingales so the idea is to apply the optional stopping theorem. Before proving the theorem, we define a suitable stopping time and show how the consensus event relates to the configuration of the system at this stopping time. Let

$$T_* = \inf\{t : \|\xi_t(x) - \xi_t(y)\| \notin [\tau/2, \tau] \text{ for all } x, y \in \mathscr{V}\}.$$

Note that time T_* is a stopping time for the natural filtration of the process. Time T_* is also almost surely finite according to Lemma 3.4, so we have the following result.

Lemma 4.1. *Time T_* is an almost surely finite stopping time.*

We now identify a collection of configurations at the stopping time T_* that always lead the population to consensus eventually. More precisely, we let

$$\mathscr{A} = \bigcup_{x \in \mathscr{V}} \left\{ \sup_{c \in \Delta} \|\xi_{T_*}(x) - c\| < \tau \right\}$$

be the event that, at the stopping time, there is (at least) one "centrist" individual whose opinion is within distance τ of all other possible opinions. Then, we have the following inclusion showing that the event $\mathscr{A}$ plays the role of an attractor in the sense that, whenever this event occurs, the process will almost surely evolve to a consensus.

Lemma 4.2 (Attractor)**.** *We have the inclusion $\mathscr{A} \subset \mathscr{C}$.*

Proof. The definition of T_* implies that

$$\xi_{T_*}(y) \in B(\xi_{T_*}(x), \tau) \ \Rightarrow \ \xi_{T_*}(y) \in B(\xi_{T_*}(x), \tau/2). \tag{4.1}$$

In addition, by the proof of Lemma 3.5 (convexity argument),

$$\xi_{T_*}(y) \in B(c, \tau/2) \text{ for all } y \in \mathscr{V} \ \Rightarrow \ \xi_s(y) \in B(c, \tau/2) \text{ for all } y \in \mathscr{V} \text{ and } s > T_*. \tag{4.2}$$

This, together with Lemma 3.5 itself, gives the implications

$$\begin{aligned}
&\sup\nolimits_{c \in \Delta} \|\xi_{T_*}(x) - c\| < \tau \text{ for some } x \in \mathscr{V} \\
&\quad \Rightarrow \ (\xi_{T_*}(y) \in B(\xi_{T_*}(x), \tau) \text{ for all } y \in \mathscr{V}) \text{ for some } x \in \mathscr{V} \\
&\quad \Rightarrow \ (\xi_{T_*}(y) \in B(\xi_{T_*}(x), \tau/2) \text{ for all } y \in \mathscr{V}) \text{ for some } x \in \mathscr{V} \quad \text{(by (4.1))} \\
&\quad \Rightarrow \ (\xi_s(y) \in B(c, \tau/2) \text{ for all } y \in \mathscr{V} \text{ and } s > T_*) \text{ for some } c \in \Delta \quad \text{(by (4.2))} \\
&\quad \Rightarrow \ \lim\nolimits_{s \to \infty} \|\xi_s(y) - \xi_s(z)\| = 0 \text{ for all } y, z \in \mathscr{V} \quad \text{(by (P2) and choice of } \tau/2).
\end{aligned}$$

This completes the proof. □

Proof of Theorem 2.1. By Lemma 3.2, for all $c \in \Delta$, the process $(X_t(c))$ is bounded and almost surely nonincreasing. In particular, the process is a bounded supermartingale with respect to the natural filtration of the opinion model. By Lemma 4.1, we also have that the random time T_* is an almost surely finite stopping time with respect to the same filtration therefore it follows from the optional stopping theorem that

$$E(X_{T_*}(c)) \le E(X_0(c)) = E\left(\sum_{x \in \mathscr{V}} \|\xi_0(x) - c\| \right) = \operatorname{card}(\mathscr{V}) \cdot E\,\|X - c\| \tag{4.3}$$

for all $c \in \Delta$. Now, on the complement of $\mathscr{A}$,

$$\text{for all } x \in \mathscr{V}, \quad \text{there exists } c_x \in \Delta \quad \text{such that} \quad \|\xi_{T_*}(x) - c_x\| \ge \tau.$$

This and the triangle inequality imply that

$$\|\xi_{T_*}(x) - \mathbf{c}\| \ge \|\xi_{T_*}(x) - c_x\| - \|c_x - \mathbf{c}\| \ge \tau - \mathbf{r} \quad \text{for all} \quad x \in \Delta.$$

This gives the following bound for the conditional expectation:

$$E(X_{T_*}(\mathbf{c}) \,|\, \mathscr{A}^c) = E\left(\sum_{x \in \mathscr{V}} \|\xi_{T_*}(x) - \mathbf{c}\| \,\middle|\, \mathscr{A}^c \right) \ge (\tau - \mathbf{r}) \cdot \operatorname{card}(\mathscr{V}). \tag{4.4}$$

Combining (4.3) with $c = \mathbf{c}$ and (4.4), we deduce that

$$(\tau - \mathbf{r})(1 - P(\mathscr{A})) \leq \frac{E(X_{T_*}(\mathbf{c}) \mid \mathscr{A}^c) \, P(\mathscr{A}^c)}{\operatorname{card}(\mathscr{V})} \leq \frac{E(X_{T_*}(\mathbf{c}))}{\operatorname{card}(\mathscr{V})} \leq E \, \|X - \mathbf{c}\|$$

which, together with Lemma 4.2, implies that

$$P(\mathscr{C}) \geq P(\mathscr{A}) \geq 1 - \frac{E \, \|X - \mathbf{c}\|}{\tau - \mathbf{r}} \quad \text{for all} \quad \tau > \mathbf{r}.$$

This completes the proof of the theorem. □

References

[1] R. Axelrod. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, 41(2):203–226, 1997.

[2] M. Bramson and D. Griffeath. Renormalizing the 3-dimensional voter model. *Ann. Probab.*, 7(3):418–432, 1979. MR-0528320

[3] M. Bramson and D. Griffeath. Clustering and dispersion rates for some interacting particle systems on $\mathbb{Z}$. *Ann. Probab.*, 8(2):183–213, 1980. MR-0566588

[4] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics*, 81(2):591, 2009.

[5] P. Clifford and A. Sudbury. A model for spatial conflict. *Biometrika*, 60:581–588, 1973. MR-0343950

[6] J. Theodore Cox and David Griffeath. Occupation time limit theorems for the voter model. *Ann. Probab.*, 11(4):876–893, 1983. MR-0714952

[7] J. Theodore Cox and David Griffeath. Diffusive clustering in the two-dimensional voter model. *Ann. Probab.*, 14(2):347–370, 1986. MR-0832014

[8] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(01n04):87–98, 2000.

[9] N. Gantert, M. Heydenreich, and T. Hirscher. Strictly weak consensus in the uniform compass model on $\mathbb{Z}$. *Bernoulli*, 26(2):1269–1293, 2020. MR-4058367

[10] O. Häggström. A pairwise averaging procedure with application to consensus formation in the Deffuant model. *Acta Appl. Math.*, 119:185–201, 2012. MR-2915577

[11] O. Häggström and T. Hirscher. Further results on consensus formation in the Deffuant model. *Electron. J. Probab.*, 19:no. 19, 26, 2014. MR-3164772

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

[13] T. Hirscher. *Consensus formation in the Deffuant model*. 2014. Thesis (Ph.D.)–Chalmers University of Technology.

[14] T. Hirscher. The Deffuant model on $\mathbb{Z}$ with higher-dimensional opinion spaces. *ALEA Lat. Am. J. Probab. Math. Stat.*, 11(1):409–444, 2014. MR-3265084

[15] R. A. Holley and T. M. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probability*, 3(4):643–663, 1975. MR-0402985

[16] N. Lanchier. The Axelrod model for the dissemination of culture revisited. *Ann. Appl. Probab.*, 22(2):860–880, 2012. MR-2953571

[17] N. Lanchier. The critical value of the Deffuant model equals one half. *ALEA Lat. Am. J. Probab. Math. Stat.*, 9(2):383–402, 2012. MR-3069370

[18] N. Lanchier. *Stochastic modeling*. Universitext. Springer, Cham, 2017. MR-3617451

[19] N. Lanchier and P.-H. Moisson. Fixation results for the two-feature Axelrod model with a variable number of opinions. *J. Theoret. Probab.*, 29(4):1554–1580, 2016. MR-3571254

[20] N. Lanchier and S. Scarlatos. Fixation in the one-dimensional Axelrod model. *Ann. Appl. Probab.*, 23(6):2538–2559, 2013. MR-3127944

[21] N. Lanchier and S. Scarlatos. Clustering and coexistence in the one-dimensional vectorial Deffuant model. *ALEA Lat. Am. J. Probab. Math. Stat.*, 11(1):541–564, 2014. MR-3274645

[22] N. Lanchier and S. Scarlatos. Limiting behavior for a general class of voter models with confidence threshold. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1):63–92, 2017. MR-3612323

[23] N. Lanchier and J. Schweinsberg. Consensus in the two-state Axelrod model. *Stochastic Process. Appl.*, 122(11):3701–3717, 2012. MR-2965921

[24] J. Li. Axelrod's model in two dimensions. ProQuest LLC, Ann Arbor, MI, 2014. Thesis (Ph.D.)–Duke University. MR-3232281

[25] T. M. Liggett. Stochastic interacting systems: contact, voter and exclusion processes, volume 324 of Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, Berlin, 1999. MR-1717346

[26] S. Scarlatos. *Voter models with confidence parameter*. 2013. Thesis (Ph.D.)–University of Patras, Greece.

[27] F. Vazquez, P. L. Krapivsky, and S. Redner. Constrained opinion dynamics: freezing and slow evolution. *J. Phys. A*, 36(3):L61–L68, 2003. MR-1959419

Acknowledgments. The authors would like to thank two anonymous referees for valuable comments that help improve this work, and a third referee for suggesting relevant references and pointing out a couple of mistakes in the preliminary version.