Adv. Appl. Prob. **47,** 668–692 (2015)
Printed in Northern Ireland

GALAM'S BOTTOM-UP HIERARCHICAL SYSTEM AND PUBLIC DEBATE MODEL REVISITED

N. LANCHIER * ** AND
N. TAYLOR,* *Arizona State University*

Abstract

This paper is concerned with the bottom-up hierarchical system and public debate model proposed by Galam (2008), as well as a spatial version of the public debate model. In all three models, there is a population of individuals who are characterized by one of two competing opinions, say opinion -1 and opinion $+1$. This population is further divided into groups of common size s. In the bottom-up hierarchical system, each group elects a representative candidate, whereas in the other two models, all the members of each group discuss at random times until they reach a consensus. At each election/discussion, the winning opinion is chosen according to Galam's majority rule: the opinion with the majority of representatives wins when there is a strict majority, while one opinion, say opinion -1, is chosen by default in the case of a tie. For the public debate models we also consider the following natural updating rule that we call proportional rule: the winning opinion is chosen at random with a probability equal to the fraction of its supporters in the group. The three models differ in term of their population structure: in the bottom-up hierarchical system, individuals are located on a finite regular tree, in the nonspatial public debate model, they are located on a complete graph, and in the spatial public debate model, they are located on the d-dimensional regular lattice. For the bottom-up hierarchical system and nonspatial public debate model, Galam studied the probability that a given opinion wins under the majority rule and, assuming that individuals' opinions are initially independent, making the initial number of supporters of a given opinion a binomial random variable. The first objective of this paper is to revisit Galam's result, assuming that the initial number of individuals in favor of a given opinion is a fixed deterministic number. Our analysis reveals phase transitions that are sharper under our assumption than under Galam's assumption, particularly with small population size. The second objective is to determine whether both opinions can coexist at equilibrium for the spatial public debate model under the proportional rule, which depends on the spatial dimension.

Keywords: Voting system; public debate; interacting particle system; martingale

2010 Mathematics Subject Classification: Primary 60K35
Secondary 91D25

1. Introduction

Galam's bottom-up hierarchical system and public debate model [3] are used to understand voting behaviors of two competing opinions in democratic societies. In his models, Galam assumes that initially individuals in the population are independently in favor of one opinion with a fixed probability, making the initial number of that type of opinion a binomial random variable.

Received 30 August 2013; revision received 19 May 2014.
* Postal address: School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA.
** Email address: nicolas.lanchier@asu.edu

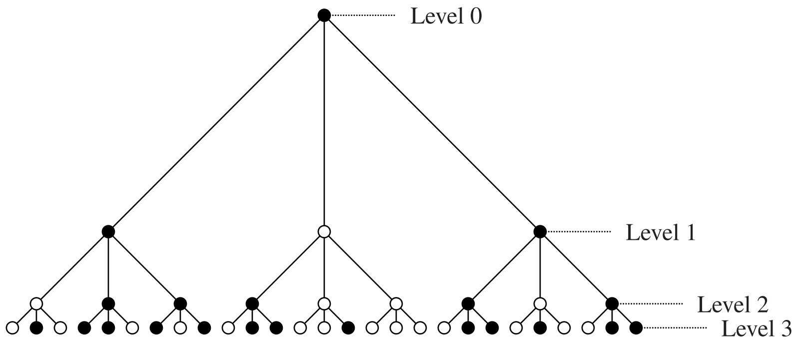

FIGURE 1: Schematic illustration of the bottom-up hierarchical system with $s = 3$ and $N = 3$. Black dots represent individuals supporting opinion $+1$ and white dots individuals supporting opinion -1.

This analysis revisits Galam's models by assuming that the initial number of individuals in favor of an opinion is a fixed deterministic number, which is more realistic when analyzing small populations. In this paper we are also concerned with a spatial version of Galam's public debate model introduced in [6]. Before stating our results, we start with a detailed description of these three models.

Bottom-up hierarchical system. The bottom-up hierarchical system [3] is a stochastic process that depends on two parameters: the group size s and the number of voting steps N, which are both positive integers. The structure of this model, which is shown in Figure 1, begins with a population of s^N individuals in favor of either opinion $+1$ or opinion -1 on the bottom level. This population is further divided into groups of size s, and local majority rules determine a representative candidate of each group who then ascends to another group at the next lowest level. This process continues until a single winner at level 0 is elected. When the group size s is odd, majority rule is well defined, whereas when the group size s is even, a bias is introduced favoring a predetermined type, say opinion -1, if there is a tie. That is, the representative candidates are determined at each step according to the majority rule whenever there is a strict majority but is chosen to be the one in favor of opinion -1 in case of a tie. This assumption is justified by Galam [3], based on the principle of social inertia. More formally, one can think of the model as a rooted regular tree with degree s and N levels plus the root. Denote by

$$X_n(i) \quad \text{for } n = 0, 1, \dots, N, \; i = 1, 2, \dots, s^n$$

the opinion of the ith node/individual at level n. Then the opinion of each node is determined from the configuration of opinions X_N at the bottom level and the recursive rule

$$X_n(i) := \operatorname{sign}\left(\sum_{j=1}^{s} X_{n+1}(s(i-1)+j) - \frac{1}{2}\right) \quad \text{for all } i = 1, 2, \dots, s^n.$$

Note in particular that this recursive rule is deterministic, making the process stochastic only through its configuration at the bottom level. Galam [3] assumes that nodes at the bottom level are independently in favor of a given opinion with a fixed probability. In contrast, we will assume that the configuration at the bottom level is a random permutation with a fixed number of nodes in favor of a given opinion.

https://doi.org/10.1239/aap/1444308877 Published online by Cambridge University Press

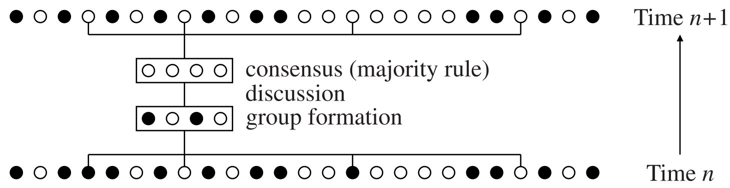

FIGURE 2: One time step in the nonspatial public debate model with $s = 4$ and $N = 25$. Black dots represent individuals supporting opinion $+1$ and white dots individuals supporting opinion -1.

Nonspatial public debate model. The second model under consideration in this paper is Galam's public debate model that examines the dynamics of opinion shifts. This process again depends on the same two parameters but now evolves in time. There is a population of N individuals each with either opinion $+1$ or opinion -1. At each time step, a random group of size s, called the discussion group, is chosen from the population, which results in all the individuals in the group adopting the same opinion. The updating rule considered in [3] is again the majority rule: if there are opposing opinions in the discussion group, then the opinion with the majority of supporters dominates the other opinion causing the individuals who initially supported the minority opinion to change his/her opinion to the majority opinion. As previously, when the group size s is even, ties may occur, in which case a bias is introduced in favor of opinion -1; see Figure 2 for a schematic representation of this process. In this paper we will also consider another natural updating rule that we will call proportional rule, which assumes that all the individuals in the group adopt opinion ± 1 with a probability equal to the fraction of supporters of this opinion in the group before discussion. To define these processes more formally, we now let

$$X_n(i) \quad \text{for } n \in \mathbb{N},\ i = 1, 2, \ldots, N$$

be the opinion of individual i at time n. In both processes, a set of s individuals, say B_s, is chosen uniformly from the population at each time step. Under the majority rule, we set

$$X_n(i) := \operatorname{sign}\biggl(\sum_{j \in B_s} X_{n-1}(j) - \frac{1}{2}\biggr) \quad \text{for all } i \in B_s,$$

while under the proportional rule, we set

$$X_n(i) := \begin{cases} +1 & \text{for all } i \in B_s \text{ with probability } s^{-1} \sum_{j \in B_s} \mathbf{1}_{\{X_n(j)=+1\}}, \\ -1 & \text{for all } i \in B_s \text{ with probability } s^{-1} \sum_{j \in B_s} \mathbf{1}_{\{X_n(j)=-1\}}. \end{cases}$$

In both processes, individuals outside B_s are not affected by the discussion and the evolution rule is iterated until everyone in the population has the same opinion. We will see later that the process that keeps track of the number of individuals with opinion $+1$ rather than the actual configuration is itself a discrete-time Markov chain. As for the bottom-up hierarchical system, we will assume that the configuration at time 0 has a fixed number of individuals in favor of a given opinion, whereas Galam studied the (majority rule) public debate model under the assumption that initially individuals are independently in favor of a given opinion with a fixed probability.

Spatial public debate model. The third model studied in this paper is a spatial version of the public debate model introduced in [6]. The spatial structure is represented by the infinite

d-dimensional regular lattice. Each site of the lattice is occupied by one individual who is again characterized by their opinion: either opinion $+1$ or opinion -1. The population being located on a geometrical structure, space can be included by assuming that only individuals in the same neighborhood can interact. More precisely, we assume that the set of discussion groups is

$$x + B_s \quad \text{for } x \in \mathbb{Z}^d,$$

where $B_s := \{0, 1, \dots, s-1\}^d$. Since the number of discussion groups is infinite and countable, the statement 'choosing a group uniformly at random' is no longer well defined. Therefore, we define the process in continuous time using the framework of interacting particle systems assuming that discussion groups are updated independently at rate one, i.e. at the arrival times of independent Poisson processes with intensity 1. The analysis in [6] is concerned with the spatial model under the majority rule, whereas we focus on the proportional rule: all the individuals in the same discussion group adopt the same opinion with a probability equal to the fraction of supporters of this opinion in the group before discussion. Formally, the state of the process at time t is now a function

$$\eta_t : \mathbb{Z}^d \longrightarrow \{-1, +1\}$$

with $\eta_t(x)$ denoting the opinion at time t of the individual located at site x, and the dynamics of the process is described by the Markov generator

$$\begin{aligned} Lf(\eta) = & \sum_x \sum_{z \in B_s} s^{-d}\, \mathbf{1}_{\{\eta(x+z)=+1\}} [f(\tau_x^+ \eta) - f(\eta)] \\ & + \sum_x \sum_{z \in B_s} s^{-d}\, \mathbf{1}_{\{\eta(x+z)=-1\}} [f(\tau_x^- \eta) - f(\eta)], \end{aligned}$$

where τ_x^+ and τ_x^- are the operators defined on the set of configurations by

$$(\tau_x^+ \eta)(z) := \begin{cases} +1 & \text{for } z \in x + B_x, \\ \eta(z) & \text{for } z \notin x + B_x, \end{cases} \qquad (\tau_x^- \eta)(z) := \begin{cases} -1 & \text{for } z \in x + B_x, \\ \eta(z) & \text{for } z \notin x + B_x. \end{cases}$$

The first part of the generator indicates that, for each x, all the individuals in $x + B_s$ switch simultaneously to opinion $+1$ at a rate equal to the fraction of individuals with opinion $+1$ in the group, while the second part provides similar transition rates for opinion -1. Basic properties of Poisson processes imply that each group is indeed updated at rate 1 according to the proportional rule. Note that the process no longer depends on N since the population size is infinite but we will see that its behavior strongly depends on the spatial dimension d.

2. Main results

For the bottom-up hierarchical system and the nonspatial public debate model, the main problem is to determine the probability that a given opinion, say opinion $+1$, wins as a function of the density or number of individuals holding this opinion in the initial configuration. For the spatial public debate model, since the population is infinite, the time to reach a configuration in which all the individuals share the same opinion is almost surely infinite when starting from a configuration with infinitely many individuals of each type. In this case, the main problem is to determine whether the system clusters or opinions can coexist at equilibrium.

Galam's results. Galam studied the bottom-up hierarchical system and the nonspatial public debate model under the majority rule. As previously explained, the assumption in [3] about the initial configuration of each model is that individuals are independently in favor of opinion $+1$ with some fixed probability. Under this assumption, the analysis is simplified because the probability of an individual being in favor of a given opinion at one level for the bottom-up hierarchical system or at one time step for the public debate model can be computed explicitly in a simple manner from its counterpart at the previous level or time step. More precisely, focusing on the bottom-up hierarchical system for concreteness and letting p_n be the common probability of any given individual being in favor of opinion $+1$ at level n, we have the recursive formula

$$p_n = Q_s(p_{n+1}), \quad \text{where } Q_s(X) := \sum_{j=s'}^{s} \binom{s}{j} X^j (1-X)^{s-j},\ s' := \left\lceil \frac{s+1}{2} \right\rceil.$$

The probability that a given opinion wins the election can then be computed explicitly. For both models, in the limit as the population size tends to ∞, the problem reduces to finding the fixed points of the polynomial Q_s. When $s = 3$,

$$Q_3(X) - X = 3X^2(1-X) + X^3 - X = -X(X-1)(2X-1)$$

and therefore one half is a fixed point. It follows that, with probability close to 1 when the population size is large, the winning opinion is the one that has initially the largest frequency of representatives, a result that easily extends to all odd sizes. The case of even sizes is more intriguing: when the group size $s = 4$, we have

$$Q_4(X) - X = 4X^3(1-X) + X^4 - X = -3X(X-1)(X-c_-)(X-c_+), \tag{1}$$

where the roots c_- and c_+ are given by

$$c_- := \left(\tfrac{1}{6}\right)(1-\sqrt{13}) \approx -0.434, \qquad c_+ := \left(\tfrac{1}{6}\right)(1+\sqrt{13}) \approx 0.768.$$

This implies that, when the population is large, the probability that opinion $+1$ wins is near 0 if the initial frequency of its representatives is below $c_+ \approx 0.768$. It can be proved that the same result holds for the nonspatial public debate model when the population size is large. Because opinions are initially independent and of a given type with a fixed probability, the initial number of individuals with opinion $+1$ is a binomial random variable, and the main reason behind the simplicity of Galam's results is that the dynamics of his models preserves this property: at any level/time, the number of individuals with opinion $+1$ is again binomial. The first objective of this paper is to revisit Galam's results under the assumption that the initial number of individuals with opinion $+1$ is a fixed deterministic number rather than binomially distributed. This assumption is more realistic for small populations but the analysis is also more challenging because the numbers of individuals with a given opinion in nonoverlapping groups are no longer independent.

Bottom-up hierarchical system. For the bottom-up hierarchical system, we start with a fixed deterministic number x of individuals holding opinion $+1$ at the bottom level. The main objective is then to determine the winning probability

$$\begin{aligned} p_x(N,s) &:= \text{probability that opinion } +1 \text{ wins} \\ &:= \mathbb{P}(X_0(1) = +1 \mid \operatorname{card}\{i : X_N(i) = +1\} = x), \end{aligned} \tag{2}$$

where s is the group size and N is the number of voting steps. Assuming that individuals holding the same opinion are identical, there are s^N who choose x possible configurations at the bottom level of the system. To compute the probability (2), the most natural approach is to compute the number of such configurations that result in the election of candidate $+1$. This problem, however, is quite challenging so we use instead a different strategy. The main idea is to count configurations that are compatible with the victory of $+1$ going backwards in the hierarchy: we count the number of configurations at level 1 that result in the election of candidate $+1$, then the number of configurations at level 2 that result in any of these configurations at level 1, and so on. To compute the number of such configurations, for each size-level pair (s, n), we set

$$s' := \lceil (\tfrac{1}{2})(s+1) \rceil, \qquad I_{s,n} := \{0, 1, \ldots, (s'-1)(s^n - x) + (s - s')x\}. \tag{3}$$

Then, for all $y \in I_{s,n}$, we define

$$c_n(s, x, s'x + y) = \sum_{z_0, \ldots, z_s} \binom{s^n}{x}^{-1} \binom{s^n}{z_0, z_1, \ldots, z_s} \prod_{j=0}^{s} \binom{s}{j}^{z_j}, \tag{4}$$

where the sum is over all $z_0, z_1, \ldots, z_s$ such that

$$z_0 + z_1 + \cdots + z_{s'-1} = s^n - x, \qquad z_{s'} + z_{s'+1} + \cdots + z_s = x,$$

and such that

$$\sum_{j=1,2,\ldots,s'-1} j(z_j + z_{s'+j}) = y \quad \text{if } s \text{ is odd},$$

$$\sum_{j=1,2,\ldots,s'-2} j(z_j + z_{s'+j}) + (s'-1)z_{s'-1} = y \quad \text{if } s \text{ is even}.$$

We will prove that the number of configurations with $s'x + y$ individuals holding opinion $+1$ at level $n+1$ that result in a given configuration with x individuals holding opinion $+1$ at level n is exactly given by (4). The fact that the evolution rules are deterministic also implies that different configurations at a given level cannot result from the same configuration at a lower level. In particular, the number of configurations at the bottom level that result in the victory of opinion $+1$ can be deduced from a simple summation as in the proof of Chapman–Kolmogorov's equations in the theory of Markov chains. More precisely, we have the following theorem.

Theorem 1. (Bottom-up hierarchical system.) *For all $s \geq 3$, we have*

$$p_x(N, s) = \binom{s^N}{x}^{-1} \sum_{x_1=0}^{s} \sum_{x_2=0}^{s^2} \cdots \sum_{x_{N-1}=0}^{s^{N-1}} \prod_{n=1}^{N} c_n(s, x_{n-1}, x_n), \tag{5}$$

where $x_0 = 1$ and $x_N = x$.

The expression for the probabilities (5) cannot be simplified but for any fixed parameter it can be computed explicitly. In the case of groups of size 3, (4) reduces to

$$c_n(3, x, 2x + y) = \sum_{i+j=y} \binom{x}{i} \binom{3^n - x}{j} 3^{x-i+j} \quad \text{for } y \leq 3^n, \tag{6}$$

while in the case of groups of size 4, this reduces to

$$c_n(4, x, 3x + y) = \sum_{i+2j+k=y} \binom{x}{i} \binom{4^n - x}{j} \binom{4^n - x - j}{k} 4^{x-i+k} 6^j \quad \text{for } y \leq 2^{2n+1} - x.$$

In Figure 3, which shows the probabilities (5) for different values of the number of levels and group size along with the corresponding probabilities under Galam's assumption, it is revealed that the phase transition is sharper when starting from a fixed number rather than a binomially distributed number of supporters. This property is expected because our model is 'more deterministic'. To comment on this aspect, we observe that the minimal initial configurations needed to win an election are the ones such that each group at each level has either no more than just a majority of supporters or no supporters at all, from which it follows that

$$p_x(N, s) = 0 \quad \text{whenever } x < \left\lceil \left(\tfrac{1}{2}\right)(s+1) \right\rceil^N.$$

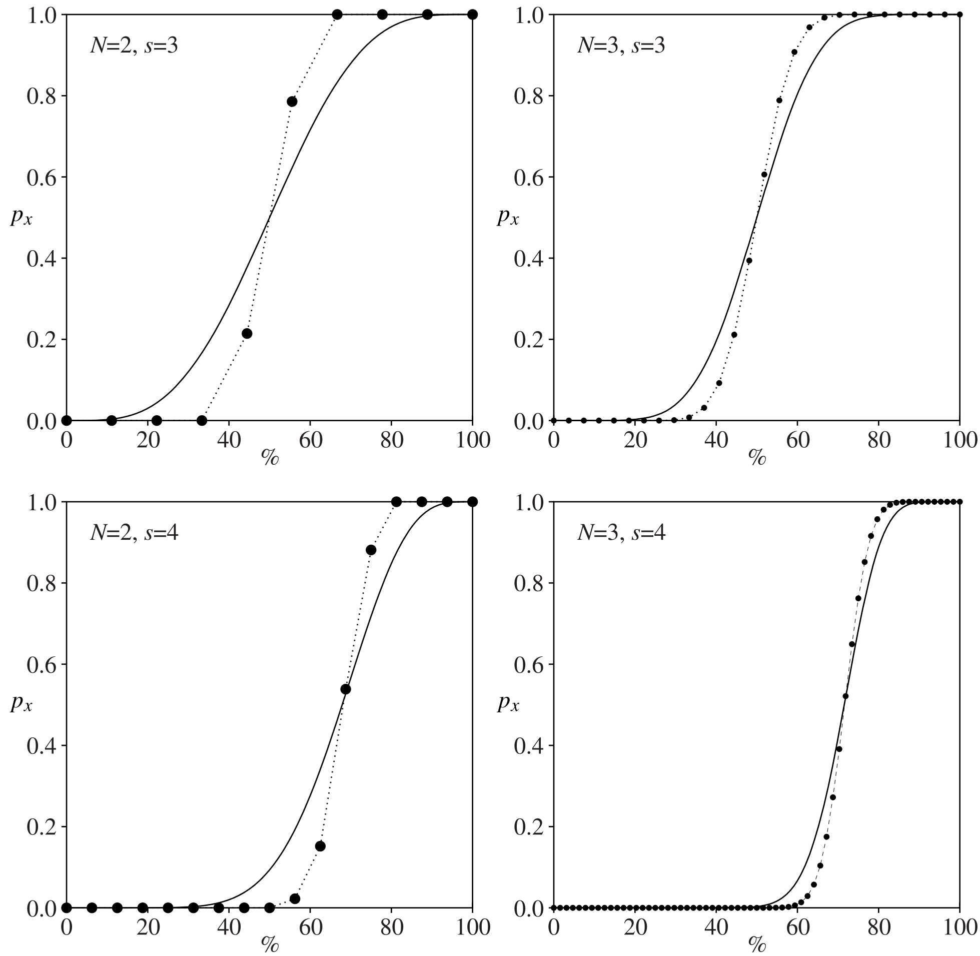

FIGURE 3: Probability that opinion +1 wins as a function of the initial density/number of its supporters at the bottom level of the bottom-up hierarchical system for different values of the number of levels and group size. The continuous black curve is the graph of the function $p \mapsto Q_s^N(p)$ corresponding to the winning probability when assuming that individuals at the bottom level hold independently opinion +1 with probability p. The black dots are the probabilities computed from Theorem 1 when starting from a fixed number of individuals holding opinion +1.

This property can also be deduced from our equation for the winning probabilities but is somewhat hidden. In contrast, starting from a binomial distribution, even when the probability of an individual being in favor of $+1$ at the bottom of the hierarchy is low, there is still a positive probability that the initial number of supporters of that opinion deviates enough from its average to make $+1$ the winner. This explains why the winning probability needs more supporters to take off and, due to the symmetry of the model, why the phase transition is sharper in our context, especially for small population sizes where deviation from the mean is more likely.

Nonspatial public debate model. For the nonspatial public debate model, our main objective is again to determine the winning probability when starting from a fixed number of individuals holding opinion $+1$. Since at each time step all the individuals are equally likely to be part of the chosen discussion group, the actual label on each individual is unimportant. In particular, we simply define X_n as the number of individuals with opinion $+1$ at time n rather than the vector of opinions. With this new definition, the winning probability can be written as

$$\begin{aligned} p_x(N,s) &:= \text{probability that opinion } +1 \text{ wins} \\ &:= \mathbb{P}(X_n = N \text{ for some } n > 0 \mid X_0 = x), \end{aligned}$$

where s is the group size and N is the total number of individuals. We start with the model under the majority rule. In this case, we have the following result, where we use the convention that an empty sum is equal to 0 and that (n choose k) $= 0$ whenever $n < k$.

Theorem 2. (Nonspatial public debate model.) *Under the majority rule,*

$$p_x(N,3) = 2^{-(N-3)} \sum_{z=0}^{x-2} \binom{N-3}{z} \quad \textit{for } x = 0, 1, \ldots, N.$$

In addition, there exists $a_0 > 0$ such that, for all $\epsilon > 0$,

$$\begin{aligned} p_x(N,4) &\le \exp(-a_0 \epsilon N) \quad \textit{for all large } N,\ x < N(c_+ - \epsilon), \\ p_x(N,4) &\ge 1 - \exp(-a_0 \epsilon N) \quad \textit{for all large } N,\ x > N(c_+ + \epsilon). \end{aligned}$$

Note that the first part of the theorem implies that the large deviation estimates also hold when the group size is 3. Indeed, letting $Z = \text{binomial}(N-3, \frac{1}{2})$, we have

$$p_x(N,3) = \sum_{z=0}^{x-2} \binom{N-3}{z} \left(\frac{1}{2}\right)^z \left(\frac{1}{2}\right)^{N-3-z} = \sum_{z=0}^{x-2} \mathbb{P}(Z = z) = \mathbb{P}(Z \le x-2).$$

In particular, it follows from the Chernoff bound that, for all $x < N(\frac{1}{2} - \epsilon)$,

$$p_x(N,3) = \mathbb{P}(Z \le x-2) \le \mathbb{P}\big(Z \le (N-3)\big(\tfrac{1}{2} - \epsilon\big)\big) \le \exp(-2\epsilon^2 (N-3)).$$

Note also that this probability can be written as

$$p_x(N,3) = \frac{\text{card}\,\{A \colon A \subset \{1,2,\ldots,N-3\} \text{ and card}\,(A) \le x-2\}}{\text{card}\,\{A \colon A \subset \{1,2,\ldots,N-3\}\}}.$$

We do not know why the winning probability has this neat combinatorial interpretation but this is what follows from our calculation, which is based on a first-step analysis, a standard

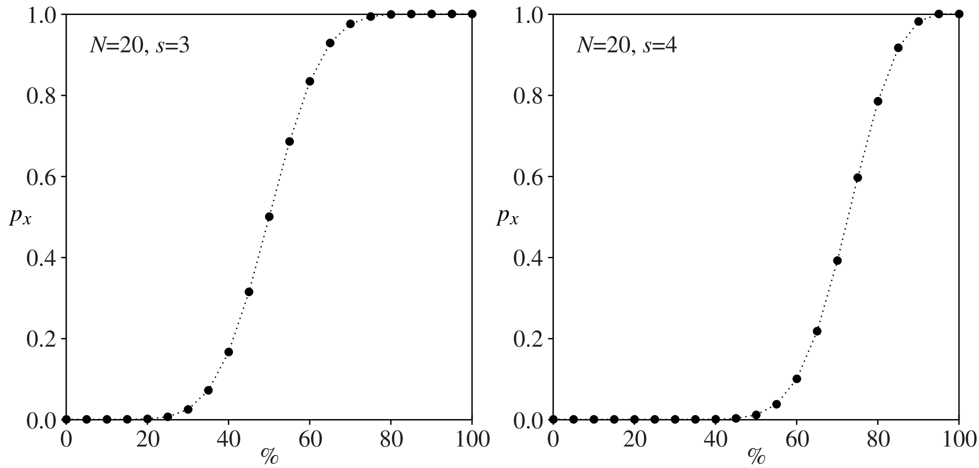

FIGURE 4: Probability that opinion $+1$ wins as a function of the initial number of its supporters in the nonspatial public debate model. The probabilities for $s = 3$ (*left*) are computed from the first part of Theorem 2, whereas the probabilities for $s = 4$ (*right*) are computed recursively from a first-step analysis.

technique in the theory of Markov chains. This technique can also be used to determine the winning probabilities for larger s recursively, which is how the right-hand side of Figure 4 is obtained. However, for $s > 3$, the algebra becomes too complicated to obtain an explicit solution. Based on the first part of the theorem, a natural candidate for the winning probability when $s = 5$ would be

$$\frac{\operatorname{card}\{A\colon A \subset \{1, 2, \ldots, N-5\} \text{ and } \operatorname{card}(A) \leq x-3\}}{\operatorname{card}\{A\colon A \subset \{1, 2, \ldots, N-5\}\}}$$

but we have checked that this expression is not a solution of the order 5 recurrence relation derived from the first-step analysis. Interestingly, from the second part of the theorem, we see that the critical threshold $c_+ \approx 0.768$ obtained under Galam's assumption appears again under our assumption on the initial configuration, though it comes from a different calculation. This result follows in part from an application of the optimal stopping theorem. Turning to the nonspatial public debate model under the proportional rule, first-step analysis is again problematic when the group size exceeds 3. Nevertheless, the winning probabilities can be computed explicitly.

Theorem 3. (Nonspatial public debate model.) *Under the proportional rule,*

$$p_x(N, s) = \frac{x}{N} \quad \textit{for all } s > 1.$$

In words, under the proportional rule, the probability that opinion $+1$ wins is simply equal to the initial fraction of individuals holding this opinion. The proof relies on an application of the optimal stopping theorem after observing that, under the proportional rule, the expected number of individuals in favor of a given opinion is invariant under the dynamics, thus showing that the process is a martingale. Since this property is not sensitive to the group size, we point out that our theorem holds in fact for more general models in which, for instance, the group size is chosen at each time step according to some distribution with values in $\{2, 3, \ldots, N\}$.

Spatial public debate model. Contrary to the nonspatial public debate model, for the spatial version starting with a finite number of individuals with opinion $+1$, the number of such individuals does not evolve according to a Markov chain because the actual locations of these individuals matters. However, under the proportional rule, the auxiliary process that keeps track of the number of individuals with opinion $+1$ is a martingale with respect to the natural filtration of the spatial model. Since it is also integer valued and the population is infinite, it follows from the martingale convergence theorem that opinion $+1$ dies out with probability 1. Therefore, to avoid trivialities, we return to Galam's assumption for the spatial model: we assume that individuals independently support opinion $+1$ with probability $\theta \in (0, 1)$. Since the population is infinite, both opinions are present at any time, and the main objective is now to determine whether they can coexist at equilibrium. The answer depends on the spatial dimension d, as for the voter model [1], [5].

Theorem 4. (Spatial public debate model.) *Under the proportional rule:*

- *the system clusters in $d \leq 2$, i.e.*

$$\lim_{t\to\infty} \mathbb{P}(\eta_t(x) \neq \eta_t(y)) = 0 \quad \text{for all } x, y \in \mathbb{Z}^d;$$

- *both opinions coexist in $d \geq 3$, i.e. η_t converges in distribution to an invariant measure in which there is a positive density of both opinions.*

The proof relies on a certain duality relationship between the spatial model and coalescing random walks, just as for the voter model, though this relationship is somewhat hidden in the case of the public debate model. Before proving our theorems, we point out that the spatial public debate model under the majority rule has also been recently studied in [6]. There it is proved that the one-dimensional process clusters when the group size s is odd, whereas opinion -1 invades the lattice and outcompetes the other opinion when the group size is even. It is also proved, based on a rescaling argument, that opinion -1 wins in two dimensions when $s^2 = 2 \times 2 = 4$.

3. Proof of Theorem 1 (bottom-up hierarchical system)

The main objective is to count the number of configurations at level N with x individuals with opinion $+1$ that will deterministically result in the election of type $+1$ presidents after N consecutive voting steps. Even though the evolution rules of the voting system are deterministic (recall that the model is only stochastic through its random initial configuration), our approach is somewhat reminiscent of the theory of Markov chains. The idea is to reverse time by thinking of the type of the president at level 0 as the initial state, and more generally the configuration at level n as the state at time n. In the theory of discrete-time Markov chains, the distribution at time n, given the initial state, can be computed by looking at the nth power of the transition matrix, which keeps track of the probabilities of all possible sample paths that connect two particular states in n time steps. To this extend, the right-hand side of (5) can be seen as the analog of the nth power of a transition matrix, or Chapman–Kolmogorov's equation, with, however, two exceptions. First, (5) is more complicated because the number of individuals per level is not constant and therefore the evolution rules are not homogeneous in time. Second, and more importantly, the transition probability from $x \to z$ at time n is replaced by an integer, namely

$$c_n(s,x,z) := \text{the number of configurations with } z \text{ individuals holding opinion } +1 \text{ at level } n+1 \text{ that results in a given configuration with } x \text{ individuals holding opinion } +1 \text{ at level } n. \quad (7)$$

By thinking of the bottom-up hierarchical system going backwards in time, the question becomes: how many configurations with x individuals of type $+1$ at time/level N result from the initial configuration +1 at time/level 0, which corresponds to the victory of type $+1$ presidents. To make the argument rigorous and to prove (5), we first define

$$\text{card}\, X := \text{card}\,\{i \in \{1,2,\dots,s^n\} : X(i) = 1\} \quad \text{for all } X \in \Lambda_{s^n} := \{-1,+1\}^{s^n}.$$

Recall that if $Z \in \Lambda_{s^{n+1}}$ then the configuration X at level n is given by

$$X(i) := \text{sign}\left(\sum_{j=1}^{s} Z(s(i-1)+j) - \frac{1}{2}\right) \quad \text{for all } i = 1,2,\dots,s^n,$$

which we write as $Z \to X$. We also say that configuration Z induces configuration X. More generally, we say that configuration $Z \in \Lambda_{s^m}$ induces configuration $X \in \Lambda_{s^n}$ if for all $i \in \{n, n+1, \dots, m-1\}$ there exists $X^i \in \Lambda_{s^i}$ such that $X^{i+1} \to X^i$, where $X^m = Z$ and $X^n = X$, which we again write as $Z \to X$. Finally, we let

$$c_n(s,X,z) := \text{card}\,\{Z \in \Lambda_{s^{n+1}} : Z \to X \text{and} \text{card}\, Z = z\} \quad \text{for all } X \in \Lambda_{s^n}$$

denote the number of configurations with z individuals of type $+1$ at level $n+1$ that induce configuration X at level n. The first key is that $c_n(s,X,z)$ depends only on the number of type $+1$ individuals in configuration X, which is proved in the following lemma.

Lemma 1. *Let $X, Y \in \Lambda_{s^n}$. Then,*

$$\text{card}\, X = \text{card}\, Y \Longrightarrow c_n(s,X,z) = c_n(s,Y,z).$$

Proof. Since $\text{card}\, X = \text{card}\, Y$, there exists $\sigma \in \mathfrak{S}_{s^n}$ such that

$$Y(i) = X(\sigma(i)) \quad \text{for } i = 1,2,\dots,s^n,$$

where $\mathfrak{S}_{s^n}$ denotes the permutation group. Using the permutation σ, we then construct an endomorphism on the set of configurations at level $n+1$ by setting

$$(\phi(Z))(s(i-1)+j) := Z(s(\sigma(i)-1)+j) \quad \text{for } i = 1,2,\dots,s^n,\ j = 1,2,\dots,s.$$

In words, partitioning configurations into s^n consecutive blocks of size s, we set the ith block of $\phi(Z) :=$ the $\sigma(i)$th block of Z for all $i = 1,2,\dots,s^n$. Now, we observe that

$$Z \longrightarrow X \Longleftrightarrow \begin{cases} X(i) = \text{sign}\left(\sum_{j=1}^{s} Z(s(i-1)+j) - \frac{1}{2}\right) & \text{for all } i, \\ X(\sigma(i)) = \text{sign}\left(\sum_{j=1}^{s} Z(s(\sigma(i)-1)+j) - \frac{1}{2}\right) & \text{for all } i, \\ Y(i) = \text{sign}\left(\sum_{j=1}^{s} (\phi(Z))(s(i-1)+j) - \frac{1}{2}\right) & \text{for all } i, \\ \phi(Z) \longrightarrow Y. \end{cases}$$

Since in addition card $Z = \operatorname{card} \phi(Z)$, which directly follows from the fact that $\phi(Z)$ is obtained from a permutation of the blocks of size s in Z, we deduce that

$$\phi(\{Z \colon Z \to X \text{ and card } Z = z\}) \subset \{Z \colon Z \to Y \text{ and card } Z = z\}.$$

That is, for all Z in the first set, $\phi(Z)$ is a configuration in the second set. To conclude, we observe that the function ϕ is an injection from the first set to the second set. We have

$$Z \neq Z' \implies \begin{cases} Z(s(i-1)+j) \neq Z'(s(i-1)+j) & \text{for some } i, j, \\ Z(s(\sigma(i)-1)+j) \neq Z'(s(\sigma(i)-1)+j) & \text{for some } i, j, \\ (\phi(Z))(s(i-1)+j) \neq (\phi(Z'))(s(i-1)+j) & \text{for some } i, j, \\ \phi(Z) \neq \phi(Z'). \end{cases}$$

The injectivity of ϕ implies that

$$\begin{aligned} c_n(s, X, z) &= \operatorname{card}\{Z \in \Lambda_{s^{n+1}} \colon Z \to X \text{ and card } Z = z\} \\ &\leq \operatorname{card}\{Z \in \Lambda_{s^{n+1}} \colon Z \to Y \text{ and card } Z = z\} \\ &= c_n(s, Y, Z). \end{aligned}$$

In particular, the lemma follows from the obvious symmetry of the problem.

In view of Lemma 1, for all $x \in \{0, 1, \ldots, s^n\}$, we can write

$$c_n(s, X, z) := c_n(s, x, z) \quad \text{for all } X \in \Lambda_{s^n} \text{ with card } X = x.$$

The interpretation of $c_n(s, x, z)$ is given in (7). The next step in establishing (5) is given by the following lemma, which follows from the deterministic nature of the evolution rules.

Lemma 2. *Let $X, Y \in \Lambda_{s^n}$. Then*

$$X \neq Y \implies \{Z \in \Lambda_{s^{n+1}} \colon Z \to X\} \cap \{Z \in \Lambda_{s^{n+1}} \colon Z \to Y\} = \varnothing.$$

Proof. To begin with, observe that the assumption implies that

$$X(i) \neq Y(i) \quad \text{for some } i = 1, 2, \ldots, s^n.$$

In particular, if $Z \to X$ and $Z' \to Y$, then, for this specific i, we have

$$\begin{aligned} X(i) &= \operatorname{sign}\biggl(\sum_{j=1}^{s} Z(s(i-1)+j) - \frac{1}{2}\biggr) \\ &\neq \operatorname{sign}\biggl(\sum_{j=1}^{s} Z'(s(i-1)+j) - \frac{1}{2}\biggr) \\ &= Y(i), \end{aligned}$$

which in turn implies that

$$Z(s(i-1)+j) \neq Z'(s(i-1)+j) \quad \text{for some } j = 1, 2, \ldots, s.$$

In conclusion, $Z \neq Z'$. This completes the proof.

https://doi.org/10.1239/aap/1444308877 Published online by Cambridge University Press

Recalling (7) and using the fact that there is only one configuration at level 0 in which type $+1$ is president, as well as the previous lemma, we deduce that the product

$$c_1(s, x_0, x_1)c_2(s, x_1, x_2)\cdots c_{N-1}(s, x_{N-2}, x_{N-1})c_N(s, x_{N-1}, x_N)$$

is the number of configurations with x_N type $+1$ individuals at level N that consecutively induce a configuration with x_n type $+1$ individuals at level n. The number of configurations with x type $+1$ individuals at level N that result in the election of type $+1$ is then obtained by setting $x_0 = 1$ and $x_N = x$ and by summing over all the possible values of the other x_n. Therefore,

$$\operatorname{card}\{X \in \Lambda_{s^N} : X \to (1) \text{ and } \operatorname{card} X = x\} = \sum_{x_1=0}^{s} \sum_{x_2=0}^{s^2} \cdots \sum_{x_{N-1}=0}^{s^{N-1}} \prod_{n=1}^{N} c_n(s, x_{n-1}, x_n).$$

As previously explained, this equation can be seen as the analog of Chapman–Kolmogorov's equation for time-heterogeneous Markov chains, though it represents a number of configurations rather than transition probabilities. Finally, since there are s^N choose x configurations with exactly x type $+1$ individuals at level N, we deduce that the conditional probability that type $+1$ is elected, given that there are x type $+1$ individuals at the bottom of the hierarchy, is

$$p_x(N, s) = \binom{s^N}{x}^{-1} \sum_{x_1=0}^{s} \sum_{x_2=0}^{s^2} \cdots \sum_{x_{N-1}=0}^{s^{N-1}} \prod_{n=1}^{N} c_n(s, x_{n-1}, x_n).$$

To complete the proof of the theorem, the last step is to compute $c_n(s, x, z)$. We start by proving (6), the special case when $s = 3$.

Lemma 3. *For all $y \in \{0, 1, \dots, 3^n\}$, we have*

$$c_n(3, x, 2x + y) = \sum_{i+j=y} \binom{x}{i} \binom{3^n - x}{j} 3^{x-i+j}.$$

Proof. Fix $X \in \Lambda_{s^n}$ with $\operatorname{card} X = x$. Assume that $Z \to X$ and let z_j denote the number of blocks of size 3 with exactly j type $+1$ individuals, i.e.

$$z_j := \operatorname{card}\left\{i : \sum_{k=1}^{3} Z(3(i-1)+k) = j - (3-j)\right\} \quad \text{for all } j = 0, 1, 2, 3.$$

The fact that $\operatorname{card} X = x$ imposes

$$z_0 + z_1 = 3^n - x, \qquad z_2 + z_3 = x \tag{8}$$

implies that, for configuration Z,

- there are x choose z_3 permutations of the blocks with 2 or 3 type $+1$ individuals,
- there are $3^n - x$ choose z_1 permutations of the blocks with 0 or 1 type $+1$ individuals,
- there are 3 choose j possible blocks of size 3 with j type $+1$ individuals.

https://doi.org/10.1239/aap/1444308877 Published online by Cambridge University Press

In particular, the number of $Z \to X$ with z_j blocks with j type $+1$ individuals is

$$\binom{x}{z_3}\binom{3^n - x}{z_1}\prod_{j=0}^{3}\binom{3}{j}^{z_j} = \binom{x}{z_3}\binom{3^n - x}{z_1}3^{z_1+z_2}. \tag{9}$$

Using again (8) and the definition of z_j also implies that

$$\begin{aligned}\operatorname{card} Z &= z_1 + 2z_2 + 3z_3\\ &= z_1 + 2(x - z_3) + 3z_3\\ &= 2x + z_1 + z_3\\ &\in \{2x, 2x+1, \dots, 2x + 3^n\},\end{aligned}$$

which provides the range for y in the statement of the lemma and

$$y := (\operatorname{card} Z) - 2x = z_1 + z_3.$$

From this, together with (9) and $z_1 + z_2 = z_1 + x - z_3$, we finally obtain

$$c_n(3, x, 2x + y) = \sum_{z_1+z_3=y}\binom{x}{z_3}\binom{3^n - x}{z_1}3^{z_1+z_2} = \sum_{z_1+z_3=y}\binom{x}{z_3}\binom{3^n - x}{z_1}3^{x-z_3+z_1}.$$

This completes the proof.

Following the same approach, we now prove the general case (4).

Lemma 4. *For all (s, n) and all $y \in I_{s,n}$ as defined in (3), we have*

$$c_n(s, x, s'x + y) = \sum_{z_0,\dots,z_s}\binom{s^n}{x}^{-1}\binom{s^n}{z_0, z_1, \dots, z_s}\prod_{j=0}^{s}\binom{s}{j}^{z_j},$$

where the sum is over all $z_0, z_1, \dots, z_s$ such that

$$z_0 + z_1 + \cdots + z_{s'-1} = s^n - x, \qquad z_{s'} + z_{s'+1} + \cdots + z_s = x,$$

and such that

$$\sum_{j=1,2,\dots,s'-1} j(z_j + z_{s'+j}) = y \quad \textit{if } s \textit{ is odd},$$

$$\sum_{j=1,2,\dots,s'-2} j(z_j + z_{s'+j}) + (s'-1)\, z_{s'-1} = y \quad \textit{if } s \textit{ is even}.$$

Proof. Again, we fix $X \in \Lambda_{s^n}$ with card $X = x$, let $Z \to X$ and

$$z_j := \operatorname{card}\left\{i : \sum_{k=1}^{s} Z(s(i-1)+k) = j - (s-j)\right\} \quad \text{for all } j = 0, 1, \dots, s.$$

Card $X = x$ now imposes

$$z_0 + z_1 + \cdots + z_{s'-1} = s^n - x, \qquad z_{s'} + z_{s'+1} + \cdots + z_s = x. \tag{10}$$

https://doi.org/10.1239/aap/1444308877 Published online by Cambridge University Press

This and the definition of z_j imply that

$$\begin{aligned}
\text{card } Z &= (z_1 + 2z_2 + \cdots + (s'-1)z_{s'-1}) + (s'z_{s'} + \cdots + sz_s) \\
&= (z_1 + 2z_2 + \cdots + (s'-1)z_{s'-1}) + s'(x - z_{s'+1} - \cdots - z_s) \\
&\quad + ((s'+1)z_{s'+1} + \cdots + s\, z_s) \\
&= s'x + z_1 + 2z_2 + \cdots + (s'-1)z_{s'-1} + z_{s'+1} + 2\, z_{s'+2} + \cdots + (s-s')z_s, \\
\text{card } Z &\in s'x + \{0, 1, \ldots, (s'-1)(s^n - x) + (s-s')x\},
\end{aligned}$$

which provides the range for y. Rearranging the terms, we obtain

$$y \;:=\; (\text{card } Z) - s'x = \begin{cases} \displaystyle\sum_{j=1,2,\ldots,s'-1} j(z_j + z_{s'+j}) & \text{if } s \text{ is odd,} \\ \displaystyle\sum_{j=1,2,\ldots,s'-2} j(z_j + z_{s'+j}) + (s'-1)z_{s'-1} & \text{if } s \text{ is even.} \end{cases} \tag{11}$$

Now, using again (10), the number of permutations of the blocks with at least s' type $+1$ individuals is equal to

$$\binom{x}{z_{s'}, \ldots, z_s} := \frac{x!}{z_{s'}! \cdots z_s!}$$

and the number of permutations of the blocks with at most $s'-1$ type $+1$ individuals is equal to

$$\binom{s^n - x}{z_0, \ldots, z_{s'-1}} := \frac{(s^n - x)!}{z_0! \cdots z_{s'-1}!}.$$

Since there are s choose j possible blocks of size s with j type $+1$ individuals, the number of configurations with z_j blocks with j type $+1$ individuals that induce X is then

$$\binom{x}{z_{s'}, \ldots, z_s}\binom{s^n - x}{z_0, \ldots, z_{s'-1}} \prod_{j=0}^{s} \binom{s}{j}^{z_j} = \binom{s^n}{x}^{-1} \binom{s^n}{z_0, z_1, \ldots, z_s} \prod_{j=0}^{s} \binom{s}{j}^{z_j}.$$

This implies that, for all suitable y,

$$c_n(s, x, s'x + y) = \sum_{z_0, \ldots, z_s} \binom{s^n}{x}^{-1} \binom{s^n}{z_0, z_1, \ldots, z_s} \prod_{j=0}^{s} \binom{s}{j}^{z_j},$$

where the sum is over all $z_0, z_1, \ldots, z_s$ such that (10) and (11) hold.

4. Proof of Theorems 2 and 3 (nonspatial public debate model)

This section is devoted to the proof of Theorems 2 and 3 which deal with the nonspatial public debate model. There is no more hierarchical structure and the evolution rules are now stochastic. At each time step, s distinct individuals are chosen uniformly at random to form a discussion group, which results in all the individuals within the group reaching a consensus. The new opinion is chosen according to either the majority rule or the proportional rule.

Majority rule and size 3. In this case, the process can be understood by simply using a first-step analysis whose basic idea is to condition on all the possible outcomes of the first update and then use the Markov property to find a relationship among the winning probabilities for the process starting from different states. We point out that this approach is tractable only when $s = 3$, due to a small number of possible outcomes at each update.

Lemma 5. *Under the majority rule, we have*

$$p_x(N,3) = 2^{-(N-3)} \sum_{z=0}^{x-2} \binom{N-3}{z} \quad \textit{for all } x = 0, 1, \ldots, N.$$

Proof. Note that the winning probability is obvious when $x = 0, 1, N-1, N$, therefore we focus only on the other cases. The first step is to exhibit a relationship among the probabilities to be found by conditioning on all the possible outcomes of the first update. Recall that

$$p_x := p_x(N,3) = \mathbb{P}(X_n = N \text{ for some } n \mid X_0 = x)$$

and, for $x = 2, 3, \ldots, N-2$, let $\mu_x := q_{-1}(x)/q_1(x)$, where

$$q_j(x) := \mathbb{P}(X_{n+1} = x + j \mid X_n = x) \quad \text{for } j = -1, 1.$$

Conditioning on the possible values for X_1 and using the Markov property, we obtain

$$\begin{aligned} p_x &= \mathbb{P}(X_n = N \text{ for some } n \mid X_1 = x-1)\mathbb{P}(X_1 = x-1 \mid X_0 = x) \\ &\quad + \mathbb{P}(X_n = N \text{ for some } n \mid X_1 = x)\mathbb{P}(X_1 = x \mid X_0 = x) \\ &\quad + \mathbb{P}(X_n = N \text{ for some } n \mid X_1 = x+1)\mathbb{P}(X_1 = x+1 \mid X_0 = x) \\ &= q_{-1}(x)p_{x-1} + (1 - q_{-1}(x) - q_1(x))p_x + q_1(x)p_{x+1}. \end{aligned}$$

In particular, $q_1(x)(p_{x+1} - p_x) = q_{-1}(x)(p_x - p_{x-1})$, so from a simple induction, we obtain

$$\begin{aligned} p_{x+1} - p_x &= \mu_x\,(p_x - p_{x-1}) \\ &= \mu_x \mu_{x-1}(p_{x-1} - p_{x-2}) \\ &= \cdots \\ &= \mu_x \mu_{x-1} \cdots \mu_2(p_2 - p_1) \\ &= \mu_x\, \mu_{x-1} \cdots \mu_2 p_2. \end{aligned}$$

Using again $p_1 = 0$, it follows that

$$p_x = \sum_{z=1,2,\ldots,x-1} (p_{z+1} - p_z) = \Bigg(1 + \sum_{z=2,3,\ldots,x-1} \mu_2\mu_3 \cdots \mu_z\Bigg) p_2. \tag{12}$$

Now, using $p_{N-1} = 1$, we obtain

$$p_{N-1} = \Bigg(1 + \sum_{z=2,3,\ldots,N-2} \mu_2\mu_3 \cdots \mu_z\Bigg) p_2 = 1. \tag{13}$$

Combining (12) and (13), we obtain

$$\begin{aligned} p_x &= \Bigg(1 + \sum_{z=2,\ldots,x-1} \mu_2\mu_3 \cdots \mu_z\Bigg) p_2 \\ &= \Bigg(1 + \sum_{z=2,\ldots,x-1} \mu_2\mu_3 \cdots \mu_z\Bigg)\Bigg(1 + \sum_{z=2,\ldots,N-2} \mu_2\mu_3 \cdots \mu_z\Bigg)^{-1}. \end{aligned} \tag{14}$$

To find an explicit expression for (14), the last step is to compute $q_{-1}(x)$ and $q_1(x)$. Observing that these two probabilities are respectively the probability of selecting a group with one type $+1$ individual and the probability of selecting a group with two type $+1$ individuals, we obtain

$$q_{-1}(x) = \binom{N}{3}^{-1} \binom{x}{1} \binom{N-x}{2}, \qquad q_1(x) = \binom{N}{3}^{-1} \binom{x}{2} \binom{N-x}{1}.$$

From this, we have the following expression for the ratio:

$$\mu_x := \frac{q_{-1}(x)}{q_1(x)} = \frac{x(N-x)(N-x-1)}{x(x-1)(N-x)} = \frac{N-x-1}{x-1}$$

for $x = 2, 3, \ldots, N-2$. For the product, we have the following expression:

$$\begin{aligned} \mu_2 \mu_3 \cdots \mu_z &= \frac{N-3}{1} \, \frac{N-4}{2} \cdots \frac{N-z-1}{z-1} \\ &= \frac{(N-3)!}{(z-1)!(N-z-2)!} \\ &= \binom{N-3}{z-1} \end{aligned} \tag{15}$$

for $z = 2, 3, \ldots, N-2$. Finally, combining (14) and (15), we obtain

$$\begin{aligned} p_x &= \left(1 + \sum_{z=2}^{x-1} \binom{N-3}{z-1}\right) \left(1 + \sum_{z=2}^{N-2} \binom{N-3}{z-1}\right)^{-1} \\ &= \left(\sum_{z=0}^{N-3} \binom{N-3}{z}\right)^{-1} \sum_{z=0}^{x-2} \binom{N-3}{z} \\ &= 2^{-(N-3)} \sum_{z=0}^{x-2} \binom{N-3}{z} \end{aligned}$$

for all $x \in \{2, \ldots, N-2\}$. This completes the proof.

Majority rule and size 4. Increasing the common size of the discussion groups, a first-step analysis can again be used to find a recursive formula for the winning probabilities but the algebra becomes too messy to deduce an explicit formula. Instead, we prove lower and upper bounds for the winning probabilities using the optimal stopping theorem for supermartingales. To describe more precisely our approach, consider the transition probabilities

$$q_j(x) := \mathbb{P}(X_{n+1} - X_n = j \mid X_n = x) \quad \text{for } j = -2, -1, 0, 1$$

as well as the new Markov chain (Z_n) with transition probabilities

$$p(0,0) = p(N,N) = 1, \qquad p(x, x+j) = q_j(x)\,(q_1(x) + q_{-1}(x) + q_{-2}(x))^{-1}$$

for all $x = 1, 2, \ldots, N-1$ and all $j = -2, -1, 1$. The process (Z_n) can be seen as the random sequence of states visited by the public debate model until fixation. In particular,

$$p_x(N,4) := \mathbb{P}(X_n = N \text{ for some } n > 0) = \mathbb{P}(Z_n = N \text{ for some } n > 0).$$

The main idea of the proof is to first identify exponentials of the process (Z_n) that are supermartingales and then apply the optimal stopping theorem to these processes. We start by proving that the drift of the Markov chain is either negative or positive, depending on whether the number of individuals in favor of the $+1$ opinion is smaller or larger than $c_+ \approx 0.768$. In particular, using a different calculation, we recover the critical threshold c_+ obtained by Galam.

Lemma 6. *For all* $\epsilon > 0$,

$$\mathbb{E}(Z_{n+1} - Z_n \mid Z_n = x) \begin{cases} \leq -\left(\frac{1}{2}\right)(\sqrt{13} - 1)\epsilon + O(N^{-1}) & \textit{for } x < N(c_+ - \epsilon), \\ \geq +\sqrt{13}\epsilon + O(N^{-1}) & \textit{for } x > N(c_+ + \epsilon). \end{cases}$$

Proof. Observing that $q_{-2}(x)$ is the probability that a randomly chosen group of size 4 has two individuals in favor and two individuals against the $+1$ opinion, we obtain

$$q_{-2}(x) = \binom{N}{4}^{-1} \frac{x(x-1)}{2} \frac{(N-x)(N-x-1)}{2} = 6c^2(1-c)^2 + O(N^{-1})$$

provided that $x = \text{round}(cN)$. Similarly, we show that

$$q_1(x) = 4c^3(1-c) + O(N^{-1}), \qquad q_{-1}(x) = 4c(1-c)^3 + O(N^{-1}),$$

from which it follows that

$$\begin{aligned} q_1(x) - q_{-1}(x) - 2q_{-2}(x) &= 4c(1-c)(3c^2 - c - 1) + O(N^{-1}), \\ q_1(x) + q_{-1}(x) + q_{-2}(x) &= 2c(1-c)(c^2 - c + 2) + O(N^{-1}). \end{aligned}$$

Taking the ratio of the previous two estimates, we have

$$\mathbb{E}(Z_{n+1} - Z_n \mid Z_n = x) = 6(c - c_-)(c - c_+)(c^2 - c + 2)^{-1} + O(N^{-1}),$$

from which we deduce that

$$\begin{aligned} \mathbb{E}(Z_{n+1} - Z_n \mid Z_n = x) &\leq 3(-c_-)(-\epsilon) + O(N^{-1}) \\ &= -\left(\tfrac{1}{2}\right)(\sqrt{13} - 1)\epsilon + O(N^{-1}) \quad \text{for } x < N(c_+ - \epsilon), \\ \mathbb{E}(Z_{n+1} - Z_n \mid Z_n = x) &\geq 3(c_+ - c_-)\epsilon + O(N^{-1}) \\ &= \sqrt{13}\epsilon + O(N^{-1}) \quad \text{for } x > N(c_+ + \epsilon). \end{aligned}$$

This completes the proof.

Lemma 7. *There exists* $\bar{a} > 0$ *such that, for all* $\epsilon > 0$,

$$\begin{aligned} \mathbb{E}(\exp(+\bar{a}Z_{n+1}) - \exp(+\bar{a}Z_n) \mid Z_n = x) \leq 0 \quad &\textit{for } x < N(c_+ - \epsilon), \\ \mathbb{E}(\exp(-\bar{a}Z_{n+1}) - \exp(-\bar{a}Z_n) \mid Z_n = x) \leq 0 \quad &\textit{for } x > N(c_+ + \epsilon) \end{aligned}$$

for all sufficiently large N.

Proof. To begin with, we define the functions

$$\phi_x(a) := \mathbb{E}(\exp(aZ_{n+1}) - \exp(aZ_n) \mid Z_n = x).$$

https://doi.org/10.1239/aap/1444308877 Published online by Cambridge University Press

Differentiating then applying Lemma 6, we obtain

$$\phi_x'(a) = \mathbb{E}(Z_{n+1}\exp(aZ_{n+1}) - Z_n\exp(aZ_n) \mid Z_n = x),$$
$$\phi_x'(0) = \mathbb{E}(Z_{n+1} - Z_n \mid Z_n = x) \le -\left(\tfrac{1}{2}\right)(\sqrt{13}-1)\epsilon + O(N^{-1}) < 0$$

for all $x < (c_+ - \epsilon)N$ and large N. Since $\phi_x(0) = 0$, there is $a_+ > 0$ such that

$$\phi_x(+a) \le 0 \quad \text{for all } a \in (0, a_+) \text{ and all } x < N(c_+ - \epsilon).$$

Differentiating $a \mapsto \phi_x(-a)$ and using Lemma 6, we also have

$$\phi_x(-a) \le 0 \quad \text{for all } a \in (0, a_-) \text{ and all } x > N(c_+ + \epsilon)$$

for some $a_- > 0$. In particular, for $\bar{a} := \min(a_+, a_-) > 0$,

$$\phi_x(+\bar{a}) \le 0 \quad \text{for all } x < N(c_+ - \epsilon), \qquad \phi_x(-\bar{a}) \le 0 \quad \text{for all } x > N(c_+ + \epsilon),$$

which, recalling the definition of ϕ_x, is exactly the statement of the lemma.

With Lemma 7 in hand, we are now ready to prove the upper and lower bounds for the winning probabilities using the optimal stopping theorem.

Lemma 8. *Let $a_0 = (\frac{1}{2})\bar{a}$. Then, for all $\epsilon > 0$,*

$$p_x(N, 4) \le \exp(-a_0\epsilon N) \quad \textit{for all } N \textit{ large and } x < N(c_+ - \epsilon).$$

Proof. First, we introduce the stopping times

$$\tau_0 := \inf\{n \colon Z_n = 0\}, \qquad \tau_- := \inf\{n \colon Z_n > N(c_+ - \epsilon)\}$$

as well as $T_- := \min(\tau_0, \tau_-)$. Since the process $\exp(\bar{a}\, Z_n)$ stopped at time T_- is a supermartingale according to the first assertion in Lemma 7, and the stopping time T_- is almost surely finite, the optimal stopping theorem implies that

$$\begin{aligned} \mathbb{E}(\exp(\bar{a}Z_{T_-}) \mid Z_0 = x) &\le \mathbb{E}(\exp(\bar{a}Z_0) \mid Z_0 = x) \\ &\le \exp(\bar{a}(c_+ - 2\epsilon)N) \quad \text{for all } x < N(c_+ - 2\epsilon). \end{aligned} \tag{16}$$

In addition,

$$\begin{aligned} \mathbb{E}(\exp(\bar{a}Z_{T_-})) &= \mathbb{E}(\exp(\bar{a}Z_{T_-}) \mid T_- = \tau_0)\mathbb{P}(T_- = \tau_0) \\ &\quad + \mathbb{E}(\exp(\bar{a}Z_{T_-}) \mid T_- = \tau_-)\mathbb{P}(T_- = \tau_-) \\ &\ge \mathbb{P}(T_- = \tau_0) + \exp(a(c_+ - \epsilon)N)\mathbb{P}(T_- \ne \tau_0) \\ &= 1 - (1 - \exp(\bar{a}(c_+ - \epsilon)N))\mathbb{P}(T_- \ne \tau_0). \end{aligned} \tag{17}$$

Noting that opinion +1 wins only if $T_- \ne \tau_0$ and combining (16) and (17), we have

$$\begin{aligned} p_x(N, 4) &\le \mathbb{P}(T_- \ne \tau_0) \\ &\le (\exp(\bar{a}(c_+ - 2\epsilon)N) - 1)(\exp(\bar{a}(c_+ - \epsilon)N) - 1)^{-1} \\ &\le \exp(\bar{a}(c_+ - 2\epsilon)N)(\exp(\bar{a}(c_+ - \epsilon)N))^{-1} \\ &= \exp(-\bar{a}\epsilon N) \quad \text{for all } x < N(c_+ - 2\epsilon) \end{aligned}$$

and all sufficiently large N.

Lemma 9. *Let* $a_0 = (\frac{1}{2})\bar{a}$. *Then, for all* $\epsilon > 0$,

$$p_x(N, 4) \geq 1 - \exp(-a_0 \epsilon N) \quad \text{for all large } N \text{ and } x > N(c_+ + \epsilon).$$

Proof. This follows from the same arguments as in the proof of Lemma 8 but applying the optimal stopping theorem at the stopping time $T_+ := \min(\tau_N, \tau_+)$, where

$$\tau_N := \inf\{n : Z_n = N\}, \qquad \tau_+ := \inf\{n : Z_n < N(c_+ + \epsilon)\}$$

instead of T_- and using the second instead of the first assertion in Lemma 7.

Proportional rule. We now prove Theorem 3, which deals with the nonspatial public debate model under the proportional rule. As previously, a first-step analysis does not allow us to find an explicit expression for the winning probabilities, but the result can be easily deduced from the optimal stopping theorem, observing that the number of individuals in favor of a given opinion is a martingale with respect to the natural filtration of the process.

Lemma 10. *Under the proportional rule, we have* $p_x(N, s) = x/N$.

Proof. To begin with, we observe that each time a group is chosen uniformly from the population, regardless of the random number j of individuals with opinion +1 the group contains, the expected value of the variation in the number of +1 individuals is always

$$\begin{aligned}\mathbb{E}(X_{n+1} - X_n) &= (s - j) \times (\text{fraction of } +1 \text{ individuals in the group}) \\ &\quad - j \times (\text{fraction of } -1 \text{ individuals in the group}) \\ &= (s - j) j s^{-1} - j(s - j) s^{-1} \\ &= 0,\end{aligned}$$

which shows that the process (X_n) is a martingale. Now, let

$$T_+ := \inf\{n \colon X_n = N\}, \qquad T_- := \inf\{n \colon X_n = 0\}$$

and observe that the stopping time $T := \min(T_+, T_-)$ is almost surely finite. Since, in addition, the process is bounded, an application of the optimal stopping theorem implies that

$$\begin{aligned}\mathbb{E}(X_T \mid X_0 = x) &= \mathbb{E}(X_0 \mid X_0 = x) \\ &= x \\ &= N\mathbb{P}(T = T_+) + 0\mathbb{P}(T = T_-) \\ &= N p_x(N, s)\end{aligned}$$

from which it follows that the winning probability $p_x(N, s) = x/N$.

5. Proof of Theorem 4 (spatial public debate model)

To conclude, we study the spatial version of the public debate model introduced in [6] but replacing the majority rule with the proportional rule. The key to our analysis is similar to the approach used in previous works [1], [5] regarding the voter model. The idea is to construct the process from a so-called Harris' graphical representation and then use the resulting graphical structure to exhibit a relationship between the process and a system of coalescing random walks.

https://doi.org/10.1239/aap/1444308877 Published online by Cambridge University Press

Graphical representations. We first provide a possible graphical representation from which the spatial public debate model can be constructed starting from any initial configuration. Though natural, this graphical representation does not allow us to derive a useful duality relationship between the process and coalescing random walks. We then introduce an alternative way to construct the process leading to such a duality relationship. Recall that

$$\{x + B_s : x \in \mathbb{Z}^d\}, \quad \text{where } B_s := \{0, 1, \ldots, s-1\}^d$$

represents the collection of discussion groups. Each of these groups is updated in continuous time at rate 1, i.e. at the arrival times of independent Poisson processes with intensity 1. In addition, since the new opinion of the group after an update is chosen to be +1 with its probability being the fraction of +1 individuals in the group just before the update, the new opinion can be determined by comparing the fraction of +1 individuals with a uniform random variable over the unit interval. In particular, a natural way to construct the spatial public debate model graphically is to

- let $T_n(x) :=$ the nth arrival time of a Poisson process with rate 1 and
- let $U_n(x) :=$ a uniform random variable over the interval (0, 1)

for all $x \in \mathbb{Z}^d$ and $n > 0$. At time $t := T_n(x)$, all the individuals in $x + B_s$ are simultaneously updated as a result of a discussion and we set

$$\eta_t(y) := 2\,\mathbf{1}_{\{U_n(x) < s^{-d} \sum_{z \in x + B_s} \mathbf{1}\{\eta_{t-}(z) = +1\}\}} - 1 \quad \text{for all } y \in x + B_s, \tag{18}$$

while the configuration outside $x + B_s$ stays unchanged. An idea of Harris [4] implies that the process starting from any initial configuration can be constructed using this rule. We now construct another process (ξ_t) with the same state space as follows: the times at which individuals in the same discussion group interact are defined as above from the same collection of independent Poisson processes, but to determine the outcome of the discussion we now

- let $W_n(x) :=$ a uniform random variable over the set $x + B_s$

for all $x \in \mathbb{Z}^d$ and $n > 0$. At time $t := T_n(x)$, all the individuals in $x + B_s$ are simultaneously updated as a result of a discussion and we set

$$\xi_t(y) := \xi_{t-}(W_n(x)) \quad \text{for all } y \in x + B_s, \tag{19}$$

while the configuration outside $x + B_s$ stays unchanged. The next lemma, whose proof is simply based on a rewriting of events under consideration, shows that both rules (18) and (19) define in fact the same process: the processes (η_t) and (ξ_t) are stochastically equal.

Lemma 11. *The constructions (18) and (19) are equivalent:*

$$\eta_{t-} = \xi_{t-} \Longrightarrow \mathbb{P}(\eta_t(x) = 1) = \mathbb{P}(\xi_t(x) = 1) \quad \textit{for all } x \in \mathbb{Z}^d.$$

Proof. This is only nontrivial for pairs $(x, t) \in \mathbb{Z}^d \times \mathbb{R}_+$ such that

$$t := T_n(z), \qquad x \in z + B_s \quad \text{for some } (z, n) \in \mathbb{Z}^d \times \mathbb{N}^*.$$

In this case, we have

$$\begin{aligned}
\mathbb{P}(\xi_t(x) = 1) &= \mathbb{P}(\xi_{t-}(W_n(z)) = 1) \\
&= \mathbb{P}(W_n(z) \in \{y \in z + B_s : \xi_{t-}(y) = 1\}) \\
&= \frac{\operatorname{card}\{y \in z + B_s : \xi_{t-}(y) = 1\}}{\operatorname{card}(z + B_s)} \\
&= s^{-d} \sum_{y \in z + B_s} \mathbf{1}_{\{\xi_{t-}(y)=1\}} \\
&= \mathbb{P}(U_n(z) < s^{-d} \sum_{y \in z + B_s} \mathbf{1}_{\{\xi_{t-}(y)=1\}}) \\
&= \mathbb{P}(U_n(z) < s^{-d} \sum_{y \in z + B_s} \mathbf{1}_{\{\eta_{t-}(y)=1\}}) \\
&= \mathbb{P}(\eta_t(x) = 1).
\end{aligned}$$

This completes the proof of the lemma.

Duality with coalescing random walks. The duality relationship between the voter model and coalescing random walks results from keeping track of the ancestors of different space-time points going backwards in time through the graphical representation. In the case of the public debate model $\eta.$, the opinion of an individual just after an interaction depends on the opinion of all the individuals in the corresponding discussion group just before the interaction. Therefore, to define the set of ancestors of a given space-time point, we draw an arrow

$$z_1 \longrightarrow z_2 \quad \text{at time } t := T_n(z) \text{ for all } z_1, z_2 \in z + B_s,\ (z, n) \in \mathbb{Z}^d \times \mathbb{N}^*$$

to indicate that the opinion at (z_2, t) depends on the opinion at $(z_1, t-)$, and say that there is an η-path connecting two space-time points, which we write as

$$(y, t - s) \longrightarrow_\eta (x, t) \quad \text{for } x, y \in \mathbb{Z}^d,\ s, t > 0,$$

whenever there are sequences of times and spatial locations

$$t - s < s_1 < s_2 < \cdots < s_{n-1} < t, \qquad z_1 := y, z_2, \ldots, z_n := x \in \mathbb{Z}^d$$

such that there is an arrow

$$z_j \longrightarrow z_{j+1} \quad \text{at time } s_j \text{ for } j = 1, 2, \ldots, n - 1.$$

The set of ancestors of (x, t) at time $t - s$ is then encoded in the set-valued process

$$\hat{\eta}_s(x, t) := \{y \in \mathbb{Z}^d : (y, t - s) \longrightarrow_\eta (x, t)\}. \tag{20}$$

Note that the opinion at (x, t) can be deduced from the graphical representation of $\eta.$ and the initial opinion at sites that belong to $\hat{\eta}_t(x, t)$. Note also that the process (20) grows linearly going backwards in time, i.e. increasing s. See the left-hand side of Figure 5 for a representation. This makes the process $\eta.$ mathematically intractable to prove clustering and coexistence. To establish the connection between the spatial process and coalescing random walks, we use instead the other, mathematically equivalent, version $\xi.$ of the spatial public debate model. For

https://doi.org/10.1239/aap/1444308877 Published online by Cambridge University Press

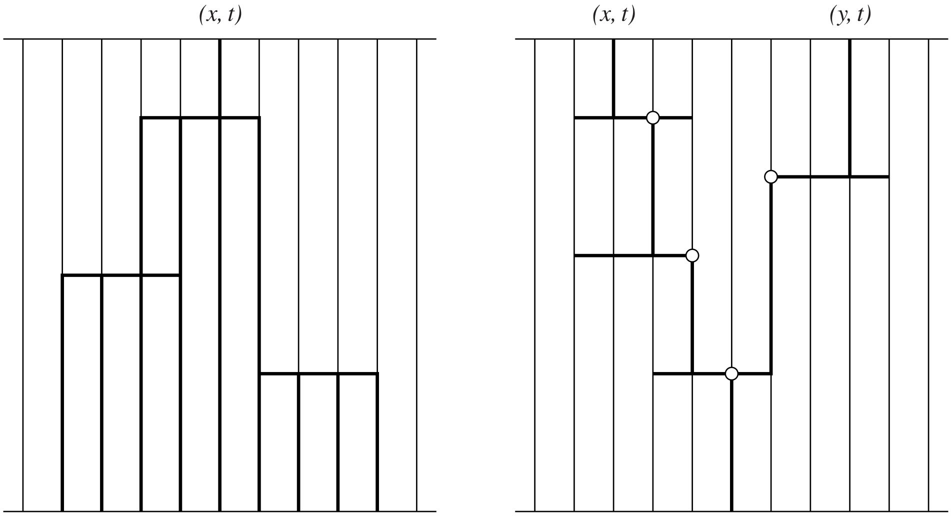

FIGURE 5: The graphical representation for a set of ancestors. In both cases, $s = 4$ and the times at which discussion groups are updated (time goes up) are represented by horizontal line segments (bold), while the set of ancestors is represented by vertical line segments (bold). Shown is an illustration of the set-valued process (20) *(left)* and an illustration of the dual process (21) *(right)* starting from $A = \{x, y\}$. The open circles *(right)* correspond to the value of the uniform W random variables.

this version, the opinion of an individual just after an interaction depends on the opinion of only one individual in the corresponding discussion group just before the interaction. The location of this individual is given by the value of the uniform W random variables. Therefore, to define the set of ancestors of a given space-time point, we now draw an arrow

$$W_n(z) \longrightarrow z' \quad \text{at time } t := T_n(z) \text{ for all } z' \in z + B_s,\ (z, n) \in \mathbb{Z}^d \times \mathbb{N}^*$$

to indicate that the opinion at (z', t) depends on the opinion at $(W_n(z), t-)$. We then define ξ-paths, which we now write '$\longrightarrow_\xi$', as previously but using this new random set of arrows. The set of ancestors of (x, t) at time $t - s$ is now encoded in

$$\hat{\xi}_s(x, t) := \{y \in \mathbb{Z}^d : (y, t - s) \longrightarrow_\xi (x, t)\}. \tag{21}$$

More generally, for $A \subset \mathbb{Z}^d$ finite, we define the dual process starting at (A, t) as

$$\begin{aligned} \hat{\xi}_s(A, t) &:= \{y \in \mathbb{Z}^d : y \in \hat{\xi}_s(x, t) \text{ for some } x \in A\} \\ &:= \{y \in \mathbb{Z}^d : (y, t - s) \longrightarrow_\xi (x, t) \text{ for some } x \in A\}. \end{aligned}$$

See the right-hand side of Figure 5 for a representation. Note that (21) is reduced to a singleton for all times $s \in (0, t)$ and that we have the duality relationship

$$\xi_t(x) = \xi_{t-s}(Z_s(x)) = \xi_0(Z_t(x)) \quad \text{for all } s \in (0, t),$$

where $Z_s(x) := \hat{\xi}_s(x, t)$. In the next lemma, we prove that $Z_s(x)$ is a symmetric random walk, which makes the dual process itself a system of coalescing symmetric random walks with one walk starting from each site in the finite set A.

https://doi.org/10.1239/aap/1444308877 Published online by Cambridge University Press

Lemma 12. *The process $Z_s(x) := \hat{\xi}_s(x, t)$ is a symmetric random walk.*

Proof. By construction of the dual process, for $t - s := T_n(z)$,

$$Z_s(x) := \hat{\xi}_s(x, t) = \begin{cases} Z_{s-}(x) & \text{when } Z_{s-}(x) \notin z + B_s, \\ W_n(z) & \text{when } Z_{s-}(x) \in z + B_s. \end{cases}$$

Since, in addition, discussion groups are updated at rate 1 and

$$\mathbb{P}(W_n(z) = y) = s^{-d} \quad \text{for all } y \in z + B_s,$$

we obtain the following transition rates:

$$\begin{aligned} &\lim_{h \to 0} h^{-1} \mathbb{P}(Z_{s+h}(x) = y + \boldsymbol{w} \mid Z_s(x) = y) \\ &\qquad = s^{-d} \operatorname{card} \{z \in \mathbb{Z}^d : y \in z + B_s \text{ and } y + \boldsymbol{w} \in z + B_s\}. \end{aligned} \tag{22}$$

In addition, since for all $\boldsymbol{w} \in \mathbb{Z}^d$ the translation operator $y \mapsto y + \boldsymbol{w}$ is a one-to-one correspondence from the set of discussion groups to itself and since

$$y, y + \boldsymbol{w} \in z + B_s \iff y - \boldsymbol{w}, y \in (z - \boldsymbol{w}) + B_s,$$

we have the equality

$$\begin{aligned} &\operatorname{card} \{z \in \mathbb{Z}^d : y \in z + B_s \text{ and } y + \boldsymbol{w} \in z + B_s\} \\ &\qquad = \operatorname{card} \{z \in \mathbb{Z}^d : y \in z + B_s \text{ and } y - \boldsymbol{w} \in z + B_s\}. \end{aligned} \tag{23}$$

Combining (22) and (23), we conclude that

$$\begin{aligned} &\lim_{h \to 0} h^{-1} \mathbb{P}(Z_{s+h}(x) = y + \boldsymbol{w} \mid Z_s(x) = y) \\ &\qquad = \lim_{h \to 0} h^{-1} \mathbb{P}(Z_{s+h}(x) = y - \boldsymbol{w} \mid Z_s(x) = y) \quad \text{for all } y, \boldsymbol{w} \in \mathbb{Z}^d, \end{aligned}$$

which completes the proof.

In fact, from some basic geometry, we have

$$\lim_{h \to 0} h^{-1} \mathbb{P}(Z_{s+h}(x) = y + \boldsymbol{w} \mid Z_s(x) = y) = s^{-d} \prod_{j=1}^{d} (s - |w_j|),$$

where w_j is the jth coordinate of the vector $\boldsymbol{w}$. The theorem can be deduced from the previous two lemmas using the exact same approach as for the voter model. For the details on how to deduce clustering and coexistence, see [2, Theorems 3.1] and [2, Theorems 3.2], respectively.

Acknowledgements

The authors would like to thank an anonymous referee for useful suggestions to improve the presentation of the results and the clarity of some of the proofs. N. Lanchier's research is supported in part by the National Science Foundation (grant no. DMS-10-05282).

References

[1] Clifford, P. and Sudbury, A. (1973). A model for spatial conflict. *Biometrika* **60,** 581–588.

[2] Durrett, R. (1995). Ten lectures on particle systems. In *Lectures on Probability Theory* (Saint-Flour, 1993; Lecture Notes Math. **1608**), Springer, Berlin, pp. 97–201.

[3] Galam, S. (2008). Sociophysics: a review of Galam models. *Internat. J. Mod. Phys. C* **19,** 409–440.

[4] Harris, T. E. (1972). Nearest-neighbor Markov interaction processes on multidimensional lattices. *Adv. Math.* **9,** 66–89.

[5] Holley, R. A. and Liggett, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Prob. 3,* 643–663.

[6] Lanchier, N. and Neufer, J. (2013). Stochastic dynamics on hypergraphs and the spatial majority rule model. *J. Statist. Phys.* **151,** 21–45.