

Stochastic Dynamics on Hypergraphs and the Spatial Majority Rule Model

N. Lanchier · J. Neuffer

Received: 11 March 2012 / Accepted: 6 July 2012 / Published online: 18 July 2012
© Springer Science+Business Media, LLC 2012

Abstract This article starts by introducing a new theoretical framework to model spatial systems which is obtained from the framework of interacting particle systems by replacing the traditional graphical structure that defines the network of interactions with a structure of hypergraph. This new perspective is more appropriate to define stochastic spatial processes in which large blocks of vertices may flip simultaneously, which is then applied to define a spatial version of the Galam's majority rule model. In our spatial model, each vertex of the lattice has one of two possible competing opinions, say opinion 0 and opinion 1, as in the popular voter model. Hyperedges are updated at rate one, which results in all the vertices in the hyperedge changing simultaneously their opinion to the majority opinion of the hyperedge. In the case of a tie in hyperedges with even size, a bias is introduced in favor of type 1, which is motivated by the principle of social inertia. Our analytical results along with simulations and heuristic arguments suggest that, in any spatial dimensions and when the set of hyperedges consists of the collection of all $n \times \cdots \times n$ blocks of the lattice, opinion 1 wins when n is even while the system clusters when n is odd, which contrasts with results about the voter model in high dimensions for which opinions coexist. This is fully proved in one dimension while the rest of our analysis focuses on the cases when $n = 2$ and $n = 3$ in two dimensions.

Keywords Interacting particle systems · Hypergraph · Social group · Majority rule · Voter model

1 Introduction

There has been recently a growing interest in spatial models of social and cultural dynamics, where space is modeled through an underlying graph whose vertices represent individuals and edges potential dyadic interactions (see Castellano *et al.* [4] for a review). Such systems

Research supported in part by NSF Grant DMS-10-05282.

N. Lanchier (✉) · J. Neuffer
School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA
e-mail: nicolas.lanchier@asu.edu

are commonly called agent-based models while the mathematical term is interacting particle systems. There has been an increasing effort of applied scientists to understand such systems based on either heuristic arguments or numerical simulations. This effort must be complemented by analytical studies of interacting particle systems of interest in social sciences, mainly because stochastic spatial simulations are known to be difficult to interpret and might lead to erroneous conclusions. However, while the mathematical study of interacting particle systems over the past forty years has successfully led to a better understanding of a number of physical and biological systems, much less attention has been paid to the field of social sciences, with the notable exception of the popular voter model introduced independently in [5, 13]. The first objective of this paper is to extend the traditional framework of interacting particle systems by replacing the underlying graph that represents the network of interactions with the more general structure of a hypergraph, which is better suited to model social dynamics. The second objective is to use this new framework to construct a spatial version of the majority rule model proposed by Galam [11] to describe public debates, and also to initiate a rigorous analysis of this model of interest in social sciences. Even though our proofs rely mostly on new techniques, our model is somewhat reminiscent of the voter model, so we start with a description of the voter model and a review of its main properties.

1.1 The Voter Model

In the voter model, individuals are located on the vertex set of a graph and are characterized by one of two possible competing opinions, say opinion 0 and opinion 1. In particular, the state of the system at time t is a so-called spatial configuration η_t that can be viewed either as a function that maps the vertex set into $\{0, 1\}$, in which case the value of $\eta_t(x)$ represents the opinion of vertex x at time t , or as a subset of the vertex set, i.e., the set of vertices with opinion 1 at time t . The system evolves as follows: individuals independently update their opinion at rate one, i.e., at the times of independent Poisson processes with intensity one, by mimicking one of their nearest neighbors chosen uniformly at random. Since the times between consecutive updates are independent exponential random variables, the probability of two simultaneous updates is equal to zero, therefore the process is well-defined when the number of individuals is finite. On infinite graphs, since there are infinitely many Poisson processes, the time to the first update does not exist, and to prove that the voter model is well-defined, one must show in addition that the opinion at any given space-time point results from only finitely many interactions. This follows from an argument due to Harris [12]. The idea is that for $t > 0$ small the vertex set can be partitioned into finite islands that do not interact with each other by time t , which allows to construct the process independently on each of those islands up to time t , and by induction at any time.

We now specialize in the d -dimensional regular lattice in which each vertex is connected by an edge to each of its $2d$ nearest neighbors. The argument of Harris [12] mentioned above allows to construct the voter model graphically from a collection of independent Poisson processes, but also to prove a certain duality relationship between the voter model and a system of coalescing random walks on the lattice. Using this duality relationship and the fact that simple symmetric random walks are recurrent in $d \leq 2$ but transient in $d \geq 3$, one can prove that clustering occurs in one and two dimensions, i.e., any two vertices eventually share the same opinion which translates into the formation of clusters that keep growing indefinitely, whereas coexistence occurs in higher dimensions, i.e., the process converges to an equilibrium in which any two vertices have a positive probability of having different opinions [5, 13].

In view of the dichotomy clustering/coexistence depending on the dimension, a natural question about the voter model is: how fast do clusters grow in low dimensions? In one

dimension, Bramson and Griffeath [3] proved that the size of the clusters scales asymptotically like the square root of time: assuming that two vertices are distance t^a apart at time t , as $t \rightarrow \infty$, both vertices have independent opinions when $a > 1/2$ whereas they are totally correlated, i.e., the probability that they share the same opinion tends to one, when $a < 1/2$. In contrast, there is no natural scale for the cluster size in two dimensions, a result due to Cox and Griffeath [7]. More precisely, both vertices again have independent opinions when $a > 1/2$ but the probability that they share the same opinion when $a < 1/2$ is no longer equal to one but linear in a .

In higher dimensions, though coexistence occurs, the local interactions that dictate the dynamics of the voter model again induce spatial correlations, and a natural question is: how strong are the spatial correlations at equilibrium? To answer this question, the idea is to look at the random number of vertices with opinion 1 at equilibrium in a large cube minus its mean in order to obtain a centered random variable. Then, according to the central limit theorem, if opinions at different vertices were independent, this random variable rescaled by the square root of the size of the cube would be well approximated by a centered Gaussian. However, such a convergence is obtained for a larger exponent in the renormalization factor, a result due to Bramson and Griffeath [2] in three dimensions and extended by Zähle [15] to higher dimensions. This indicates that, even in high dimensions, spatial correlations are still significant.

Another natural question about the voter model is: does the opinion of any given vertex stay the same after a finite random time? To answer this question, we look at the fraction of time a given vertex is of type 1 in the long run, a random variable called the occupation time. When coexistence occurs, vertices change their opinion infinitely often therefore it is expected that the occupation time converges almost surely to the initial density of type 1. In contrast, when clustering occurs, it is expected that any given vertex will eventually be covered by a cluster and that the previous law of large number does not hold. However, Cox and Griffeath [6] proved almost sure convergence of the occupation time to the initial density of type 1 in dimensions $d \geq 2$. This does not hold in one dimension but it can be proved that, even in this case, the type of any given vertex keeps changing indefinitely. From the combination of the previous results, we obtain the following description of the voter model on the one- and two-dimensional lattices. Clusters form and appear to grow indefinitely so only one type is present at equilibrium. However, any given vertex flips infinitely often, which also indicates that clusters are not fixed in space but move around, and thus may give the impression of local transience though, strictly speaking, coexistence does not occur.

1.2 The Majority Rule Model

The main objective of this article is to initiate a rigorous analysis of another interacting particle system of interest in social sciences: a spatial version of the majority rule model proposed by Galam [11] to describe public debates. In the original nonspatial model, the population is finite and each agent is either in state zero or in state one, representing two differing opinions. At each time step, a positive integer, say n , is randomly chosen according to a given distribution, then n agents are chosen uniformly at random from the population. These agents form a discussion group which results in all n agents changing simultaneously their opinion to the majority opinion of the group. The majority rule is well-defined when n is odd while, when n is even and a tie occurs, a bias is introduced in favor of one opinion, say opinion 1, which is motivated by the principle of social inertia. Note that most (if not all) models of interacting particle systems studied in the mathematics literature are naturally defined through an underlying graph which encodes the pairs of vertices that may interact

using edges. However, to define spatial versions of the majority rule model, one needs a more complex network of interactions since the dynamics do not reduce to dyadic interactions: vertices interact by blocks. The most natural mathematical structure that can model such a network is the structure of hypergraph. In order to define spatial versions of the majority rule model, we thus extend the traditional definition of interacting particle systems by replacing the underlying graph structure with that of a hypergraph, an approach that we propose as a new modeling framework to describe social and cultural dynamics.

1.3 Stochastic Dynamics on Hypergraphs

In a number of coordinated sociological systems, people's opinions are subject to change in large groups due to, e.g., the influence of an opinion leader or public debates. To model such systems, we replace the underlying graph structure of traditional particle systems with one of a hypergraph, which has a set of vertices, as does a graph, but the set of edges is replaced by a set of hyperedges, which are no longer limited to a connection between two vertices modeling dyadic interactions: hyperedges are nonempty subsets of vertices which can be arbitrarily large. In the context of social dynamics, each hyperedge can be thought of as a social group such as a family unit, a team of co-workers, or a group of classmates, in which members interact simultaneously. To define the framework mathematically, we let $\mathbb{H} = (V, H)$ be a hypergraph and, for simplicity, restrict ourselves to spin systems, i.e., the individual at each vertex is characterized by one of only two possible opinions. In particular, as for the voter model described above, the state of the system at time t is a so-called spatial configuration η_t that can be viewed either as a function that maps the vertex set into $\{0, 1\}$, in which case the value of $\eta_t(x)$ represents the opinion of vertex x at time t , or as a subset of the vertex set, i.e., the set of vertices with opinion 1 at time t . Then, we define the dynamics using a Markov generator of the form

$$Lf(\eta) = \sum_{h \in H} \sum_{A \subset h} c_A(\eta \cap h) [f(\eta_{A,h}) - f(\eta)] \quad \text{where } \eta_{A,h} = (\eta \setminus h) \cup A. \quad (1)$$

Equation (1) simply means that for every hyperedge h and every subset $A \subset h$ of the hyperedge, the spins of all vertices in A become 1 while the spins of all the other vertices in the hyperedge become 0, at a rate that only depends on A and the configuration in the hyperedge, which is thus denoted by the function $c_A(\eta \cap h)$. In particular, we point out that each update of the system corresponds to the simultaneous update of all the vertices in a given hyperedge, rather than a single vertex, and that the rate at which a hyperedge is updated only depends on the configuration in this hyperedge. Also, we recall that an update occurs at rate c if the waiting time for this update is exponentially distributed with mean $1/c$. Returning to the majority rule model, to design a spatial analog that also accounts for the social structure of the population—in which discussions occur among agents that indeed belong to a common social group rather than agents chosen at random—it is natural to use the framework of stochastic dynamics on hypergraphs, where each hyperedge represents a social group. The dynamics of the majority rule are described by the Markov generator

$$\begin{aligned} Lf(\eta) = & \sum_{h \in H} \mathbf{1}\{\text{card}(\eta \cap h) < (\text{card } h)/2\} [f(\eta \setminus h) - f(\eta)] \\ & + \sum_{h \in H} \mathbf{1}\{\text{card}(\eta \cap h) \geq (\text{card } h)/2\} [f(\eta \cup h) - f(\eta)] \end{aligned} \quad (2)$$

where $\mathbf{1}$ is the indicator function, which is equal to one if its argument is true and zero otherwise, and where card stands for the cardinality. In particular, $\text{card}(\eta \cap h)$ is the number

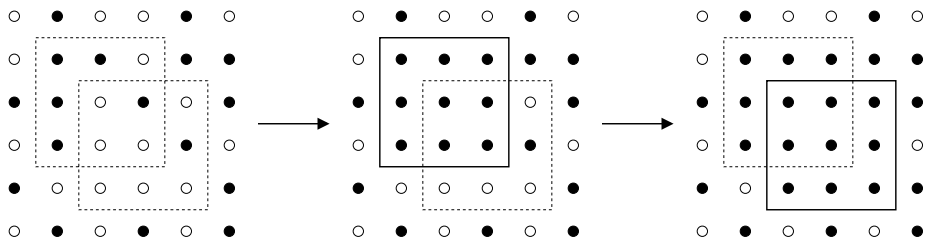


Fig. 1 Dynamics on hypergraph

of individuals with opinion 1 in the hyperedge h . Note that (2) simply is a particular case of (1) with

$$\begin{aligned} c_A(\eta \cap h) &= 0 \quad \text{for } A \notin \{\emptyset, h\}, \\ c_\emptyset(\eta \cap h) &= \mathbf{1}\{\text{card}(\eta \cap h) < (\text{card } h)/2\}, \\ c_h(\eta \cap h) &= \mathbf{1}\{\text{card}(\eta \cap h) \geq (\text{card } h)/2\}. \end{aligned}$$

Note also that the first indicator function is equal to one, and the second one equal to zero, if and only if there is a strict majority of type 0 in the hyperedge h . In particular, the expression of the generator (2) means that the configuration in each hyperedge is updated at rate one with all the vertices changing simultaneously their opinion to the majority opinion in the hyperedge. In case of a tie, opinion 1 is adopted. Since hyperedges are updated at the times of independent Poisson processes with intensity one, the probability of two overlapping hyperedges being updated simultaneously is equal to zero, therefore the process is well-defined on finite hypergraphs. Moreover, the argument of Harris [12] described above to justify the existence of the voter model on infinite graphs also applies to the majority rule model, therefore our process is well-defined on finite hypergraphs as well. For simplicity, we assume from now on that the vertex set is \mathbb{Z}^d , we let $n > 1$ be a nonrandom integer, and we define the collection of hyperedges by

$$H = \{h_x : x \in \mathbb{Z}^d\} \quad \text{where } h_x = x + \{0, 1, 2, \dots, n-1\}^d.$$

Each social group is thus represented by a $n \times \dots \times n$ block on the lattice. Note that, even though this might look simplistic, any two vertices are connected by paths of overlapping hyperedges, which results in a system that exhibits spatial correlations and nontrivial dynamics. Figure 1 gives an example of realization when $d = 2$ and $n = 3$ with two consecutive updates. Time goes from left to right and at each step the hyperedge which is updated is framed with thick continuous lines. Black and white dots refer to opinion 1 and opinion 0, respectively.

2 The Majority Rule in Space

Numerical simulations of the majority rule model in one and two dimensions suggest that the asymptotic behavior of the process is somewhat similar to that of the voter model when n is odd in the sense that the system clusters but similar to that of the biased voter model, i.e., the voter model modified so that individuals with opinion 0 update their opinion at a larger than individuals with opinion 1, when n is even. As explained at the end of this section, our last theorem together with some heuristic arguments also suggests that, in contrast with the

voter model, the majority rule model with n odd clusters in any dimension. Hence, we state the following conjecture.

Conjecture 1 *Clustering occurs when n is odd, i.e., starting from any configuration,*

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

whereas opinion 1 wins when n is even, i.e., starting from any configuration that has infinitely many hyperedges with a majority of type 1 vertices,

$$P(\eta_t(x) = 0) \rightarrow 0 \quad \text{as } t \rightarrow \infty \text{ for all } x \in \mathbb{Z}^d.$$

The conclusion for n even can be understood intuitively by observing that each tie in a discussion group, i.e., each tie at the time of the update of a hyperedge, results in a set of $n^d/2$ type 0 vertices changing their opinion to type 1, while if the process is modified so that ties do not affect the configuration of the system, the rules are symmetric. The case when the parameter n is odd is more interesting: in contrast with the results of Cox and Griffeath [7] which indicate that there is no natural scale for the asymptotics of the cluster size in the two-dimensional voter model, numerical simulations suggest that spatial correlations emerge much faster and that interface dynamics follow motion by mean curvature. This behavior is somewhat reminiscent of the behavior of some threshold voter models. Threshold voter models denote a class of stochastic processes which, similarly to the voter model, describe opinion dynamics on the regular lattice. As in the voter model, individuals are characterized by one of two possible competing opinions and update their opinion independently at rate one based on their neighbors' opinion. But unlike in the voter model, the new opinion is not chosen uniformly at random from the neighborhood. Instead, individuals change their opinion if and only if the number of their neighbors with the opposite opinion exceeds a parameter θ , called the threshold. The majority vote model is the special case in which the threshold θ equals half of the number of neighbors, therefore vertices are updated individually at rate one by adopting the majority opinion of their neighborhood. Although this model seems to be the spin system excluding the simultaneous update of several vertices the most closely related to our spatial version of the majority rule model, it does not cluster: instead, fixation occurs, i.e., any given individual stops changing opinion after a finite random time, as proved by Durrett and Steif [10]. In fact, the behavior of the majority rule model is similar to that of threshold voter models with threshold parameter slightly smaller than half of the number of neighbors.

The asymptotic behavior in the one-dimensional case is fully analyzed in this paper. Based on random walk estimates, we first prove that opinion 1 wins when n is even.

Theorem 2 *Assume that $d = 1$ and n is even. Then, opinion 1 wins.*

To study the process when n is odd, we rely on duality techniques following the standard approach introduced for the voter model. In the case of the majority rule model, the set of vertices that one has to keep track to determine the opinion of a space-time point grows linearly going backwards in time, which is the main difficulty to establish clustering. The trick is to prove that the correlation between two space-time points only depends on a space-time region which is delimited by a semblance of the centers of the corresponding dual processes. Recurrence of symmetric random walks, which is the key to proving clustering of the voter model in one and two dimensions, is invoked in order to show that this space-time region is almost surely bounded.

Theorem 3 *Assume that $d = 1$ and n is odd. Then, the process clusters.*

The two-dimensional system is more difficult to study mainly because of the underlying hypergraph structure and the lack of mathematical tools in this context. Even though duality techniques are available in two dimensions as well, the dual process is hardly tractable due to an abundance of branching events. For simplicity, we mainly focus on the cases when $n = 2$ and $n = 3$ rather than the more general even/odd dichotomy.

One key to analyzing the two-dimensional system is to look at a dual representation of the system that consists of keeping track of the disagreements along the edges of the lattice rather than the actual opinion at each vertex. This approach is partly motivated by the fact that the two-dimensional lattice seen as a planar graph is self dual. More precisely, we introduce a spin system coupled with the process and defined on the edge set by

$$\xi_t(e) = \xi_t((x, y)) = \mathbf{1}\{\eta_t(x) \neq \eta_t(y)\} \quad \text{for each edge } e = (x, y).$$

To visualize the state space of this process, it is convenient to delete all the edges in state 0 and rotate all the edges in state 1 of a quarter turn as shown in Fig. 3. Motivated by the resulting picture, we call this process the *contour* associated with the original spin system. This representation can as well be obtained by replacing every vertex $x \in \eta$ by the unit square centered at x and taking the topological boundary of the union of these squares. We point out that the contour associated with a configuration η can also be seen as a random subgraph of the dual lattice

$$\mathbb{D}^2 := \left\{ x + \left(\frac{1}{2}, \frac{1}{2} \right) : x \in \mathbb{Z}^2 \right\}.$$

Following the terminology of percolation theory, we say that an edge of the dual lattice is open if it belongs to the contour, and closed if it does not belong to the contour.

To prove invasion of type 1 when $n = 2$, the key is to look at the majority rule model modified to have only type 0 outside a horizontal slice of height three and so that all vertices to the right of a type 0 also are in state 0. The reason for looking at such a process is that the profile of its contour can be simply characterized by a two-dimensional vector that keeps track of the distance between the rightmost type 1 vertices at all three levels. Moreover, the evolution rules of the distance between these vertices as well as the rate at which the vertices of type 1 are added to or removed from the system can be expressed in a simple manner based on certain geometric properties of the contour. Relying in addition on a block construction leads to the following result.

Theorem 4 *Assume that $d = 2$ and $n = 2$. Then, opinion 1 wins.*

Finally, we look at the two-dimensional majority rule when the set of hyperedges consists of the set of all three by three squares as a test model to understand the general case when n is odd. To motivate our last result, consider the traditional voter model starting with a finite number of vertices of type 1. The process that keeps track of the number of type 1 vertices is a martingale since each time two vertices in different states interact, both vertices are equally likely to flip. In particular, the expected number of type 1 vertices is preserved by the dynamics, though the martingale convergence theorem implies almost sure extinction of the type 1 vertices. One of the most interesting aspects of the majority rule model, which again is reminiscent of threshold voter models with appropriate threshold parameter, is that, when starting from a finite initial configuration, the expected number of type 1 vertices is not constant. Our last theorem gives, for a class of configurations that we call regular clusters, an

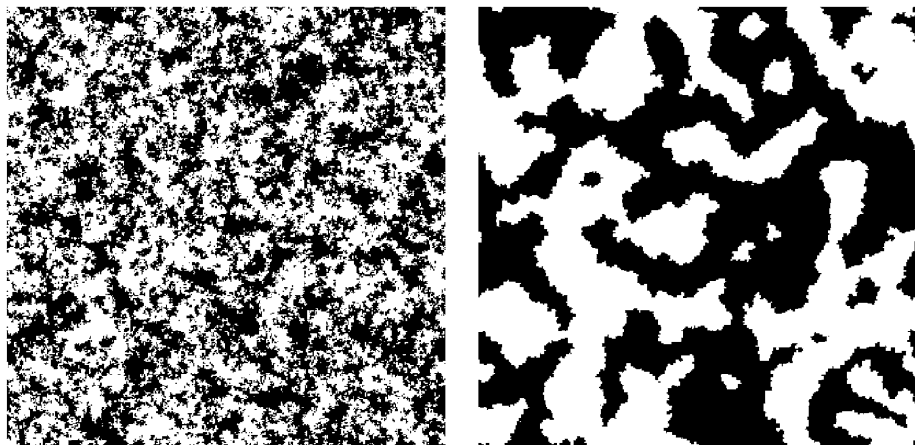


Fig. 2 Simulation pictures of the voter model and majority rule model at time 20, respectively. Both processes evolve on a 400×400 lattice with periodic boundary conditions, and start from a Bernoulli product measure with an equal density of *white* and *black* vertices

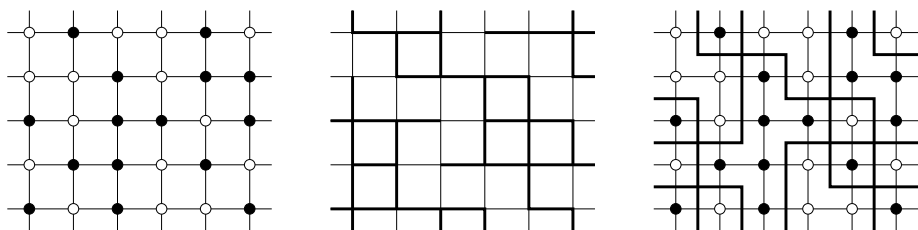


Fig. 3 Dual representation between the majority rule model and the contour process

explicit expression of the variation rate of the number of type 1 as a function of the geometry of the cluster. This result supplemented with a heuristic argument strongly suggests that, as for some threshold voter models, the process with n odd clusters and explains the reason why the snapshot of the majority rule model on the right hand side of Fig. 2 differs significantly from that of the voter model on the left hand side. To state our result, we need a few more definitions: given a configuration η , we call vertex x a *corner* whenever

$$\eta(x - e_1 - e_2) = \eta(x + e_1 + e_2) \neq \eta(x) \quad \text{or} \quad \eta(x - e_1 + e_2) = \eta(x + e_1 - e_2) \neq \eta(x)$$

where e_1 and e_2 are the first and second unit vectors of the Euclidean plane. A corner x is said to be a *positive corner* if $\eta(x) = 1$ and a *negative corner* if $\eta(x) = 0$. Also, we call η a *cluster* if its contour Γ is a Jordan curve, i.e., a non-self-intersecting loop, and a *regular cluster* if in addition

1. the set $(x + D_2) \cap \Gamma$ is connected for all $x \in \mathbb{D}^2$ and
2. if vertex x is a corner and $(x + D_2) \cap (y + D_2) \neq \emptyset$ then vertex y is not a corner

where $D_2 = [-1, 1]^2$. Without loss of generality, we assume that vertices located in the bounded region delimited by the Jordan curve are of type 1, which forces vertices in the unbounded region to be of type 0. Condition 1 above essentially says that the microscopic structure of the boundary of the cluster is not too complicated, i.e., the Jordan curve does

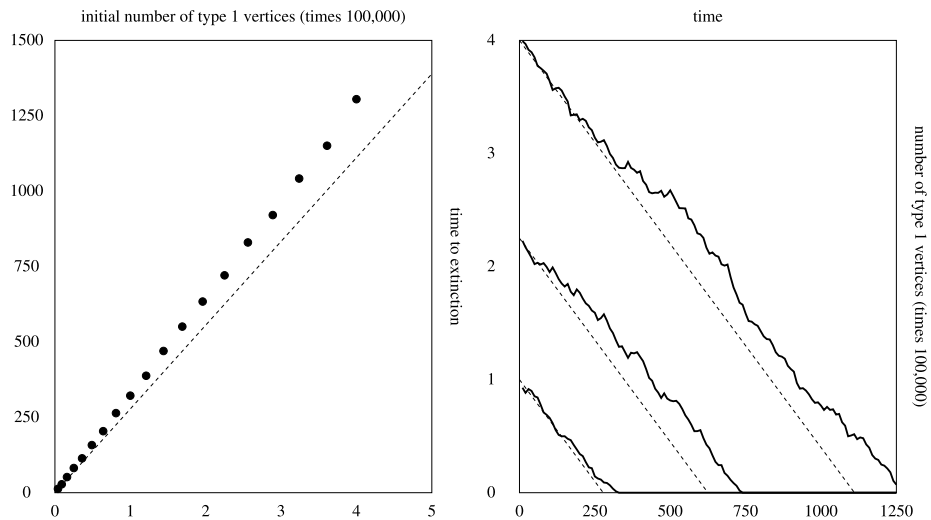


Fig. 4 Simulation results for the process starting from a square cluster. The *dots* on the *left picture* give the time to extinction of type 1 vertices as a function of the initial number of type 1 averaged over 100 realizations. The *right picture* gives the evolution of the number of type 1 vertices for three realizations. In both pictures, the *dashed lines* represent the corresponding expected values assuming a loss of 36 vertices of type 1 per unit of time

not zigzag too much, while condition 2 simply indicates that corners cannot be too close to each other. Finally, we let c_+ and c_- denote the number of positive and negative corners, respectively, and let $\phi(\eta, x)$ denote the variation of the number of type 1 vertices after the three by three square centered at x is updated. Note that, when configuration η is a cluster, $\phi(\eta, x) = 0$ for all but a finite number of vertices.

Theorem 5 Assume that η is a regular cluster with at least 11 vertices. Then

$$\sum_{x \in \mathbb{Z}^2} \phi(\eta, x) = 9(c_- - c_+).$$

In other words, the rate of variation of the number of type 1 vertices can be easily expressed as a function of the number of positive and negative corners, which directly implies that the expected number of type 1 vertices is not constant.

To conclude this section, we give some heuristic arguments which, together with the previous theorem, supports the first part of Conjecture 1 when n is odd. The main purpose of Theorem 5 is to support the idea that the time to extinction of a finite cluster scales like the original size of the cluster. To this extent, the assumption that the cluster must have at least 11 vertices is not a limitation since clusters with an even smaller size are destroyed quickly. To relate the theorem to the time to extinction of the type 1 opinion, we first observe that, traveling around the Jordan curve clockwise, the number of right turns is equal to the number of left turns plus four. The result directly follows by using a simple induction over the number of type 1 vertices. This suggests that, when averaged over time from time 0 to the time to extinction, the difference between the number of negative and positive corners should be about -4 , further suggesting that the time to extinction is equal to about the initial number of type 1 vertices divided by $9 \times 4 = 36$. Figure 4 compares our speculative

argument with simulation results for the process starting with a square cluster. Even though these do not fit perfectly, the numerical results strongly suggest that the time to extinction is indeed linear in the initial number of type 1 vertices, which drastically contrasts with the voter model, and that our 36 is not far from the truth. This heuristic argument also indicates that the majority rule dynamics quickly destroy small clusters, thus resulting in a clustering more pronounced than in the voter model, which explains the striking difference between the two pictures of Fig. 2. Finally, we point out that the intuitive ideas behind the proof of Theorem 5 are not sensitive to the spatial dimension. Also, we conjecture that the theorem holds in higher dimensions with the constant 9 replaced by 3^d , and that the majority rule model with n odd clusters in any spatial dimensions, as mentioned at the beginning of this section in Conjecture 1.

3 Proof of Theorems 2 and 3 ($d = 1$)

This section is devoted to the analysis of the one-dimensional majority rule model for which we prove that clustering occurs for all sizes n of the hyperedges while opinion 1 wins under the additional assumption that n is even. The latter is based on simple random walk estimates while the former further relies on duality techniques, which consists of keeping track of the ancestry of a finite number of vertices going backwards in time.

In order to define the dual process, the first step is to construct the process graphically. The construction is similar in any spatial dimension. To each hyperedge h_x we attach a Poisson process with parameter 1 whose j th arrival time is denoted by $T_j(x)$. Poisson processes attached to different hyperedges are independent. At time $t = T_j(x)$, we have the following alternative:

1. If $\text{card}(\eta_{t-} \cap h_x) < n^d/2$ then all vertices in h_x become of type 0.
2. If $\text{card}(\eta_{t-} \cap h_x) \geq n^d/2$ then all vertices in h_x become of type 1.

Results due to Harris [12] which apply to traditional interacting particle systems on lattices but extend directly to the hypergraph \mathbb{H} guarantee that the majority rule model starting from any initial configuration $\eta_0 \subset \mathbb{Z}^d$ can be constructed using the collection of independent Poisson processes and the majority rule at the arrival times introduced above. To visualize this in one dimension, we draw a line segment from vertex x to vertex $x + n - 1$ at the arrival times of the Poisson process attached to the hyperedge h_x . Note that this line segment connects all the vertices in h_x . To study the process when n is even, we first assume that $\eta_0 = (-\infty, 0] \cap \mathbb{Z}$. Therefore

$$\eta_t = (-\infty, X_t] \cap \mathbb{Z} \quad \text{for all } t > 0 \text{ where } X_t := \max\{x \in \mathbb{Z} : x \in \eta_t\}.$$

The key to proving Theorem 2 is the following lemma.

Lemma 6 *With probability one, $X_t \rightarrow \infty$ as $t \rightarrow \infty$.*

Proof Note that there are exactly n hyperedges that contain vertex X_t therefore $n - 1$ possible events that affect the position of the rightmost 1. From the leftmost to the rightmost, these updates create/remove respectively the following numbers of type 1 vertices:

$$\text{create } 1, 2, \dots, \frac{n}{2} - 1, \frac{n}{2} \quad \text{remove } \frac{n}{2} - 1, \frac{n}{2} - 2, \dots, 2, 1.$$

In other words, we have the transition rates

$$X_t \rightarrow X_t + j \quad \text{at rate one for all } j \in \left\{1 - \frac{n}{2}, 2 - \frac{n}{2}, \dots, \frac{n}{2}\right\}.$$

Summing over all the possible values of the increment, we get

$$E(X_1 - X_0) = \sum_{j=1}^n \left(j - \frac{n}{2} \right) = \frac{n(n+1)}{2} - \frac{n^2}{2} = \frac{n}{2} > 0.$$

The expected value can be understood intuitively as follows. There are $n - 2$ updates that can be paired off in such a way that each pair consists of one update that causes k vertices of type 0 to flip and one update that causes k vertices of type 1 to flip. The remaining update corresponds to a tie that causes $n/2$ vertices of type 0 to flip, which gives the expected value above. In particular, an application of the Law of Large Numbers implies that X_t converges almost surely to infinity as time goes to infinity, which completes the proof. \square

It is straightforward to deduce from Lemma 6 that

$$\begin{aligned} P(\eta_t(x) \rightarrow 1 \text{ as } t \rightarrow \infty \text{ for all } x \in \mathbb{Z} \mid \eta_0 = h_0) \\ \geq P(X_t - X_0 \geq 0 \text{ for all } t \geq 0) \times P(X_0 - X_t \leq 0 \text{ for all } t \geq 0) > 0. \end{aligned}$$

Theorem 2 follows directly from the previous estimate since the latter implies that, starting with infinitely many hyperedges with a majority of type 1, there exists with probability one a cluster of vertices of type 1 that expands indefinitely.

We now turn to the proof of Theorem 3 which relies on duality techniques. The ancestry of a given space-time point, i.e., the set of vertices at earlier times that determine the opinion of the point under consideration, grows linearly going backwards in time. While the whole structure of the ancestry, which keeps growing indefinitely, is necessary to determine the opinion of a given vertex based on the initial configuration, given two vertices, only a finite space-time region is relevant in proving that they share ultimately the same opinion. In order to define this space-time region and the dual process starting at a given point, we first introduce

$$T(u) = \{T_j(u) : j \geq 1\} \quad \text{and} \quad c(u) = u + \frac{n-1}{2} \quad \text{for all } u \in \mathbb{Z}.$$

Note that $c(u)$ is simply the center of the hyperedge h_u . The dual process starting at a given space-time point (x, T) is the set-valued process initiated at $\hat{\eta}_0(x, T) = \{x\}$ and defined recursively as follows: assuming that the dual process has been defined up to time s , we let

$$\tau(s) = T - \sup\{T(v) \cap (0, T - s) : v \in \mathbb{Z} \text{ and } \hat{\eta}_s(x, T) \cap h_v \neq \emptyset\}.$$

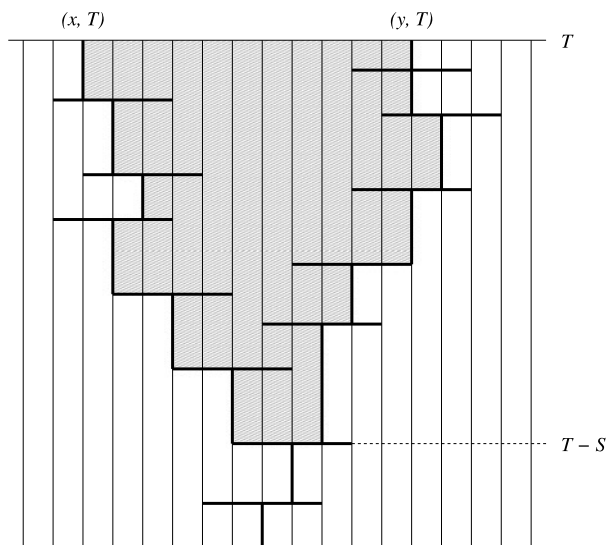
There is a unique vertex $w \in \mathbb{Z}$ such that $T - \tau(s) \in T(w)$. Then, we define

$$\hat{\eta}_t(x, T) = \hat{\eta}_s(x, T) \quad \text{for all } t \in (s, \tau(s)) \quad \text{and} \quad \hat{\eta}_{\tau(s)}(x, T) = \hat{\eta}_s(x, T) \cup h_w.$$

In words, going backwards in time, each time the dual process “encounters” a line segment in the graphical representation, the corresponding hyperedge is added to the process. Therefore, the dual process consists of an interval of vertices that grows linearly going backwards in time. The graphical representation restricted to the space-time region induced by the dual process together with the initial configuration in $\hat{\eta}_T(x, T)$ allows to determine the opinion of (x, T) . However, we can prove that two given vertices share ultimately the same opinion without looking at their opinion or the whole structure of their dual processes. To do so, we define a new process $c_s(x, T)$ that we shall call the *center path* of space-time point (x, T) . Again, $c_0(x, T) = x$ and the process is defined recursively based on the Poisson events: assuming that the path has been defined until time s , let

$$\sigma(s) = T - \sup\{T(v) \cap (0, T - s) : v \in \mathbb{Z} \text{ and } c_s(x, T) \in h_v\}.$$

Fig. 5 Picture of the center paths when $n = 5$



There is a unique vertex $w \in \mathbb{Z}$ such that $T - \sigma(s) \in T(w)$ and we define

$$c_t(x, T) = c_s(x, T) \quad \text{for all } t \in (s, \sigma(s)) \quad \text{and} \quad c_{\sigma(s)}(x, T) = c(w).$$

In words, going backwards in time, each time the center path “encounters” a line segment in the graphical representation, it jumps to the center of this line segment. To complete the construction, we now let $x < y$ be two vertices, and define the space-time region Ω which is delimited by their respective center paths by setting

$$S = \inf\{s > 0 : c_s(x, T) = c_s(y, T)\},$$

$$\Omega = \{(z, t) \in \mathbb{Z} \times (\max(T - S, 0), T) : c_{T-t}(x, T) \leq z \leq c_{T-t}(y, T)\}.$$

We refer to Fig. 5 for a picture where Ω is represented by the hatched polygonal region. The key to proving Theorem 3 is that, provided the center paths intersect by time 0, all space-time points in the region Ω share the same opinion, which is established in the following lemma.

Lemma 7 Assume that $S < T$. Then, the function $\Phi(z, t) := \eta_t(z)$ is constant on Ω .

Proof Define $\Lambda = \{(u, t) \in \mathbb{Z} \times \mathbb{R}_+ : t \in T(u)\}$ and the collection

$$H^* = \{h(u, t) := h_u \times \{t\} : (u, t) \in \Lambda \text{ and } h(u, t) \cap \Omega \neq \emptyset\}$$

which can be seen as the set of all line segments of the graphical representation that intersect the space-time region Ω . Note that this corresponds to the set of all Poisson events that may affect the configuration of the process in Ω . First, by definition of S , there is $h(u, t) \in H^*$ such that

$$t = T - S \quad \text{and} \quad c_{S-}(x, T), c_{S-}(y, T) \in h_u$$

from which it follows that Φ is constant on $\Omega \cap (\mathbb{Z} \times \{T - S\})$ and equal to the majority type in the hyperedge h_u at time $T - S$. To prove that this property is retained at later times, let

$$h(u, t) \in H^* \quad \text{such that} \quad s := T - t \neq S$$

and observe that we have the following alternative:

1. $c_{s-}(x, T), c_{s-}(y, T) \notin h_u$ and then

$$h_u \subset (c_s(x, T), c_s(y, T)) \cap \mathbb{Z} \quad \text{and} \quad \text{card}\{h_u \cap [c_s(x, T), c_s(y, T)]\} = n.$$

2. $c_{s-}(x, T) \in h_u$ and then $c_s(x, T) = c(u)$ and

$$\text{card}\{h_u \cap [c_s(x, T), c_s(y, T)]\} = \text{card}\{c(u), c(u) + 1, \dots, u + n - 1\} > n/2.$$

3. $c_{s-}(y, T) \in h_u$ and then $c_s(y, T) = c(u)$ and

$$\text{card}\{h_u \cap [c_s(x, T), c_s(y, T)]\} = \text{card}\{u, u + 1, \dots, c(u)\} > n/2.$$

In all three cases, we have that

$$\eta_{t-}(z) = i \quad \text{for } c_s(x, T) \leq z \leq c_s(y, T) \quad \text{implies} \quad \text{card}\{z \in h_u : \eta_{t-}(z) = i\} > n/2$$

from which it follows that $\eta_t(z) = i$ for all $c_{s-}(x, T) \leq z \leq c_{s-}(y, T)$. This indicates that the property to be proved is retained going forward in time through the Poisson events in H^* . Since the other Poisson events do not affect the space-time region Ω , the lemma follows. \square

In view of the previous lemma, we have $\eta_T(x) = \eta_T(y)$ whenever $S < T$. In other respect, the same argument as in Lemma 6 implies that both center paths evolve according to independent symmetric random walks until they intersect. More precisely,

$$c_s(x, T) \rightarrow c_s(x + T) + j \quad \text{at rate one for all } j \in \left\{ -\frac{n-1}{2}, \dots, \frac{n-1}{2} \right\}.$$

Since symmetric random walks are recurrent in one dimension, the probability that they intersect by time 0, that is $S < T$, approaches one as time $T \rightarrow \infty$. This proves Theorem 3.

4 Proof of Theorem 4 ($d = 2$ and $n = 2$)

This section is devoted to Theorem 4 whose proof is based on a rescaling argument. This technique is also known as block construction and was introduced by Bramson and Durrett [1] and further refined by Durrett [9]. Even though the block construction is now a standard tool in the field of interacting particle systems, its application is rarely straightforward and requires additional nonstandard arguments, especially in the case of the majority rule model.

In preparation for the application of a block construction, we first investigate a new process that we shall call the *slice process* which is the 2-dimensional 4-neighborhood majority rule model modified in the following two ways. First, the process is restricted to the horizontal slice

$$S_3 = \{x = (x_1, x_2) \in \mathbb{Z}^2 : |x_2| \leq 1\}$$

in the sense that all vertices in the complement of S_3 are unchangeably in state 0. Second, the process is modified so that all vertices to the right of a vertex in state 0 and with the same second coordinate flip instantaneously to state 0, which implies that updates that result in the existence of a vertex in state 1 to the right of a vertex in state 0 are suppressed. In particular, state 1 is instantaneously driven to extinction when starting from a random initial condition for which

$$P(\text{for all } (x_1, x_2) \in S_3 \text{ there exists } z_1 \leq x_1 \text{ such that } (z_1, x_2) \text{ is in state 0}) = 1.$$

Therefore, to avoid trivialities, we assume that the slice process starts from the deterministic configuration in which all the vertices in the horizontal slice S_3 with a nonpositive first coordinate are in state 1 and all other vertices are in state 0. Although our verbal description of the slice process is probably clear enough, for the sake of rigor we also give its Markov generator

$$\begin{aligned} L_3 f(\bar{\eta}) = & \sum_x \mathbf{1}\{\text{card}(\bar{\eta} \cap h_x) < 2\} [f(\bar{\eta} \setminus h_x) - f(\bar{\eta})] \\ & + \sum_x \mathbf{1}\{\text{card}(\bar{\eta} \cap h_x) \geq 2, h_x \subset S_3, (-\infty, x_1) \times \{x_2, x_2 + 1\} \subset \bar{\eta}\} \\ & \times [f(\bar{\eta} \cup h_x) - f(\bar{\eta})] \end{aligned}$$

where x_1 and x_2 denote the first and second coordinates of vertex $x \in \mathbb{Z}^2$. Note that the slice process is stochastically smaller than the original majority rule model, i.e., the processes starting from the same initial configuration can be coupled in such a way that, at all times, the set of type 1 vertices of the majority rule model contains the set of type 1 vertices of the slice process. The reason for introducing the two modification rules that define the slice process is that they simplify the dynamics to make them more tractable mathematically without however preventing opinion 1 from invading the slice S_3 so that Theorem 4 can be eventually deduced from stochastic domination and a block construction.

To investigate the slice process and prove that it invades the slice S_3 we note that the second modification rule implies that, for $x_2 = -1, 0, 1$, all vertices with second coordinate x_2 to the left of the rightmost vertex in state 1 and also with second coordinate x_2 are in state 1. In particular, the configuration of the slice process is uniquely defined by the position of its three rightmost vertices in state 1 with second coordinate $-1, 0, 1$, or equivalently the Markov process

$$X_t = (X_t(x_2) : x_2 = -1, 0, 1) \quad \text{where } X_t(x_2) = \max\{x_1 : \bar{\eta}_t((x_1, x_2)) = 1\}.$$

Note also that the dynamics of the slice process induced by the 4-neighborhood majority rule model imply that the middle component of the process X_t cannot be simultaneously smaller than the other two components. By invasion in the slice we mean almost sure convergence of all three components to infinity. We first introduce the following functional associated to the slice process:

$$\begin{aligned} D(\Sigma_t) &= \lim_{h \rightarrow 0} h^{-1} (\Sigma_{t+h} - \Sigma_t) \quad \text{where } \Sigma_t = X_t(-1) + X_t(0) + X_t(1), \\ D(G_t) &= \lim_{h \rightarrow 0} h^{-1} (G_{t+h} - G_t) \quad \text{where } G_t = |X_t(1) - X_t(0)| + |X_t(-1) - X_t(0)| \end{aligned}$$

that we call the *sum's drift* and the *gap's drift*, respectively. The analysis of these two processes indicate that the sum Σ_t drifts to infinity whereas the gap G_t is uniformly bounded in time, from which it follows that all three components of X_t converge almost surely to infinity. The analysis of the sum's and gap's drifts relies on asymptotic properties of the functional

$$\iota(X_t) = (X_t^+, X_t^-) \quad \text{where } X_t^+ = X_t(1) - X_t(0) \text{ and } X_t^- = X_t(-1) - X_t(0)$$

that we shall call for obvious reasons the *interface process*. Letting (a, b) denotes the state of the interface, we always have $a \leq 0$ or $b \leq 0$ because the middle component of the process X_t cannot be simultaneously smaller than its other two components. We also observe that the value of the sum's drift and the value of the gap's drift are not affected by the symmetry about the x -axis. In particular, we identify interfaces that can be deduced from one another

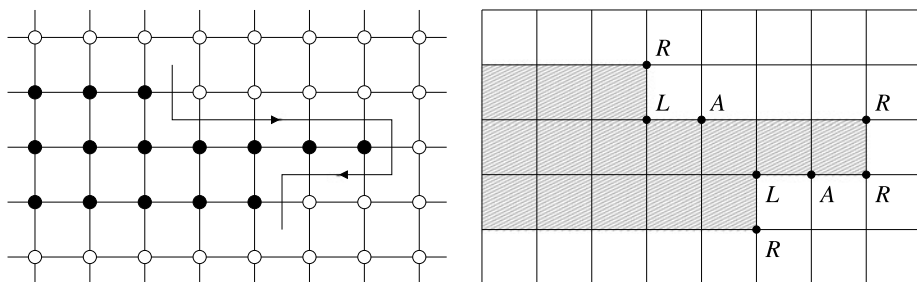


Fig. 6 Pictures related to the proof of Lemma 8

by this axial symmetry, that is we identify states (a, b) and (b, a) . Therefore, the interface process can be seen as a continuous-time random walk on a certain connected graph with vertex set

$$V = \{(a, b) : a \leq 0 \text{ and } a \leq b\}.$$

We have represented the transition rates of the interface process on a portion of this connected graph around vertex $(0, 0) \in V$ in Fig. 7. Information about the dynamics of the interface given in this figure are employed frequently in some of the following lemmas.

Lemma 8 *We have $D(\Sigma_t) = 2(N(X_t) - 1)$ where $N(X_t) = \mathbf{1}\{|X_t^+| \neq 1\} + \mathbf{1}\{|X_t^-| \neq 1\}$.*

Proof The proof relies on a series of simple geometric arguments. First of all, we observe that each point z of the dual lattice of \mathbb{Z}^2 has exactly four nearest neighbors in \mathbb{Z}^2 . These four neighbors define a size 4 neighborhood that is updated at rate 1 and that we call the neighborhood with center z . For each configuration $\bar{\eta}$ of the slice process, we let $\Gamma = \Gamma(\bar{\eta})$ be the contour associated with $\bar{\eta}$ defined as in Sect. 2. Note that this contour is a doubly infinite self-avoiding path on the dual lattice. We let $\vec{\gamma}$ be the finite portion of this path that connects the points

$$\gamma^+ = \left(X_t(1) + \frac{1}{2}, \frac{3}{2}\right) \quad \text{and} \quad \gamma^- = \left(X_t(-1) + \frac{1}{2}, -\frac{3}{2}\right)$$

and orient this portion from point γ^+ to point γ^- . The first picture of Fig. 6 gives an example of configuration of the slice process where black dots refer to vertices in state 1 and white dots to vertices in state 0, together with the corresponding oriented path $\vec{\gamma}$ represented in thick lines. Note that the oriented path $\Gamma(\bar{\eta}_t)$ has exactly $G_t + 4$ vertices so we write

$$\vec{\gamma} = (\gamma(1) = \gamma^+, \gamma(2), \gamma(3), \dots, \gamma(G_t + 4) = \gamma^-)$$

in the direction of the orientation, and let $\epsilon(j)$ be the edge connecting $\gamma(j)$ and $\gamma(j + 1)$. Note also that any update in a neighborhood whose center does not belong to the oriented path does not yield any change in the configuration of the slice process, either because this neighborhood already contains four vertices in the same state, or because it contains two vertices in each state but is not included in S_3 . To compute the drift of Σ_t , we introduce the following classification.

1. Point $z \in \vec{\gamma}$ is called a *right turn* if the neighborhood with center z contains exactly one vertex in state 1. In this case, an update of the slice process in this neighborhood always results in one vertex changing from state 1 to state 0.

The figure shows a grid graph with 10 nodes and 20 edges. The nodes are arranged in a 3x3 grid with an additional node at the bottom center. The edges are labeled with weights: 1, 2, 3, and 4. The graph is used to illustrate the construction of a shortest path tree.

- In the second picture of Fig. 6, right turns, active straight points, and left turns corresponding to the configuration in the first picture are marked with the letters R , A , L , respectively. Since neighborhoods are updated independently and at rate one, the drift can be computed based on the number of left/right turns and active straight points. To count the number of points in each class, we first observe that γ^+ and γ^- are always right turns, while to determine the class of the other points, we distinguish between the following two cases.

- The class of points $\gamma(j)$, $j = |X_t^+| + 3, \dots, G_t + 3$, can be determined similarly. In particular, the number of right turns minus the number of left turns always equals two, which allows to quantify the drift in terms of the number of active straight points exclusively:

$$D(\Sigma_t) = 2 \times \text{number of active straight points} - 2.$$

Finally, there is one active straight point with second coordinate $1/2$ if and only if $X_t^+ \neq 1$, and one with second coordinate $-1/2$ if and only if $X_t^- \neq 1$. Therefore, we conclude that the number of active straight points is simply equal to $N(X_t)$, which completes the proof of the lemma. \square

Lemma 9 *There exist constants $C_1 < \infty$, $\gamma_1 > 0$ and $c > 0$ such that*

$$P(\Sigma_{cN} < 7N) + P(\Sigma_t < -N \text{ for some } t < cN) \leq C_1 \exp(-\gamma_1 N).$$

Proof Since the drift $D(\Sigma_t) \in \{-2, 0, 2\}$ according to Lemma 8, the key step is to prove that the fraction of time spent on good interfaces is in average strictly larger than the fraction of time spent on bad interfaces, where good interfaces refer to the ones for which the sum's drift is positive and bad interfaces refer to the ones for which the sum's drift is negative. To compare these two quantities, we let $e(a, b)$ for all $(a, b) \in V$ denote the expected time spent on good interfaces before hitting a bad interface when starting from interface (a, b) , that is,

$$e(a, b) = E \left[\int_0^T \mathbf{1}\{D(\Sigma_t) = 2\} dt \mid \iota(X_0) = (a, b) \right]$$

where $T = \inf\{t > 0 : D(\Sigma_t) = -2\}$. In the picture of Fig. 7, good interfaces are marked with a black dot and bad interfaces with a white dot. Based on the transition rates given in this picture and using successive first-step analyses, we obtain that

$$e(0, 0) = \frac{1}{4} + e(-1, 0) \geq \frac{1}{4} + \frac{1}{5}e(0, 0) + \frac{2}{5}e(-2, 0) \geq \frac{1}{4} + \frac{1}{5} \times \frac{1}{4} + \frac{2}{5} \times \frac{1}{6} = \frac{11}{30}$$

from which it follows that

$$\begin{aligned} e(0, 1) &\geq \frac{3}{5}e(0, 0) \geq \frac{3}{5} \times \frac{11}{30} = \frac{11}{50}, \\ e(0, 2) &\geq \frac{1}{6} + \frac{1}{6}e(-2, 0) + \frac{1}{6} \frac{3}{5}e(0, 0) \geq \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{3}{5} \times \frac{11}{30} = \frac{52}{225}, \\ e(-1, 0) &\geq \frac{1}{5}e(0, 0) + \frac{2}{5}e(-2, 0) \geq \frac{1}{5} \times \frac{11}{30} + \frac{2}{5} \times \frac{1}{6} = \frac{7}{50}, \\ e(-2, 0) &\geq \frac{1}{6} + \frac{1}{6}e(0, 0) + \frac{1}{3}e(-3, 0) \geq \frac{1}{6} + \frac{1}{6} \times \frac{11}{30} + \frac{1}{3} \times \frac{1}{6} = \frac{17}{60}. \end{aligned}$$

Using the previous lower bounds, we obtain that the expected time spent on good interfaces after leaving the bad interface $(-1, -1)$ is bounded from below by

$$\frac{e(0, 0)}{3} + \frac{e(-1, 0)}{3} \geq \frac{1}{3} \times \frac{11}{30} + \frac{1}{3} \times \frac{7}{50} = \frac{38}{225} > \frac{38}{228} = \frac{1}{6}$$

while the analog for the bad interface $(-1, 1)$ is bounded from below by

$$\frac{e(0, 1)}{6} + \frac{e(0, 2)}{6} + \frac{e(-1, 0)}{3} + \frac{e(-2, 0)}{6} \geq \frac{913}{5400} > \frac{913}{5478} = \frac{1}{6}.$$

Since the expected time spent on each of the two bad interfaces at each visit is equal to $1/6$, and the previous lower bounds indicate that the time spent on good configurations between two consecutive visits of a bad interface is strictly larger than $1/6$, we deduce that

$$\liminf_{t \rightarrow \infty} P(D(\Sigma_t) = 2) - \limsup_{t \rightarrow \infty} P(D(\Sigma_t) = -2) = a > 0.$$

In particular, we obtain the inequality

$$\liminf_{t \rightarrow \infty} t^{-1} \Sigma_t \geq 2 \liminf_{t \rightarrow \infty} P(D(\Sigma_t) = 2) - 2 \limsup_{t \rightarrow \infty} P(D(\Sigma_t) = -2) = 2a > 0.$$

Let $c = 7a^{-1}$ and $\epsilon = a > 0$. Since Σ_t is asymptotically bounded from below by $2at$, standard large deviation estimates for the Poisson distribution imply that

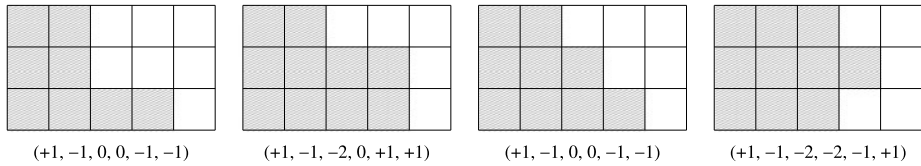


Fig. 8 Pictures related to the proof of Lemma 10

$$\begin{aligned}
 &P(\Sigma_{cN} < 7N) + P(\Sigma_t < -N \text{ for some } t < cN) \\
 &\leq P(\Sigma_{cN} < (2a - \epsilon)cN) + P(\Sigma_t < -N \text{ for some } t > 0) \\
 &\leq C_2 \exp(-\gamma_2 N) + C_3 \exp(-\gamma_3 N)
 \end{aligned}$$

for suitable constants $C_2, C_3 < \infty$ and $\gamma_2, \gamma_3 > 0$. This completes the proof. \square

Lemma 10 Assume that $G_t \geq 2$. Then $D(G_t) \leq -2 \times \mathbf{1}\{X_t^+ X_t^- \neq 0\}$.

Proof First, we note that, when $G_t = 2$, there are only four possible interfaces provided one identifies pairs of interfaces that can be deduced from one another by an axial symmetry. Note also that there are $G_t + 4 = 6$ possible updates for each of these four interfaces. Figure 8 gives a picture of these interfaces. The six numbers at the bottom represent the variation of G_t for each of the six possible updates. Since each update occurs at rate one, the drift $D(G_t)$ is simply equal to the sum of these six numbers. More generally, when $G_t \geq 2$, we have the following alternative.

1. In the case $X_t^+ = 0$ or $X_t^- = 0$, there are only six possible updates of the interface, each of which gives the same variation of $D(G_t)$ as in one of the first two pictures. Therefore,

$$D(G_t) \leq 1 - 1 - 2 + 1 + 1 = 0 \quad \text{when } X_t^+ X_t^- = 0.$$

2. The case $X_t^+ \neq 0$ and $X_t^- \neq 0$ is similar to one of the last two pictures except that there might be one or two active straight points in addition to the six turns. Updates at these straight points cannot increase the value of the gap process, therefore,

$$D(G_t) \leq 1 - 1 - 1 - 1 = -2 \quad \text{when } X_t^+ X_t^- \neq 0.$$

The lemma follows. \square

Lemma 11 Let $c > 0$ as in Lemma 9. Then there exist $C_4 < \infty$ and $\gamma_4 > 0$ such that

$$P(G_t > \sqrt{N} \text{ for some } t < cN) \leq C_4 \exp(-\gamma_4 \sqrt{N}).$$

Proof For all times $s > 0$, we introduce the two stopping times

$$T^-(s) = \inf\{t > s : G_t \leq 1\} \quad \text{and} \quad T^+(s) = \inf\{t > s : G_t > \sqrt{N}\}.$$

Since each time the slice process visits an interface such that $X_t^+ X_t^- = 0$ there is a strictly positive probability that $X_t^+ X_t^- \neq 0$ at the next jump, we have

$$\liminf_{t \rightarrow \infty} P(X_t^+ X_t^- \neq 0 \mid G_t \geq 2) = b > 0.$$

This, together with Lemma 10, gives

$$\liminf_{t \rightarrow \infty} E[D(G_t) \mid G_t \geq 2] \leq -2 \liminf_{t \rightarrow \infty} P(X_t^+ X_t^- \neq 0 \mid G_t \geq 2) = -2b < 0$$

therefore, standard large deviation estimates imply that

$$P(T^+(s) < T^-(s) \mid G_s = 2) \leq C_5 \exp(-\gamma_5 \sqrt{N}) \quad (3)$$

for suitable constants $C_5 < \infty$ and $\gamma_5 > 0$. Let $v_t(2)$ and J_t denote respectively the number of times the gap process visits state 2 and the number of times it jumps by time t . Since the process jumps at rate at most 8, large deviation estimates for the Poisson distribution imply that

$$P(v_{cN}(2) > 5cN) \leq P(J_{cN} > 10cN) \leq C_6 \exp(-\gamma_6 N) \quad (4)$$

for appropriate $C_6 < \infty$ and $\gamma_6 > 0$. Combining (3) and (4), we obtain

$$\begin{aligned} P(G_t > \sqrt{N} \text{ for some } t < cN) &\leq P(v_{cN}(2) > 5cN) + 5cN P(T^+(s) < T^-(s) \mid G_s = 2) \\ &\leq C_6 \exp(-\gamma_6 N) + 5cN \times C_5 \exp(-\gamma_5 \sqrt{N}), \end{aligned}$$

which proves the lemma. \square

Corollary 12 *There exist $C_7 < \infty$ and $\gamma_7 > 0$ such that, for $x_2 = -1, 0, 1$,*

$$P(X_{cN}(x_2) \leq 2N) + P(X_t(x_2) \leq -N \text{ for some } t < cN) \leq C_7 \exp(-\gamma_7 \sqrt{N}).$$

Proof This follows directly from the previous lemmas. First, since

$$\Sigma_t = X_t(-1) + X_t(0) + X_t(1) \leq 3X_t(x_2) + 2G_t \quad \text{for all } t \geq 0 \text{ and } x_2 = -1, 0, 1,$$

a straightforward application of Lemmas 9 and 11 gives

$$P(X_{cN}(x_2) \leq 2N) \leq P(\Sigma_{cN} < 7N) + P(G_{cN} > \sqrt{N}) \leq C_8 \exp(-\gamma_8 \sqrt{N})$$

for suitable $C_8 < \infty$ and $\gamma_8 > 0$ and all N sufficiently large. Similarly,

$$\begin{aligned} P(X_t(x_2) \leq -N \text{ for some } t < cN) \\ \leq P(\Sigma_t < -N \text{ for some } t < cN) \\ + P(G_t > \sqrt{N} \text{ for some } t < cN) \leq C_9 \exp(-\gamma_9 \sqrt{N}) \end{aligned}$$

for suitable $C_9 < \infty$ and $\gamma_9 > 0$ and all N sufficiently large. \square

To complete the proof of Theorem 4, we now return to the majority rule model. To compare the process properly rescaled in space and time with oriented site percolation, we let $T = cN$ where c is the positive constant introduced in Lemma 9, and define for all $w \in \mathbb{Z}^2$

$$B_w = (2N + 1)w + [-N, N]^2 \quad \text{and} \quad \mathcal{G} = \{(w, j) \in \mathbb{Z}^2 \times \mathbb{Z}_+ : w_1 + w_2 + j \text{ is even}\}.$$

Site $(w, j) \in \mathcal{G}$ is said to be good whenever all vertices in B_w are in state 1 at time jT for the original majority rule model. Then, we have the following lemma.

Lemma 13 *For all N sufficiently large,*

$$P((e_1, 1) \text{ is not good} \mid (0, 0) \text{ is good}) \leq C_7(6N - 3) \exp(-\gamma_7 \sqrt{N}).$$

Proof The idea is to observe that the majority rule model is stochastically larger than a certain union of bidirectional slice processes. More precisely, for all integers $z \in \mathbb{Z}$, we let r_t^z denote the process obtained from the majority rule model in the same manner as the slice process introduced above but applying the translation of vector (N, z) to both the

evolution rules and the initial configuration. Also, we let l_t^z denote the process obtained from r_t^z by applying the symmetry about the vertical axis to the evolution rules and the initial configuration. In particular, we have

$$r_t^z \stackrel{d}{=} \bar{\eta}_t + (N, z) \quad \text{and} \quad l_t^z \stackrel{d}{=} -\bar{\eta}_t + (-N, z) \quad \text{for all } z \in \mathbb{Z} \quad (5)$$

where $\stackrel{d}{=}$ means equal in distribution. Having the majority rule model and all these processes constructed from the same collection of independent rate one Poisson processes, and identifying each spin system with its set of vertices in state 1, on the event that

$$(-\infty, 0] \times \{z-1, z, z+1\} \subset r_t^z \quad \text{and} \quad [0, +\infty) \times \{z-1, z, z+1\} \subset l_t^z \quad (6)$$

for all $t \leq T$ and all $z \in \{-(N-1), \dots, N-1\}$, we have

$$\bigcup_{z=-(N-1)}^{N-1} (r_T^z \cap l_T^z) \subset \eta_T \quad \text{provided} \quad \eta_0 = \bigcup_{z=-(N-1)}^{N-1} (r_0^z \cap l_0^z) = B_0. \quad (7)$$

Combining (5)–(7) with Corollary 12, we obtain

$$\begin{aligned} P((e_1, 1) \text{ is not good} \mid (0, 0) \text{ is good}) &\leq P(B_{e_1} \not\subset \eta_T \mid \eta_0 = B_0) \\ &\leq P((-\infty, 0] \times \{z-1, z, z+1\} \not\subset r_t^z \\ &\quad \text{for some } (z, t) \in \{-(N-1), \dots, N-1\} \times (0, T)) \\ &\quad + P((-\infty, 3N] \times \{z-1, z, z+1\} \not\subset r_T^z \text{ for some } z \in \{-(N-1), \dots, N-1\}) \\ &\leq (2N-1) \times P((-\infty, -N] \times \{-1, 0, 1\} \not\subset \bar{\eta}_t \text{ for some } t \in (0, T)) \\ &\quad + (2N-1) \times P((-\infty, 2N] \times \{-1, 0, 1\} \not\subset \bar{\eta}_{cN}) \\ &\leq C_7(6N-3) \exp(-\gamma_7 \sqrt{N}) \end{aligned}$$

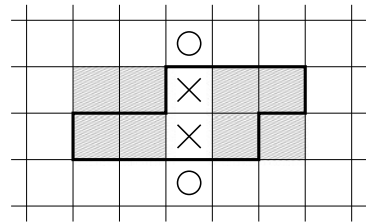
for all N sufficiently large, as desired. \square

Since the probability in the statement of Lemma 13 can be made arbitrarily small by choosing the parameter N sufficiently large, Theorem 4.3 in Durrett [9] implies that the set of good sites dominates stochastically the set of wet sites of an oriented site percolation process on \mathcal{G} where sites are open with probability arbitrarily close to one. This only proves survival of the type 1 opinion since the percolation process has a positive density of closed sites, and thus a positive density of dry sites, i.e., sites which are not wet. To conclude, we apply Lemma 15 of [14], which relies on ideas from Durrett [8] and proves the lack of percolation of the dry sites for oriented percolation on a certain directed graph with vertex set \mathcal{G} when the density of open sites is large enough. This lemma and the construction given in its proof imply the existence of an in-all-directions expanding region which is void of vertices in state 0, so opinion 1 indeed outcompetes opinion 0 when starting with infinitely many 2×2 squares in state 1. This completes the proof of Theorem 4.

5 Proof of Theorem 5 ($d = 2$ and $n = 3$)

This last section is devoted to the proof of Theorem 5, which relates the variation rate of the number of type 1 vertices to the number of positive and negative corners when the configuration is a regular cluster. Recall that configuration η is a regular cluster whenever

Fig. 9 Picture related to the proof of Lemma 14



(R0) the contour Γ as defined in Sect. 2 is a Jordan curve,
 (R1) the set $(x + D_2) \cap \Gamma$ is connected for all $x \in \mathbb{D}^2$ and
 (R2) if vertex x is a corner and $(x + D_2) \cap (y + D_2) \neq \emptyset$ then vertex y is not a corner,
 where $D_2 = [-1, 1]^2$. The proof is divided into three steps: we first prove a geometric property of regular clusters, then establish the theorem in the particular case when there is no positive corner nor negative corner, and finally combine these two results to obtain the full theorem.

Lemma 14 Assume that η is a regular cluster with at least 11 vertices. If two nearest neighbors, say vertex x and vertex $x + e_i$ for some $i \in \{1, 2\}$, are in different states then

$$\eta(x - 2e_i) = \eta(x - e_i) = \eta(x) \neq \eta(x + e_i) = \eta(x + 2e_i) = \eta(x + 3e_i).$$

Proof Accounting for the invariance of the problem by translation and rotation, it suffices to prove that each of the following four scenarios leads to a contradiction:

1. $\eta(e_2) \neq \eta(0) = \eta(2e_2) = 1$
2. $\eta(e_2) = \eta(2e_2) \neq \eta(0) = \eta(3e_2) = 1$
3. $\eta(e_2) \neq \eta(0) = \eta(2e_2) = 0$
4. $\eta(e_2) = \eta(2e_2) \neq \eta(0) = \eta(3e_2) = 0$

We focus on conditions 2 and 4 since the remaining two conditions can be excluded based on the same approach. Using the notations of the tic-tac-toe game by denoting each state by \times or \circ where \times means either state 0 or state 1 and \circ means the other state, both conditions 2 and 4 result in the configuration of \times and \circ given in Fig. 9, which implies the existence of two horizontal line segments of length one that must be subsets of the contour. According to (R1), the hatched square on the left of the picture must contain a path γ_1 that connects the left extremities of the two segments. Similarly, the right extremities must be connected by a path γ_2 contained in the hatched square on the right of the picture. The concatenation of the two segments and the two paths defines a Jordan curve γ . To conclude, we distinguish between the following two conditions.

1. If \times means state 1 then $\gamma \subsetneq \Gamma$ contradicting the fact that Γ is a Jordan curve.
2. If \times means state 0 then $\gamma = \Gamma$ indicating that the cluster contains at most 10 vertices, which again leads to a contradiction.

This completes the proof. □

Lemma 15 Theorem 5 holds when $c_- = c_+ = 0$.

Proof The proof relies on a geometric construction much easier to visualize than to explain so we refer the reader to Fig. 10 for pictures that help to understand our approach. The basic

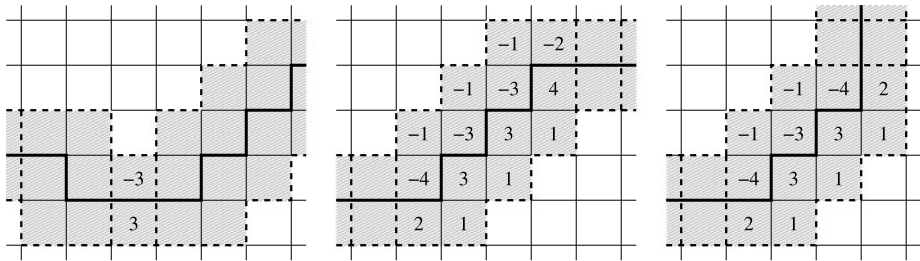


Fig. 10 Pictures related to the proof of Lemma 15

idea is to define a partition $\{\Delta_1, \Delta_2, \dots, \Delta_k\}$ of the support of $\phi(\eta, \cdot)$ such that the property to be proved holds for each member of the partition, i.e.,

$$\sum_{x \in \Delta_j} \phi(\eta, x) = 0 \quad \text{for all } j = 1, 2, \dots, k. \quad (8)$$

First, we define an oriented contour embedded in the Jordan curve Γ by letting

$$\{x_1, x_2, \dots, x_m\} = \Gamma \cap \mathbb{D}^2$$

where x_i is the i th point we encounter going around the curve starting from a given point and following a given orientation. To turn the Jordan curve into an oriented contour, we draw an arrow from vertex x_i to vertex x_j whenever $j = i + 1 \bmod m$, in which case both vertices are nearest neighbors on the dual lattice. Then, we define

$$\Lambda = \bigcup_{i=1}^m (x_i + D_2) \quad \text{and} \quad \Delta = \text{supp } \phi(\eta, \cdot) = \Lambda \cap \mathbb{Z}^2.$$

If the arrows (x_{i-1}, x_i) and (x_i, x_{i+1}) are oriented in the same direction, we draw a segment line of length two centered at x_i and perpendicular to the segment (x_{i-1}, x_{i+1}) . This induces a partition of the set Λ and a partition of the support Δ , i.e.,

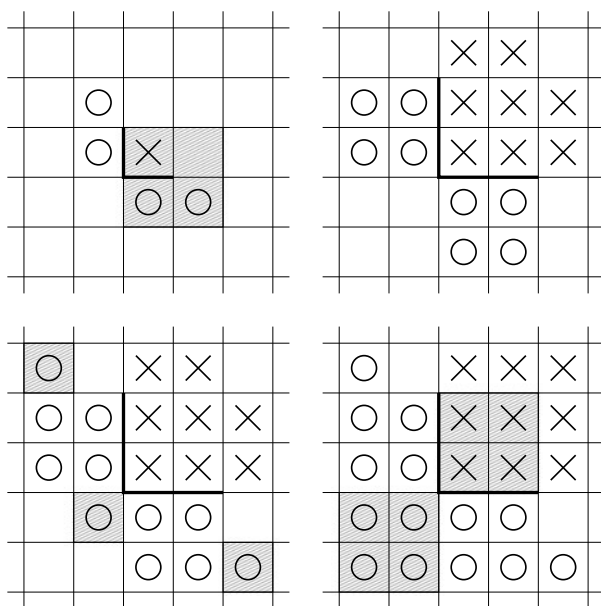
$$\Lambda = \bigcup_{j=1}^k \Lambda_j \quad \text{and} \quad \Delta = \bigcup_{j=1}^k \Delta_j = \bigcup_{j=1}^k (\Lambda_j \cap \mathbb{Z}^2)$$

where the unions are disjoint. In Fig. 10, the sets Λ_j are delimited by dashed lines. Now, thinking of the contour as a sequence of length m consisting of four different types of arrows, a direct application of Lemma 14 implies that two consecutive vertical arrows going in opposite direction must be separated by at least three horizontal arrows all going in the same direction. The same holds by exchanging the role of vertical and horizontal arrows. Each sequence of $l \geq 3$ consecutive arrows oriented in the same direction induces $l - 2$ members of the partition of the support with exactly two vertices such that

$$\Delta_j = \{x, y\} \quad \text{with } \phi(\eta, x) + \phi(\eta, y) = 3 - 3 = 0.$$

See the first picture of Fig. 10 for an illustration. Finally, the absence of positive and negative corners implies that the path $(\rightarrow, \rightarrow, \uparrow, \uparrow)$ as well as the seven other paths deduced by symmetry or rotation are not allowed. In particular, accounting again for symmetry and rotation, for all j such that $\text{card } \Delta_j \neq 2$, we must have

$$\Lambda_j \cap \Gamma = (\rightarrow, \uparrow, \rightarrow, \uparrow, \dots, \rightarrow, \uparrow) \quad \text{or} \quad \Lambda_j \cap \Gamma = (\rightarrow, \uparrow, \rightarrow, \uparrow, \dots, \rightarrow, \uparrow, \rightarrow)$$

Fig. 11 Pictures related to the proof of Lemma 16

where both paths have length at least three, both paths must be preceded by a \rightarrow , the first path must be followed by a \uparrow and the second path must be followed by a \rightarrow . Such paths are represented in the right hand side of Fig. 10 which gives the values of $\phi(\eta, \cdot)$ in the set Δ_j and shows that (8) is indeed satisfied. Since the sets Δ_j form a partition, the proof is complete. \square

Lemma 16 *Theorem 5 holds for all values of c_- and c_+ .*

Proof The first step is to characterize the configuration of the process in the 5×5 square centered at a positive or negative corner, which is illustrated in Fig. 11. Following the notations introduced above, we denote both states by \times and \circ respectively. Since the problem is invariant by translation and rotation, we may assume without loss of generality that vertex 0 is a corner with

$$\eta(0) \neq \eta(-e_1 + e_2) = \eta(e_1 - e_2)$$

as illustrated by the top left picture of the figure. We now prove that the configuration constructed step by step in the figure is indeed the only possible configuration.

Top left picture The top left corner and the bottom right corner of the unit square centered at vertex 0 must belong to the Jordan curve Γ . Invoking condition (R1), we deduce the existence of a path connecting these two points and included in the hatched square of the picture, which further implies that the Jordan curve must contain the top side or the left side of the unit square centered at 0. By symmetry, it must also contain the bottom side or the right side. A direct application of Lemma 14 shows that two opposite sides of the unit square cannot simultaneously be included in the curve Γ so we may assume without loss of generality that the left and bottom sides of the unit square, drawn in thick lines in the picture, are included in the Jordan curve, which also determines the state of the vertices to the left of and under vertex 0.

Top right picture Applying Lemma 14 repeatedly from the previous picture gives the state of a total of 16 vertices included in the 5×5 square, as shown in this picture.

Bottom left picture Invoking condition (R2), the three hatched unit squares in this picture cannot be positive or negative corners, which forces their state to be \circ rather than \times .

Bottom right picture The top right corner of the 2×2 hatched square on the left is a point of the Jordan curve Γ therefore the presence of a \times in this hatched square would contradict condition (R1). It follows that all four vertices in this square are \circ . Using a similar reasoning with the other hatched square, we prove that the vertex at the top right of the picture must be \times , while the remaining two vertices with no symbol can be of either type.

To deduce the theorem, we let $\bar{\eta}$ denote the configuration obtained from η by switching the state of vertex 0 and leaving the state of all other vertices unchanged. It is straightforward to check that, regardless of the type of the remaining two vertices, none of the nine 3×3 squares containing vertex 0 contains exactly five \times 's, indicating that the majority type in each of these nine square is not modified by switching the state of 0. In particular, each of the nine corresponding updates results in the same configuration that we start from η or from $\bar{\eta}$ which further implies that

$$\phi(\eta, x) = \begin{cases} \phi(\bar{\eta}, x) + 1 & \text{if } \eta(0) = 0 \\ \phi(\bar{\eta}, x) - 1 & \text{if } \eta(0) = 1 \end{cases} \quad (9)$$

for every vertex x in the 3×3 square centered at 0. More generally, letting $\bar{\eta}$ be the configuration obtained from η by switching the state of all corners, we get

$$\sum_{x \in \mathbb{Z}^2} \phi(\eta, x) = \sum_{x \in \mathbb{Z}^2} \phi(\bar{\eta}, x) + 9(c_- - c_+) = 9(c_- - c_+)$$

where the first equation is obtained by applying (9) at each corner, and where the second equation follows from Lemma 15 and the fact that the configuration $\bar{\eta}$ obtained by switching the state of each corner is a regular cluster with no corner. This completes the proof. \square

Acknowledgement The authors would like to thank two anonymous referees for many comments that helped to improve the clarity of this article.

References

1. Bramson, M., Durrett, R.: A simple proof of the stability criterion of Gray and Griffeath. *Probab. Theory Relat. Fields* **80**, 293–298 (1988)
2. Bramson, M., Griffeath, D.: Renormalizing the 3-dimensional voter model. *Ann. Probab.* **7**, 418–432 (1979)
3. Bramson, M., Griffeath, D.: Clustering and dispersion rates for some interacting particle systems on \mathbb{Z} . *Ann. Probab.* **8**, 183–213 (1980)
4. Castellano, C., Fortunato, S., Loreto, V.: Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**, 591–646 (2009)
5. Clifford, P., Sudbury, A.: A model for spatial conflict. *Biometrika* **60**, 581–588 (1973)
6. Cox, J.T., Griffeath, D.: Occupation time limit theorems for the voter model. *Ann. Probab.* **11**, 876–893 (1983)
7. Cox, J.T., Griffeath, D.: Diffusive clustering in the two-dimensional voter model. *Ann. Probab.* **14**, 347–370 (1986)
8. Durrett, R.: Multicolor particle systems with large threshold and range. *J. Theor. Probab.* **5**, 127–152 (1992)
9. Durrett, R.: Ten lectures on particle systems. In: *Lectures on Probability Theory*, Saint-Flour, 1993. *Lecture Notes in Math.*, vol. 1608, pp. 97–201. Springer, Berlin (1995)

10. Durrett, R., Steif, J.E.: Fixation results for threshold voter systems. *Ann. Probab.* **21**, 232–247 (1993)
11. Galam, S.: Minority opinion spreading in random geometry. *Eur. Phys. J. B* **25**, 403–406 (2002)
12. Harris, T.E.: Nearest neighbor Markov interaction processes on multidimensional lattices. *Adv. Math.* **9**, 66–89 (1972)
13. Holley, R.A., Liggett, T.M.: Ergodic theorems for weakly interacting systems and the voter model. *Ann. Probab.* **3**, 643–663 (1975)
14. Lanchier, N.: 2012, Stochastic spatial model of producer-consumer systems on the lattice. Preprint
15. Zähle, I.: Renormalization of the voter model in equilibrium. *Ann. Probab.* **29**, 1262–1302 (2001)