EVOLUTION AND STEADY STATE OF A LONG-RANGE TWO-DIMENSIONAL SCHELLING-TYPE SPIN SYSTEM∗

BY HAMED OMIDVAR AND MASSIMO FRANCESCETTI

University of California, San Diego

We consider a long-range interacting particle system in which binary particles are located at the integer points of a flat torus. Based on the interactions with other particles in its “neighborhood” and on the value of a common intolerance threshold \( \tau \), every particle decides whether to change its state after an independent and exponentially distributed waiting time. This is equivalent to a Schelling model of self-organized segregation in an open system, a zero-temperature Ising model with Glauber dynamics, or an Asynchronous Cellular Automaton (ACA) with extended Moore neighborhoods.

We first prove a shape theorem for the spread of the “affected” nodes during the process dynamics. Second, we show that when the process stops, for all \( \tau \in (\tau^*, 1 - \tau^*) \setminus \{1/2\} \) where \( \tau^* \approx 0.488 \), and when the size of the neighborhood of interaction \( N \) is sufficiently large, any particle is contained in a large “monochromatic region” of size exponential in \( N \), almost surely. When particles are placed on the infinite lattice \( \mathbb{Z}^2 \) rather than on a flat torus, for the values of \( \tau \) mentioned above, sufficiently large \( N \), and after a sufficiently long evolution time, any particle is contained in a large monochromatic region of size exponential in \( N \), almost surely.

1. Introduction.

1.1. Background. Consider the graph formed by nodes placed at the integer points of a large flat torus and edges connecting each node to all the ones located in a small square neighborhood of itself. Put a particle at each node such that its initial binary state is chosen independently and uniformly at random. Each particle is then labeled as follows. All particles have a common intolerance threshold \( 0 < \tau < 1 \), indicating the minimum fraction of

∗An extended abstract of an earlier version of this paper containing a subset of the results and with most proofs omitted appears in the proceedings of ACM Symposium on Theory of Computing (STOC ’18). This work was partially supported by Army Research Office (ARO), award number W911NF-15-1-0253.

MSC 2010 subject classifications: Primary 60K35, 82C22, 82C43, 82B43

Keywords and phrases: Interacting Particle System (IPS), Agent-Based Model, Asynchronous Cellular Automaton, Zero-Temperature Ising Model, Unperturbed Schelling Segregation, Distributed Algorithm, Self-Organized Segregation, Percolation Theory, First Passage Percolation, Exponential Segregation

arXiv:1804.00358v2  [math.PR]  18 May 2018
particles in their same state that must be in their neighborhood to make them “stable.” Each particle is assigned an independent and identical Poisson clock, and when the clock rings the particle’s state is flipped if the particle is “unstable” and the flip will make it stable. This change is then immediately detected by the neighbors who update their labels accordingly.

In social sciences and economy, this model is known as the Schelling model in an “open” system [33, 34]. In computation theory, mathematics, physics, complexity theory, theoretical biology and material sciences, it is known as a two-dimensional, two-state Asynchronous Cellular Automaton (ACA) with extended Moore neighborhoods and exponential waiting times [11]. Related models appeared in epidemiology [19, 12], economics [22], engineering and computer sciences [26, 14]. Mathematically, all of them fall in the general area of interacting particle systems [27, 28]. For an intolerance value of 1/2, the model corresponds to the Ising model with zero temperature, which exhibits spontaneous magnetization as spins align along the direction of the local field [36, 10].

The dynamics of these processes can be roughly divided into two classes. Kawasaki dynamics assume that pairs of unstable particles swap their locations if this will make both of them stable. Glauber dynamics assume unstable particles to simply flip their state if this makes them stable. This flipping action indicates that the particle has moved out of the system and a new particle has occupied its location. The Kawasaki dynamics correspond to a “closed” system where the number of particles of each state is fixed, while the Glauber dynamics correspond to an “open” system where the number of particles of each state can change over time. In this paper, we consider the Glauber dynamics, and assume unstable particles to flip their state only if this makes them stable. Other variants are possible, including having unstable particles swap (or flip) regardless of whether this makes them stable or not, or to assume that particles have a small probability of acting differently than what the general rule prescribes, have multiple intolerance levels, multiple states, different distributions, and time-varying intolerance [39, 40, 41, 42, 30, 29, 7, 35, 4, 5].

A common effect observed by simulating several variants of the model is that when the system reaches a steady state, large monochromatic areas of particles with the same state are formed, for a wide range of the intolerance threshold. This corresponds to observing spontaneous self-organization resulting from local interactions. See Figure 1 for a simulation of this behavior.

1.2. Our contribution and related works. Although simulation results have been available for a long time, rigorous results appeared only recently
Fig 1. Self-organization arising over time for a value of the intolerance $\tau = 0.42$ on a flat torus of size $1000 \times 1000$ and neighborhood size 441. Green and blue indicate areas of “stable” particles in states (+1) and (-1), respectively. White and yellow indicate areas of “unstable” particles at states (+1) and (-1) respectively. Initial configuration (a), transient configurations (b)-(c), steady state (d). When the process terminates all particles are stable and large monochromatic regions can be observed.

Fig 2. For $\tau \in (\tau_*, 1 - \tau_*) \setminus \{1/2\}$, where $\tau_* \approx 0.433$, we prove a shape theorem for the spread of the “affected” nodes during the process dynamics (gray plus black region) and then show that in the steady state, for all $\tau \in (\tau^*, 1 - \tau^*) \setminus \{1/2\}$ where $\tau^* \approx 0.488$, and when the size of the neighborhood of interaction $N$ is sufficiently large, any particle is contained in a large “monochromatic region” of size exponential in $N$, almost surely (gray region).

even for the one-dimensional version of the model. Brandt et al. [8] considered a ring graph for the Kawasaki model of evolution. In this setting, letting the neighborhood of a particle be the set of nearby particles that is used to determine whether the particle is stable or not, they showed that for an intolerance level $\tau = 1/2$, the expected size of the largest monochromatic region containing an arbitrary particle in steady state is polynomial in the size of the neighborhood. Barmpalias et al. [3] showed that there exists a value of $\tau^* \approx 0.35$, such that for all $\tau < \tau^*$ the initial configuration remains static with high probability (w.h.p.), while for all $\tau^* < \tau < 1/2$ the size of the monochromatic region in steady state becomes exponential in the size of the neighborhood w.h.p. On the other hand, for all $\tau > 1/2$ the system evolves w.h.p. towards a state with only two monochromatic components. For the Glauber model the behavior is similar, but symmetric around $\tau = 1/2$, with a first transition from a static configuration to
exponential monochromatic regions occurring at \( \tau \approx 0.35 \), a special point \( \tau = 1/2 \) with the monochromatic regions of expected polynomial size, then again exponential monochromatic regions until \( \tau \approx 0.65 \), and finally a static configuration for larger values of \( \tau \). Holden and Sheffield [20] have considered the case \( \tau = 1/2 \) and studied the dynamical scaling limit as the size of the neighborhood tends to infinity and the lattice is correspondingly rescaled.

In the two-dimensional model, the case \( \tau = 1/2 \) is open. Immorlica et al. [21] have shown for the Glauber dynamics the existence of a value \( \tau^* < 1/2 \), such that for all \( \tau^* < \tau < 1/2 \) the expected size of the monochromatic region is exponential in the size of the neighborhood. This shows that exponential monochromatic regions are expected in the small interval \( \tau \in (1/2-\epsilon, 1/2) \). Barmpalias et al. [6] considered a model in which particles in different states have different intolerance parameters, i.e., \( \tau_1 \) and \( \tau_2 \). For the special case of \( \tau_1 = \tau_2 = \tau \), they have shown that when \( \tau > 3/4 \), or \( \tau < 1/4 \), the initial configuration remains static w.h.p.

In a previous work by the authors [32], the intolerance interval that leads to the formation of large monochromatic regions has been enlarged from the size \( \epsilon > 0 \) to size \( \approx 0.134 \), namely when \( 0.433 < \tau < 1/2 \) (and by symmetry \( 1/2 < \tau < 0.567 \)), the expected size of the largest monochromatic region is exponential in the size of the neighborhood. In addition, the interval leading to large monochromatic regions has been further extended to size \( \approx 0.312 \).

In this case, “almost monochromatic” regions have been considered, namely regions where the ratio of the number of particles in one state and the number of particles in the other state quickly vanishes as the size of the neighborhood grows, and it has been shown that for \( 0.344 < \tau \leq 0.433 \) (and by symmetry for \( 0.567 \leq \tau < 0.656 \)) the expected size of the largest almost monochromatic region is exponential in the size of the neighborhood.

The first contribution of this paper is the development of a shape theorem for the spread of “affected” nodes – namely nodes on which a particle would be unstable in exactly one of its states – during the process dynamics. The theorem gives a precise geometrical description of the set of affected nodes at any given time, and is a consequence of a concentration bound that we develop for the spreading time. It is the first result that precisely describes the transient dynamics of the spreading process. Our second contribution is determining the size of the monochromatic region in the steady state, for a given interval of \( \tau \). A weakness of all current results for the two-dimensional case is that they obtain lower bounds on the expected size of the monochromatic region containing a given particle, but they do not show that in the steady state any particle ends up in an exponentially large monochromatic region almost surely, nor with high probability. A possibility that is
consistent with the results in the literature (but inconsistent with the simulation results) is that only an exponentially small fraction of the nodes are contained in large monochromatic regions at the end of the process, but that those regions are so large that the expected radius of the monochromatic region containing any node is exponentially large. For this reason, current results leave a large gap in our qualitative understanding of the two-dimensional process. We close this gap by showing that when the process stops, for all \( \tau \in (\tau^*, 1 - \tau^*) \setminus \{1/2\} \) where \( \tau^* \approx 0.488 \), and when the size of the neighborhood of interaction \( N \) is sufficiently large, any particle is contained in a large “monochromatic region” of size exponential in \( N \), almost surely. When particles are placed on the infinite lattice \( \mathbb{Z}^2 \) rather than on a flat torus, for the values of \( \tau \) mentioned above, sufficiently large \( N \), and after a sufficiently long evolution time, any particle is contained in a large monochromatic region of size exponential in \( N \), almost surely (see Figure 2).

We now mention some additional related works. For the case of constant neighborhood, Fontes et al. [16] have shown the existence of a critical probability \( 1/2 < p^* < 1 \) for the initial Bernoulli distribution of the particle states such that for \( \tau = 1/2 \) and \( p > p^* \) the Glauber model on the \( d \)-dimensional grid converges to a state where only particles in one state are present. This shows that complete monochromaticity occurs w.h.p. for \( \tau = 1/2 \) and \( p \in (1 - \epsilon, 1) \). Morris [31] has shown that \( p^* \) converges to \( 1/2 \) as \( d \to \infty \). Caputo and Martinelli [9] have shown the same result for \( d \)-regular trees, while Kanoria and Montanari [23] derived it for \( d \)-regular trees in a synchronous setting where flips occur simultaneously, and obtained lower bounds on \( p^*(d) \) for small values of \( d \). The case \( d = 1 \) was first investigated by Erdős and Ney [15], and Arratia [2] has proven that \( p^*(1) = 1 \).

The rest of the paper is organized as follows. In section 2 we introduce the model, state our results, and give a summary of the proof construction. In section 3 we provide a few preliminary results along with some results from previous works. In section 4 we develop the concentration bound for the spreading time of the affected nodes. In section 5 we use this concentration bound to obtain a shape theorem for the spread of affected nodes. In section 6 we prove a size theorem for the configuration of particles in the steady state.

2. Model and Main Results.

2.1. The Model.

Initial Configuration. Consider a node at each integer point of a flat torus \( T = [-h, h] \times [-h, h] \) where \( h \in \mathbb{N} \) and let the horizon \( w \in o(\sqrt{\log h}) \) be a natural number such that each node is connected to all the nodes located within an \( l_\infty \) neighborhood of radius \( w \) of itself. We place a particle at each
node of the resulting graph $G_w$. Each particle has a binary state which is chosen independently at random to be (+1) or (-1) according to a Bernoulli distribution of parameter $p = 1/2$. We use $\theta$ to denote an unknown or unspecified state and $\bar{\theta}$ to denote its complement state and by a $\theta$-particle we mean a particle that is in state $\theta$.

A neighborhood is a connected sub-graph of $G_w$. A neighborhood of radius $\rho$, denoted by $\rho$-neighborhood or $\mathcal{N}_\rho$, is the set of all particles with $l_\infty$ distance at most $\rho$ from a central node. The size of a neighborhood is the number of particles in it. The neighborhood of a particle $u$ is a neighborhood of radius equal to the horizon and centered at $u$, and is denoted by $\mathcal{N}(u)$. Without loss of generality, we derive our result for an arbitrary particle located at the origin $0 \equiv (0,0) \in \mathbb{T}$.

**Dynamics.** We let the rational $\tau$, called intolerance, be $\lceil \tilde{\tau}N \rceil / N$, where $\tilde{\tau} \in [0,1]$ and $N = (2w+1)^2$ is the size of the neighborhood of a particle. The integer $\tau N$ represents the minimum number of particles in the same state as $u$ that must be present in $\mathcal{N}(u)$ to make $u$ stable. More precisely, for every particle $u$, we let $s(u)$ be the ratio between the number of particles that are in the same state as $u$ in its neighborhood and the size of the neighborhood. At any point in continuous-time, if $s(u) \geq \tau$ then $u$ is labeled **stable**, otherwise it is labeled **unstable**. We assign independent and exponentially distributed waiting times to unstable particles, and after every waiting time, the state of the particle is flipped if and only if the particle is still unstable and this flip will make the particle stable. Two observations are now in order. First, for $\tau < 1/2$ flipping its state will always make an unstable particle stable, but this is not the case for $\tau > 1/2$. Second, the process dynamics are equivalent to a discrete-time model where at each discrete time step one unstable particle is chosen uniformly at random and its state is flipped if this will make the particle stable.

**Affected Nodes.** A node in $G_w$ is called $\theta$-affected whenever a $\theta$-particle located there would be unstable and a $\bar{\theta}$-particle would be stable.

**Steady State.** The process continues until there are no unstable particles left, or there are no unstable particles that can become stable by flipping their state. By defining a Lyapunov function to be the sum over all particles $u$ of the number of particles of the same state as $u$ present in its neighborhood, it is easy to argue that the process indeed terminates. We call this final state the **steady state**.
Monochromatic Region. At each point in time, the monochromatic regions of a particle \( u \) are the neighborhoods with largest radius that contain particles in a single state and that also contain \( u \). The monochromatic region of \( u \) is one of these regions chosen arbitrarily.

Throughout the paper we use the terminology with high probability (w.h.p.) meaning that the probability of an event is \( 1 - o(w^{-2}) \) as \( w \to \infty \). Almost surely (a.s.) instead indicates with probability equal to one (w.p.1).

2.2. Main Results. To state our first result, we let \( \tau_* \approx 0.433 \) be the solution of

\[
\frac{3}{4} \left( 1 - H \left( \frac{4}{3} \tau_* \right) \right) - (1 - H(\tau_*)) = 0,
\]

where \( H(\tau) = -\tau \log_2(\tau) - (1 - \tau) \log_2(1 - \tau) \) is the binary entropy function. Also, let \( \tau \in (\tau_*, 1/2) \cup (1/2, 1 - \tau_*) \), \( 0 < c_1 < c_2 < 0.5(1 - H(\tau)) \), and \( 2^{c_1 N} \leq n \leq 2^{c_2 N} \). Consider our process on \( G_w \). Let \( l \) be a norm on \( \mathbb{R}^2 \) and \( B_l(0, n) \) denote a ball of radius \( n \) in this norm, centered at the origin. Finally, let \( A_F(0, n) \) denote the set of \( \theta \)-affected nodes at time \( n \) in a neighborhood with radius \( 2^{c_2 N + 1} \) centered at the origin when initially all the particles in a \( w/2 \)-neighborhood centered at the origin are \( \theta \)-affected and when the independent and identical distribution of the waiting times of unstable particles (called flipping times) is \( F \). Since we consider Glauber dynamics this means \( F \) is exponential, however, all of our results hold for distributions that satisfy the following conditions:

\[
F(x) = 0 \text{ for } x \leq 0, \\
F \text{ is not concentrated on one point,} \\
\exists \gamma > 0, \text{ such that } \int e^{\gamma x} F(dx) < \infty.
\]

The described neighborhood centered at the origin, starts w.h.p. a cascading process leading to more and more affected nodes. The following theorem shows w.h.p. the existence of a norm \( l_* \) such that the affected nodes at time \( n \) are contained in \( B_{l_*}(0, n + o(n)) \), and the nodes in \( B_{l_*}(0, n - o(n)) \cap \mathbb{Z}^2 \) are all affected (see Figure 3).
H. OMIDVAR AND M. FRANCESCHETTI.

Theorem 1 (Shape Theorem —Transient). \(W.h.p.\) there exists a norm \(l_\ast\) on \(\mathbb{R}^2\), and \(c > 0\), such that,

\[
B_{l_\ast}(0, n - N^c n^{1/2} \log^{3/2} n) \cap \mathbb{Z}^2 \subset A_F(0, n) \subset B_{l_\ast}(0, n + N^c n^{1/2} \log^{3/2} n).
\]

Remarks. Here, without loss of generality, we have assumed that the linear time scale is chosen such that the re-scaled limit shape is a unit \(l_\ast\)-ball.

Similar shape theorems have been proven in the literature for percolation models and other processes [24, 25, 1, 13, 38], but none of them applies to our model.

To state our second result, we let \(\tau^* \approx 0.488\) be the solution of

\[
5 \left(1 + f(\tau)\right)^2 - 6 = 0,
\]

where

\[
f(\tau) = \frac{3(\tau - 0.5) + \sqrt{9(\tau - 0.5)^2 - 7(\tau - 0.5)(3\tau + 0.5)}}{2(3\tau + 0.5)}.
\]

Theorem 2 (Size Theorem —Steady State). Let \(M\) denote the size of the monochromatic region of the particle at the origin in the steady state. For all \(\tau \in (\tau^*, 1 - \tau^*) \setminus \{1/2\}\) and for sufficiently large \(N\), almost surely,

\[
2^a(\tau)^N \leq M \leq 2^{b(\tau)N}.
\]

The numerical values for \(a(\tau)\) and \(b(\tau)\) derived in the proof of the above theorem are plotted in Figure 4. For \(\tau \in (\tau^*, 1 - \tau^*) \setminus \{1/2\}\), as the intolerance gets farther from one half in both directions, larger monochromatic regions are formed \(w.h.p.\).
2.3. The Infinite Lattice Case. We can consider similar dynamics occurring on the infinite lattice $\mathbb{Z}^2$ instead of the finite setting described above. Using Theorem B3 of [28, p. 3] it is easy to verify that the process on $\mathbb{Z}^2$ exists, and is unique, and is a Feller Markov process on $\{-1, +1\}^\mathbb{Z}^2$. The following corollary follows from the proof of Theorem 1.

**Corollary 1 (Size Theorem —Infinite Lattice).** Let $M_n$ denote the size of the monochromatic region of the particle at the origin at time $n$. For all $\epsilon > 0$, $\tau \in (\tau^*, 1 - \tau^*) \setminus \{1/2\}$, let $n^* = 2^{(a(\tau) + \epsilon)N}$. For sufficiently large $N$, and all $n \geq n^*$, almost surely,

$$2^{a(\tau)N} \leq M_n \leq 2^{b(\tau)N}. \quad (2.9)$$

2.4. Proof Outline.

The shape theorem. To prove the shape theorem, we adapt a strategy developed by Tessera [38] for first passage percolation (FPP) to our process on $G_w$. However, while Tessera’s result relies on Talagrand’s concentration inequality [37] for the spreading time in FPP, our shape theorem is based on a concentration bound that we develop independently for the spread of affected nodes in our process on $G_w$, by extending some results of Kesten’s [24, 25]. This bound is a key step in our proof.
We show that the set of affected nodes at time $n$ is close to the set of nodes whose expected time of becoming affected is at most $n$, and that there exists a norm $l_*$ such that the latter set is also close to an $l_*$-ball of radius $n$. The former statement is proved using the concentration bound (our Theorem 4), and the latter statement is proved showing that the expected time of the spread from one point to another in $G_w$ has an approximate geodesic (again using Theorem 4).

To derive the bound on the spreading time of the affected nodes, we represent the difference between the random spreading time and the mean spreading time between any two nodes as a sum of martingale differences and, after estimating the sum of squares of these differences, apply a martingale inequality developed by Kesten (re-stated as Theorem 5 of this paper). Along the way, to bound the martingale differences, we use a modified result from [24] to compare our process with two FPP processes on $\mathbb{Z}^d$.

The size theorem. The main idea of the proof in this case is to show that w.h.p. while the spread of the $\theta$-affected nodes (i.e., nodes on which a $\theta$-particle would be unstable) reaches the origin, the $\theta$-affected nodes are still at distances at least exponential in $N$ from the origin. Once the origin is reached, the unstable particles around it will w.h.p. lead to the formation of an exponentially large “firewall” that is indestructible by other spreading processes. The interior of this firewall will then become monochromatic, so that in the steady state there will be w.h.p. an exponentially large monochromatic region around the origin.

To elaborate on this main idea, we define an expandable region, that is composed of a local configuration of particles and a possible set of flips inside it, that can lead to at least one new affected node outside of it. We consider
the expandable region closest to the origin in the $l_*$ norm, and denote its type by $\theta$. We denote its $l_*$-distance to the origin by $X$, and consider an $l_*$-ball of radius $X$ at the origin. We then argue that, since there are no expandable regions in this ball, any spreading of affected nodes inside this ball dies out quickly, while the expandable region starts a spreading of $\theta$-affected nodes towards the origin w.h.p. We then find an upper bound $X \leq \rho$, where $\rho = \rho(\tau, N)$, that holds w.h.p., and choose $\rho' = \rho'(\tau, N)$ such that w.h.p. there is no $\theta$-expandable region inside the annulus $B_{l_*}(0, X + \rho') \setminus B_{l_*}(0, X)$. We consider the worst case $X = \rho$ and study the “race” between the possible spreads of the $\theta$-affected nodes from outside $B_{l_*}(0, \rho + \rho')$ and the spread of the expandable region at distance $\rho$ towards the origin, see Figure 5.

We consider the gradual spreading of the expandable region in time intervals of $\rho'/4$ towards the origin. Using the shape theorem, we argue that w.h.p. the origin will eventually be contained in a neighborhood such that all of its nodes are $\theta$-affected. We then show that the origin is quickly surrounded by an exponentially large firewall while any spreading of affected nodes started from outside $B_{l_*}(0, \rho + \rho')$ is still at large distances from it. This firewall is a indestructible monochromatic annulus which isolates the origin from the outside flips, see Figure 6. It will thus protect the cascading process which w.h.p. leads to the formation of a monochromatic region of size exponential in $N$ containing the origin. This shows that the lower bound occurs w.h.p. We note that in a large enough exponential size neighborhood around the origin, w.h.p. the origin will be surrounded by exponentially large monochromatic regions of particles in both states protected by firewalls. Finally, we let $A_w$ for $w = 1, 2, ...$ be the sequence of events of having the origin contained in monochromatic regions of size exponential in $N$ in an appropriate probability space. We note that these events occur w.h.p. (so their complements occur with probability $o(w^{-2})$). The proof now follows from the fact that $\sum_w w^{-2} < \infty$.

In our proofs throughout the paper, we focus on the case where $\tau < 1/2$. The results for $\tau > 1/2$ follow by a simple symmetry argument provided in Section 6.1.
3. Preliminary and Previous Results. We begin with the following elementary lemma giving lower and upper bounds for the probability of a node being affected.

**Lemma 1.** Let $p_u$ be the probability of being $\theta$-affected for an arbitrary node in the initial configuration. There exist positive constants $c_l$ and $c_u$ which depend only on $\tau$ such that

$$c_l \frac{2^{-[1-H(\tau')]}N}{\sqrt{N}} \leq p_u \leq c_u \frac{2^{-[1-H(\tau')]}N}{\sqrt{N}},$$

where $\tau' = \frac{\tau N - 2}{N-1}$, and $H$ is the binary entropy function.

**Proof.** We have

$$p_u = \frac{1}{2^{N-1}} \sum_{k=0}^{\tau N/2} \binom{N-1}{k},$$

where the two unit reduction is to account for the strict inequality and the particle at the node itself at the center of the neighborhood. Let $\tau' = \frac{\tau N - 2}{N-1}$.

After some algebra, we have

$$\binom{N-1}{\tau'(N-1)} \leq \sum_{k=0}^{\tau'(N-1)} \binom{N-1}{k} \leq \frac{1 - \tau'}{1 - 2\tau'} \binom{N-1}{\tau'(N-1)},$$

and using Stirling’s formula, there exist constants $c,c' \in \mathbb{R}^+$ such that

$$c \frac{2^{H(\tau')(N-1)}}{\sqrt{(N-1)\tau'(1-\tau')}} \leq \binom{N-1}{\tau'(N-1)} \leq c' \frac{2^{H(\tau')(N-1)}}{\sqrt{(N-1)\tau'(1-\tau')}},$$

The result follows by combining the above inequalities.

The following lemma is a consequence of Lemma 1.

**Lemma 2.** Let $\rho = 2^{[1-H(\tau')]/N/2}$. The following event occurs w.h.p.

$$A = \{ \text{\# $\theta$-affected node in } \mathcal{N}_\rho \}.$$

$m$-block. We define an $m$-block to be a neighborhood of radius $m/2$. A monochromatic block is a block whose particles are all at the same state. When $m$ is not specified, by a block we mean a $w$-block.
Region of expansion. We call a region of expansion (of type $\theta$) any neighborhood whose configuration is such that by placing a monochromatic $w$-block with all particles in state $\theta$ anywhere inside it, all the $\theta$-particles on its outside boundary (i.e., the set of particles in the set $(w+2)$-block co-centered with the $w$-block excluding the $w$-block itself) become unstable w.p.1. By a region of expansion we mean a region of expansion of only one type ($\theta$).

The next lemma, which is a restatement of Lemma 8 in [32], shows that as long as $\tau \in (\tau^*, 1/2)$, w.h.p. a monochromatic block on $G_w$ can make an exponentially large area monochromatic.

**Lemma 3 ([32]).** Let $\tau \in (\tau^*, 1/2)$ and let $N_r$ be a neighborhood with radius $r < 2^{0.5[H(\tau')]}N^{-o(N)}$ in the initial configuration. $N_r$ is a region of expansion w.h.p.

**Remark** Equation (2.1) is derived in the proof of the above lemma.

Let $\mathcal{I}_x$ be the collection of sets of particles in $w/2$-neighborhoods in an $m$-block ($m \geq w$) on $G_w$ in the initial configuration. Also, let $W_I$ be the random variable representing the number of particles in state $\theta$ in $I \in \mathcal{I}_x$, and $N_I$ be the total number of particles in $I \in \mathcal{I}_x$.

Good block. For any $\epsilon \in (0, 1/2)$, a good $m$-block of type $\theta$ is an $m$-block, such that for all $I \in \mathcal{I}_x$ we have $W_I - N_I/2 < N^{1/2+\epsilon}$. An $m$-block that does not satisfy this property is called a bad $m$-block (see Figure 7). By the following lemma (which is a restatement of Lemma 11 in [32]), an $N$-block is a good block w.h.p. Since the number of different particles in a good block is balanced enough for sufficiently large $N$, a node whose entire neighborhood is contained in a good block cannot be a $\theta$-affected node (we assume throughout that $N$ is large enough such that this is the case).
Lemma 4 ([32]). Let \( \epsilon \in (0, 1/2) \). There exists a constant \( c > 0 \), such that, for all \( I \in \mathcal{I} \) we have
\[
W_I - N_I/2 < N^{1/2+\epsilon}
\]
with probability at least
\[
1 - e^{-cN^{1/2+o(1)}}.
\]

We now want to review a result from percolation theory. Without loss of generality, we assume \( G_w \) is defined on \( \mathbb{Z}^2 \). Let us re-normalize \( G_w \) into \( m \)-blocks. Let \( S(k) \) be the ball of radius \( k \) with center at the origin, i.e., \( S(k) \) is the set of all \( m \)-blocks \( x \) on the re-normalized \( G_w \) for which \( \Delta(0, x) \leq k \), where \( \Delta \) denotes the \( l_\infty \) distance on the re-normalized \( G_w \). Let \( \partial S(k) \) denote the surface of \( S(k) \), i.e., the set of all \( x \) such that \( \Delta(0, x) = k \). We define a path of \( m \)-blocks as an ordered set of \( m \)-blocks such that each pair of consecutive \( m \)-blocks are in the Moore neighborhood (set of nearest \( l_\infty \)-neighbors) of each other and no \( m \)-block appears more than once in the set.

Let \( A_k \) be the event that there exists a path of bad \( m \)-blocks joining the origin to some vertex in \( \partial S(k) \). Let the radius of a bad cluster (i.e., cluster of bad \( m \)-blocks) be defined as
\[
\sup\{\Delta(0, x) : x \in \text{bad cluster}\}.
\]

Let \( p \) denote the probability of having a bad \( m \)-block. It is noted that an \( m \)-block is a bad \( m \)-block independently of the others. Let \( p_c \) denote the critical probability in the above percolation setting. The following result is Theorem 5.4 in [18].

Theorem 3. (Exponential tail decay of the radius of a bad cluster.) If \( p < p_c \), there exists \( \psi(p) > 0 \) such that
\[
P_p(A_k) < e^{-k\psi(p)}, \quad \text{for all } k.
\]

Firewall. A firewall of radius \( r \) and center \( u \) is a monochromatic annulus
\[
A_r(u) = \left\{ y : r - \sqrt{2}w \leq \|u - y\|_2 \leq r \right\},
\]
where \( \|\cdot\|_2 \) denotes Euclidean distance and \( r \geq 3w \).

Consider a disc of radius \( r \), centered at a particle such that all the particles inside the disc are in the same state. It is easy to see that if \( r \) is sufficiently large then all the particles inside the disc will remain stable regardless of the configuration of the particles outside the disc. Lemma 6 in [21] shows that for \( r > w^3 \) this would be the case for sufficiently large \( w \). Here we state a similar lemma but for a firewall, without proof.
Lemma 5 ([21]). Let \( A_r(u) \) be the set of particles contained in an annulus of outer radius \( r \geq w^3 \) and of width \( \sqrt{2}w \) centered at \( u \). For all \( \tau \in (\tau^*,1/2) \) and for a sufficiently large constant \( N \), if \( A_r(u) \) is monochromatic at time \( n \), then it will remain monochromatic at all times \( n' > n \).

By Lemma 5, once formed a firewall of sufficiently large radius remains static, and since its width is \( \sqrt{2}w \) the particles inside the inner circle are not going to be affected by the configurations outside the firewall.

We now review some of the definitions from [32]. In that paper, the goal is to identify a configuration that can trigger a cascading process leading to monochromatic regions w.h.p.

**Radical region.** For any \( \epsilon, \epsilon' \in (0,1/2) \) let \( \hat{\tau} = \tau[1 - 1/(\tau N^{1/2-\epsilon})] \) and define a radical region to be a neighborhood of radius \( (1 + \epsilon')w \) containing less than \( \hat{\tau}(1 + \epsilon')^2N \) particles in state \( \bar{\theta} \).

We define an unstable region to be a neighborhood of radius \( \epsilon'w \), containing at least \( \lfloor \tau \epsilon^2 N - N^{1/2+\epsilon} \rfloor \) unstable particles in state \( \theta \). The following is Lemma 4 in [32].

**Lemma 6 ([32]).** A radical region \( N_{(1+\epsilon')w} \) in the initial configuration contains an unstable region \( N_{\epsilon'w} \) at its center w.h.p.

Now consider a geometric configuration where a radical region, and neighborhoods \( N_{\epsilon'w}, N_{w/2} \) and \( N_{\rho} \) with \( \rho > 3w \), are all co-centered. Let

\[
T(\rho) = \inf\{n : \exists v \in N_{\rho}, v \text{ is a } \bar{\theta}-\text{affected node}\}.
\]

**Expandable radical region.** A radical region is called an expandable radical region of type \( \theta \) if there is a sequence of at most \( (w + 1)^2 \) possible flips inside it that can make the neighborhood \( N_{w/2} \) at its center monochromatic with particles in state \( \bar{\theta} \). The next lemma, which is a restatement of Lemma 19 in [32], shows that a radical region in this configuration is an expandable radical region of type \( \theta \) w.h.p., provided that \( \epsilon' \) is large enough and there is no \( \bar{\theta} \)-affected node in \( N_{\rho} \). The main idea is that the \( \bar{\theta} \) particles in the unstable region at the center of the radical region can trigger a process that leads to a monochromatic region of radius \( w \).

**Lemma 7 ([32]).** For all \( \epsilon' > f(\tau) \) there exists w.h.p. a sequence of at most \( (w + 1)^2 \) possible flips in \( N_{(1+\epsilon')w} \) such that if they happen before \( T(\rho) \), then all the particles inside \( N_{w/2} \) will have the same state.
Fig 8. The function $f(\tau)$ gives the infimum of $\epsilon'$ to potentially trigger a cascading process.

The function $f$ is plotted in Figure 8. The following lemma, which is Lemma 20 in [32], gives a lower bound and an upper bound for the probability that an arbitrary neighborhood with the size of a radical region is a radical region in the initial configuration.

**Lemma 8 ([32]).** Let $\epsilon'$ be positive constants. There exist positive constants $c_l$ and $c_u$ which depend only on $\tau$ such that in the initial configuration, an arbitrary neighborhood with radius $(1 + \epsilon')w$ is a radical region with probability $p_{\epsilon'}$,

$$c_l 2^{\left[1 - H(\tau'')\right] (1 + \epsilon')^2 N - o(N)} \leq p_{\epsilon'} \leq c_u 2^{\left[1 - H(\tau'')\right] (1 + \epsilon')^2 N + o(N)},$$

where $\tau'' = (\left[\hat{\tau}(1 + \epsilon')^2 N\right] - 1)/(1 + \epsilon')^2 N$, $\hat{\tau} = (1 - 1/(\tau N^{1/2 - \epsilon}))\tau$, and $H$ is the binary entropy function.

4. **The Concentration Bound.** The concentration bound developed in this section is a main step required for proving our main results and it resembles the one developed by Kesten in the context of FPP. The beauty of Kesten’s proof is that he uses a martingale representation for the difference between the spreading time and its mean that is independent of the
underlying geometrical structure. We exploit such a property in our proof. In this section, we consider $G_w$ on an infinite integer lattice.

4.1. Notational conventions. Let $t_1, t_2, \ldots$ be independent random variables, each with distribution $F$. For a vertex or vector $v = (v(1), \ldots, v(d))$ we shall use the $l_\infty$, $l_1$ and the $l_2$ norm. These are denoted by

$$
||v||_\infty = \max_{1 \leq i \leq d} |v(i)|, \quad ||v||_1 = \sum_{i=1}^{d} |v(i)| \text{ and } ||v|| = \left\{ \sum_{i=1}^{d} (v(i))^2 \right\}^{1/2},
$$

respectively. $\lceil a \rceil$ ($\lfloor a \rfloor$) is the largest (smallest) integer $\leq a$ ($\geq a$), $a \wedge b = \min(a, b)$, $a \vee b = \max(a, b)$.

4.2. Setup and definitions.

Flipping time. Let us call the waiting time of an unstable particle $i$ the flipping time and denote it by $t_i$. Although we work with exponentially distributed waiting times, the results of this section holds for any distribution that satisfies (2.2), (2.3), and (2.4).

Affected* node and block: We define an affected* node as a node such that regardless of the states of its neighbors a $\theta$-particle on it is labeled as unstable. An affected* block is a $w$-block such that all the nodes on it are affected*.

Path and first passage time. Let a path $r$ be a set of particles such that they can flip their states in a sequence $v_1, v_2, \ldots, v_k$. Let

$$
T(r) = \sum_{i=1}^{k} t_i.
$$

Also let $\mathcal{N}$ be a neighborhood containing at least one affected* block, and $u$ an arbitrary node on $G_w$. Then, we define the first passage time from $\mathcal{N}$ to $u$ as

$$
T(\mathcal{N}, u) = \inf_{r \in \mathcal{P}} \{ T(r) \},
$$

where $\mathcal{P}$ is the set of possible sequences of flips started by the affected* block in $\mathcal{N}$ that will lead to an affected node at $u$, when we assume that in the initial configuration the entire $G_w$ is a region of expansion and there are no affected nodes on $G_w$ (with these assumptions $\mathcal{P}$ is always nonempty and each particle will only flip once).
Finally, we define the first passage time or distance from 0 to \( u \) as
\[
a_{0,u} := T(0^*, u),
\]
where \( 0^* \) denotes an affected* block centered at the origin.

4.3. The Concentration Bound.

**Theorem 4.** Let \( c > 0 \) be any constant and let \( u \) be any node on \( G_w \) whose \( l_2 \)-distance from the origin is at least \( 2^cN \). There exist a constant \( c' > 0 \) (independent of \( N \)) such that for \( N \) sufficiently large, when \( G_w \) is a region of expansion, and there are no affected nodes on \( G_w \), for all \( \lambda \leq ||u|| \), we have
\[
P\left\{ |a_{0,u} - \mathbb{E}[a_{0,u}]| \geq \lambda \right\} \leq e^{-c'\lambda/\sqrt{||u|| \log ||u||}}.
\] (4.1)

In the following, we introduce two intermediate processes which we later use in the proof of Theorem 4. We decorate the variables pertaining to these two processes by prime and double prime respectively throughout this section.

4.4. First intermediate FPP process. Consider a FPP process describing the spread of affected nodes on \( \mathbb{Z}^2 \) where all the particles are of type \( \theta \) and each unstable particle will make all the nodes that have at most \( l_\infty \) distance \( w \) from them \( \theta \)-affected when it makes a flip (i.e. after its passage time). For the passage time of each node, consider i.i.d. random variables with distribution \( F \). The set of possible flips in this FPP process, is a superset for the set of possible flips which can start from an affected* block and lead to an affected node in a specific location on \( G_w \) in the original process.

Now consider a \( d \)-dimensional lattice, where \( d = \log_2 N \). Consider a FPP process on the nodes of this lattice where after its passage time, an affected node can make its \( l_1 \)-nearest neighbors \( \theta \)-affected. We refer to this FPP process as our first intermediate FPP process.

**Path and first passage time.** Let a path \( r' \) be a set of nodes in the first intermediate FPP setting such that the sequence \( v_1, v_2, ..., v_k \) of passages is possible. Let
\[
T'(r') = \sum_{i=1}^{k} t_i.
\]
Let $N$ be a neighborhood containing at least one $\theta$-affected block and $N'$ an arbitrary neighborhood. Then we define the first passage time from $N$ to $N'$ as

$$T'(N,N') = \inf_{r' \in \mathcal{P}'} \{T'(r')\}$$

where $\mathcal{P}'$ is the set of possible sequences of passages started by an $\theta$-affected block in $N$ that will lead to an affected node in $N'$.

Finally, we write $\xi_i$ for the $i$th coordinate vector and define the passage time from 0 to $n\xi_1$ as

$$a'_{0,n} := T'(0^*, n\xi_1),$$

where $0^*$ denotes an affected block centered at 0.

It is easy to argue that the event that the passage time of a path of at least $K$ nodes in this FPP process is smaller than a given value, has a higher probability than the corresponding event in the previous setting and hence it has a higher probability than the corresponding event in the setting of Theorem 4.

Now we recall a proposition from [24]. Let $p_c(\mathbb{Z}^d)$ be the critical probability of site percolation on $\mathbb{Z}^d$.

**Proposition 1 (Kesten [24]).** If

$$F(0) = P\{t(v) = 0\} < p_c(\mathbb{Z}^d),$$

then there exist constants $0 < B, C, D < \infty$, depending on $d$ and $F$ only, such that

$$P\{\exists \text{ self-avoiding path } r' \text{ from the origin which contains at least } n \text{ edges but has } T'(r') < Bn\} \leq Ce^{-Dn}.$$ 

We now slightly modify the above proposition to account for the dimension $d$ in the right hand side. This proposition is used in the proof of Theorem 4. The proof follows easily from the proof of the original statement and is omitted.

**Proposition 2 (Kesten – modified).** Let $c > 0$ be a constant, $n \geq 2^{cn}$, and $d = \log_2 N$. If (2.2) holds then for sufficiently large $N$, there exist
constants $0 < B, C, D < \infty, 1 \leq E < \infty$, depending on $F$ only, such that
\[
P\{\exists \text{ self-avoiding path } r' \text{ from the origin which contains at least } n \text{ vertices but has } T'(r') < Bn\} \leq Ce^{-Dn/(\log n)^E}.
\]

4.5. Second intermediate FPP process. We now introduce the second FPP process. For this process, we consider a lattice and put $\theta$-particles on all the nodes. We divide the lattice into $w$-blocks starting from the block centered at the origin, and we consider the passages which consists of flips of all the agents in each $w$-block that are horizontally or vertically adjacent to a $w$-block on which all the particles have already flipped to state $\bar{\theta}$. We assume that passage times are sums of flipping times of all the particles in each $w$-block which are i.i.d. with distribution $F$ (it is noted that on $G_w$ in a region of expansion a monochromatic block can always make its adjacent block monochromatic).

Path and first passage time. Let a path $r''$ be a set of particles such that they can flip their states in a sequence $v_1, v_2, ..., v_k$ that satisfies the above passage assumptions. Let
\[
T''(r'') = \sum_{i=1}^{k} t_i.
\]
Also let $\mathcal{N}$ be a neighborhood containing at least one affected* block and $\mathcal{N}'$ an arbitrary neighborhood. Then we define the first passage time from $\mathcal{N}$ to $\mathcal{N}'$ as
\[
T''(\mathcal{N}, \mathcal{N}') = \inf_{r'' \in \mathcal{P}''} \{T''(r'')\}
\]
where $\mathcal{P}''$ is the set of possible sequences of flips started by the affected* block in $\mathcal{N}$ that will lead to monochromatic blocks of $\bar{\theta}$-particles containing $\mathcal{N}'$.

Finally we write $\xi_i$ for the $i$'th coordinate vector and define the passage time from 0 to $n\xi_1$ as
\[
a''_{0,n} := T''(0^*, n\xi_1).
\]

4.6. The Bound. We work with flipping times with i.i.d. distributions $F$ that satisfy (2.2), (2.3), and (2.4).
Theorem 5 (Kesten [25]). Let \( \{ \mathcal{F}_k \}_{0 \leq k \leq V} \) be an increasing family of \( \sigma \)-fields of measurable sets and let \( \{ U_k \}_{0 \leq k \leq V} \) be a family of positive random variables that are \( \mathcal{F}_V \)-measurable. (We do not assume that \( U_k \) is \( \mathcal{F}_k \)-measurable.) Let \( \{ Z_k \}_{0 \leq k \leq V} \) be a martingale with respect to \( \{ \mathcal{F}_k \}_{0 \leq k \leq V} \). (We allow \( V = \infty \), in which case \( \mathcal{F}_V = \bigvee_{0 \leq k \leq \infty} \mathcal{F}_k \) and we merely assume that \( \{ Z_k \}_{0 \leq k < \infty} \) is a martingale.) Assume that the increments \( \Delta_k := Z_k - Z_{k-1} \) satisfy

\[
|\Delta_k| \leq c \text{ for some constant } c, \tag{4.2}
\]

and

\[
E\{ \Delta_k^2 | \mathcal{F}_{k-1} \} \leq E\{ U_k | \mathcal{F}_{k-1} \}. \tag{4.3}
\]

Assume further that for some constants \( 0 < C_1, C_2 < \infty \) and \( x_0 \) with

\[
x_0 \geq e^2 c^2, \tag{4.4}
\]

we have

\[
P\left\{ \sum_{k=1}^V U_k > x \right\} \leq C_1 e^{-C_2 x} \text{ when } x \geq x_0. \tag{4.5}
\]

Then, in the case where \( V = \infty \), \( Z_V = \lim_{k \to \infty} Z_k \) exists w.p.1. Moreover, irrespective of the value of \( V \), there exist universal constants \( 0 < C_3, C_4 < \infty \) that do not depend on \( V, C_1, C_2, c \) and \( x_0 \), nor on the distribution of \( \{ Z_k \} \) and \( \{ U_k \} \), such that

\[
P\{ Z_V - Z_0 \geq x \} \leq C_3 \left( 1 + C_1 + \frac{C_1}{C_2 x_0} \right) \times \exp \left( -\frac{C_4 x}{x_0^{1/2} + C_2^{-1/3} x_0^{1/3}} \right). \tag{4.6}
\]

In particular, for \( x \leq C_2 x_0^{3/2} \),

\[
P\{ Z_V - Z_0 \geq x \} \leq C_3 \left( 1 + C_1 + \frac{C_1}{C_2 x_0} \right) \exp \left( -\frac{C_4 x}{2 \sqrt{x_0}} \right). \tag{4.7}
\]

With a bit of abuse of notation, let

\[ a_{0,n} := T(0^*, n \xi_1). \]
We order the vertices of $G_w$ in some arbitrary way, $v_1, v_2, ...$ which remains fixed throughout. We work with $\mathbb{Z}^d$ where $d = 2$ but the results can also be generalized to any dimension. In this section, we work with any initial configuration that satisfies the assumptions of Theorem 4 (and because of that, for the proof of Theorem 4, we work with a slightly simpler probability space which does not take into account the initial configuration of the particles $\{−1,+1\}^{\mathbb{Z}^2}$, nor the possibility of having multiple flipping times for each particle). Our probability space is

$$\Omega = \mathbb{R}_+ \times \mathbb{R}_+ \times \ldots, \quad \mathbb{R}_+ = [0, \infty),$$

and a generic point of $\Omega$ is denoted by $\omega = (\omega_1, \omega_2, ...)$. In the configuration $\omega$, the flipping time of $v_i$ is

$$t(v_i) = t(v_i, \omega) = \omega_i.$$

When it is necessary to indicate the dependence on $\omega$ we write $a_{0,k}(\omega)$ instead of $a_{0,k}$, $T(r, \omega)$ instead of $T(r)$, etc. We shall use the following $\sigma$-fields of subsets of $\Omega$:

$$\mathcal{F}_0 = \text{the trivial } \sigma\text{-field } = \{\emptyset, \Omega\},$$

$$\mathcal{F}_k = \sigma\text{-field generated by } \omega_1, \ldots, \omega_k, \quad k \geq 1.$$

The martingale representation of $a_{0,n} - \mathbb{E}[a_{0,n}]$ is

$$a_{0,n} - \mathbb{E}[a_{0,n}] = \sum_{k=1}^{\infty} \left( \mathbb{E}\{a_{0,n}|\mathcal{F}_k\} - \mathbb{E}\{a_{0,n}|\mathcal{F}_{k-1}\} \right). \quad (4.8)$$

This representation is valid because $Z_0 := 0$ and

$$Z_l := \sum_{k=1}^{l} \left( \mathbb{E}\{a_{0,n}|\mathcal{F}_k\} - \mathbb{E}\{a_{0,n}|\mathcal{F}_{k-1}\} \right) \quad (4.9)$$

$$= \mathbb{E}\{a_{0,n}|\mathcal{F}_l\} - \mathbb{E}[a_{0,n}], \quad l \geq 1, \quad (4.10)$$

defines an $\{\mathcal{F}_l\}$-martingale that converges w.p.1 to $a_{0,n} - \mathbb{E}[a_{0,n}]$. The increments of $\{Z_l\}$ are denoted by

$$\Delta_k = \Delta_{k,n}(\omega) = \mathbb{E}\{a_{0,n}|\mathcal{F}_k\} - \mathbb{E}\{a_{0,n}|\mathcal{F}_{k-1}\}. \quad (4.11)$$

The principal step is to estimate

$$\mathbb{E}\{\Delta_k^2|\mathcal{F}_{k-1}\}.$$
To this end we write
\[ a_{0,n}(\omega) = f(t(v_1, \omega), t(v_2, \omega), ...) = f(\omega_1, \omega_2, ...) \]
for some Borel function \( f: \Omega \to \mathbb{R}_+ \). Also, the following notation is useful. If \( \omega = (\omega_1, \omega_2, ...) \) and \( \sigma = (\sigma_1, \sigma_2, ...) \) are points of \( \Omega \), then
\[ [\omega, \sigma]_k = (\omega_1, ..., \omega_k, \sigma_{k+1}, ...) \]
is the point that agrees with \( \omega \) and \( \sigma \) on the first \( k \) coordinates and the coordinates after \( k \), respectively. \( \nu_{k+1} \) will be the product measure
\[ \nu_{k+1} = \prod_{k+1}^\infty F_i \]
on the obvious \( \sigma \)-field in
\[ \Omega_{k+1} = R_{k+1} \times R_{k+2} \times ... \]
when each \( R_i = \mathbb{R}_+ \) and \( F_i = F \). We can think of \( \Omega \) as \( R_1 \times ... \times R_k \times \Omega_{k+1} \) and if \( G_w \) is a function from \( \Omega \to \mathbb{R} \), then if we fix \( \sigma_1, ..., \sigma_k, g(\sigma) \) can be viewed as a function of \( \sigma_{k+1}, \sigma_{k+2}, ... \); that is, as a function on \( \Omega_{k+1} \). Correspondingly,
\[ \int_{\Omega_{k+1}} \nu_{k+1}(d\sigma)g(\sigma) := \int \prod_{k+1}^\infty F_i(d\sigma_i)g(\sigma_1, ..., \sigma_k, \sigma_{k+1}, ...) \]
is the integral over all coordinates \( \sigma_i \), with \( i \geq k + 1 \), and is a function of \( \sigma_1, ..., \sigma_k \). By the independence of the \( t(v_i, \omega) = \omega_i, i \geq 1 \), we have
\begin{equation}
(4.12) \quad \mathbb{E}\{a_{0,n}|\mathcal{F}_k\}(\omega) = \int_{\Omega_{k+1}} \nu_{k+1}(d\sigma)f([\omega, \sigma]_k).
\end{equation}
This is a function of \( t(v_i, \omega) = \omega_i, 1 \leq i \leq k \), only. It also equals
\begin{equation}
(4.13) \quad \int_{\Omega_k} \nu_k(d\sigma)f([\omega, \sigma]_k),
\end{equation}
because \( [\omega, \sigma]_k \) does not involve \( \sigma_k \) and the integration over \( \sigma_k \) has no effect. Using (4.13) for \( \mathbb{E}\{a_{0,n}|\mathcal{F}_k\} \) and (4.12) with \( k \) replaced by \( (k - 1) \) for \( \mathbb{E}\{a_{0,n}|\mathcal{F}_k\} \), we find
\begin{equation}
(4.14) \quad \Delta_k = \int_{\Omega_k} \nu_k(d\sigma)\{f[\omega, \sigma]_k - f([\omega, \sigma]_{k-1})\}.
\end{equation}
Our task now is to estimate
\[ g_k(\omega, \sigma) := |f([\omega, \sigma]_k) - f([\omega, \sigma]_{k-1})|. \]

Note that
\[ t(v_i, [\omega, \sigma]_k) = t(v_i, [\omega, \sigma]_{k-1}) = \begin{cases} t(v_i, \omega), & \text{if } i \leq k - 1, \\ t(v_i, \sigma), & \text{if } i \geq k + 1. \end{cases} \]

Only for \( i = k \) do we obtain different values for the flipping time of \( v_i \) in the two configurations \([\omega, \sigma]_k\) and \([\omega, \sigma]_{k-1}\):
\[ t(v_k, [\omega, \sigma]_k) = t(v_k, \omega), \quad t(v_k, [\omega, \sigma]_{k-1}) = t(v_k, \sigma). \]

We claim that this implies
\[ g_k(\omega, \sigma) \leq |t(v_k, \omega) - t(v_k, \sigma)|. \]

Indeed, for any path \( r \),
\[ |T(r, [\omega, \sigma]_k) - T(r, [\omega, \sigma]_{k-1})| \leq \sum_{v \in r} |t(v, [\omega, \sigma]_k) - t(v, [\omega, \sigma]_{k-1})| \leq |t(v_k, \omega) - t(v_k, \sigma)|. \]

Therefore, the same estimate holds for
\[ |a_{0,n}([\omega, \sigma]_k) - a_{0,n}([\omega, \sigma]_{k-1})| = |\inf_r T(r, [\omega, \sigma]_k) - \inf_r T(r, [\omega, \sigma]_{k-1})|. \]

This proves (4.16). Let \( \pi_n(\omega) \) be the optimal path from 0 to \( n\xi_1 \) in the configuration \( \omega \); that is, \( \pi_n(\omega) \) is a path from 0 to \( n\xi_1 \) with
\[ a_{0,n}(\omega) = T(\pi_n(\omega), \omega). \]

Due to the assumptions of the theorem, such a path always exists. There could, however, be several paths with this property. To define \( \pi_n(\omega) \) uniquely in case of ties, we order all paths from 0 to \( n\xi_1 \) in some arbitrary way, and take for \( \pi_n(\omega) \) the first path in this ordering that satisfies (4.18). We write \( v \in \pi \) to denote that \( v \) is a vertex in the path \( \pi \). Then, if
\[ v_k \notin \pi_n([\omega, \sigma]_k), \]
(4.15) and (4.17) show that
\[ T(\pi_n([\omega, \sigma]_k), [\omega, \sigma]_k) = T(\pi_n([\omega, \sigma]_k), [\omega, \sigma]_{k-1}). \]
Thus, under (4.19),

\[ a_{0,n}([\omega, \sigma]_{k-1}) = \inf \{ T(r, [\omega, \sigma]_{k-1}) : r \text{ a path from 0 to } n \xi_1 \} \]

\[ \leq T(\pi_n([\omega, \sigma]_k, [\omega, \sigma]_{k-1})) \]

\[ = T(\pi_n([\omega, \sigma]_k, [\omega, \sigma]_k)) \]

\[ = a_{0,n}([\omega, \sigma]_k). \]

Similarly, if

(4.20) \quad v_k \notin \pi_n([\omega, \sigma]_{k-1}),

then

\[ a_{0,n}([\omega, \sigma]_k) \leq a_{0,n}([\omega, \sigma]_{k-1}). \]

It follows that \( g_k(\omega, \sigma) = 0 \) if (4.19) and (4.20) both hold, and by virtue of (4.16),

(4.21) \quad g_k(\omega, \sigma) \leq |t(v_k, \omega) - t(v_k, \sigma)| I[v_k \in \pi_n([\omega, \sigma]_{k-1}) \text{ or } v_k \in \pi_n([\omega, \sigma]_k)].

where \( I_k(.) \) is the indicator function. This is then an estimate for \( g_k \). Let \( I_k(\omega, \sigma) \) denote the indicator function in the above inequality. Now using Schwarz’s inequality and (4.14) we have

(4.22) \quad \mathbb{E} \left[ \Delta_k^2 \left| \mathcal{F}_{k-1} \right. \right] \leq \mathbb{E} \left[ \left( \int_{\Omega_k} \nu_k(d\sigma) g_k(\omega, \sigma) \right)^2 \left| \mathcal{F}_{k-1} \right. \right]

\[ \leq \mathbb{E} \left[ \left( \int_{\Omega_k} \nu_k(d\sigma) |t(v_k, \omega) - t(v_k, \sigma)| I(\omega, \sigma) \right)^2 \left| \mathcal{F}_{k-1} \right. \right]

\[ \leq \mathbb{E} \left[ \int_{\Omega_k} \nu_k(d\sigma) |t(v_k, \omega) - t(v_k, \sigma)|^2 I(\omega, \sigma) \right. \]

\[ \times \left. \int_{\Omega_k} \nu_k(d\sigma) I(\omega, \sigma) \left| \mathcal{F}_{k-1} \right. \right]

\[ \leq \mathbb{E} \left[ \int_{\Omega_k} \nu_k(d\sigma) |t(v_k, \omega) - t(v_k, \sigma)|^2 I(\omega, \sigma) \left| \mathcal{F}_{k-1} \right. \right]. \]
Now
\[
\mathbb{E} \left[ \int_{\Omega_k} \nu_k(d\sigma) |t(v_k, \omega) - t(v_k, \sigma)|^2 I(\omega, \sigma) \mid \mathcal{F}_{k-1} \right],
\]
is a function of $\omega_1, \ldots, \omega_k$ only; the $\sigma$-variables all have been integrated out. Similar to (4.12) we have
\[
\mathbb{E} \left[ \int_{\Omega_k} \nu_k(d\sigma) |t(v_k, \omega) - t(v_k, \sigma)|^2 I(\omega, \sigma) \mid \mathcal{F}_{k-1} \right] = \int F(d\omega) \int_{\Omega_k} \nu_k(d\sigma) |t(v_k, \omega) - t(v_k, \sigma)|^2 I_k(\omega, \sigma).\]
where we have used the fact that $\nu_k$ can be written as the product measure $F \times \nu_{k+1}$ on $R_k \times \Omega_{k+1} = \mathbb{R}^+ \times \Omega_{k+1}$. Let us write
\[
J_k(\omega) = I[v_k \in \pi_n(\omega)].
\]
Then
\[
I_k(\omega, \sigma) = J_k([\omega, \sigma]_{k-1}) \vee J_k([\omega, \sigma]_k) = J_k(\omega_1, \ldots, \omega_{k-1}, \sigma_k, \sigma_{k+1}, \ldots) \vee J_k(\omega_1, \ldots, \omega_{k-1}, \omega_k, \sigma_{k+1}, \ldots).
\]
Recall that $t(v_k, \omega) = \omega_k$, $t(v_k, \sigma) = \sigma_k$, so that
\[
|t(v_k, \omega) - t(v_k, \sigma)|^2 I_k(\omega, \sigma) \leq |\omega_k - \sigma_k|^2 \left\{ J_k(\omega_1, \ldots, \omega_{k-1}, \sigma_k, \sigma_{k+1}, \ldots) \right. \\
\left. \vee J_k(\omega_1, \ldots, \omega_{k-1}, \omega_k, \sigma_{k+1}, \ldots) \right\}.
\]
Clearly the right-hand side is symmetric in $\omega_k$ and $\sigma_k$ for fixed
\[
\omega_1, \ldots, \omega_{k-1}, \sigma_{k+1}, \sigma_{k+2}, \ldots,
\]
in (fact, this is true also for the left-hand-side). It is also clear that on \{ $\sigma_k \leq \omega_k$ \} or on \{ $t(v_k, \sigma) \leq t(v_k, \omega)$ \} we have
\[
|t(v_k, \omega) - t(v_k, \sigma)| \leq t(v_k, \omega),
\]
(4.25)
(4.26) \[ J_k(\omega_1, ..., \omega_{k-1}, \sigma_k, \sigma_{k+1}, ...) \lor J_k(\omega_1, ..., \omega_{k-1}, \omega_k, \sigma_{k+1}, ...) = J_k([\omega, \sigma]_{k-1}). \]

(4.26) simply says that if \( v_k \) belongs to the optimal path in configuration \([\omega, \sigma]_{k-1}\) or in configuration \([\omega, \sigma]_k\), then it will belong to the optimal path in the configuration that gives the lower value to \( t(v_k) \). Substituting (4.23)-(4.26) into (4.22) we find

(4.27) \[ \mathbb{E}[\Delta^2_k|\mathcal{F}_{k-1}] \]
\[ \leq 2 \int_{\Omega_k+1} \nu_{k+1}(d\sigma) \int_{\sigma_k \leq \omega_k} F(d\omega_k) F(d\sigma_k) t^2(v_k, \omega) J_k([\omega, \sigma]_{k-1}) \]
\[ \leq 2 \int F(d\omega_k) t^2(v_k, \omega) \int F(d\sigma_k) \int_{\Omega_k+1} \nu_{k+1}(d\sigma) J_k([\omega, \sigma]_{k-1}) \]
\[ = 2 \int x^2 dF(x) P\{v_k \in \pi_n(\omega)|\mathcal{F}_{k-1}\}. \]

Let \(|\pi|\) denote the number of vertices in \( \pi \). For any \( a > 0, y > 0, \)

(4.29) \[ P\{|\pi_n| \geq yn\} \leq P\{a_0, n \geq ayn\} \]
\[ + P\{\exists \text{ self-avoiding path } r \text{ starting at } 0 \]
\[ \text{of at least } yn \text{ steps but with } T(r) < ayn\}. \]

Now we want to use our intermediate FPP processes to bound the probabilities on the right hand side of (4.29). Let us decorate the variables corresponding to our first and second intermediate FPP processes by a prime and a double prime respectively. We are going to find upper bounds for the right hand side terms of (4.29) using these intermediate processes.

(4.30) \[ P\{a_{0,n} \geq ayn\} \leq P\{a_{0,n}'' \geq ayn\} \leq P\{T''(r_{n}'') \geq ayn\} \leq P \left\{ \sum_{i=1}^{w_n} t_i \geq ayn \right\}, \]

where \( w \) adjusts the number of particles for the renormalized graph (our second intermediate process). We also have

(4.31) \[ P\{\exists \text{ self-avoiding path } r \text{ starting at } 0 \text{ of at least } yn \text{ steps} \]
\[ \text{but with } T(r) < ayn\} \]
\[ \leq P\{\exists \text{ self-avoiding path } r' \text{ starting at } 0 \text{ of at least } yn \text{ steps} \]
\[ \text{but with } T'(r') < ayn\}. \]
For a suitable \( a \), using Proposition 2, and for sufficiently large \( N \), we have
\[
P\{ \exists \text{ self-avoiding path } r' \text{ starting at } 0 \text{ of at least } yn \text{ steps} \}
\leq C_2 \exp(-C_3 yn / (\log n)^{C_4}).
\]

Hence, we have
\[
(4.32) \quad P\{|\pi_n| \geq yn\} \leq P \left\{ \sum_{i=1}^{\ln n} t_i \geq ayn \right\} + C_2 \exp(-C_3 yn / (\log n)^{C_4}).
\]

**Proof of Theorem 4.** We show that
\[
(4.33) \quad P\{|a_{0,n} - E[a_{0,n}]| \geq x\} \leq C_3 e^{-C_4 x / \sqrt{n \log n}} \quad \text{for } x \leq n,
\]
when \( n \geq 2^cN \) for some constant \( c > 0 \) and \( N \) sufficiently large. Then, due to the fact that our proof does not rely on any geometrical property of the model nor the initial configuration other than the assumptions of the theorem, we conclude that (4.1) follows.

We want to use Theorem 5, for which we need a truncation argument. Let \( n \) be fixed. Define
\[
(4.34) \quad \hat{t}(v_i) = t(v_i) \wedge \frac{4d}{\gamma} \log n,
\]
with \( \gamma \) as in (2.4). Passage times and related quantities, when defined in terms of the \( \hat{t} \) instead of the \( t \), will be denoted by the old symbols decorated with a caret. For example, if \( r = (v_1, ..., v_k) \), then
\[
\hat{T}(r) = \sum_{i=1}^{k} \hat{t}(v_i);
\]
\[
\hat{a}_{0,n} = \inf \{ \hat{T}(r) : r \text{ a path from } 0 \text{ to } n\xi \},
\]
\[
\hat{\pi}_n = \text{optimal path for } \hat{a}_{0,n}.
\]

**Lemma 9.** If (2.2) and (2.4) hold, then there exist constants \( 0 < C_i < \infty \) such that
\[
(4.35) \quad P\{\hat{\pi}_n \not\subseteq [-n(\log n)^{C_1}, n(\log n)^{C_1}]^d\} \leq 3e^{-C_2 n},
\]
\[
(4.36) \quad P\{|a_{0,n} - \hat{a}_{0,n}| \geq x\} \leq 3e^{-C_2 n} + C_3 e^{-(\gamma/2)x}, \quad x \geq 0,
\]
and
\[
(4.37) \quad |E[a_{0,n}] - E[\hat{a}_{0,n}]| \leq C_4.
\]
Proof. The probability in the left-hand side of (4.35) is bounded by

\[
P\{|\hat{\pi}_n| \geq n(\log n)^{C_1}\} \leq P\{\hat{a}_{0,n} \geq aC_1n^{\log n}\}
\]

\[+ P\{\exists \text{ self-avoiding path } r \text{ starting at } 0 \text{ of at least } C_1n^{\log n} \text{ steps but with } \hat{T}(r) < aC_1n^{\log n}\}
\]

similar to (4.32). Note that the last inequality is true for some constants \( C_5, C_6 \) that are independent of \( n \), although the distribution of \( \hat{t} \) does depend on \( n \). This is so because for any constant \( C \), \( \hat{t}(v) \geq \hat{t}(v) \wedge C \) for large enough \( n \). Thus also \( \hat{T}(t) \geq \sum_{v \in r} \{t(v) \wedge C\} \), and it suffices to apply Proposition 2 and choose \( C_1 \) to be equal to \( C_4 \).

Now we use (2.4) for the following standard large deviation estimate:

\[
P\left\{ \sum_{i=1}^{\binom{n}{2}} t_i \geq an(\log n)^{C_1}\right\} \leq e^{-\gamma an(\log n)^{C_1}}\left(E[e^{\gamma t_1}]\right)^{\binom{n}{2}}.
\]

Now, using the fact that \( n \geq 2^{cN} \), for sufficiently large \( N \), the above expression is bounded above by \( \exp(-n) \). Now, the inequality in (4.35) follows.

To prove (4.36) and (4.37) we note that

\[
0 \leq a_{0,n} - \hat{a}_{0,n} \leq T(\hat{\pi}_n) - \hat{T}(\hat{\pi}_n)
\]

\[= \sum_{v \in \hat{\pi}_n} \{t(v) - \hat{t}(v)\} \leq \sum_{v \in \hat{\pi}_n} t(v)1\{t(v) > \frac{4d}{\gamma} \log n\}.
\]

If \( \hat{\pi}_n \subseteq [-n(\log n)^{C_1}, n(\log n)^{C_1}]^d \), then the last member of (4.40) is at most

\[
\sum_{v \in [-n(\log n)^{C_1}, n(\log n)^{C_1}]^d} t(v)1\{t(v) > \frac{4d}{\gamma} \log n\}.
\]

Hence,

\[
P\{|a_{0,n} - \hat{a}_{0,n}| \geq x\} \leq P\{\hat{\pi}_n \not\subseteq [-n(\log n)^{C_1}, n(\log n)^{C_1}]^d\}
\]

\[+ P\left\{ \sum_{i=1}^{M'} t_i1\{t_i > \frac{4d}{\gamma} \log n\} \geq x\right\},
\]

(4.41)
where

\( M' = \text{number of nodes in } [-C_1 n \log n, C_1 n \log n]^d \sim d(2C_1 n \log n)^d. \)

Therefore

\[
P \left\{ \sum_{i=1}^{M'} t_i 1_{\{ t_i > \frac{2d}{\gamma} \log n \}} \geq x \right\} \\
\leq e^{-\left(\frac{\gamma}{2}\right)x} \left[ 1 + \int_{y \geq 4(\frac{d}{\gamma}) \log n} \left( e^{\left(\frac{\gamma}{2}\right)y} - 1 \right) F(dy) \right]^{M'} \\
\leq e^{-\left(\frac{\gamma}{2}\right)x} \left[ 1 + e^{-2d \log n} \int e^{\gamma y} F(dy) \right]^{M'} \\
\leq \exp \left\{ -\left(\frac{\gamma}{2}\right)x + M'n^{-2d} \int e^{\gamma y} F(dy) \right\} \\
\leq C_3 e^{-\left(\frac{\gamma}{2}\right)x}. \tag{4.43} \]

(4.36) follows from (4.41), (4.35) and (4.43). (4.37) follows from (4.36) plus the additional estimates

\[
0 \leq \hat{a}_{0,n} \leq a_{0,n} \]

and, there exists \( y_0 \) such that for \( y \geq y_0, \)

\[
P\{a_{0,n} - \hat{a}_{0,n} \geq yn \log n\} \leq P\{a_{0,n} \geq yn \log n\} \leq P\{a''_{0,n} \geq yn \log n\} \\
\leq e^{-y \log n} \int e^{\gamma x} F(dx) \leq e^{-\left(\frac{\gamma}{2}\right)y \log n}. \tag{4.44} \]

Now we can prove (4.1). By Lemma 9 we have for \( x\sqrt{n} \geq 2C_4, \)

\[
P\{a_{0,n} - \mathbb{E}[a_{0,n}] \geq x\sqrt{n}\} \leq P\{a_{0,n} - \hat{a}_{0,n} \geq \frac{x}{4} \sqrt{n}\} + P\{\hat{a}_{0,n} - \mathbb{E}[\hat{a}_{0,n}] \geq \frac{x}{4} \sqrt{n}\} \leq 3 \exp(-C_2 n) + C_3 \exp\left(-\frac{\gamma}{8}x\sqrt{n}\right) + P\{\hat{a}_{0,n} - \mathbb{E}[\hat{a}_{0,n}] \geq \frac{x}{4} \sqrt{n}\}. \]
The first term in the last member is at most \(3 \exp(-C_2x)\) for \(x \leq n\). Hence to prove (4.1), all we need is to prove

\[
(4.45) \quad P\left\{ |\hat{a}_{0,n} - \mathbb{E}[\hat{a}_{0,n}]| \geq \frac{x}{4} \sqrt{n} \right\} \leq C_7 \exp(-C_8x/\sqrt{\log n}) \quad \text{for } x \leq n.
\]

The remaining part of the proof deduces (4.45) from Theorem 5. As in (4.8)-(4.11),

\[
\hat{a}_{0,n} - \mathbb{E}[\hat{a}_{0,n}] = \sum_{k=1}^{\infty} \hat{\Delta}_k
\]

with

\[
\hat{\Delta}_k = \mathbb{E}[\hat{a}_{0,n} | \mathcal{F}_k] - \mathbb{E}[\hat{a}_{0,n} | \mathcal{F}_{k-1}].
\]

Moreover,

\[
Z_0 = 0, \quad Z_l = \sum_{k=1}^{l} \hat{\Delta}_k, \quad l \geq 1,
\]

defines a martingale. We shall now verify the hypotheses of Theorem 5 for the martingale. Note that replacing \(t(v_i), a_{0,n}\) and \(\Delta_k\) by \(\hat{t}(v_i), \hat{a}_{0,n}\) and \(\hat{\Delta}_k\) merely amounts to changing the distribution \(F\) to

\[
\hat{F}(x) = F(x) \vee 1_{\{x \geq \frac{4d\gamma}{\log n}\}}.
\]

Therefore, by (4.16), (4.14) and the definition (4.34) of \(\hat{t}\),

\[
|\hat{\Delta}_k| \leq 2 \max(\text{supp } \hat{F}) \leq \frac{8d}{\gamma} \log n.
\]

This is (4.2) for

\[
c = \frac{8d}{\gamma} \log n.
\]

Furthermore, by (4.28),

\[
\mathbb{E}[\hat{\Delta}^2 | \mathcal{F}_{k-1}] \leq 2 \int x^2 \hat{F}(dx) P\{v_k \in \hat{\pi}_n(\omega) | \mathcal{F}_{k-1}\}
\leq 2 \int x^2 F(dx) P\{v_k \in \hat{\pi}_n(\omega) | \mathcal{F}_{k-1}\}.
\]
Thus (4.3) holds with
\[ U_k = D \hat{J}_k, \]

where
\[ D = 2 \int x^2 F(dx), \quad \hat{J}_k(w) = 1_{\{v_k \in \hat{\pi}_n(\omega)\}}. \]

We next take \( C > 0 \) as in Proposition 2 and
\[(4.46) \quad x_0 = n \log n \frac{2D}{\gamma C} \log \left\{ \int e^{\gamma u} F(du) \right\}.\]

Clearly this satisfies (4.4) for large \( n \). Finally, we verify (4.5). We have
\[(4.47) \quad \sum_{k=1}^{\infty} U_k = D \sum_{k=1}^{\infty} 1_{\{v_k \in \hat{\pi}(\omega)\}} = D |\hat{\pi}_n(\omega)| = D \times \text{length of } \hat{\pi}_n(\omega).\]

Moreover, as in (4.38) and (4.39),
\[(4.48) \quad P\{|\hat{\pi}_n(\omega)| \geq y\} \leq P\{\hat{a}_{0,n} \geq Cy\}
+ P\{\exists \text{ self-avoiding path } r \text{ starting at 0 of at least } y \text{ steps but with } \hat{T}(r) < Cy\}
\leq e^{-\gamma Cy} \left[ \int e^{\gamma u} F(du) \right]^{\frac{\text{length of } \hat{\pi}_n(\omega)}{n}} + C_9 \exp(-C_{10}y/(\log n)^{C_{11}}).\]

For \( y \geq x_0/D \) and \( x_0 \) as in (4.46), the right-hand side of (4.48) is at most
\[ e^{-\gamma Cy} + C_9 e^{-C_{10}y/(\log n)^{C_{11}}}. \]

Therefore, (4.5) holds with \( C_1 = (1 + C_9) \) and \( C_2 = \frac{\gamma C}{2D} \land \frac{C_{10}}{D(\log n)^{C_{11}}}. \)

Thus, by (4.7) (applied to \( \hat{a}_{0,n} \) and to \( -\hat{a}_{0,n} \)), for \( n \geq 2^{cN}, x \leq n \) when \( N \) is sufficiently large, we have
\[ P\{|\hat{a}_{0,n} - E\hat{a}_{0,n}| \geq x\} \leq 2C_3(1 + C_1)(1 + \frac{C_1}{C_2 x_0}) \exp(-\frac{C_4 x}{2 \sqrt{x_0}})
\leq \exp\left(-C_{12} \frac{x}{\sqrt{n \log n}}\right),\]

which proves (4.1) for our process on \( G_w. \)
5. Proof of the Shape Theorem. In this section, we consider our graph $G_w$ on an infinite integer lattice. We work with the obvious probability space defined by our process on the lattice (i.e., product space for the initial configuration and the waiting times). We introduce the following definition from Tessera [38] with some modifications.

Strong Asymptotic Geodesicity (SAG) for $E[a]$: Let $Q : R_+ \to R_+$ be an increasing function such that
\[
\lim_{\alpha \to \infty} Q(\alpha) = \infty.
\]
Let $N$ be a neighborhood and let $c'_1 \in (0, c_1)$ be a constant. $E[a]$ is called SAG($Q$) in $N$ when for all integer $m \geq 1$, for all $x, y \in N$ such that $E[a_{x,y}]/m \geq 2c'_1N$, there exists a sequence $x = x_0, ..., x_m = y$ in $N$ satisfying,
\[
\alpha \left( 1 - \frac{1}{Q(\alpha)} \right) \leq E[a_{x_i,x_{i+1}}] \leq \alpha \left( 1 + \frac{1}{Q(\alpha)} \right),
\]
where $\alpha = E[a_{x,y}]/m$; and for $r \geq 2c'_1N$,
\[
\tilde{A} \left( 0, \left( 1 + \frac{1}{Q(r)} \right) r \right) \subset \tilde{A}(0, r) \frac{\alpha}{Q(r)},
\]
where $\tilde{A}(0, r) = \{ x \mid E[a_{0,x}] \leq r \}$ and $\tilde{A}_t$ denotes the $t$-neighborhood of the subset $\tilde{A}$ with respect to $E[a]$.

Proposition 3. Let $2c'_1N \leq n \leq 2c_2N$ and
\[
A'_F(0, n) = \{ x \mid a_{0,x} \leq n \}.
\]
For sufficiently large $N$, there exists a norm $l_*$ on $R^2$, and $c > 0$, such that almost surely,
\[
B_{l_*}(0, n - N^c n^{1/2} \log^{3/2} n) \cap Z^2 \subset A'_F(0, n)
\subset B_{l_*}(0, n + N^c n^{1/2} \log^{3/2} n).
\]

The proof of this proposition, as outlined in Section 2, consists of two main parts. We now start by preparing the stage for the first part of the proof.

Let $d_1$ and $d_2$ be the metrics defined by the expected value of first passage times in the first and second intermediate FPPs when we choose the nodes
to be on $G_w$ defined on a two dimensional integer lattice. Let $x, y$ denote two arbitrary nodes on $G_w$. Assuming that $G_w$ is a region of expansion and there are no affected nodes on $G_w$, it is easy to see that there exists $c, c' > 0$ such that for sufficiently large $N$ we can write

\begin{equation}
\frac{d_{12}(x, y)}{N^c} \leq \bar{d}_1(0, \xi_1 || x - y ||_1 / N^{c'}) \leq \mathbb{E}[a_{x,y}] \leq \bar{d}_2(x, y) \leq N^c d_{12}(x, y).
\end{equation}

(5.4)

**Lemma 10.** Let $c_2 < c'_2 < c''_2 \in (c_2, 0.5(1 - H(\tau)))$ be constants and let $N_\rho$ and $N_{\rho'}$ be two neighborhoods with radii $\rho = 2^c_2 N$ and $\rho' = 2^c_3 N$. Let

$$A = \{ \# \theta - \text{affected node in } N_{\rho'} \text{ in the initial configuration} \}$$

There exists a constant $c \in \mathbb{R}^+$, such that conditional on event $A$ and for sufficiently large $N$,

$$P(T(\rho) > \rho) > 1 - e^{-\sqrt{\rho}/N^c},$$

where $T(\rho)$ is defined in (3.2).

**Proof.** Let $x, y$ be two nodes on $N_\rho$ and $N_{\rho'}$ respectively such that they have the largest $l_\infty$ distance from the origin (i.e., they are on the boundaries of these neighborhoods). Let $a'$ denote the first passage time between two nodes located on $\mathbb{Z}^2$, using (5.4) there exists a constant $c' > 0$ such that for sufficiently large $N$ we can write

$$P(a_{x,y} \leq \rho) \leq P\left(a'_0, \xi_1(\rho - \rho')/N^{c'} \leq \rho \right).$$

Now due to (5.4) there exists a constant $c'' > 0$ such that we can write

$$P(a_{x,y} \leq \rho) \leq P\left(a'_0, \xi_1(\rho - \rho')/N^{c'} - \mathbb{E}\left[a'_0, \xi_1(\rho - \rho')/N^{c'} \right] \leq \rho - \frac{\rho' - \rho}{N^{c''}} \right).$$

Now using the fact that there are less than $64\rho'^2$ pairs of nodes on the boundaries of $N_\rho$ and $N_{\rho'}$ we can use the Talagrand concentration bound [37] to conclude that there exists a constant $c > 0$ such that for sufficiently large $N$ we have

$$P(T(\rho) \leq \rho) \leq \exp(-\sqrt{\rho}/N^c).$$

\[\blacksquare\]
Proposition 4 (Tessera – modified). Let $2^{c_1N} \leq r_w \leq 2^{c_2N}$ be a sequence and $N_{r_w}$ be a sequence of neighborhoods with radii $r_w$ centered at the origin. There exists $C > 0$ and $w_0 \in \mathbb{N}$ such that for $w \geq w_0$ (hence $N \geq N_0$), almost surely,

$$\sup_{x,y \in N_{r_w}} |a_{x,y} - \mathbb{E}[a_{x,y}]| \leq Cr_w^{1/2} \log^{3/2} r_w. \quad (5.5)$$

Proof. Let $c \in [c_1, c_2]$, $c_2 < c'_2 < c''_2 < 0.5(1 - H(\tau)))$ be constants and let $N'_\rho$ be a neighborhood with radius $\rho = 2^{c''_2N}$ centered at the origin. Let $A_1$ denote the event of not having any affected nodes in this neighborhood. Using Lemma 2, $A_1$ occurs w.h.p. Let $A_2$ denote the event that this neighborhood is a region of expansion. Using Lemma 3 this event occurs w.h.p. Let $A_3$ denote the event that there are no affected nodes in this neighborhood before the time $2^{c'_2N}$. Using Lemma 10, this event also occurs w.h.p.

We now need the following lemma whose proof follows the same lines as of Lemma 9 and is omitted.

Lemma 11. Let $N_{r_w}$ be a neighborhood with radius $r_w$ centered at the origin. If (2.2) and (2.4) hold, then conditional on $A_1, A_2$ and $A_3$, there exist constants $0 < C_1, C_2 < \infty$ such that for all $x,y \in N_{r_w}$

$$P\{\pi_{x,y} \not\subseteq N_{r_w(\log r_w)} c_{1}\} \leq e^{-C_2 r_w}, \quad (5.6)$$

where $\pi_{x,y}$ denotes the optimal path between $x, y$ and $N_{r_w(\log r_w)} c_{1}$ denotes a neighborhood with radius $r_w(\log r_w) C_1$ centered at the origin.

Let $A_4$ denote the event that the optimal path of every pair of nodes in $N_{r_w}$ are contained in $N_{r_w(\log r_w)} c_{1}$. Using Lemma 11 this event occurs w.h.p.

Let $A_5$ be the event that for every pair of nodes in $N_{r_w}$, the first passage time is at most $2^{c'_2N}$. To see this event occurs w.h.p., let $x, y \in N_{r_w}$, we first note that we can write

$$P(a_{x,y}'' > 2^{c'_2N}) \leq P(a_{x,y}' > 2^{c'_2N}),$$

where $a''$ is the first passage time of the second intermediate FPP. Now it is easy to see that standard concentration bounds imply that there exists a constant $C''$ such that the last term in the above inequality is at most $\exp(-C'' 2^{c'_2N})$. It follows that event $A_5$ occurs w.h.p.

Let $C_3$ be a constant to be determined later. Let $w$ be sufficiently large so that $P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5)$ is greater than $1/2$. Let $E = c/4$ and let $w$ be large enough such that for $||x - y|| \geq 2^{EN}$, (4.1) holds. Using (4.1) for these
nodes and standard concentration bounds for nodes such that \(|x - y| < 2^EN|

it is easy to see that for all \(x, y\), there exists a constant \(C_4\), such that for sufficiently large \(N\),

\[
P\left( \left| a_{x,y} - \mathbb{E}[a_{x,y}] \right|^2 \geq C_3 \log^3 r \right) \leq 2 \exp(-c'C_3 \log r).
\]

where \(c'\) is the constant in (4.1). Now letting \(C_3 = 6/c'\), we have that for large enough \(w\), all \(x, y\) such that \(|x - y| < r w\),

\[
P\left( \left| a_{x,y} - \mathbb{E}[a_{x,y}] \right|^2 \geq C_3 r w \log^3 r w \right) \leq 2^{-6 w}.
\]

Hence for \(w\) large enough we have

\[
P\left( \sup_{x,y \in B(o_w, r w)} \left| a_{x,y} - \mathbb{E}[a_{x,y}] \right|^2 \geq C_3 r w \log^3 r w \right) \leq 2^{-6 \sum_{w=1}^{\infty} w^{-2}} < \infty.
\]

Now let \(2^{c_1 N} \leq n \leq 2^{c_2 N}\). It follows from Proposition 4 that there exists \(C'\), such that for sufficiently large \(N\), we have

\[
\bar{A}(0, n - C'n^{1/2} \log^{3/2} n) \subset A'_F(0, n) \subset \bar{A}(0, n + C'n^{1/2} \log^{3/2} n).
\]

We now focus on the second part of the proof of Proposition 3.

Theorem 6. Let \(2^{c_1 N} \leq n \leq 2^{c_2 N}\). There exits \(c > 0\) and \(N_0\) such that, for every \(N > N_0\), there exists a norm \(l^*_w\) such that, almost surely,

\[
B_{l^*_w}(0, n - N^c n^{1/2} \log^{3/2} n) \cap \mathbb{Z}^2 \subset \bar{A}(0, n) \subset B_{l^*_w}(0, n + N^c n^{1/2} \log^{3/2} n).
\]

The proof of the above theorem follows from the following results. We first prove the following lemma.

Lemma 12. Consider a good 10\(w\)-block denoted by \(N\) in the initial configuration. Now assume that only \(\theta\)-particles can flip their states and at time \(n\) there is a \(\theta\)-affected node \(u\) at the center of \(N\). Then, at time \(n\), w.h.p. there exists a sequence of flips that if they happen while the particles outside \(N\) maintain their configuration, there will be a \(\theta\)-affected \(w\)-block inside \(N\).
Proof. Let $N_{a,b}$ denote a rectangle of sides $a$ and $b$. Let $I$ be the collection of sets of particles in the possible intersections of rectangles $N_{0.5, w^{1/4}}$ and $N_{w^{1/4}, 0.5}$ with the neighborhood $N$ on $G_w$ in the initial configuration. Also, let $W_I$ be the random variable representing the number of particles in state $\bar{\theta}$ in $I \in I$, and $N_I$ be the total number of particles in $I \in I$. For any $\epsilon \in (0, 1/2)$, and for all $I \in I$ we have $W_I - N_I/2 < w^{1/8+\epsilon}$ w.h.p. (with a similar argument as in Lemma 4). Let us denote this property as the balance property.

Now assume that the balance property is true for $N$. Consider a $w$-block (henceforth referred to as a block) inside $N$ such that the affected node is right in the lower-left corner of it. We then re-normalize $G_w$ into blocks starting from this block. Now since node $u$ is an affected node, the balance property implies that there has been $N' = (1/2 - \tau)N + o(N)$ flips in the four blocks around $u$ that are part of its neighborhood. This in turn implies that there has been $N'$ unstable particles in these two blocks that have flipped so that node $u$ has become affected. Now, the worst case is when these unstable particles that have flipped are divided equally in these blocks. This means that each block has had $0.25N' + o(N')$ unstable particles. If these flips had happened close to the node $u$, this makes our argument only easier since the previous flips making these flips possible have also been close to $u$. Hence, the worst case scenario is when the flips in the four blocks have been in the farthest corner of each of them from $u$.

Now consider the first unstable particle in one of these four blocks. For this particle to become unstable, since it has been the first unstable particle in the block, this implies that the flips leading to this unstable particle have been inside the blocks consisting the rest of its neighborhood. This means that in the worst case of considering a corner block there has been $(1/3)N' + o(N')$ flips in each of these blocks. Now, as a result of the balance property, and the fact that there are going to be at least $(1/3)N' + o(N')$ unstable particles around the first unstable particle that we considered, one can check that the flips of all these unstable particles will lead to $O(1/3N')$ additional unstable blocks whose flips are enough to lead to the formation of a block such that all of its nodes are affected w.h.p., and this completes the proof. ■

Proposition 5. There exists a constant $c > 0$ such that for sufficiently large $N$, $E[a]$ is $SAG(Q)$ in $N_\rho$ where $\rho = 2c^2N$ and

$$Q(\alpha) = \frac{\alpha^{1/2}}{N^c \log^{3/2} \alpha}.$$ (5.9)
Proof. Let $r_w = 2^{c_1^N}$ be a sequence where $c_1^N$ is the constant in the definition of SAG, and $c_2^N > c_2 > c_2$ be constants. Let $\mathcal{N}_\rho$ be a sequence of neighborhoods with radius $\rho = 2^{c_2^N}$. Let events $A_1, \ldots, A_5$ be as defined in the proof of Proposition 4. Let $A_6$ be the event that the first flipping time of all the particles inside $\mathcal{N}_\rho$ are less than $\sqrt{r_w}$. Let $t_x$ denote the first flipping time of particle at node $x \in \mathcal{N}_\rho$. Standard concentration bounds imply that for sufficiently large $N$ there exists a constant $c > 0$ such that

$$P(t_x > \sqrt{r_w}) \leq \exp(-c\sqrt{r_w}).$$

It follows that event $A_6$ occurs w.h.p.

Let us define a *very good block* as a 10$w$-block that satisfies the balance property and a *very bad block* as a 10$w$-block that does not satisfy this property. Let us divide the lattice into 10$w$-blocks starting from the block centered at the origin. Since each of these blocks is a very good block w.h.p., it follows from Theorem 3 that in $\mathcal{N}_\rho$ there are no clusters of very bad blocks with radius larger than $N^2$ in the initial configuration w.h.p. Now using Lemma 12 and the fact that the event described in this lemma and the event of having a $\theta$-affected node are both increasing in the change of a $\theta$-particle to a $\bar{\theta}$-particle, using the FKG inequality [17] we can conclude that these events are positively correlated. Hence we can conclude that when a node becomes affected there is an affected block in a neighborhood of radius $N^3$. Let us denote this event by $A_7$.

Now we can conclude that the intersection of the above events occurs w.h.p. In particular, for $w$ sufficiently large

$$\sup_{x,y \in \mathcal{N}_{r_w}} |a_{x,y} - \mathbb{E}[a_{x,y}]| \leq C_j^1 r_w^{1/2} \log^{3/2} r_w,$$

and the optimal path between any two nodes inside $\mathcal{N}_{r_w}$ is contained in $\mathcal{N}_\rho$, and the first flipping time of all the particles inside $\mathcal{N}_\rho$ is less than $r_w^{1/3}$ w.h.p.

Consider an optimal path $\gamma$ between $x, y$ where $x, y \in \mathcal{N}_{r_w}$. It follows that the maximum flipping time over all the particles on $\gamma$ is at most $N^2$ w.h.p. Therefore w.h.p. one can find a node $z$ in $\gamma$ such that we have

$$|\lambda a_{x,y} - a_{x,z}| \leq N^2,$$

and

$$|(1 - \lambda)a_{x,y} - (a_{x,y} - a_{x,z})| \leq N^2.$$
Now let us denote \( a_{x,y} - a_{x,z} \) by \( a^*_{z,y} \). We now show that w.h.p.

\[
(5.11) \quad a^*_{z,y} \leq a_{z,y} + N^3 r_w^{1/3}.
\]

To see this, we note that the event of \( a_{z,y} \) being smaller than some value and the event of having an affected block centered at \( z \) are both increasing in the change of a \( \theta \)-particle to a \( \bar{\theta} \)-particle so using the FKG inequality these events are positively correlated. Now we note that since \( z \) has become affected after \( a_{x,z} \) this means that after at most \( N^4 \) time, this node will be inside an affected block and hence we can conclude (5.11).

Next we show that there exists a constant \( c > 0 \) such that \( a^*_{z,y} \geq a_{z,y} - N^3 r_w^{1/2} \log^{3/2} r_w \) w.h.p. To see this, first we note that due to (5.10) and (5.4) we can conclude that there exists \( c' > 0 \) such that there are no affected nodes in \( N_\rho \setminus [A(x,a_{x,y})]_{N' r_w^{1/2} \log^{3/2} r_w} \), where \( [A]_l \) denotes the \( l_\infty \) neighborhood of \( A \). Consider a neighborhood \( N(z) \) with radius \( N^{1/2} r_w^{1/2} \log^{3/2} r_w \) centered at \( z \). We want to argue that w.h.p. the formation of the optimal path from \( x \) to \( y \) from the time \( a_{x,z} \) only involves the spread of affected nodes from \( N(z) \). To see this, assume that the spread of affected nodes started from a node \( z' \) contained in \( A'(0,a_{x,z}) \setminus N(z) \) has also participated in the formation of the optimal path. This implies that w.h.p. there had been a sequence of possible flips leading to an unstable particle that was needed for the formation of the optimal path before the spread of affected nodes from \( N(z) \) had reached this particle. However, this implies that w.h.p. the node \( z \) is not on the optimal path from \( x \) to \( y \) which is a contradiction. Hence we can conclude that there exists \( c > 0 \) such that \( a^*_{z,y} \geq a_{z,y} - N^3 r_w^{1/2} \log^{3/2} r_w \) w.h.p.

Now we can conclude that there exists a constant \( C'' > 0 \) such that w.h.p. we can write

\[
|\lambda a_{x,y} - a_{x,z}| \leq N^2,
\]

and

\[
|(1 - \lambda)a_{x,y} - a_{z,y}| \leq N^{C''} r_w^{1/2} \log^{3/2} r_w.
\]

Combined with (5.10) we can conclude that there exists a constant \( D > 0 \) such that for all \( x, y \in N_\rho \) there exists a node \( z \) for which we have

\[
|\lambda E[a_{x,y}] - E[a_{x,z}]| \leq N^D r_w^{1/2} \log^{3/2} r_w,
\]

and

\[
|(1 - \lambda)E[a_{x,y}] - E[a_{z,y}]| \leq N^D r_w^{1/2} \log^{3/2} r_w.
\]

The result now follows from Sections 3 and 4 in [38]. ■
Now, in order to use the above definition to conclude the shape theorem we need the following result which is a modified version of Proposition 1.8 in [38].

**Proposition 6** (Tessera – modified). If \( \mathbb{E}[a] \) is SAG\( (Q) \) with

\[
Q(\alpha) = \frac{\alpha^{1/2}}{N^{c} \log^{3/2} \alpha},
\]

then there exists a norm \( l_* \) on \( \mathbb{R}^d \) and \( C > 0 \) such that for all \( 2^{c_1 N} \leq n \leq 2^{c_2 N} \), when \( N \) is sufficiently large we have

\[
B_{l_*} \left( 0, n - N^{c_1/2} \log^{3/2} n \right) \cap \mathbb{Z}^d \subset \bar{A}(0, n) \subset B_{l_*} \left( 0, n + N^{c_1/2} \log^{3/2} n \right).
\]

**Proof.** The proof is similar to the proof of Proposition 1.8 in [38], except for the fact that we need to prove a statement similar to Lemma 5.1 in [38] for our setting (since we do not have triangular inequality for \( \mathbb{E}[a] \)). We want to show that, there exists a constant \( c'' > 0 \) such that for all \( W \in \mathbb{N} \) and all \( r/W \geq 2^{c_1 N} \) where \( c'_1 \) is the constant in the definition of SAG, and for \( N \) sufficiently large we have,

\[
d_H \left( \frac{1}{r} \bar{A}(0, r/W)^W, \frac{1}{r} \bar{A}(0, r) \right) \leq \frac{N^{c''}}{Q(r/W)},
\]

where \( d_H \) denotes the Hausdorff distance with respect to the \( l_2 \) norm. We note that due to (5.1) we can write

\[
\bar{A}(0, (1 - \epsilon)r/W)^W \subset \bar{A}(0, r) \subset \bar{A}(0, (1 + \epsilon)r/W)^W,
\]

where

\[
\epsilon = \frac{1}{Q(r/W)}
\]

Since there exists a \( c'' > 0 \) such that \( ||x - y|| \leq N^{c''} \mathbb{E}[a_{x,y}] \) for all \( x, y \in \mathcal{N} \) we have,

\[
\bar{A}(0, (1 + \epsilon)r/W)^W \subset \bar{A}(0, (1 - \epsilon)r/W)^W_{12\epsilon/W} \subset \bar{A}(0, (1 - \epsilon)r/W))B_{l_*}(0, 12N^{c''} \epsilon r/W)^W
\]

\[
= (\bar{A}(0, (1 - \epsilon)r/W))^W B_{l_*}(0, 12N^{c''} \epsilon r)
\]

where the notation \([\bar{A}(0, r/W)]_{12\epsilon/W}\) stands for the \( 12\epsilon/W \)-neighborhood of \( A(0, r/W) \). The rest of the proof follows from Section 5.3 in [38].
Proof of Proposition 3. The proof follows combining Propositions 4, 5, and 6. Using Proposition 4, we can conclude that there exists \( C' > 0 \) such that for sufficiently large \( N \), almost surely we have
\[
\bar{A}(0, n - C'n^{1/2} \log^{3/2} n) \subset A_F(0, n) \subset \bar{A}(0, n + C'n^{1/2} \log^{3/2} n).
\]

Using Proposition 5 we can conclude that \( \mathbb{E}[a] \) is \( \text{SAG}(\alpha^{1/2}/N^c \log^{3/2} \alpha) \).

Finally, the proof follows from Proposition 6 with \( Q(n) = n^{1/2}/N^c \log^{3/2} n \) when \( N \) is sufficiently large.

Proof of Theorem 1 (Shape Theorem). Let \( \mathcal{N}_N \) and \( \mathcal{N}_{2N} \) be two neighborhoods with radii \( N \) and \( 2N \) respectively that are centered at the origin. Using Lemma 4 w.h.p. all the blocks in \( \mathcal{N}_{2N} \setminus \mathcal{N}_N \) are good blocks. Let \( c_2' \in (c_2, 0.5(1 - H(\tau))) \) be a constant and let \( \mathcal{N}_\rho \) be a neighborhood with radius \( \rho \) centered at the origin. Using Lemma 2 there are no affected nodes in \( \mathcal{N}_\rho \setminus \mathcal{N}_{2N} \) w.h.p. Using Lemma 3, this region is a region of expansion w.h.p.

Consider a time interval of size \( n \). It follows from Proposition 3 (by choosing \( c_2' \) in that proposition to be equal to \( c_2' \)), that for sufficiently large \( N \) there exists \( c > 0 \), such that the spreads of possible affected nodes outside of \( \mathcal{N}_\rho \) towards the origin in the \( l_\infty \) norm is less than \( n + N^c \sqrt{n} \log^{3/2} n \) w.h.p. Since for sufficiently large \( N \), \( n + N^c \sqrt{n} \log^{3/2} n < \rho - 2c_2N^1 \) this means that these affected nodes will not be in a neighborhood with radius \( 2c_2N^1 \) at time \( n \). Now assume that all the blocks that share at least one node with \( \mathcal{N}_{2N} \) are affected* blocks. It follows from Proposition 3 and the fact that size of \( \mathcal{N}_{2N} \) is \( o(\sqrt{n}) \) that there exists a constant \( c > 0 \) such that w.h.p.
\[
A_F(0, n) \subset B_{\ast\omega}(0, n + N^c n^{1/2} \log^{3/2} n).
\]

To see the other inclusion, let \( A \) be the event of having a \( \theta \)-affected block at the origin. We have
\[
P\left( B_{\ast\omega}(0, n - N^c n^{1/2} \log^{3/2} n) \cap \mathbb{Z}^2 \subset A_F(0, n) \right) = P\left( B_{\ast\omega}(0, n - N^c n^{1/2} \log^{3/2} n) \cap \mathbb{Z}^2 \subset A'_F(0, n) \mid A \right).
\]

Now since event \( A \) and the event of having \( B_{\ast\omega}(0, n - N^c n^{1/2} \log^{3/2} n) \cap \mathbb{Z}^2 \subset A'_F(0, n) \) are both increasing events in the change of a \( \theta \)-particle to a \( \bar{\theta} \)-particle using the Fortuin-Kasteleyn-Ginibre (FKG) inequality [17] and using Proposition 3 we can conclude that the following event occurs w.h.p.
\[
B_{\ast\omega}(0, n - N^c n^{1/2} \log^{3/2} n) \cap \mathbb{Z}^2 \subset A_F(0, n).
\]
6. Proof of the Size Theorem. Without loss of generality, we assume that the set of nodes of $G_w$ is a subset of the nodes on $\mathbb{Z}^2$ and we work with the obvious probability space. We first define the expandable region; to do so we need the following lemma.

Lemma 13. Let $c$ be an arbitrary positive constant. W.h.p. there are no clusters of bad blocks with radius greater than $N^3$ in a neighborhood with radius $\rho = O(2^{cN})$ in the initial configuration.

Proof. Let $p_*$ be the probability of having a bad block, and let $k = N^3$. By Theorem 3, it follows that w.h.p. there is no cluster of bad blocks containing a bad block with $l_1$-distance from its center greater than $N^3$ in a neighborhood with exponential radius in $N$. \hfill \blacksquare

Now re-normalize $G_w$ into $N$-blocks and consider the union of particles inside a cluster of bad $N$-blocks and the set of particles outside the cluster whose $l_\infty$ distance to at least one node in the cluster is less than or equal to $N/4$. We denote this set by $\mathcal{X}_1$. Note that for sufficiently large $N$, the probability of having a bad $N$-block is below the critical probability of percolation, and each $N$-block is a bad $N$-block independently of the others, hence by Lemma 13, w.h.p. there is no cluster of bad $N$-blocks with radius larger than $N^3$ in a neighborhood with exponential size in $N$ on $G_w$. Also, consider the set of all the particles outside $\mathcal{X}_1$ whose $l_\infty$ distance to at least one particle in $\mathcal{X}_1$ is less than or equal to $N/4$ and denote it by $\mathcal{X}_2$.

Expandable region. $\mathcal{X}_1$ is called a $\theta$-expandable region whenever there exists a set of flips of $\theta$-particles inside $\mathcal{X}_1$ leading to a $\theta$-affected node in $\mathcal{X}_2$.

Fig 9. Expandable region. $\mathcal{X}_1$ is called a $\theta$-expandable region whenever there exists a set of flips of $\theta$-particles inside $\mathcal{X}_1$ leading to a $\theta$-affected node in $\mathcal{X}_2$. It is noted that if $\mathcal{X}_1$ is not an expandable region, the possible spread of $\theta$-affected nodes started in it will die out before reaching $\mathcal{X}_2$. 
The center of an expandable region is the node at the center of the smallest neighborhood that contains the expandable region.

We now want to argue about the distance of the closest expandable region to the origin. The following lemma shows how far the closest expandable region to the origin can be. We show this by establishing a relationship between radical regions and expandable regions. The following lemma exploits the fact that the closest radical region to the origin is also an expandable region w.h.p. (note that an expandable radical region is defined differently from an expandable region.)

**Lemma 14.** Let $\epsilon > 0$. W.h.p. the $l_*$-distance of the origin from the node at the center of the closest expandable region in the initial configuration is at most

$$\rho = 2^{0.5(1-H(\tau)+\epsilon)(1+\epsilon')^2N}.$$ 

**Proof.** Consider a neighborhood with radius $2^{0.5(1-H(\tau)+\epsilon/2)(1+\epsilon')^2N}$ centered at the origin. Using Lemma 8, w.h.p. there exists a radical region in this neighborhood. Using Lemma 7 a radical region is expandable w.h.p. Using Lemma 13, w.h.p. there is no cluster of bad blocks with radius larger than $N^3$ in a neighborhood with radius $\rho$ centered at the origin in the initial configuration. Now consider a neighborhood with radius $3N^3$ centered at the center of the radical region. Since the event of having an expandable radical region at the center of this neighborhood and the event of this region begin a region of expansion are both increasing events in the change of a $\theta$-particle to a $\overline{\theta}$-particle, by an application of Fortuin-Kasteleyn-Ginibre (FKG) inequality [17] and using Lemma 3 this neighborhood is a region of expansion w.h.p. and this completes the proof. $\blacksquare$

The following lemma shows that in the initial configuration w.h.p. there is no part of an expandable region in an annulus around the origin whose width is $\rho'$.

**Lemma 15.** Let $\epsilon > 0$. W.h.p. there is no node that belongs to an expandable region in

$$B_{l_*}(0, \rho + \rho') \setminus B_{l_*}(0, \rho),$$

for

$$\rho = 2^{0.5(1-H(\tau)+\epsilon/2)(1+\epsilon')^2N},$$

$$\rho' = 2^{(1-H(\tau)-\epsilon)(1-0.5(1+\epsilon')^2)N+o(N)}.$$
Proof. By Lemma 13, w.h.p. there is no cluster of bad blocks with radius larger than $N^3$ in a neighborhood with radius $2(\rho + \rho')$ centered at the origin in the initial configuration. Also, for large $N$ we have,

$$\text{Number of nodes in } B_{l^*}(0, \rho + \rho' + 2N^3) \setminus B_{l^*}(0, \rho) \leq 8\rho\rho'.$$

Also, if we have

$$\text{Number of nodes in } B_{l^*}(0, \rho + \rho' + 2N^3) \setminus B_{l^*}(0, \rho) = 2^{(1 - H(\tau) - \epsilon) + o(N)},$$

then w.h.p. there will be no location inside $B_{l^*}(0, \rho + \rho' + 2N^3) \setminus B_{l^*}(0, \rho)$ for which a particle would be unstable. Here we have used the fact that $(H(\tau) - H(\tau'')) \in o(1)$, and the fact that having an unstable particle is a necessary condition for having an expandable region. ■

The following lemma shows that w.h.p. an expandable region can lead to the formation of a $\theta$-affected $w$-block.

Lemma 16. W.h.p. there exists a sequence of possible flips in $X_1 \cup X_2$ that can lead to a $\theta$-affected $w$-block centered in $X_1$.

Proof. Since $X_1$ is an expandable region, there exists a sequence of flips leading to a new affected node inside a good $N/4$-block. Using Lemma 12 we can conclude that, w.h.p. there exists a sequence of possible flips in the $N/4$-block that can lead to a monochromatic $w$-block in it. Now with an application of Lemma 3 and using the fact that the event of having $X_1 \cup X_2$ inside a $N^3$-block and the event of having that block being a region of expansion are positively correlated, we can conclude that the later event occurs w.h.p. and this completes the proof. ■

We are now ready to begin the proof of Theorem 2.

Proof of Theorem 2. We first show that the size of the monochromatic region is at least exponential in $N$ w.h.p. Let $\epsilon > 0$, and

$$a(\tau) = (1 - H(\tau) - \epsilon) \left(2 - (1 + \epsilon')^2\right),$$

where $\epsilon' > f(\tau)$. Let $n^* = 2^{a(\tau)}N$. We wish to show that for all $n \geq n^*$,

$$M_n \geq 2^{a(\tau)N} \text{ w.h.p.}$$

By Lemma 13, w.h.p. there is no cluster of bad blocks with radius larger than $N^3$ in a neighborhood with radius $2^N$ centered at the origin in the initial configuration (event $A_0$).
EVOLUTION AND STEADY STATE OF A LONG-RANGE SPIN SYSTEM

Let
\[ \rho = 2^{0.5(1-H(\tau) + \epsilon/2)(1+\epsilon')^2N}, \]
\[ \rho' = 2^{(1-H(\tau) - \epsilon)(1-0.5(1+\epsilon')^2)N + 2 \log_2 N}, \]
\[ \rho'' = 2^{(1-H(\tau) + \epsilon)((1+\epsilon')^2 - 1)N}. \]

We let \( N \) be sufficiently large so that there exists a norm \( l_* \) and \( C > 0 \) such that (5.3) in Proposition 3 is satisfied for \( n = \rho^{1/3}, \rho'/N > \omega^3 \), and we have \( LL' < \rho'/4 - \rho'^{1/3} - (N^4 + N)\rho'' - N \) where

\[ L = \left\lfloor \frac{\rho}{\rho'/4 - 2NC\sqrt[3]{\rho'} \log^{3/2} \rho'' - N} \right\rfloor, \]
\[ L' = \left\lfloor 2NC\sqrt[3]{\rho'} \log^{3/2} \rho' + N \right\rfloor. \]

Let the closest expandable region to the origin in the \( l_* \) norm be a \( \theta \)-expandable region and let \( \mathcal{N}_{\rho'/4} \) be a neighborhood at the origin with radius \( \rho'/4 \), now let

\[ T(\rho'/4) = \inf \left\{ n \mid \exists \text{ a } \theta \text{-affected node in } \mathcal{N}_{\rho'/4} \right\}. \]
\[ A = \left\{ \text{The origin is contained in a firewall of radius } \rho'/N \text{ before } T(\rho'/4) \right\}. \]

We now want to show

(6.1) \( A \) occurs w.h.p.

To show this, we condition this event on a few events that occur w.h.p. and argue that since the conditional probability occurs w.h.p., this event also occurs w.h.p.

Let \( X \) denote the \( l_* \)-distance from the origin to the closest node in an expandable region and let

\[ A_1 = \left\{ X \leq \rho, \text{ at } n = 0 \right\}, \]
\[ A_2 = \left\{ \not\exists \text{ a } \theta \text{-expandable region in } B_{l_*}(0, X + \rho') \setminus B_{l_*}(0, X) \text{ at } n = 0 \right\}. \]

Consider an \( l_* \)-ball of radius \( \rho \). According to Lemma 14, w.h.p. there is an expandable region in this ball (see Figure 10). This implies

\( A_1 \) occurs w.h.p.

Using the fact that the existence of a \( \theta \)-expandable region in \( B_{l_*}(0, X + \rho') \setminus B_{l_*}(0, \rho) \) can only increase the probability of event \( A_2 \) (since they are both
increasing events in the change of a $\theta$-particle to a $\bar{\theta}$-particle, by an application of FKG inequality [17] for the initial configuration they are positively correlated), and the fact that conditional on event $A_1$, event $A_2$ would have the smallest probability when $X = \rho$, we let

$$A'_2 = \{ \not\exists \text{ a } \bar{\theta}\text{-expandable region in } B_{l^*}(0, \rho + \rho') \setminus B_{l^*}(0, \rho) \text{ at } n = 0 \},$$

and by an application of FKG inequality [17] for the initial configuration we have

$$P(A_2) \geq P(A_2 \mid A_1) P(A_1) \geq P(A_2 \mid X = \rho) P(A_1) \geq P(A'_2) P(A_1).$$

Using Lemma 15, event $A'_2$ occurs w.h.p., hence we have

$$A_2 \text{ occurs w.h.p.}$$

Now consider the line segment from the center of the closest expandable region to the origin. Let $\mathcal{N}$ denote the set of particles such that their $l^*$-distances from at least one point on the line segment is less than or equal to $2N'\rho'$ where $c'$ is the constant in Lemma 11. Let

$$A_3 = \{ \not\exists \text{ a } \bar{\theta}\text{-affected node in } \mathcal{N} \text{ and it is a region}
\text{ of expansion at } n = 0 \}. $$

Since event $A_3$ has the smallest probability when $X = \rho$, with an application of Lemma 1 and Lemma 3, we can conclude that

$$A_3 \text{ occurs w.h.p.}$$
Fig 11. The gradual spreads of affected nodes from the expandable region towards the origin in time intervals of $\rho'/4$. After each time interval, all the nodes inside the corresponding $l_*$-ball of radius $\rho'/4 - o(\rho')$ are going to be affected.

We are now going to consider the gradual spread of the affected nodes from the expandable region towards the origin in $\rho'/4$ time intervals (see Figure 5).

To consider the growth of the $\theta$-affected nodes we first notice that using Lemma 16, w.h.p. there exists a sequence of less than $N^4$ flips that can create a $\theta$-affected $w$-block centered at the center of the expandable region. Let us denote this event by $A_4$. Hence we have

$$A_4 \text{ occurs w.h.p.}$$

Let $T_N$ denote the time it takes until $N^4$ flips occur one by one. Let $A'_5 = \{T_N < \rho^{1/3}\}$. Standard concentration bounds imply that there exist $c > 0$ such that this event occurs with probability at least $1 - \exp(-c\rho^{1/3})$. Let $A_5$ denote the event that the time that it takes until we have a $\theta$-affected block at the center of the expandable region is less than $\rho^{1/3}$. We have

$$P(A_5) \geq P(A'_5) \geq 1 - \exp\left(-c\rho^{1/3}\right).$$

Hence we have

$$A_5 \text{ occurs w.h.p.}$$

Let $A_6$ denote the event of having less than $\rho''$ affected nodes in the $l_*$-ball of radius $\rho$. Standard concentration bounds imply that this event also occurs w.h.p., hence we have

$$A_6 \text{ occurs w.h.p.}$$
Now we want to consider the spread of possible $\theta$-affected nodes from outside of $B_l((0, \rho+\rho')$ towards the origin. Since there are at most $\rho''$ affected nodes in this ball w.h.p., if we remove all the annuli from this ball that contain an affected node along with their clusters of bad blocks and also a margin of good blocks around these clusters we would have a ball of radius at least $\rho - (N^4 + N)\rho''$. Furthermore, we argue that since the event of having larger growths of $\theta$-affected nodes in a given time interval and the event of having this ball being a region of expansion of type $\theta$ are both increasing in the change of a $\theta$-particle to a $\theta$-particle, by assuming that this ball is a region of expansion of type $\theta$ we would only get an upper bound for the speed of the spread of $\theta$-affected nodes. Hence, from this point forward, to consider the spread of $\theta$-affected nodes towards the origin we consider this ball and assume it is a region of expansion and does not contain any affected nodes.

Let $A_7$ denote the event that the growths of the $\theta$-affected nodes in a time interval of $\rho'^{1/3}$ – that conditional on $A_i$, $i = 1, 2, \ldots, 6$ is at most needed for the formation of the $\theta$-affected $w$-blocks in the first gradual growth – is less than $\rho'^{1/3} + N^C \rho'^{1/6} \log^{3/2} \rho'$ in the $l_*$ norm in all directions in the annulus around the origin. To show that this event occurs w.h.p. we consider the $l_*$-ball of radius $\rho - (N^4 + N)\rho''$ described above. It follows from the proof of Proposition 3 that the growths of all the $\theta$-affected nodes in this ball will be less than $\rho'^{1/3} + N^C \rho'^{1/6} \log^{3/2} \rho'$ hence we can conclude that

\[ A_7 \text{ occurs w.h.p.} \]

Now let $A_8$ denote the event that the origin is contained in a $w$-block of $\theta$-affected nodes before there are any $\theta$-affected nodes in an $l_*$-ball with radius $\rho'/2$ around the origin. To show that this event occurs w.h.p. we consider $L$ time intervals of size $\rho'/4$ and argue that first of all in every one of these time intervals the growth of $\theta$-affected nodes from the closest $\theta$-affected $w$-block towards the origin (first started from the expandable region) is at least $\rho'/4 - N^C (\rho')^{1/2} \log^{3/2} \rho'$. (Note that we have already argued in the proof of the shape theorem that the event of having a $\theta$-affected block and having larger growths are positively correlated.) To see this, consider a line segment of length $\rho$ and let $N'$ denote the set of particles such that their $l_*$-distances from at least one point on the line segment is less than or equal to $N^C \rho'$ where $c'$ is the constant in Lemma 11. We argue that not having $\theta$-affected nodes in $N'$ and having smaller growths in a given time interval in this neighborhood are both increasing events in the change of a $\theta$-particle to a $\theta$-particle hence using the FKG inequality they are positively correlated. Let us consider a neighborhood with the shape of $N'$ that does not contain
any affected nodes and is a region of expansion. It follows from the proof of Proposition 3 that assuming only particles in this neighborhood can make a flip, all the growths of an affected* block in this neighborhood will be at least $\rho'/4 - N^C(\rho')^{1/2} \log^{3/2} \rho'$. It follows that all the gradual growths of $\theta$-affected nodes started by the expandable region towards the origin will be at least $\rho'/4 - N^C(\rho')^{1/2} \log^{3/2} \rho'$.

Second of all we need to show that the possible growths of $\bar{\theta}$-affected nodes started from outside of $B_{l^*}(0, \rho + \rho' - w)$ are not going to interfere with any of these growths and will not reach the $B_{l^*}(0, \rho'/2)$ before having the origin contained in a $\theta$-affected $w$-block (see Figure 11). To show this, since we are conditioning on events $A_i$, $i = 1, 2, ..., 7$, and since $LL' < \rho'/4 - \rho'^{1/3} - (N^4 + N)\rho'' - N$ it suffices to show that the growths of $\bar{\theta}$-affected nodes started from outside a $B_{l}(0, \rho - \rho'^{1/3} - N^c\rho'' - N)$ which does not contain any affected nodes and is assumed to be a region of expansion in every time interval of size $\rho'/4$ is at most $\rho'/4 + N^C(\rho')^{1/2} \log^{3/2} \rho'$, which guarantees that not only these growths will not interfere with the growths of $\theta$-affected nodes but also they will not reach $B_{l^*}(0, \rho'/2)$ before having the origin contained in a $\theta$-affected $w$-block. It follows from the proof of Proposition 3 that this event also occurs w.h.p. Hence we can conclude that

$$A_8 \text{ occurs w.h.p.}$$

This also implies that by the time the origin is contained in a $\theta$-affected $w$-block, the $\theta$-affected nodes are still in $l^*$-distance of more than $\rho'/2$ from the origin w.h.p.

Now let $r$ be proportional to $\rho'/N$. Let us denote the event that the time it takes until a number of affected nodes equal to the number of all the particles in a firewall with radius $r$ centered at the origin and a line of width $2\sqrt{N}$ from the origin to the firewall make a flip one by one being smaller than $\rho'/4$ by $A'_9$ (see Figure 12). Standard concentration bounds imply $A'_9$ occurs w.h.p.

Let $A_9$ denote the event that this firewall is formed in a time interval smaller than $\rho'/4$. We have $P(A_9) \leq P(A'_9)$ and since $A'_9$ occurs w.h.p. we have

$$A_9 \text{ occurs w.h.p.}$$

With a similar argument for event $A_8$, w.h.p. the growth of all the possible $\theta$ affected nodes will be less than $\rho'/3$ for this interval (event $A_{10}$). Hence, we have

$$A_{10} \text{ occurs w.h.p.}$$
Finally we can write
\[ P(A) \geq P \left( A \mid A_0, A_1, \ldots, A_{10} \right) P(A_0 \cap A_1 \cap \ldots \cap A_{10}), \]
hence, we have that
\[ A \text{ occurs w.h.p.} \]

Now, using Lemma 3 w.h.p. the interior of the firewall is a region of expansion in the initial configuration and since w.h.p. only \( \theta \)-affected nodes have reached this region by the time of the formation of the firewall, it is still a region of expansion for the state \( \bar{\theta} \). Now, since the sum of the time for the gradual growths, formation of the firewall, and the time that it takes until the interior of the firewall becomes monochromatic (by a standard concentration bound) is less than \( n^* \) w.h.p., for all \( n \geq n^* \) it will be monochromatic w.h.p. and this proves the lower bound.

Next, we show the corresponding upper bound. Let
\[ b(\tau) = (1 + \epsilon')^2 (1 - H(\tau) + \epsilon). \]

Consider four neighborhoods with radius \( N(\rho + \rho') \) such that each of them share the origin as a different corner node. Divide the union of these neighborhoods into neighborhoods of radius \( \rho + \rho' \) in an arbitrary way and consider the nodes at the center of each of these neighborhoods. Now using the above result we have that for \( n \geq n^* \), w.h.p., all these central nodes will have a monochromatic region of size at least \( 2^{a(\tau)} N \). Also it is easy to see that for \( n \geq n^* \), w.h.p. all the four neighborhoods defined above will have particles with exponentially large monochromatic regions of both states. This implies that for all \( n \geq n^* \) the size of the monochromatic region of the origin is at most \( 4N^2(\rho + \rho')^2 \). Finally let
\[ A_w = \left\{ 2^{a(\tau)} N \leq M_n \leq 2^{b(\tau)} N \text{ on } G_w \text{ for } n \geq n^* \right\}. \]

We have that \( P(A^C_w) = o(w^{-2}) \). The result now follows from the fact that \( \sum_{w=1}^{\infty} P(A^C_w) < \infty. \)
6.1. Extension to $\tau > 1/2$. We call super-unstable particles the unstable particles that can potentially become stable once they flip their state. While for $\tau < 1/2$ unstable particles can always become stable by flipping their state, for $\tau > 1/2$ this is only true for the super-unstable particles. It follows that for $\tau > 1/2$ super-unstable particles act in the same way as unstable particles do for $\tau < 1/2$.

We let $\bar{\tau} = 1 - \tau + 2/N$. A super-unstable particle of type $\theta$ is a particle for which $W < \bar{\tau}N$ where $W$ is the number of $\theta$ particles in its neighborhood. The reason for adding the term $2/N$ in the definition is to account for the strict inequality that is needed for being unstable and the flip of the particle at the center of the neighborhood which adds one particle of its type to the neighborhood. A super-radical region is a neighborhood $N_S$ of radius $S = (1 + \epsilon')w$ such that $W_S < \bar{\tau}'(1 + \epsilon')^2N$, where $\epsilon \in (0, 1/2)$ and

$$\bar{\tau}' = \left(1 - \frac{1}{\bar{\tau}N^{1/2 - \epsilon}}\right) \bar{\tau}.$$ 

By replacing $\tau$ with $\bar{\tau}$, “unstable particle” with “super-unstable particle” and “radical region” with “super-radical region,” it can be checked that all proofs extend to the interval $1/2 < \tau < 1 - \tau^*$ for the shape theorem and the interval $1/2 < \tau < 1 - \tau^*$ for the size theorem.

Acknowledgment. The authors thank Prof. Jason Schweinsberg of the Mathematics Department of University of California at San Diego for providing invaluable feedback and suggestions on earlier drafts of the paper.

References.
[1] Alves, O., Machado, F. P., Popov, S. Y. et al. (2002). The shape theorem for the frog model. *The Annals of Applied Probability* **12** 533–546.
[2] Arratia, R. (1983). Site Recurrence for Annihilating Random Walks on $\mathbb{Z}^d$. *Ann. Probab.* **11** 706–713.
[3] Barmpalias, G., Elwes, R. and Lewis-Pye, A. (2014). Digital morphogenesis via Schelling segregation. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on* 156–165. IEEE.
[4] Barmpalias, G., Elwes, R. and Lewis-Pye, A. (2015a). Minority population in the one-dimensional Schelling model of segregation. *arXiv preprint arXiv:1508.02497*.
[5] Barmpalias, G., Elwes, R. and Lewis-Pye, A. (2015b). Tipping points in 1-dimensional Schelling models with switching agents. *Journal of Statistical Physics* **158** 806–852.
[6] Barmpalias, G., Elwes, R. and Lewis-Pye, A. (2016). Unperturbed Schelling Segregation in Two or Three Dimensions. *Journal of Statistical Physics* **164** 1460–1487.
[7] Bhakta, P., Miracle, S. and Randall, D. (2014). Clustering and mixing times for segregation models on $\mathbb{Z}^d$. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms* 327–340. Society for Industrial and Applied Mathematics.
[8] Brandt, C., Immorlica, N., Kamath, G. and Kleinberg, R. (2012). An analysis of one-dimensional Schelling segregation. In Proceedings of the forty-fourth annual ACM symposium on Theory of computing 789–804. ACM.

[9] Caputo, P. and Martinelli, F. (2006). Phase ordering after a deep quench: the stochastic Ising and hard core gas models on a tree. Probability theory and related fields 136 37–80.

[10] Castellano, C., Fortunato, S. and Loreto, V. (2009). Statistical physics of social dynamics. Reviews of modern physics 81 591.

[11] Chopard, B. and Droz, M. (1998). Cellular automata. Springer.

[12] Draief, M. and Massouli, L. (2010). Epidemics and rumours in complex networks. Cambridge University Press.

[13] Drewitz, A., Ráth, B. and Sapozhnikov, A. (2014). On chemical distances and shape theorems in percolation models with long-range correlations. Journal of Mathematical Physics 55 083307.

[14] Easley, D. and Kleinberg, J. (2010). Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press.

[15] Erdős, P. and Ney, P. (1974). Some problems on random intervals and annihilating particles. The Annals of Probability 2 828–839.

[16] Fontes, L. R., Schonmann, R. and Sidoravicius, V. (2002). Stretched exponential fixation in stochastic Ising models at zero temperature. Communications in mathematical physics 228 495–518.

[17] Fortuin, C. M., Kasteleyn, P. W. and Ginibre, J. (1971). Correlation inequalities on some partially ordered sets. Communications in Mathematical Physics 22 89–103.

[18] Grimmett, G. (1999). Percolation, second ed. Springer.

[19] Hethcote, H. W. (2000). The mathematics of infectious diseases. SIAM review 42 599–653.

[20] Holden, N. and Sheffield, S. (2018). Scaling limits of the Schelling model. arXiv preprint arXiv:1802.03346.

[21] Immorlica, N., Kleinberg, R., Lucier, B. and Zadomichaddam, M. (2017). Exponential Segregation in a Two-dimensional Schelling Model with Tolerant Individuals. 984–993.

[22] Jackson, M. O. and Watts, A. (2002). On the formation of interaction networks in social coordination games. Games and Economic Behavior 41 265–291.

[23] Kanoria, Y., Montanari, A. et al. (2011). Majority dynamics on trees and the dynamic cavity method. The Annals of Applied Probability 21 1694–1748.

[24] Kesten, H. (1986). Aspects of first passage percolation. École d’Été de Probabilités de Saint Flour XIV-1984 125–264.

[25] Kesten, H. (1993). On the Speed of Convergence in First-Passage Percolation. Ann. Appl. Probab. 3 296–338.

[26] Kleinberg, J. (2007). Cascading behavior in networks: Algorithmic and economic issues. Algorithmic game theory 24 613–632.

[27] Liggett, T. (2012). Interacting particle systems 276. Springer Science & Business Media.

[28] Liggett, T. M. (2013). Stochastic interacting systems: contact, voter and exclusion processes 324. Springer Science & Business Media.

[29] Meyer-Ortmanns, H. (2003). Immigration, integration and ghetto formation. International Journal of Modern Physics C 14 311–320.

[30] Mobius, M. M. and Rosenblat, T. (2000). The formation of ghettos as a local interaction phenomenon. Unpublished manuscript, Harvard University.

[31] Morris, R. (2011). Zero-temperature Glauber dynamics on $\mathbb{Z}^d$. Probability theory
and related fields 149 417–434.

[32] Omidvar, H. and Franceschetti, M. (2018). Self-organized Segregation on the Grid. Journal of Statistical Physics 170 748–783.

[33] Schelling, T. C. (1969). Models of Segregation. The American Economic Review 59 488–493.

[34] Schelling, T. C. (1971). Dynamic models of segregation. Journal of mathematical sociology 1 143–186.

[35] Schulze, C. (2005). Potts-like model for ghetto formation in multi-cultural societies. International Journal of Modern Physics C 16 351–355.

[36] Stauffer, D. and Solomon, S. (2007). Ising, Schelling and self-organising segregation. The European Physical Journal B-Condensed Matter and Complex Systems 57 473–479.

[37] Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. Publications Mathématiques de l’IHES 81 73–205.

[38] Tessera, R. (2017). Speed of convergence in first passage percolation and geodesicity of the average distance. Annales de l’institut Henri Poincaré.

[39] Young, H. P. (2001). Individual strategy and social structure: An evolutionary theory of institutions. Princeton University Press.

[40] Zhang, J. (2004a). A dynamic model of residential segregation. Journal of Mathematical Sociology 28 147–170.

[41] Zhang, J. (2004b). Residential segregation in an all-integrationist world. Journal of Economic Behavior & Organization 54 533–550.

[42] Zhang, J. (2011). Tipping and Residential Segregation: A Unified Schelling Model*. Journal of Regional Science 51 167–193.