The Annals of Probability
1994, Vol. 22, No. 3, 1121–1139

THE THRESHOLD VOTER AUTOMATON AT A CRITICAL POINT[1]

BY JEFFREY E. STEIF

Chalmers University of Technology

We consider the threshold voter automaton in one dimension with threshold $\tau > n/2$, where n is the number of neighbors and where we start from a product measure with density $\frac{1}{2}$. It has recently been shown that there is a critical value $\theta_c \approx 0.6469076$, so that if $\tau = \theta n$ with $\theta > \theta_c$ and n is large, then most sites never flip, while for $\theta \in (\frac{1}{2}, \theta_c)$ and n large, there is a limiting state consisting mostly of large regions of points of the same type. Using a supercritical branching process, we show that the behavior at θ_c differs from both the $\theta > \theta_c$ regime and the $\theta < \theta_c$ regime and that, in some sense, there is a discontinuity both from the left and from the right at this critical value.

1. Introduction. In this paper we will consider a discrete-time model, the *threshold voter automaton*, in which the state at time t is $\xi_t\colon \mathbb{Z} \to \{0, 1\}$ (or equivalently $\{0, 1\}^{\mathbb{Z}}$) and we think of $\xi_t(x)$ as giving the opinion of the voter at location x at time t. See [5] for recent results concerning this model and for a continuous-time analogue.

The threshold voter automaton is a two-parameter (r and τ) deterministic discrete-time process with state space $\{0, 1\}^{\mathbb{Z}}$ in which at each time n, the voter at x examines the opinions of her neighbors $\{y\colon |y - x| \le r\}$ and changes her opinion if and only if at least τ neighbors have the opposite opinion (r and τ here are nonnegative integers). More precisely, we have a transformation $T^{r,\tau}$ from $\{0, 1\}^{\mathbb{Z}}$ to itself given by

$$T^{r,\tau}\eta(x) = 1 - \eta(x) \quad \text{if and only if} \sum_{y=x-r}^{x+r} I_{\{\eta(y) \ne \eta(x)\}} \ge \tau.$$

Here η denotes a typical element of the state space $\{0, 1\}^{\mathbb{Z}}$. Throughout this paper, we always start the system with a product measure with density $\frac{1}{2}$ and want to investigate what happens after many iterations of $T^{r,\tau}$. This system is an example of what is called a *cellular automaton*. (If τ is not an integer, $T^{r,\tau}$ will be taken to mean $T^{r,\lfloor\tau\rfloor}$.)

The first result we mention is due to Fisch and Gravner [6].

THEOREM 1.1. *If $\tau = r + 1$ or $\tau \ge 5r/4$, then the system fixates a.s.* (*in that each point changes its value only finitely many times*).

Received March 1993.
[1]Supported by a grant from the Swedish National Science Foundation.
AMS 1991 *subject classifications*. Primary 60K35; secondary 60J80, 60F10.
Key words and phrases. Cellular automata, critical value, Chen–Stein method, branching processes.

The main result of [5] gives qualitative properties of this “limiting state.” We put the phrase limiting state in quotes since the limit is not known to exist if $r+1 < \tau < 5r/4$.

To formulate this result, let

$$c(a) = \log 2 + a\log a + (1-a)\log(1-a).$$

(All logs in this paper are natural.) The reason for our interest in this quantity is that if S_n is the sum of n independent random variables that are 0 or 1 with equal probability, then for $a > \frac{1}{2}$,

$$\lim_{n\to\infty} \frac{1}{n}\log P(S_n \geq na) = -c(a) \tag{1.1}$$

and

$$P(S_n \geq na) \leq e^{-c(a)n} \quad \text{for all } n. \tag{1.2}$$

For this large deviations result, see, for example, Section 9 in Chapter 1 of [4]. The main result of [5] is the following theorem.

THEOREM 1.2. *Let B_k be the event that the voters at all sites x with $|x| \leq k$ never change. Let A_k be the event that all the voters at all sites x with $|x| \leq k$ fixate in the same state. Suppose $\tau = \theta(2r+1)$. Let θ_c be the unique solution in $(\frac{1}{2}, \frac{3}{4})$ of the equation $2c(\theta) = c(2\theta - 1)$ and let $P^{r,\theta}$ denote probabilities with respect to the $T^{r,\theta(2r+1)}$ dynamics.*

(i) *If $\theta > \theta_c$, then for all k, $P^{r,\theta}(B_k) \to 1$ as $r \to \infty$.*
(ii) *If $\theta \in (\frac{1}{2}, \theta_c)$, then for all k, $P^{r,\theta}(A_k) \to 1$ as $r \to \infty$.*

(Note that the events B_k and A_k depend implicitly on the parameters r and θ.) It is easy to see that for any r, θ and k, $P^{r,\theta}(A_k \cap B_k) \leq 2/2^k$, which implies that for any θ, $\lim_{k\to\infty} \limsup_{r\to\infty} P^{r,\theta}(A_k \cap B_k) = 0$. This together with Theorem 1.2 easily implies that

$$\lim_{k\to\infty} \limsup_{r\to\infty} P^{r,\theta}(A_k) = 0 \quad \text{for } \theta > \theta_c \tag{1.3}$$

and

$$\lim_{k\to\infty} \limsup_{r\to\infty} P^{r,\theta}(B_k) = 0 \quad \text{for } \theta \in \left(\tfrac{1}{2}, \theta_c\right). \tag{1.4}$$

Theorem 1.2 together with (1.3) and (1.4) tells us that the system behaves differently above and below θ_c.

The reason for interest in this result is that it is a prototype for conclusions that one would like to prove for related systems with more than two states and in higher dimensions. See [7]. Theorem 1.2 can be explained on the basis of the following heuristic arguments. Suppose $\theta < \frac{3}{4}$. Call an interval of 1’s of length greater than r a *1-blob*, and call an interval of length $r+1$ in which there are less than $(2r+1)\theta - r$ 1's a *1-blockade*. It is easy to check that 1-blobs tend to grow but that a 1-blockade will stop a 1-blob. To prove (i), it suffices to observe that in order for a blob to form we need a point to be *1-unsatisfied*, that is, to have at least $\theta(2r+1)$ 1's in its neighborhood. We think of a point as *satisfied* if the density of 1's (and hence of 0's) in its r-neighborhood lies in $(1-\theta, \theta)$, for this guarantees that the point will not change its value at the next iteration. Of course, most points will (by the weak law of large numbers) be satisfied initially. If a point is unsatisfied, either the density of 1's or of 0's is larger than θ. The 1 in front of "unsatisfied" refers to the fact that it is the 1's whose density is larger than θ. Finally, if $2c(\theta) > c(2\theta - 1)$, then as $r \to \infty$ blockades are much more numerous than unsatisfied points and $P^{r,\theta}(B_k) \to 1$.

To prove the converse in (ii), one needs to show that if $2c(\theta) < c(2\theta - 1)$, then (a) for large r, blobs are more numerous than blockades and (b) blobs grow until they run into each other. To prove (a), one observes that unsatisfied points are more numerous than blockades and one shows that if $\sigma > 0$ and there are at least $(\theta+\sigma)(2r+1)$ 1's in the neighborhood of a point, then a 1-blob will form with high probability. To prove (b), one shows that if a collection of unsatisfied sites does not grow into a blob, then the number of sites that flip is sufficiently small with high probability so that a blockade does not form.

The main result of the present paper gives information as to what happens at the critical value θ_c.

THEOREM 1.3. *Let A_k and B_k be as in Theorem* 1.2. *Then there is a universal constant $\gamma > 0$ (independent of r and k) so that*:

(i) *for all k*, $\liminf_{r\to\infty} P^{r,\theta_c}(B_k) \ge \gamma$ *and*
(ii) *for all k*, $\liminf_{r\to\infty} P^{r,\theta_c}(A_k) \ge \gamma$.

Theorem 1.3 immediately gives

$$\liminf_{k\to\infty} \liminf_{r\to\infty} P^{r,\theta_c}(A_k) \ge \gamma$$

and

$$\liminf_{k\to\infty} \liminf_{r\to\infty} P^{r,\theta_c}(B_k) \ge \gamma.$$

Comparing these with (1.3) and (1.4) tells us that the behavior at θ_c distinguishes itself both from the $\theta > \theta_c$ regime and the $\theta < \theta_c$ regime and is rather a "convex combination" of these two types of behavior. Moreover, this result says that in a certain sense, there is a discontinuity both from the left and from the right at the critical value.

One of the key lemmas in this analysis is that there is a δ which does not depend on r so that a 1-unsatisfied point makes a 1-blob with probability at least δ. This result, which is of interest in itself, is proved by extracting a certain process from the system which dominates a supercritical branching process.

LEMMA 1.4. *Let $\theta \in (0.5, 0.68)$. Then there is a $\delta > 0$ such that for all r,*

$$P^{r,\theta}\big([-r, r] \text{ is a 1-blob by time } 2r|0 \text{ is 1-unsatisfied}\big) \geq \delta.$$

The definitions of 1-unsatisfied and 1-blob are given more precisely in the next section.

Once one has this, using the fact (obtained by Stirling's formula) that unsatisfied points and blockades have more or less the same probabilities [in that the ratio of these probabilities is bounded (uniformly in r) away from 0 and ∞], one has that blobs and blockades also have more or less the same probabilities. One then uses the Chen–Stein method to show that both blobs and blockades are "independently and evenly distributed" in space and then (modulo a number of details) one has obtained Theorem 1.3.

In Section 2, we prove the main theorem, Theorem 1.3, assuming Lemma 1.4. In Section 3 we prove Lemma 1.4.

2. Proof of Theorem 1.3. In this section we prove Theorem 1.3 assuming Lemma 1.4. We first introduce some notation and conventions which will be used throughout this paper. If $k, \ell \in \mathbb{Z}$, we use $[k, \ell]$ to denote $[k, \ell] \cap \mathbb{Z}$. $\{X_i\}_{i\in\mathbb{Z}}$ will always denote the initial (random) element of $\{0, 1\}^{\mathbb{Z}}$. Therefore, $\{X_i\}_{i\in\mathbb{Z}}$ are i.i.d. with $P(X_0 = 1) = \frac{1}{2} = P(X_0 = 0)$. We will also let $S_{k,\ell} \equiv \sum_{i=k}^{\ell} X_i$ be the number of 1's in $[k, \ell]$. In addition, $P^{r,\theta}$ will denote probabilities with respect to the $T^{r,\theta(2r+1)}$ dynamics. Finally, a number x which might not be an integer but only makes sense as an integer will always be interpreted as the greatest integer less than or equal to x, $\lfloor x \rfloor$. Of course, $\lceil x \rceil$ will denote the smallest integer greater than or equal to x.

We begin with some definitions.

DEFINITION 2.1. We say $z \in \mathbb{Z}$ is 1-unsatisfied for η if

$$\sum_{y=z-r}^{z+r} I_{\{\eta(y)=1\}} \geq (2r + 1)\theta;$$

0-unsatisfied for η if

$$\sum_{y=z-r}^{z+r} I_{\{\eta(y)=0\}} \geq (2r + 1)\theta;$$

and unsatisfied if it is either 1-unsatisfied or 0-unsatisfied.

DEFINITION 2.2. We say $z \in \mathbb{Z}$ is a 1-blockade for η if $\sum_{y=z}^{z+r} I_{\{\eta(y)=1\}} < (2r + 1)\theta - r$; a 0-blockade for η if $\sum_{y=z}^{z+r} I_{\{\eta(y)=0\}} < (2r + 1)\theta - r$; and a blockade if it is either a 1-blockade or a 0-blockade.

Note that these definitions also depend implicitly on the parameters r and θ. The point of a 1-blockade (and similarly a 0-blockade) can be seen in the following easily verified fact:

(2.1) If $z \in \mathbb{Z}$ is a 1-blockade for η, then (under the $T^{r,\theta(2r+1)}$ dynamics) every 0 in $[z, z+r]$ remains 0 for all time independent of what happens outside. Conversely, if z is not a 1-blockade and all sites in $[z-r, z-1]$ are 1, then (under the $T^{r,\theta(2r+1)}$ dynamics) z will flip to 1 in the next iteration.

An analogous statement holds for 0-blockades by just interchanging the 0's and 1's. It follows from the large deviations result quoted in the Introduction that

$$\lim_{r\to\infty}\left(\frac{1}{r}\right)\log P^{r,\theta}(z \text{ is unsatisfied at time } 0) = -2c(\theta) \tag{2.2}$$

and

$$\lim_{r\to\infty}\left(\frac{1}{r}\right)\log P^{r,\theta}(z \text{ is a 1-blockade at time } 0) = -c(2\theta-1) \tag{2.3}$$

if $\theta \le \frac{3}{4}$. Condition (2.3) fails when $\theta > \frac{3}{4}$ for then $(2\theta - 1) > \frac{1}{2}$.

We know that at θ_c blockades and unsatisfied points have the same probabilities on an exponential scale (this is the defining property of θ_c). Our first lemma gives us the stronger fact that their ratios are bounded away from 0 and ∞. If $\{u_n\}$ and $\{v_n\}$ are sequences, we write $u_n \sim v_n$ if there exist positive constants c and C such that

$$c \le \frac{u_n}{v_n} \le C$$

for all n.

Throughout this paper, phrases such as "0 is 1-unsatisfied" will mean 0 is 1-unsatisfied at time 0.

LEMMA 2.3. *Let* $a_r = P^{r,\theta_c}$(0 *is* 1 − *unsatisfied*) *and* $b_r = P^{r,\theta_c}$(0 *is a* 1 − *blockade*). *Then* $a_r \sim b_r$.

PROOF. It suffices to prove that if $\{X_i\}_{i=1}^{\infty}$ are i.i.d. with $P(X_i = 1) = \frac{1}{2} = P(X_i = 0)$, then for $\theta \in (\frac{1}{2}, 1)$,

$$P\left(\sum_{i=1}^{n} X_i \ge n\theta\right) \sim \frac{1}{n^{1/2}} e^{-c(\theta)n},$$

where $c(\theta) = \log 2 + \theta\log\theta + (1-\theta)\log(1-\theta)$. However, this result is a simple case of the main theorem in [2]. □

DEFINITION 2.4. An interval I of length at least $r+1$ is a 1-blob for η if $\eta \equiv 1$ on I. A 0-blob is defined similarly and a blob is either a 1-blob or a 0-blob.

DEFINITION 2.5. We say $x \in \mathbb{Z}$ is a 1-UB point (UB stands for unsatisfied and blobified) if x is 1-unsatisfied at time 0, $[x-r, x+r]$ is a 1-blob at time $2r$ (and therefore from that time onwards) and there are no 0-unsatisfied points in $[x-r^5, x+r^5]$ at time 0. A 0-UB point is defined similarly and a UB-point is either a 1-UB point or a 0-UB point.

LEMMA 2.6. *Let* $a_r = P^{r,\theta_c}(0$ *is a 1-UB point) and* $b_r = P^{r,\theta_c}(0$ *is a 1-blockade). Then* $a_r \sim b_r$.

PROOF. In view of Lemma 2.3, it suffices to show that

$$P^{r,\theta_c}\big(0 \text{ is a 1-UB point}|0 \text{ is 1-unsatisfied}\big) \geq \gamma$$

for some γ which is independent of r. Lemma 1.4 tells us that

$$P^{r,\theta_c}\big([-r, r] \text{ is a 1-blob by time } 2r|0 \text{ is 1-unsatisfied }\big) \geq \delta$$

for some δ which is independent of r. Lastly, letting L be the event that there are no 0-unsatisfied points in $[-r^5, r^5]$ at time 0, Harris' inequality (see page 129 in [3]) gives us that

$$P^{r,\theta_c}\big(L|0 \text{ is 1-unsatisfied}\big) \geq P^{r,\theta_c}(L),$$

which goes to 1 as $r \to \infty$. The result follows. □

We now know that UB points and blockades have relative densities which are bounded away from 0 and ∞ as $r \to \infty$. We want to know that they are "independently and evenly distributed" in space so that starting from the origin, the probability that one sees a blockade before seeing a blob or vice versa is bounded away from 0. The following proposition assures this. It takes a fair amount of work to prove this proposition and we therefore delay its proof until the end of this section.

PROPOSITION 2.7. *Let* $N_r = \lfloor r^{1/2} e^{2rc(\theta_c)} \rfloor$. *Let* $\widetilde{B}_k$ *be the event that there are no unsatisfied points in* $[-N_r, N_r]$, *there is both a 0-blockade and a 1-blockade in* $[-N_r, -k]$ *and both a 0-blockade and a 1-blockade in* $[k, N_r]$. *Let* $\widetilde{A}_k$ *be the event that there are no blockades in* $[-N_r, N_r]$, *there is either a 1-UB or a 0-UB point (or both) in* $[-N_r, -k]$ *and either a 1-UB or a 0-UB point (or both) in* $[k, N_r]$. *Then there is a universal constant* $\varepsilon > 0$ *independent of* r *and* k *such that*

$$\liminf_{r\to\infty} P^{r,\theta_c}(\widetilde{B}_k) \geq \varepsilon$$

and

$$\liminf_{r\to\infty} P^{r,\theta_c}(\widetilde{A}_k) \geq \varepsilon.$$

At this point, one can go directly to the proof of Theorem 1.3 (i). For Theorem 1.3 (ii), the following development is needed.

The point of blockades is that they stop blobs from expanding. Proposition 2.7 tells us that there is positive probability (not going to 0 as $r \to \infty$) that there will be blobs surrounding the origin and no blockades between them and the origin to stop them. A problem which one needs to deal with is that perhaps a blockade will form (which was not there originally) but will not blobify.

To show that this problem does not occur (with high probability), we need to show that unsatisfied sites which do not turn into blobs do not produce blockades (with high probability). Let

$$D_1 = \{\text{no point in } [-2r, -0.01r) \cup (0.01r, 2r] \text{ changes by time } r\},$$
$$D_2 = \{S_{x,x+r} < (1.97 - 2\theta)(r+1) \text{ for all } x \in [-r, 0.01r]\},$$
$$D_3 = \{\text{no } x \in [-4r^2, 4r^2] \text{ is 0-unsatisfied}\}$$

and $D = \cap_{i=1}^3 D_i$. To see the reason for interest in D, observe that (a) D_3 guarantees that no 1 in $[-2r, 2r]$ will flip to 0 by time $2r$ and (b) on $D_1 \cap D_2$ each interval $[x, x+r]$ with $x \in [-r, 0.01r]$ has fewer than $(1.99 - 2\theta)(r+1)$ 1's at time r. Statement (b) says that there is no 0-blockade, and (a) says that there is no 1-blockade (unless there was one at time 0).

Let

$$H = \{S_{-r,r} \geq (2r+1)\theta\} = \{0 \text{ is 1-unsatisfied}\}$$

and

$$E = \{[-r, r] \text{ is a 1-blob at time } 2r\}.$$

LEMMA 2.8. *Let $\theta \in (0.50, 0.65)$ be fixed. Then there are constants $0 < \lambda$, $C < \infty$ (depending on θ but independent of r) so that*

$$P^{r,\theta}(E \cup D | H) \geq 1 - Ce^{-\lambda r}.$$

REMARK. If $H \cap (E \cup D)$ occurs, then we say that the unsatisfied site at 0 is *well behaved*. In words, this says that at θ_c, points which are not well behaved have exponentially smaller probability than unsatisfied points (or equivalently blockades). We give no proof of this since this is Lemma 5.3. in [5].

PROOF OF THEOREM 1.3. Let N_r be as in Proposition 2.7.

(i) Fix k. By Proposition 2.7, it suffices to show that for any $r > k$, we have that $\widetilde{B}_k \subseteq B_k$. Let C be a 1-blockade in $[-N_r, -k]$ and let D be a 1-blockade in $[k, N_r]$. Since all 0's in $C \cup D$ remain 0 forever [in view of (2.1)] and no point in $[-N_r, N_r]$ is unsatisfied, it follows that no 0 in between C and D (and hence in $[-k, k]$) ever changes to a 1. Similarly, no 1 in $[-k, k]$ ever changes to a 0.

(ii) Fix k. Letting F_x be the event that $[x-k, x+k]$ fixate in the same state, we want to show that for large r, $P^{r,\theta_c}(F_0) \geq \gamma$ for some universal constant $\gamma > 0$ independent of r and k. It suffices to show that there exists an event G such that for large r, $P^{r,\theta_c}(G) \geq \varepsilon$ for some universal constant $\varepsilon > 0$ and if G occurs, then, under the $T^{r,\theta_c(2r+1)}$ dynamics, $(1/2r^3)\sum_{x=-r^3}^{r^3-1} I_{F_x} \geq 1 - 6/r$.

Let $G = \widetilde{A}_k \cap G_1 \cap G_2 \cap G_3$, where $\widetilde{A}_k$ is given in Proposition 2.7,

$$G_1 = \{\text{all points in } [-N_r, N_r] \text{ that are unsatisfied are well behaved}\},$$
$$G_2 = \{\text{there are not two unsatisfied sites } x \text{ and } y \text{ in } [-N_r, N_r] \text{ with } 2r+1 \le |x-y| \le 2r^3 + 4r\}$$

and

$$G_3 = \{\text{all } x \in [-r^4, r^4] \text{ are satisfied}\}.$$

Lemma 2.8 tells us that G_1 has probability approaching 1 as $r \to \infty$.

The event G_2 has probability approaching 1 since, when $|x-y| \ge (2r+1)$, the events that x and y are unsatisfied are independent. Hence the expected number of unsatisfied pairs satisfying the indicated inequalities is at most

$$(2r^3 + 4r)(2N_r + 1)\exp\bigl(-4c(\theta_c)r\bigr) \to 0 \quad \text{as } r \to \infty.$$

The event G_3 has probability approaching 1 since the expected number of unsatisfied sites in $[-r^4, r^4]$ is at most

$$(2r^4 + 1)\exp\bigl(-2c(\theta_c)r\bigr) \to 0 \quad \text{as } r \to \infty.$$

Finally, Proposition 2.7 together with the above implies that for large r, $P^{r,\theta_c}(G) \ge \varepsilon$ for some universal constant $\varepsilon > 0$. The last step is to show that if G occurs, then $(1/2r^3)\sum_{x=-r^3}^{r^3-1} I_{F_x} \ge 1 - 6/r$. We skip this, however, since this argument is essentially that which is carried out on pages 245 and 246 in [5]. □

The rest of this section is devoted to proving Proposition 2.7. The outline of this proof is as follows. We will show that if we divide space into blocks of size r^6, then the number of blocks (in a suitably sized region) which contain 1-UB points (or 1-unsatisfied points) and the number of blocks which contain 1-blockades become asymptotically independent Poisson random variables. If we can do this in such a way that the parameters for these Poisson random variables remain bounded away from 0 and ∞, Proposition 2.7 will follow. We first show that the relevant probabilities are of the same order of magnitude.

PROPOSITION 2.9. *Let a_r be the probability that some $x \in [1, r^6]$ is a 1-unsatisfied point, a'_r be the probability that some $x \in [1, r^6]$ is a 1-UB point and b_r be the probability that some $x \in [1, r^6]$ is a 1-blockade. Then $a_r \sim a'_r \sim b_r \sim r^{11/2}\exp[-2rc(\theta_c)]$.*

This result essentially says that the probability of the given event is of the same order of magnitude as the expected number of points in the interval with the property.

PROOF. We first prove that $a'_r \sim r^{11/2}\exp[-2rc(\theta_c)]$. The proof of Lemma 2.3 together with the statement of Lemma 2.6 tells us that

$$P^{r,\theta_c}(0 \text{ is a 1-UB point}) \sim r^{-1/2}\exp[-2rc(\theta_c)].$$

Since $a_r' \le r^6 P^{r,\theta_c}$(0 is a 1-UB point), we have that

$$a_r' \le Cr^{11/2}\exp[-2rc(\theta_c)]$$

for some constant C. We need, of course, a bound in the other direction. The first step for this is the following lemma which we prove later.

LEMMA 2.10. *Let Y_i be the indicator function for the event that i is a 1-UB point. Then for some universal constant T and large r,*

$$P^{r,\theta_c}\big(Y_i = 0\ \forall\ i \in [-r^6, -T] \cup [T, r^6] | Y_0 = 1\big) \ge \tfrac{1}{2}.$$

Consider the points $x_k = 1 + Tk$ for $k = 0, 1, \dots, r^6/T - 1$ and let A_k be the event that x_k is a 1-UB point. Then

$$\begin{aligned} a_r' &\ge P^{r,\theta_c}\left(\bigcup_{k=0}^{r^6/T-1} A_k\right) \\ &= P^{r,\theta_c}\big(A_0\big) + P^{r,\theta_c}\big(A_1 \cap A_0^c\big) + \cdots + P^{r,\theta_c}(A_{k+1} \cap A_k^c \cap \cdots \cap A_0^c) + \cdots \\ &= P^{r,\theta_c}(A_0) + P^{r,\theta_c}\big(A_0^c | A_1\big) P^{r,\theta_c}(A_1) + \cdots \\ &\quad + P^{r,\theta_c}\big(A_0^c \cap A_1^c \cdots \cap A_k^c | A_{k+1}\big) P^{r,\theta_c}(A_{k+1}) + \cdots. \end{aligned}$$

Using Lemma 2.10 together with the fact that $P^{r,\theta_c}(A_k) \ge cr^{-1/2}\exp[-2rc(\theta_c)]$ gives us that for large r each summand is $\ge (c/2)r^{-1/2}\exp[-2rc(\theta_c)]$. As the number of summands is r^6/T, we have

$$a_r' \ge \frac{r^6}{T}\frac{c}{2} r^{-1/2} e^{-2rc(\theta_c)}$$

and we have shown that $a_r' \sim r^{11/2}\exp[-2rc(\theta_c)]$.

Showing the same asymptotics for a_r is now trivial. The trivial argument that $a_r' \le Cr^{11/2}\exp[-2rc(\theta_c)]$ works also to give $a_r \le Cr^{11/2}\exp[-2rc(\theta_c)]$. On the other hand, $a_r \ge a_r'$ and so we are done. Finally, the proof that these asymptotics also hold for b_r can be carried out exactly as they were for a_r' and the proof is somewhat simpler. □

PROOF OF LEMMMA 2.10. Let Y_i' be the indicator function of the event that i is 1-unsatisfied. So $Y_i = 1$ implies that $Y_i' = 1$ while the converse is, of course, false. By the proof of Lemma 2.6 (which, of course, uses Lemma 1.4), $P^{r,\theta_c}(Y_i = 1 | Y_i' = 1) \ge \gamma$ for some constant $\gamma > 0$ independent of r. If we can show that for some constant T,

$$P^{r,\theta_c}\big(Y_i' = 0\ \forall\ i \in [-r^6, -T] \cup [T, r^6] | Y_0' = 1\big) \ge 1 - \frac{\gamma}{2} \tag{2.4}$$

for large r, then

$$\begin{aligned} &P^{r,\theta_c}\big(Y_i = 0\ \forall\ i \in [-r^6, -T] \cup [T, r^6] | Y_0 = 1\big) \\ &\quad \ge P^{r,\theta_c}\big(Y_i' = 0\ \forall\ i \in [-r^6, -T] \cup [T, r^6] | Y_0 = 1\big), \end{aligned}$$

which must be at least $\frac{1}{2}$ for large r since otherwise

$$\begin{aligned}
&P^{r,\theta_c}\big(Y_i' = 0\ \forall\ i \in [-r^6, -T] \cup [T, r^6] | Y_0' = 1\big) \\
&\quad \le P^{r,\theta_c}\big(Y_i' = 0\ \forall\ i \in [-r^6, -T] \cup [T, r^6] | Y_0 = 1\big) P^{r,\theta_c}\big(Y_0 = 1 | Y_0' = 1\big) \\
&\qquad + P^{r,\theta_c}\big(Y_i' = 0\ \forall\ i \in [-r^6, -T] \cup [T, r^6] | Y_0 = 0, Y_0' = 1\big) P^{r,\theta_c}\big(Y_0 = 0 | Y_0' = 1\big) \\
&\quad < \frac{\sigma_r}{2} + (1 - \sigma_r) = 1 - \frac{\sigma_r}{2} \le 1 - \frac{\gamma}{2},
\end{aligned}$$

where $\sigma_r = P^{r,\theta_c}(Y_0 = 1 | Y_0' = 1)$, contradicting (2.4). This would prove the result.

To prove (2.4), we first need the fact that

$$P^{r,\theta_c}\big(Y_i' = 0\ \forall\ i \in [-r^6, -0.01r] \cup [0.01r, r^6] | Y_0' = 1\big) \to 1 \quad \text{as } r \to \infty.$$

To do this, it suffices to show that there are positive constants α and c so that for all $i \in [-r^6, -0.01r] \cup [0.01r, r^6]$,

$$P^{r,\theta_c}\big(Y_i' = 1 | Y_0' = 1\big) \le ce^{-\alpha r}, \tag{2.5}$$

which is, of course, obvious for $|i| \ge 2r$. We leave this fairly elementary calculation to the reader. One can use the methods in the proofs of Lemmas 5.1 and 5.2 in [5] to do this.

The proof of (2.4) will be complete if we can show the more difficult fact that for some constant T and for large r,

$$P^{r,\theta_c}\big(Y_i' = 0\ \forall\ i \in [-0.01r, -T] \cup [T, 0.01r] | Y_0' = 1\big) > 1 - \frac{\gamma}{2},$$

which follows from

$$P^{r,\theta_c}\big(Y_i' = 0\ \forall\ i \in [T, 0.01r] | Y_0' = 1\big) > 1 - \frac{\gamma}{4} \tag{2.6}$$

for large r.

We first note that if $\{X_i\}_{i=1}^{\infty}$ are i.i.d. with $P(X_i = 1) = \frac{1}{2} = P(X_i = 0)$, then for $\theta \in (\frac{1}{2}, 1)$,

$$P\left(\sum_{i=1}^{n} X_i \ge n\theta\right) \sim P\left(\sum_{i=1}^{n} X_i = \lceil n\theta \rceil\right), \tag{2.7}$$

since both are $\sim (1/n^{1/2})\exp[-c(\theta)n]$ [where $c(\theta) = \log 2 + \theta \log \theta + (1-\theta)\log(1-\theta)$], the former by the result in [2] mentioned in the proof of Lemma 2.3 and the latter by Stirling's formula. In proving (2.7), one can actually show that there exists an integer m independent of r so that

$$P^{r,\theta_c}\left(\sum_{i=-r}^{r} X_i < (2r+1)\theta_c + m \,\middle|\, \sum_{i=-r}^{r} X_i \ge (2r+1)\theta_c\right) \ge 1 - \frac{\gamma}{8}.$$

Next, an easy calculation shows that conditioned on $\sum_{i=-r}^{r} X_i \geq (2r+1)\theta_c$, $\{X_i\}_{i=-r}^{-r+0.01r}$ dominates $\{W_i\}_{i=-r}^{-r+0.01r}$, where the W_i's are i.i.d. with $P\{W_i = 1\} = \theta_c - 0.01 = 1 - P\{W_i = 0\}$. This means that $\{X_i\}_{i=-r}^{-r+0.01r}$ conditioned on $\sum_{i=-r}^{r} X_i \geq (2r+1)\theta_c$ and $\{W_i\}_{i=-r}^{-r+0.01r}$ can be defined on the same probability space so that $X_i \geq W_i$ for all i. This calculation is done by simply noticing that when conditioning on $\sum_{i=-r}^{r} X_i \geq (2r+1)\theta_c$, for any $j \in [-r, -r+0.01r]$, the probability that $X_j = 1$, given any values for $X_{-r}, \dots, X_{j-1}$, is at least $\theta_c - 0.01$. Next let T be such that a random walk with step size distribution

$$\frac{\theta_c - 0.01}{2}\delta_{-1} + \frac{1}{2}\delta_0 + \frac{1-(\theta_c - 0.01)}{2}\delta_1$$

starting from 0 will with probability $> 1 - \gamma/8$ stay at or below $-m$ from time T onwards.

Since conditioning on $\sum_{i=-r}^{r} X_i \geq (2r+1)\theta_c$ does not affect the distribution of $\{X_{r+1}, \dots, X_{r+0.01r}\}$, the way we have chosen T together with the domination above guarantees us that, conditioned on $\sum_{i=-r}^{r} X_i \geq (2r+1)\theta_c$, we have, with probability $> 1 - \gamma/8$,

$$S_{-r+\ell, r+\ell} \leq S_{-r,r} - m$$

for $\ell = T, T+1, \dots, 0.01r$. In view of the way m was chosen, we have that conditioned on $\sum_{i=-r}^{r} X_i \geq (2r+1)\theta_c$, we have that with probability $> 1 - \gamma/4$,

$$S_{-r+\ell, r+\ell} < (2r+1)\theta_c$$

for $\ell = T, T+1, \dots, 0.01r$, which proves (2.6). □

To show that certain objects we are looking at asymptotically become independent Poisson random variables, we will use the Chen–Stein method as described in the paper by Arratia, Goldstein and Gordon [1].

PROOF OF PROPOSITION 2.7. We first show that $\liminf_{r\to\infty} P^{r,\theta_c}(\widetilde{B}_k) \geq \varepsilon$ for some universal constant $\varepsilon > 0$. After this we mention the easy modification needed to prove that $\liminf_{r\to\infty} P^{r,\theta_c}(\widetilde{A}_k) \geq \varepsilon$ for some universal constant $\varepsilon > 0$.

For all $i \in \mathbb{Z}$, let J_i^r be the interval $[r^6(i-1)+1, r^6 i]$ of length r^6. Introduce the following random variables. Let $H_{i,1}^r$ ($H_{i,2}^r$) be the event that some point in J_i^r is a 1-unsatisfied (0-unsatisfied) point. Let $H_{i,3}^r$ ($H_{i,4}^r$) be the event that some point in J_i^r is a 0-blockade (1-blockade). Let $Y_{i,j}^r$ be the indicator function for $H_{i,j}^r$.

Let $N_r' = \lfloor r^{-11/2}\exp[2rc(\theta_c)]\rfloor$ For $j = 1, 2, 3, 4$, let $S_j^r = \sum_{i=1}^{N_r'} Y_{i,j}^r$. For $j = 5, 6, 7, 8$, let $S_j^r = \sum_{i=-N_r'+1}^{0} Y_{i,j-4}^r$. The Chen–Stein method will allow us to obtain the fact that there are two universal constants c and C such that

$$(S_1^r, S_2^r, S_3^r, S_4^r, S_5^r, S_6^r, S_7^r, S_8^r) \tag{2.8}$$

are within d_r in total variation norm of eight independent Poisson random variables with means between c and C with $d_r \to 0$ as $r \to \infty$. This implies that with probability bounded (in r) away from 0,

$$S_1^r = S_2^r = S_5^r = S_6^r = 0 \quad \text{and} \quad S_3^r = S_4^r = S_7^r = S_8^r = 1,$$

which corresponds to the event that there are no unsatisfied points in $[-N_r, N_r]$, there are both a 0-blockade and a 1-blockade in $[-N_r, 0]$ and both a 0-blockade and a 1-blockade in $[1, N_r]$. Finally, noting that for fixed k the probability that there is an unsatisfied point or a blockade in $[-k, k]$ goes exponentially to 0 as $r \to \infty$, the above implies that $\liminf_{r\to\infty} P^{r,\theta_c}(\tilde{B}_k) \geq \varepsilon$ for some universal constant $\varepsilon > 0$. Actually, since $r^{-11/2}\exp[2rc(\theta_c)]$ need not be an integer, $N_r' r^6$ might be as small as $N_r - r^6$ in which case we only know that there are no unsatisfied points in $[-N_r + r^6, N_r - r^6]$. However, since unsatisfied points (as well as blockades) are contained in an interval of length r^6 with exponentially small probability, there is no problem and we can therefore always ignore this correction.

We now state Theorem 2 contained in [1], which we shall need. We actually state a weaker version which will suffice for our purposes.

THEOREM 2.11. *Let I be a finite index set and assume that for each $\alpha \in I$, V_α is a Bernoulli 0–1 random variable equal to 1 with probability p_α. Assume that for each $\alpha \in I$, we have a subset $B_\alpha \subseteq I$ such that for all $\alpha \in I$, V_α is independent of the σ-algebra generated by $\{V_u\}_{u \notin B_\alpha}$. Let $z_1 = \sum_{\alpha\in I}\sum_{u\in B_\alpha} p_\alpha p_u$ and let $z_2 = \sum_{\alpha\in I}\sum_{\alpha \neq u \in B_\alpha} E[V_\alpha V_u]$. Next let $\mathbb{W} = \{W_\alpha\}_{\alpha\in I}$ be independent Poisson random variables with W_α having mean p_α. Letting $\mathbb{V} = \{V_\alpha\}_{\alpha\in I}$ and $\|\mathbb{L}(\mathbb{V}) - \mathbb{L}(\mathbb{W})\|$ be the total variation distance between these two random vectors, we have that*

$$\|\mathbb{L}(\mathbb{V}) - \mathbb{L}(\mathbb{W})\| \leq 4z_1 + 4z_2.$$

To apply this result, we take our index set I_r to be $[-N_r', N_r'] \times \{1, 2, 3, 4\}$ and for $(i, j) \in I_r$ we have the Bernoulli random variable $Y_{i,j}^r$. This defines for us a random vector $\mathbb{Y}^r = \{Y_\alpha^r\}_{\alpha\in[-N_r', N_r']\times\{1,2,3,4\}}$. Let a_r be the expectation of $Y_{i,1}^r$ (or equivalently of $Y_{i,2}^r$) and let b_r be the expectation of $Y_{i,3}^r$ (or equivalently of $Y_{i,4}^r$). These are exactly as they were defined in Proposition 2.9, where they were shown to be asymptotic to $r^{11/2}\exp[-2rc(\theta_c)]$. We next let $\mathbb{W}^r = \{W_\alpha^r\}_{\alpha\in[-N_r', N_r']\times\{1,2,3,4\}}$ be independent Poisson random variables with $W_{i,j}^r$ having mean a_r for $j = 1, 2$ and with $W_{i,j}^r$ having mean b_r for $j = 3, 4$. Letting $F_i = \{(i,1), (i,2), (i,3), (i,4)\}$, we let $B_{(i,j)}^r$ be $I_r \cap (F_{i-1} \cup F_i \cup F_{i+1})$.

We want to apply Theorem 2.11 to $\mathbb{Y}^r$ and $\mathbb{W}^r$. The independence assumption is obvious. We also want to show that $z_1^r (= \sum_{\alpha\in I_r}\sum_{u\in B_\alpha^r} E[Y_\alpha^r]E[Y_u^r])$ and $z_2^r (= \sum_{\alpha\in I_r}\sum_{\alpha\neq u\in B_\alpha^r} E[Y_\alpha^r Y_u^r]) \to 0$ as $r \to \infty$. The fact that $z_1^r \to 0$ is trivial. For z_2^r, it suffices to show that $E[Y_\alpha^r Y_u^r]$ is of smaller order than a_r.

There are several cases to check. Let $\alpha = (i, j)$ and $u = (i', j')$. If $\{j, j'\} \in \{\{1,2\}, \{1,4\}, \{3,2\}, \{3,4\}\}$, then this follows immediately from Harris' inequality since the two events are negatively correlated.

Now assume $\{j, j'\} = \{1, 3\}$, the case $\{2, 4\}$ being similar. In this case, we show that

$$E[Y_\alpha^r Y_u^r] \leq r^{12}\exp[-2r(c(\theta_c) + \sigma)]$$

for some constant σ, which clearly implies that $E[Y^r_\alpha Y^r_u]$ is of smaller order than a_r. To see this, it suffices to show that for any x and y in $\mathbb{Z}$,

$$P^{r,\theta_c}(x \text{ is 1-unsatisfied} | y \text{ is a 0-blockade}) \leq e^{-r\sigma} \tag{2.9}$$

for some constant σ which does not depend on x and y (and, of course, is independent of r also).

It is clear that this conditional probability is largest when the blockade $[y, y+r] \subseteq [x-r, x+r]$. In this case, large deviation theory tells us that the density of 1's in $[y, y+r]$ is less then or equal to $2 - 2\theta_c + 0.01$ with exponentially high probability. Similarly, the density of 1's in $[x-r, x+r]\backslash[y, y+r]$ is at most 0.51 with exponentially high probability. It follows that the density of 1's in $[x-r, x+r]$ is less then or equal to $(2-2\theta_c+0.01)/2+0.51/2$ with exponentially high probability. Since this quantity is less than θ_c, (2.9) follows.

The last four cases to be taken care of are $\{(i,1), (i+1,1)\}$, $\{(i,2), (i+1,2)\}$, $\{(i,3), (i+1,3)\}$ and $\{(i,4), (i+1,4)\}$. Without loss of generality, we consider only $\{(i,1), (i+1,1)\}$. Let U_1 be the event that there exist $x, y \in J^r_i \cup J^r_{i+1}$ such that $|x-y| \geq r$ and both x and y are 1-unsatisfied. Let U_2 be the event that there exists x within the r rightmost points in J^r_i with x being 1-unsatisfied. Clearly, $E[Y^r_\alpha, Y^r_u] \leq P^{r,\theta_c}(U_1) + P^{r,\theta_c}(U_2)$. By an earlier argument [(2.5)], $P^{r,\theta_c}(U_1)$ has exponentially smaller probability than a_r. Next, while U_2 does not have exponentially smaller probability than a_r, we still have that $P^{r,\theta_c}(U_2) \leq rCr^{-1/2} \exp[-2rc(\theta_c)]$, which is of smaller order than a_r, as desired.

Theorem 2.11 now tells us that for large r, $\mathbb{Y}^r$ and $\mathbb{W}^r$ are very close in total variation distance which implies that

$$(S^r_1, S^r_2, S^r_3, S^r_4, S^r_5, S^r_6, S^r_7, S^r_8)$$

and

$$(T^r_1, T^r_2, T^r_3, T^r_4, T^r_5, T^r_6, T^r_7, T^r_8)$$

are very close in total variation distance, where

$$T^r_j = \sum_{i=1}^{N'_r} W^r_{i,j}, \qquad j = 1, 2, 3, 4,$$

and

$$T^r_j = \sum_{i=-N'_r+1}^{0} W^r_{i,j-4}, \qquad j = 5, 6, 7, 8.$$

By Proposition 2.9, there are positive constants c and C such that the eight Poisson random variables

$$(T^r_1, T^r_2, T^r_3, T^r_4, T^r_5, T^r_6, T^r_7, T^r_8)$$

have means between c and C. This proves the statement involving (2.8), finishing the proof that $\liminf_{r\to\infty} P^{r,\theta_c}(B_k) \ge \varepsilon$ for some universal constant $\varepsilon > 0$.

In order to also show that $\liminf_{r\to\infty} P^{r,\theta_c}(\widetilde{A}_k) \ge \varepsilon$ for some universal constant $\varepsilon > 0$, we proceed exactly as above, the only modification being that $H^r_{i,1}(H^r_{i,2})$ is now the event that some point in J^r_i is a 1-UB (0-UB) point. The rest of the proof can then be carried out in the same way. □

3. Proof of Lemma 1.4, blob formation. In this section we prove Lemma 1.4 using a branching process argument. The argument is somewhat technical and so we first describe in words the idea involved.

Let $I_1 = [-r, -r + 0.01r - 1]$, $I_2 = [0, 0.01r]$, $I_3 = [r + 1, r + 0.01r]$ and $I = I_1 \cup I_2 \cup I_3$. First, one has that the distribution on I conditioned on 0 being 1-unsatisfied is more or less i.i.d. with density θ in $I_1 \cup I_2$ and density $\frac{1}{2}$ in I_3 (this is made precise by Lemma 3.1 below.) Assuming that this is the exact conditional distribution, then as we move from the origin one step at a time to the right until reaching $0.01r$, the density profile changes randomly according to a random walk with step size distribution

$$\frac{\theta}{2}\delta_{-1} + \frac{1}{2}\delta_0 + \frac{1-\theta}{2}\delta_1,$$

since each time we move over one lattice point, the density changes by adding some new point from I_3 and deleting some point from I_1. The number of steps to the right that we need to take until the density is one less than it was at the origin is now simply the amount of time it takes the above random walk to reach -1 starting from 0, which has expected value (by Wald's theorem) $1/(\theta - \frac{1}{2})$. All of this depends only on the configuration in $I_1 \cup I_3$. If T denotes the position at which the density falls one below that at the origin, then after one iteration all the 0's in $[0, T-1]$ become 1's and so the expected number of 0's which turn to 1's after one iteration is $(1-\theta)/(\theta - \frac{1}{2})$ which is larger than 1 since $\theta < 0.75$. If we look at the number of 0's which turn to 1's at each iteration, this description looks like a supercritical branching process which then might, as we iterate, sustain itself until all points in $[0, 0.01r]$ are 1. We now turn the above ideas into a proof. Before beginning the proof of Lemma 1.4, we need the following lemma. A related but much simpler fact was previously used in the proof of Lemma 2.10.

LEMMA 3.1. *Let $\{X_i\}$ be i.i.d. with $P(X_i = 1) = \frac{1}{2} = P(X_i = 0)$ and let $\theta \in (0.5, 0.68)$. Then for large n the conditional distribution of $\{X_i\}_{i=0}^{0.01n}$ given $\{\sum_{i=0}^{n-1} X_i \ge n\theta\}$ is stochastically dominated by $\{W_i\}_{i=0}^{0.01n}$, where the W_i's are i.i.d with $P\{W_i = 1\} = \theta + 0.07 = 1 - P\{W_i = 0\}$. [This means that we can define $\{X_i\}_{i=0}^{0.01n}$ (conditioned on $\{\sum_{i=0}^{n-1} X_i \ge n\theta\}$) and $\{W_i\}_{i=0}^{0.01n}$ on the same probability space as $X_i \le W_i$ for all i.]*

PROOF. Recall that $S_{k,\ell}$ denotes $\sum_{i=k}^{\ell} X_i$. It suffices to show that for large

n, we have that for all $0 \le j \le 0.01n$ and all $(i_0, i_1, \dots, i_{j-1}) \in \{0, 1\}^j$,

$$P\{X_j = 1 | S_{0,n-1} \ge n\theta,\ X_0 = i_0,\ X_1 = i_1, \dots, X_{j-1} = i_{j-1}\} < \theta + 0.07.$$

We first choose $\alpha > 0$ such that $c(1.03\theta) + \alpha < c(\theta + 0.03)$, where $c(a) = \log 2 + a \log a + (1-a)\log(1-a)$ as in the Introduction. Next let C be such that

$$P\big(S_{0,n-1} \ge n(1.03\theta)\big) \ge (1/C)\exp[-n(c(1.03\theta) + \alpha)] \quad \text{for all } n.$$

We then have

$$P\{X_j = 1 | S_{0,n-1} \ge n\theta,\ X_0 = i_0,\ X_1 = i_1, \dots, X_{j-1} = i_{j-1}\}$$

$$= \frac{\left(\frac{1}{2}\right)^{j+1} P\left(S_{j+1,n-1} \ge n\theta - 1 - \sum_{s=0}^{j-1} i_s\right)}{\left(\frac{1}{2}\right)^{j} P\left(S_{j,n-1} \ge n\theta - \sum_{s=0}^{j-1} i_s\right)}$$

$$\le \frac{\frac{1}{2}\sum_{\ell = n\theta - 1 - \Sigma_{s=0}^{j-1} i_s}^{n(\theta+0.03)-1} \binom{n-1-j}{\ell}\left(\frac{1}{2}\right)^{n-1-j}}{\sum_{\ell = n\theta - \Sigma_{s=0}^{j-1} i_s}^{n(\theta+0.03)} \binom{n-j}{\ell}\left(\frac{1}{2}\right)^{n-j}}$$

$$+ \frac{\frac{1}{2}P\big(S_{j+1,n-1} \ge (n-1)(\theta + 0.03)\big)}{P\big(S_{S_{j,n-1}} \ge n\theta\big)}.$$

The first term equals

$$\frac{\frac{1}{n-j}\sum_{\ell = n\theta - \Sigma_{s=0}^{j-1} i_s}^{n(\theta+0.03)} \frac{\ell}{\ell!(n-j-\ell)!}}{\sum_{\ell = n\theta - \Sigma_{s=0}^{j-1} i_s}^{n(\theta+0.03)} \frac{1}{\ell!(n-j-\ell)!}} \le \frac{n(\theta + 0.03)}{n-j}.$$

Since $j \le 0.01n$, $n/(n-j) \le 1.03$ and so this last term is less than $\theta + 0.065$.

As for the second term, the numerator is bounded by $\frac{1}{2}\exp[-(n-1-j)c(\theta + 0.03)]$. For the denominator, using the fact that $n/(n-j) \le 1.03$, we have

$$P\big(S_{j,n-1} \ge n\theta\big) \ge P\big(S_{j,n-1} \ge (n-j)1.03\theta\big) \ge \big(1/C\big)\exp[-(n-j)(c(1.03\theta) + \alpha)].$$

The resulting fraction is then at most

$$\tfrac{1}{2}C\exp[c(\theta + 0.03)]\exp[-(n-j)(c(\theta + 0.03) - c(1.03\theta) - \alpha)],$$

which is less than 0.005 for large n since $j \le 0.01n$ and by the definition of α. Together with the above, this gives us the desired bound of $\theta + 0.07$ for large n. □

PROOF OF LEMMA 1.4. Let $\theta \in (0.5, 0.68)$ be fixed. It clearly suffices to prove that there is a δ such that

$$P^{r,\theta}\big([-r, r] \text{ is a 1-blob by time } 2r | 0 \text{ is 1-unsatisfied}\big) \ge \delta$$

for all large r. To analyze the blob formation process, we first need to introduce a branching process as follows. Let $\theta' = \theta + 0.07$. We choose the offspring distribution for our branching process to be

$$\theta'\delta_0 + (1 - \theta')F,$$

where F is the distribution for the time it takes a random walk with step size distribution

$$\frac{\theta'}{2}\delta_{-1} + \frac{1}{2}\delta_0 + \frac{1 - \theta'}{2}\delta_1$$

to reach -1 starting from 0. Since $\theta' > \frac{1}{2}$, the mean waiting time for the above random walk is finite and, moreover, Wald's theorem tells us that the mean of F is $1/(\theta' - \frac{1}{2})$ and hence the mean offspring size is $(1-\theta')/(\theta' - \frac{1}{2})$. Since $\theta' < 0.75$, the above branching process is supercritical, which means that starting from one individual, the branching process survives forever with positive probability. We let $\{Y_n\}_{n=0}^{\infty}$ denote this branching process where the initial size Y_0 also has the distribution $\theta'\delta_0 + (1 - \theta')F$. (So Y_n is the number of individuals alive at time n.) Let $\alpha > 0$ denote the probability that a single particle survives and so the probability that $\{Y_n\}_{n=0}^{\infty}$ survives is at least $(1 - \theta')\alpha$.

For each integer $r > 1$, we introduce a modified branching process $\{\widetilde{Y}_n^r\}_{n=0}^{\infty}$ as follows. These will simply be functions of the branching process $\{Y_n\}$. Let $T = \inf\{n: \sum_{\ell=0}^{n} Y_\ell \ge 0.01r\}$. If $T = \infty$, let $\widetilde{Y}_n^r = Y_n$ for all n. If $T < \infty$, we let $\widetilde{Y}_n^r = Y_n$ for $n < T$, $\widetilde{Y}_T^r$ chosen such that $\sum_{\ell=0}^{T} \widetilde{Y}_\ell^r = 0.01r$ and $\widetilde{Y}_n^r = 0$ for $n > T$. Note that whether $T < \infty$ or $= \infty$, $\widetilde{Y}_n^r = 0$ for large n and $\sum_{\ell=0}^{\infty} \widetilde{Y}_\ell^r$ exists and is less than or equal to $0.01r$. Note also that survival of $\{Y_n\}$ implies that for any r, $\sum_{\ell=0}^{\infty} \widetilde{Y}_\ell^r = .01r$.

The next lemma, whose proof is given later, explains the connection between blob formation and the modified branching processes introduced above.

LEMMA 3.2. *For all integers $r > 1$, there exists a stochastic process $\{Z_n^r\}_{n=0}^{\infty}$ which is measurable with respect to $\{X_i\}_{i \in [-r, -r+0.01r-1] \cup [0, 0.01r] \cup [r+1, r+0.01r]}$ with the following properties*:

(a) *If $\{X_i\}_{i \in [-r, -r+0.01r-1] \cup [0, 0.01r] \cup [r+1, r+0.01r]}$ is chosen according to a product measure with density θ' in the first two intervals and density $\frac{1}{2}$ in the last interval, then $\{Z_n^r\}_{n=0}^{\infty} = \{\widetilde{Y}_n^r\}_{n=0}^{\infty}$ in distribution where $\{\widetilde{Y}_n^r\}_{n=0}^{\infty}$ is the rth modified branching process defined above.*

(b) *If* $S_{-r,r} \geq \theta(2r+1)$, $\Sigma_{\ell=0}^{\infty} Z_\ell^r = 0.01r$ *and there are no* 0*-unsatisfied* z *in* $[-r^3, r^3]$, *then under the* $T^{r,\theta(2r+1)}$ *dynamics, we will have that at time* $0.01r$, $[0, 0.01r]$ *will be all* 1*'s.*

(c) *If* η *and* η' *are configurations on* $[-r, -r+.01r-1] \cup [0, .01r] \cup [r+1, r+0.01r]$ *with* $\eta \geq \eta'$ *and* $\eta = \eta'$, *on* $[r+1, r+0.01r]$, *then* $\Sigma_{\ell=0}^{\infty} Z_\ell^r(\eta) = 0.01r$ *implies that* $\Sigma_{\ell=0}^{\infty} Z_\ell^r(\eta') = 0.01r$.

Note that parts (b) and (c) are nonprobabilistic statements. Let, as above, $I_1 = [-r, -r+0.01r-1]$, $I_2 = [0, 0.01r]$, $I_3 = [r+1, r+0.01r]$ and $I = I_1 \cup I_2 \cup I_3$.

Let $\{{}^{\theta'}Z_n^r\}_{n=0}^{\infty}$ denote the above process $\{Z_n^r\}_{n=0}^{\infty}$ when $\{X_i\}_{i\in I}$ is chosen according to a product measure with density θ' in I_1 and I_2 and with density $\frac{1}{2}$ in I_3. By (a) above together with the fact that survival of $\{Y_n\}$ implies that $\Sigma_{\ell=0}^{\infty} \widetilde{Y}_\ell^r = 0.01r$ for any r, we have

$$P\left\{\sum_{\ell=0}^{\infty} {}^{\theta'}Z_\ell^r = 0.01r\right\} = P\left\{\sum_{n=0}^{\infty} \widetilde{Y}_n^r = 0.01r\right\} \geq (1-\theta')\alpha. \tag{3.1}$$

Let $\{W_i\}_{i\in I}$ be i.i.d. with density θ' on $I_1 \cup I_2$ and density $\frac{1}{2}$ on I_3. Next, Lemma 3.1 tells us that for large r, $\{X_i\}_{i\in I_1\cup I_2}$ conditioned on $\{S_{-r,r} \geq \theta(2r+1)\}$ (i.e., conditioned on 0 is 1-unsatisfied) is stochastically dominated by $\{W_i\}_{i\in I_1\cup I_2}$. Using independence, we have that for large r, $\{X_i\}_{i\in I}$ conditioned on $\{S_{-r,r} \geq \theta(2r+1)\}$ can be coupled with $\{W_i\}_{i\in I}$ such that $X_i \leq W_i$ for all i and $X_i = W_i$ for $i \in I_3$. Combining this with (3.1) and Lemma 3.2(a) gives us

$$P\left\{\sum_{\ell=0}^{\infty} Z_\ell^r = 0.01r \,\middle|\, S_{-r,r} \geq \theta(2r+1)\right\} \geq (1-\theta')\alpha \tag{3.2}$$

for large r. We also choose r sufficiently large so that

$$\begin{aligned} &P^{r,\theta}\{\text{no 0-unsatisfied } z \text{ in } [-r^3, r^3] \text{ at time } 0 | S_{-r,r} \geq \theta(2r+1)\} \\ &\quad \geq 1 - \frac{(1-\theta')\alpha}{2}, \end{aligned} \tag{3.3}$$

which follows from a simple application of Harris' inequality together with the fact that unsatisfied points have exponentially small probability. Combining (3.2) and (3.3), it follows from Lemma 3.2 that for large r,

$$P^{r,\theta}\{\eta \equiv 1 \text{ on } [0, 0.01r] \text{ at time } 0.01r \mid S_{-r,r} \geq \theta(2r+1)\} \geq -\frac{(1-\theta')\alpha}{2}. \tag{3.4}$$

At this point, we can follow the proof of Lemma 5.2 in [5] and obtain

$$\begin{aligned} &P^{r,\theta}\{[-r, r] \text{ is a 1-blob at time } 2r | S_{-r,r} \geq \theta(2r+1)\} \\ &\quad \geq \frac{(1-\theta')\alpha}{4} \end{aligned} \tag{3.5}$$

for large r, as desired. □

PROOF OF LEMMA 3.2. Fix $r > 1$. We first define the auxiliary random variables $G_i = X_{r+i} - X_{-r+i-1}$ for $i = 1, \dots, 0.01r$. We now define the process $\{Z_n^r\}$ inductively.

Let $U_0 = I_{\{X_0=0\}}$. Let $a_0 = \inf\{0 \le m \le 0.01r\colon \sum_{j=1}^m G_j = -U_0\}$. Let $Z_0^r = a_0$.

Let $U_1 = |\{i \in [1, a_0]\colon X_i = 0\}|$. Let $a_1 = \inf\{a_0 \le m \le 0.01r\colon \sum_{j=a_0+1}^m G_j = -U_1\}$. Let $Z_1^r = a_1 - a_0$.

Let $U_2 = |\{i \in [a_0+1, a_1]\colon X_i = 0\}|$. Let $a_2 = \inf\{a_1 \le m \le 0.01r\colon \sum_{j=a_1+1}^m G_j = -U_2\}$. Let $Z_2^r = a_2 - a_1$.

We continue by induction with two requirements to satisfy. First, if some U_i is 0, then we take Z_j^r to be 0 for $j \ge i$. (This also follows from the above construction as long as the empty sum is defined to be 0.) Second, note that a_i has not been defined if $U_i > 0$ and $\sum_{j=a_{i-1}+1}^m G_j > -U_i$ for all $m \in [a_{i-1}, 0.01r]$. In this case, we define a_i to be $0.01r$, Z_i^r to be $a_i - a_{i-1}$ and Z_j^r to be 0 for $j > i$. Note we will then have $\sum_{\ell=0}^i Z_l^r = 0.01r$. This finishes the inductive construction of the process $\{Z_n^r\}$. (The reader should note the similarity between this last part of the construction and the way we defined the modified branching processes.)

The fact that these processes satisfy property (a) is straightforward and left to the reader.

For property (b), because there are no 0-unsatisfied points in $[-r^3, r^3]$, we need not worry about a 1 switching to a 0 for the first r units of time. Let $a_{k+1}^* = \inf\{a_k \le m \le 0.01r\colon \sum_{j=a_k+1}^m G_j = -1\}$. Clearly, if $a_{k+1} > a_k$ (or equivalently $U_{k+1} \ne 0$), then

$$a_k < a_{k+1}^* \le a_{k+1}.$$

It is easy to see that after one iteration, $[0, a_0 - 1]$ will all be 1's, after two iterations, $[0, a_1^* - 1]$ will all be 1's, after three iterations, $[0, a_2^* - 1]$ will all be 1's, and so forth. Hence after k iterations, $[0, a_{k-2}]$ will be all 1's, which proves (b).

Property (c) follows from the following two observations. The first observation is that if $\eta \ge \eta'$ and $\eta = \eta'$ on $[r+1, r+0.01r]$, then $G_j(\eta') \ge G_j(\eta)$ for all j. The second is that

$$\sum_{\ell=0}^{\infty} Z_\ell^r$$

can be alternatively expressed as

$$\inf\left\{0 \le m \le 0.01r\colon \sum_{j=1}^m G_j = -\left|j \in [0, m]\colon X_j = 0\right|\right\}.$$

where the latter is taken to be $0.01r$ if equality holds for no m. □

Acknowledgments. The author would like to thank Richard Durrett for providing an important idea for this paper and the referee for a very careful reading and for providing a number of helpful suggestions.

REFERENCES

[1] ARRATIA, R., GOLDSTEIN, L. and GORDON, L. (1989). Two moments suffice for Poisson approximations: the Chen–Stein method. *Ann. Probab.* **17** 9–25.

[2] BAHADUR, R. and RANGA RAO, R. (1960). On deviations of the sample mean. *Ann. Math. Statist.* **31** 1015–1027.

[3] DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation.* Wadsworth, Pacific Grove, CA.

[4] DURRETT, R. (1990). *Probability: Theory and Examples.* Wadsworth, Pacific Grove, CA.

[5] DURRETT, R. and STEIF, J. (1993). Fixation results for threshold voter models. *Ann. Probab.* **21** 232–247.

[6] FISCH, R. and GRAVNER, J. The one dimensional Greenberg Hastings model. Unpublished manuscript.

[7] FISCH, R., GRAVNER, J. and GRIFFEATH, D. (1991). Threshold-range scaling for excitable cellular automata. *Statist. Comput.* **1** 23–39.

DEPARTMENT OF MATHEMATICS
CHALMERS UNIVERSITY OF TECHNOLOGY
S-412 96 GOTHENBURG
SWEDEN