The Annals of Applied Probability
2011, Vol. 21, No. 1, 374–395
DOI: 10.1214/10-AAP702

SHARPNESS OF THE PERCOLATION TRANSITION IN THE TWO-DIMENSIONAL CONTACT PROCESS

BY J. VAN DEN BERG[1]

CWI and VU University Amsterdam

For ordinary (independent) percolation on a large class of lattices it is well known that below the critical percolation parameter p_c the cluster size distribution has exponential decay and that power-law behavior of this distribution can only occur at p_c. This behavior is often called "sharpness of the percolation transition."

For theoretical reasons, as well as motivated by applied research, there is an increasing interest in percolation models with (weak) dependencies. For instance, biologists and agricultural researchers have used (stationary distributions of) certain two-dimensional contact-like processes to model vegetation patterns in an arid landscape (see [20]). In that context occupied clusters are interpreted as patches of vegetation. For some of these models it is reported in [20] that computer simulations indicate power-law behavior in some interval of positive length of a model parameter. This would mean that in these models the percolation transition is not sharp.

This motivated us to investigate similar questions for the ordinary ("basic") $2D$ contact process with parameter λ. We show, using techniques from Bollobás and Riordan [8, 11], that for the upper invariant measure $\bar{\nu}_\lambda$ of this process the percolation transition is sharp. If λ is such that ($\bar{\nu}_\lambda$-a.s.) there are no infinite clusters, then for all parameter values below λ the cluster-size distribution has exponential decay.

1. Introduction and statement of the main result. The contact process was introduced as a stochastic model for the spread of an infection in a population with a geometric structure, usually represented by the d-dimensional cubic lattice. Each vertex x of this lattice represents an individual whose state, infected (1) or healthy (0), at time t is denoted by $\sigma_x(t)$. The dynamic in this model is as follows: A vertex in state 0 goes to state 1 ("becomes infected") at a rate equal to λ times the number of neighbors of that vertex that are in state 1. A vertex in state 1 goes to state 0 ("recovers") at rate 1. Here λ is the parameter of the model called the infection rate. In this paper we restrict to the case $d = 2$. Depending on the applications one has in mind the terms "infected" and "healthy" are sometimes replaced by "occupied" and "vacant," respectively. In the remainder of this paper we will use this latter terminology.

Received July 2009; revised March 2010.

[1]Supported in part by the Dutch BSIK/BRICKS project.

AMS 2000 subject classifications. Primary 60K35; secondary 92D40, 92D30, 82B43.

Key words and phrases. Percolation, contact process, sharp transition, approximate zero-one law, sharp thresholds.

The configuration at time t is denoted by $\sigma(t) := (\sigma_x(t), x \in \mathbb{Z}^2)$. Let μ_t denote the distribution of σ_t when we start at time 0 with all vertices occupied. We will use the notation $|V|$ for the cardinality of a set V.

It is well known (from a standard coupling argument) that μ_t is stochastically dominated by μ_s if $s \le t$. Hence μ_t converges weakly to a probability measure denoted by $\bar{\nu}$ ($=\bar{\nu}_\lambda$) as $t \to \infty$. This measure $\bar{\nu}$ is called the upper invariant measure. It is well known (again by standard coupling arguments) that $\bar{\nu}_{\lambda_2}$ stochastically dominates $\bar{\nu}_{\lambda_1}$ if $\lambda_2 > \lambda_1$. Realizations are typically denoted by $\sigma = (\sigma_x, x \in \mathbb{Z}^2)$. The occupied cluster of a vertex x (i.e., the maximal connected component which contains x and of which every vertex is occupied) is denoted by $\mathcal{C}_x$. (If x is the origin 0, we often omit the subscript.)

In this paper we study the sizes of occupied clusters under the measure $\bar{\nu}$. Motivation comes from work by Liggett and Steif [25] who showed that for λ sufficiently large percolation occurs [i.e., $\bar{\nu}_\lambda(|\mathcal{C}| = \infty) > 0$] and from work by biologists and agricultural researchers. In this latter work (see [20]) limit distributions of contact-like processes (more complicated than the "basic process" described above) were used to model vegetation patterns in arid regions in Spain and North Africa. In this "agricultural" context an occupied cluster is interpreted as a "vegetation patch." For some of these models it was claimed in [20] that simulations suggest power-law behavior of the cluster size distribution in an interval of some parameter.

In ordinary percolation models it is known that below the percolation threshold the distribution of the cluster size has exponential decay and that power-law behavior can only occur at the percolation threshold. Triggered by the above-mentioned claim in [20] concerning very different behavior in "their" contact-like processes, we study this question for $\bar{\nu}_\lambda$. Before we state our main result, Theorem 1.1, we give a brief and somewhat informal overview of earlier work on exponential-decay results in percolation to place our result in a broader context.

The proof of exponential decay for ordinary (independent) two-dimensional percolation goes back to the celebrated paper [21] by Kesten. A crucial step in that paper is, somewhat informally and in "modern" terminology, that if the probability of the event A that there is an occupied crossing of a given, large, box (square) is neither close to 0 nor close to 1, the expected number of so-called pivotal vertices (or, for bond percolation, pivotal edges) is large. (These are vertices with the property that flipping the state of the vertex flips the occurrence/nonoccurrence of the event A.) This step was proved in a "constructive" way with a "geometric" flavor. The above-mentioned large expectation of pivotal vertices implies that the derivative (w.r.t. the parameter p) of the probability of A is large. Hence, once the probability of A is not very small, a small increase of p makes it close to 1. This property would now be called a "sharp-threshold" phenomenon.

Moreover, by separate arguments, so-called finite-size criteria hold: if the probability of A is smaller than some absolute constant ϵ, the cluster size is finite a.s.

(and its distribution has exponential decay), while if it is larger than $1-\epsilon$ the system percolates. Combining these things gives exponential decay of the cluster size for all p smaller than p_c.

Russo [28] proved a very general "approximate zero-one law" and showed that the above mentioned sharp-threshold phenomenon can be obtained from this more general law using only a minimum of percolation arguments. In this way Kesten's "constructive, geometric" arguments could be avoided, which is very useful because carrying out such arguments turns out to be (too) hard in many dependent models. We should note, however, that for independent percolation the "constructive" argument still gives the shortest self-contained proof and that in *some* dependent models (see [2]) it gives the only currently known proof.

Unfortunately, the above-mentioned finite-size criteria involved a so-called RSW result of which no ("reasonably general") extension to dependent models was known. This explains why for a long time Russo's approximate zero-one law did not receive much attention in the percolation community. In the meantime sharper and more explicit results related to Russo's approximate zero-one law were obtained (in other areas of probability and mathematics in general) by Kahn, Kalai and Linial [19], Talagrand [29] and Friedgut and Kalai [15]. (See also [13] and [27].)

The importance for percolation of these sharp-threshold results became clear much later when Bollobás and Riordan [8] proved a more robust version of the RSW theorem which, combined with a clever use of the sharp-threshold results, led to the proof of the long-standing conjecture that the critical probability for random Voronoi percolation in the plane is $1/2$ (and that below $1/2$ this model has exponential decay). The robustness of these arguments led to similar results for several other two-dimensional percolation models (see [3, 9, 11]).

The last-mentioned paper proved for 2D lattice models exponential decay below the percolation threshold under the quite general condition that, informally speaking, the model has a "nice finitary representation" (in a well-defined sense) in terms of finite-valued independent random variables (see also [5]). It turned out that under that condition only a weak (not explicitly quantitative) form, close to that of Russo's [28], of the sharp-threshold results was needed. As an example it was shown that the Ising model (with fixed $\beta < \beta_c$ and external field parameter h playing the role of p in ordinary percolation) belongs to this class thus giving an alternative, more streamlined proof of the main result in Higuchi's paper [18]. Here the role of finite-valued independent random variables was played by the "independent updates" in a suitable discrete-time dynamics. Such a dynamics was possible by (among other things) the nearest-neighbor Gibbs property of the Ising model.

This is a big difference with the contact process for which we do not know a suitable discrete-time dynamics. Therefore, we are not able to derive exponential decay for this model from Theorem 2.2 in [3] but instead exploit the full quantitative nature of the sharp-threshold results from [19] and [29] and follow more

closely the route used in [8] and [11] for the Voronoi model and the Johnson–Mehl model (which, like the Voronoi model, is a model of planar tessellations but more complicated than the Voronoi model). Yet another route, namely by using results in [16], might work if $\bar{\nu}$ would satisfy the strong FKG condition which, however (as has been shown by Liggett), it does not. We should also note here that the exponential-decay arguments in [1] and [26], which for ordinary percolation work in all dimensions, so far have (even in 2D) no suitable analog for dependent percolation.

Our main result is the following theorem.

THEOREM 1.1. *Let* λ *be such that*

$$\bar{\nu}_\lambda(|\mathcal{C}| = \infty) = 0.$$

Then, for every $\lambda' < \lambda$ *there exist* $C_1, C_2 > 0$ *such that for all* $n \geq 1$

$$\bar{\nu}_{\lambda'}(|\mathcal{C}| \geq n) \leq C_1 \exp(-C_2 n). \tag{1}$$

Section 2 states properties of the contact process and other more general ingredients needed in the proof. It also indicates (see the Remark below the proof of Lemma 2.3) an alternative proof of the earlier-mentioned result by Liggett and Steif that percolation occurs for λ large enough.

The proof of Theorem 1.1 is given in Section 3. As mentioned before, the essence is still (as it was in [21]) to show sharp-threshold behavior for certain crossing probabilities. To do this we follow the main strategy in [8] and [11]. However, the model-specific properties of the contact process lead to many nontrivial differences in the steps. Therefore, and because the contact process is one of the main random spatial models, the proof is given in detail.

We use several well-known results, techniques and terminology from percolation theory. For an introduction to, and general information on, percolation see [17] and [10] and contact processes see [23] and [24].

Throughout this paper we use the notation $V \subset\subset W$ to express that V is a finite subset of W.

2. Preliminaries.

2.1. *Contact process ingredients.* A well-known classical result for the contact process is that there is a critical value λ_c such that:

(a) If $\lambda < \lambda_c$ the contact process "dies out" and $\bar{\nu}$ is concentrated on the trivial configuration where all vertices are vacant.

(b) If $\lambda > \lambda_c$, $\bar{\nu}$ is nontrivial and μ_t converges exponentially to $\bar{\nu}$ as $t \to \infty$ (see [24], Theorem 2.30 and equation (2.31), which are based on the work by

Bezuidenhout and Grimmett [6, 7]): For all $\lambda > \lambda_c$ there exist $C_3, C_4 > 0$ such that for all $t > 0$

$$\mu_t(\sigma_0 = 1) - \bar{\nu}(\sigma_0 = 1) \leq C_3 \exp(-C_4 t). \tag{2}$$

Since $\bar{\nu}$ is dominated by μ_t, statement (b) above implies by standard arguments:

THEOREM 2.1. *For all $\lambda > \lambda_c$ there exist $C_3, C_4 > 0$ such that for all $t > 0$ and all $\Lambda \subset\subset \mathbb{Z}^2$*

$$d_{\mathrm{TV}}(\mu_{t;\Lambda}, \bar{\nu}_{\lambda;\Lambda}) \leq |\Lambda| C_3 \exp(-C_4 t), \tag{3}$$

where d_{TV} denotes variational distance and $\mu_{t;\Lambda}$ (and $\bar{\nu}_{\lambda;\Lambda}$) are the restriction of μ_t (resp. $\bar{\nu}_\lambda$) to Λ.

REMARK. It is trivial from the definition of λ_c that for λ below p_c no percolation of occupied vertices occurs, that is, $\bar{\nu}_\lambda(|\mathcal{C}| = \infty) = 0$. As we mentioned in the Introduction, Liggett and Steif [25] showed that if λ is large enough percolation does occur. It seems to be widely believed (but no proof is known yet) that the critical value for having percolation is strictly larger than λ_c. (See [25] where this problem is formulated.)

A well known and very useful way to describe the contact process is by means of a space–time diagram or graphical representation (see, e.g., [23] for historical background and references). Consider for each vertex $v \in \mathbb{Z}^2$ its "time axis" $\{v\} \times (-\infty, \infty)$ and consider five independent Poisson point processes on this time axis: one with rate λ for each of the four directions (left, right, up, down) in the lattice and one to indicate a transition from 1 to 0. The Poisson processes of the different vertices are independent of each other.

The interpretation of a Poisson point on the time axis of v at time t for (say) the direction "right" is that if v is in state 1 at time t, it "infects" the vertex $v + (1, 0)$. That is, if the latter vertex is not occupied, it becomes occupied. To visualize this we draw an arrow from (v, t) to $(v + (1, 0), t)$. We say that t is the time coordinate of the arrow. For each of the other three directions we act similarly. The interpretation of a Poisson point in the fifth process on the time axis of v at time t is that if v is occupied (i.e., in state 1) at time t^-, it becomes immediately vacant (0). In the space–time picture this is marked by the symbol $*$ at (v, t) (see, e.g., [24], Part I, Section 1).

An active space–time path is a path that is allowed to move upward in time along the time axes without hitting $*$ points and to jump from one time axis to another along, and in the direction of, an arrow. The time coordinates of the arrows followed by a space–time path will be called the jumping times of the path. For $v, w \in \mathbb{Z}^2$ and $s < t$ we denote by $(v, s) \to (w, t)$ that there is an active path from (v, s) to (w, t). For the contact process starting at time 0 with every vertex

occupied, a vertex w is occupied at time $t > 0$ if and only if (in terms of the above-mentioned space–time diagram) for some vertex v there is an active path from $(v, 0)$ to (w, t). In other words, the joint distribution of the random variables

$$I\{\exists v \in \mathbb{Z}^2 \text{ s.t. } (v,0) \to (w,t)\}, \qquad w \in \mathbb{Z}^2$$

is μ_t. Similarly, $\bar{\nu}$ is the joint distribution of the random variables

$$I\{\forall t < 0 \exists v \in \mathbb{Z}^2 \text{ s.t. } (v,t) \to (w,0)\}, \qquad w \in \mathbb{Z}^2.$$

We will often work with the following "truncated" random variables. First some more notation: The distance between two vertices $v = (i_1, j_1)$ and $w = (i_2, j_2)$ is defined as $\max(|i_1 - i_2|, |j_1 - j_2|)$ and denoted by $d(v, w)$. The distance $d(V, W)$ between two subsets V and W of $\mathbb{Z}^2$ is defined as $\min(\{d(v, w) : v \in V, w \in W\})$. For $\Lambda \subset \mathbb{Z}^2$, σ_Λ denotes the collection of random variables $(\sigma_v, v \in \Lambda)$; straightforward generalizations of this notation will also be used.
Let

$$\sigma_x^{(n)} := I\{\exists (y,t) \text{ with } d(x,y) = \lfloor \sqrt{n} \rfloor \text{ or } t = -\sqrt{n} \text{ s.t. } (y,t) \to (x,0)\} \tag{4}$$

and let $\bar{\nu}^{(n)} = \bar{\nu}_\lambda^{(n)}$ denote the joint distribution of the random variables $\sigma_x^{(n)}$, $x \in \mathbb{Z}^2$.

It is clear from this definition that if Λ and Λ' are two finite subsets of $\mathbb{Z}^2$ and $d(\Lambda, \Lambda') > 2\sqrt{n}$, then $\sigma_\Lambda^{(n)}$ and $\sigma_{\Lambda'}^{(n)}$ are independent. It is also clear that σ is stochastically dominated by $\sigma^{(n)}$.

From Theorem 2.1, and simple estimates concerning the "spatial spread of infection in a limited time interval," it follows that

$$\begin{aligned} &\forall \lambda > \lambda_c \exists C_5, C_6 > 0 \text{ s.t. } \forall \Lambda \subset\subset \mathbb{Z}^2 \\ &\qquad d_{\mathrm{TV}}\big(\sigma_\Lambda, \sigma_\Lambda^{(n)}\big) \le |\Lambda| C_5 \exp(-C_6 n^{1/2}). \end{aligned} \tag{5}$$

REMARK. In this paper we often deal with spatial boxes of length of order n and distances of order n to each other. The somewhat arbitrary choice of $\sqrt{n}$ in the definition (4) is just one of the many possible choices that are convenient in such situations.

LEMMA 2.2. *Let $\Lambda_1, \dots, \Lambda_k$, be $3n \times n$-rectangles with the property that $d(\Lambda_i, \Lambda_j) > 2\lfloor \sqrt{n} \rfloor$, $1 \le i < j \le k$. Further let $A_1, A_2, \dots, A_k$ be events that are completely determined by, and increasing in, the σ variables on $\Lambda_1, \Lambda_2, \dots, \Lambda_k$ respectively. Then, for every $\lambda > \lambda_c$,*

$$\begin{aligned} \prod_{i=1}^k \bar{\nu}_\lambda(A_i) &\le \bar{\nu}_\lambda\left(\bigcap_{i=1}^k A_i\right) \le \bar{\nu}_\lambda^{(n)}\left(\bigcap_{i=1}^k A_i\right) \\ &= \prod_{i=1}^k \bar{\nu}_\lambda^{(n)}(A_i) \le \prod_{i=1}^k \big(\bar{\nu}_\lambda(A_i) + C_5 8n^2 \exp\big(-C_6\sqrt{n}\big)\big). \end{aligned} \tag{6}$$

PROOF. The first inequality comes from the well-known positive association of $\bar{\nu}_\lambda$ (which goes back to Harris's inequality) and the last inequality comes from (5). The second inequality and the equality follow immediately from the definitions. □

Let, for a rectangular box R in the lattice, $H(R)$ denote the event that there is an occupied horizontal crossing of R. Further, let $H(n, m)$ denote the event that there is an occupied horizontal crossing of the box $[0, n] \times [0, m]$. For vertical occupied crossings we use a similar notation, with V instead of H. From now on when we write "crossing" we always mean "occupied crossing."

LEMMA 2.3 (Finite-size criterion).

$$\exists \hat{\varepsilon} > 0, \qquad \forall \lambda > \lambda_c, \qquad \exists \hat{N}, \qquad \forall N \geq \hat{N}$$

the following holds:

(a)

(7) *If* $\bar{\nu}_\lambda(V(3N, N)) < \hat{\varepsilon}$, *the distribution of* $|\mathcal{C}|$ *has exponential decay.*

(b)

(8) *If* $\bar{\nu}_\lambda(H(3N, N)) > 1 - \hat{\varepsilon}$, *then* $\bar{\nu}_\lambda(|\mathcal{C}| = \infty) > 0$.

PROOF. The analog of part (a) was proved for ordinary percolation by Kesten in [22] by a block argument. His proof can be, and has been in the literature, easily adapted to dependent models with sufficient spatial mixing (e.g., see [3], Lemma 3.8). The mixing property described by (6) above is more than enough for this purpose. Essential is that the "extra term" [here $C_5 8n^2 \exp(-C_6\sqrt{n})$] in the factors in the right-hand side of (6) goes to 0 as $n \to \infty$.

The analog of (b) was proved for ordinary percolation in [14] by giving a suitable (and now well known) lower bound for the probability of having a horizontal crossing of a $4n \times 2n$ box in terms of the probability of the analogous event for a $2n \times n$ box. If for some n this probability is sufficiently close to 1, one can then iterate this procedure and conclude that the probability, say r_k, of a crossing of a given $2^{k+1}n \times 2^k n$ box goes very fast to 1 as $k \to \infty$. (So fast that $\sum_k (1 - r_k)$ is finite.) By Borel–Cantelli it then follows that a.s. there is a K, such that for all odd $k \geq K$, there is horizontal crossing of the rectangle $[0, 2^{k+1}n] \times [0, 2^k n]$ and for all even $k \geq K$ there is a vertical crossing of $[0, 2^k n] \times [0, 2^{k+1}n]$. By pasting together these crossings one gets an infinite occupied path. Hence, the system percolates. For dependent percolation models with sufficiently strong mixing properties simple modifications of such arguments can be obtained (and have been obtained in the literature). Informally speaking, instead of blowing the rectangles up by a factor 2, this is then done by a factor 3 to obtain an extra strip in the middle of the

next rectangle in order to separate the other two strips so that the crossing events of these other two strips are almost independent. See, for instance, [4], proof of Theorem 4.8, for a case where this has been carried out in detail. In practically the same way this can be carried out in our current situation by using Lemma 2.2 above in the same way as Lemma 2.3 was used in [4], proof of Theorem 5.1. □

REMARK. For our purpose (as will become clear later in this paper) we do not need $\hat{N}$ in Lemma 2.3 to be uniform in λ if λ is bounded away from λ_c. However, although this is not explicitly stated in the literature but pointed out to me by Geoffrey Grimmett (private communication), (2) and related bounds are, by the nature of their proofs in the literature, uniform in λ, if λ is bounded away from λ_c. Now such uniformity would also give uniformity of $\hat{N}$, in the sense mentioned above. This then, in turn, would clearly give an alternative proof of the earlier mentioned result by Liggett and Steif that $\bar{\nu}_\lambda$ has percolation if λ is large enough: Take some $\lambda' > \lambda_c$. Fix N such that for all $\lambda > \lambda'$ the "if-then statement" (b) in Lemma 2.3 holds. It is easy to see that, with N fixed, if $\lambda > \lambda'$ is large enough, the *condition* in that "if-then statement" (b) holds; hence $\bar{\nu}_\lambda(|\mathcal{C}_O| = \infty) > 0$. Since this result is already known and not the main subject of this paper, we do not work out the details of such alternative proof. It should also be noted that Liggett and Steif prove more than percolation of $\bar{\nu}$. They show, for large λ, domination of high-density product measures.

The following involves what in the Introduction was called a robust version of RSW.

PROPOSITION 2.4. *Let* $\lambda > \lambda_c$. *If*

$$\text{for some } \rho > 0 \qquad \limsup_{n\to\infty} \bar{\nu}_\lambda(H(\rho n, n)) > 0,$$

then

$$\text{for all } \rho > 0 \qquad \limsup_{n\to\infty} \bar{\nu}_\lambda(H(\rho n, n)) > 0.$$

PROOF. A similar result was proved by Bollobás and Riordan [8] for the random Voronoi model (and slightly modified to the above form in [4]). As remarked in [9] (see also [3], Section 3.4, the first three paragraphs) it holds for many percolation models on $\mathbb{Z}^2$, namely, those that satisfy: (i) a sufficiently strong mixing property, (ii) a straightforward "geometric" condition about lattice paths (which enables pasting together paths that cross each other), (iii) positive association and (iv) the condition that $\bar{\nu}$ is invariant under the symmetries of $\mathbb{Z}^2$.

Lemma 2.2 above is more than needed for (i) and it is easy to see that the probability measures $\bar{\nu}_\lambda$, $\lambda > \lambda_c$ also satisfies the other conditions. □

2.2. *Influence and sharp-threshold results.* Let $\Omega = \{0,1\}^n$ and let P_p denote the product measure with parameter p on Ω. Let A be an event (i.e., a subset of Ω) and let, for $1 \le i \le n$, I_i denote the probability that i is pivotal for A. It is often called the influence of i. More precisely,

$$I_i := P_p\big(\{\omega \in \Omega : \text{ exactly one of } \omega \text{ and } \omega^{(i)} \text{ is in } A\}\big),$$

where $\omega^{(i)}$ is the configuration obtained from ω by flipping the ith component of ω. Talagrand ([29], Corollary 1.2) proved the following theorem. See also [15] and [19] for strongly related results. Note that our I_i differs a factor $1/p$ from the expression $\mu_p(A_i)$ in Talagrand's paper.

THEOREM 2.5.

$$\sum_i I_i \ge \frac{P_p(A)(1-P_p(A))}{Kp\log(2/p)} \log\Big(\frac{1}{p\max_i I_i}\Big), \tag{9}$$

where K is a universal positive constant.

REMARK.

(i) Strictly speaking Talagrand's result is slightly stronger than Theorem 2.5 above but in the case of small p (to which we will apply it), it makes essentially no difference.

(ii) If the event A is increasing (i.e., its indicator function is a coordinate-wise nondecreasing function on Ω), the left-hand side of (9) is, according to Russo's formula, equal to $d/dp\, P_p(A)$. By this it is easy to see that Theorem 2.5 implies that if, throughout some interval, say (p_1, p_2), $\max_i I_i$ is "very small" and $P_{p_1}(A)$ is "not too small," then $P_{p_2}(A)$ is "close to 1." For such reasons Theorem 2.5 and related theorems are often indicated by the name "sharp-threshold" results, in addition to names like "influence results."

Now suppose there are at least m indices i with the property that $I_i = \max_j I_j$. There are two possibilities:

(a) $\max_i I_i \le \frac{\log m}{pm}$. If this holds then, by Theorem 2.5,

$$\sum_i I_i \ge \frac{P_p(A)(1-P_p(A))}{Kp\log(2/p)} \log\Big(\frac{m}{\log m}\Big) \ge \frac{P_p(A)(1-P_p(A))}{\tilde{K}p\log(2/p)} \log m \tag{10}$$

for some universal constant $\tilde{K}$.

(b) $\max_i I_i \ge \frac{\log m}{pm}$. Then trivially,

$$\sum_i I_i \ge m \max_i I_i \ge \frac{\log m}{p}$$

which is larger than or equal to some universal constant times the right-hand side of (10). Hence, by adjusting the value K if needed, the following holds:

COROLLARY 2.6. *Let m denote the cardinality of $\{i : I_i = \max_j I_j\}$. Then*

$$\sum_i I_i \geq \frac{P_p(A)(1 - P_p(A))}{Kp \log(2/p)} \log m.$$

REMARK. The case $m = n$ of this corollary is essentially in [15] where it is derived from the results/methods in [19]. The general case, and its derivation from Theorem 2.5, was shown to me by Oliver Riordan (private communication; see also [12]).

We will use a generalization of Theorem 2.5 and Corollary 2.6 as described below.

Let Ω be as before. Let $V \subset \{1, \ldots, n\}$ and let $0 < p_1, p_2 < 1$. Let P_{p_1,p_2} denote the product measure on Ω under which each component with index in V is 1 with probability p_1 and each with index in V^c is 1 with probability p_2. The generalization of Theorem 2.5 is the following theorem.

THEOREM 2.7.

$$\sum_i I_i \geq \frac{P_{p_1,p_2}(A)(1 - P_{p_1,p_2}(A))}{K' \max(p_1, p_2) \log(2/\min(p_1, p_2))} \log\left(\frac{1}{\max(p_1, p_2) \max_i I_i}\right),$$

where K' is a universal constant.

REMARK. In [12] (see Theorem 5 in [12] and the discussion below that theorem) it is indicated how to prove Theorem 2.7 by modifications of the proofs in Talagrand's paper. An alternative way is to start from the special case for $p = 1/2$ of Theorem 2.5 above and obtain the full case (and its generalization where different coordinates may have a different parameter p) from that special case by, informally speaking, representing (approximately) the toss of a biased coin by a combination of tosses of several fair coins.

From Theorem 2.7 the following corollary is obtained in exactly the same way as Corollary 2.6 was obtained from Theorem 2.5.

COROLLARY 2.8. *Let m denote the cardinality of $\{i : I_i = \max_j I_j\}$. Then*

$$\sum_i I_i \geq \frac{P_{p_1,p_2}(A)(1 - P_{p_1,p_2}(A))}{K' \max(p_1, p_2) \log(2/(\min(p_1, p_2)))} \log m.$$

Combined with a straightforward modification of the earlier-mentioned Russo's formula this gives:

COROLLARY 2.9. *Let m be as in the previous corollary. If the event A is increasing in the coordinates with parameter p_1 and decreasing in the coordinates with parameter p_2, then*

$$\frac{\partial}{\partial p_1} P_{p_1,p_2}(A) - \frac{\partial}{\partial p_2} P_{p_1,p_2}(A) \geq \frac{P_{p_1,p_2}(A)(1 - P_{p_1,p_2}(A))}{K' \max(p_1, p_2) \log(2/(\min(p_1, p_2))} \log m. \tag{11}$$

3. Proof of Theorem 1.1. Let $\lambda_1 > \lambda_c$ be such that under $\bar{\nu}_{\lambda_1}$ the cluster size distribution does not have exponential decay. Let $\lambda_2 > \lambda_1$. We will show that $\bar{\nu}_{\lambda_2}(|\mathcal{C}_O| = \infty) > 0$. This will immediately imply Theorem 1.1.

Let λ_1 be as fixed above and let $\hat{\varepsilon}$ and $\hat{N} = \hat{N}(\lambda_1)$ be as in Lemma 2.3. Let L_n denote a specific $4n \times n$ rectangle; its precise choice does not matter but for later convenience we choose $[n, 5n] \times [n, 2n]$. By Lemma 2.3 we have that

$$\bar{\nu}_{\lambda_1}(V(3n, n)) > \hat{\varepsilon} \qquad \text{for all } n \geq \hat{N}$$

which by Proposition 2.4 implies $\limsup_{n\to\infty} \bar{\nu}_{\lambda_1}(H(L_n)) > 0$; so there exists an $\tilde{\varepsilon} > 0$ and a sequence $n_1, n_2, \ldots$ such that

$$\bar{\nu}_{\lambda_1}(H(L_{n_i})) > \tilde{\varepsilon} \qquad \text{for all } i. \tag{12}$$

From now on we consider such fixed sequence.

In the Introduction to the contact process in the beginning of Section 1 we assumed that the recovery rate is 1. Of course the contact process with infection rate λ and recovery rate δ is simply a time-rescaled version of the contact process with infection rate λ/δ and recovery rate 1. In particular, these two contact processes have exactly the same upper invariant measure. For application of the results in Section 2.2 it is more convenient to work with one-parameter Poisson processes for which at each site of the lattice the total rate of all the Poisson processes is constant, say 1. Therefore, we consider the contact process with infection rate $q/4$ and recovery rate $1 - q$, where now $q \in (0, 1)$ is the parameter. Note that in terms of the space–time diagram this means that on each time axis we have a marked Poisson point process with density 1 and each point corresponds with a $\rightarrow$, $\leftarrow$, $\downarrow$, $\uparrow$ or $*$ with probability $q/4, q/4, q/4, q/4$ and $1 - q$, respectively. With respect to this new parameter q we use the notation $\mathcal{P}_q$ for the law governing the above-marked Poisson point process and the notation $\bar{\nu}_{\langle q\rangle}$ for the upper invariant measure of the corresponding contact process. From the above it is immediate that

$$\bar{\nu}_{\langle q\rangle} = \bar{\nu}_{q/(4(1-q))}, \qquad q \in (0, 1), \tag{13}$$

or, equivalently, $\bar{\nu}_\lambda = \bar{\nu}_{\langle 4\lambda/(1+4\lambda)\rangle}$, for $\lambda \in (0, \infty)$. In particular, by (12),

$$\bar{\nu}_{\langle q_1\rangle}(H(L_{n_i})) > \tilde{\varepsilon} \qquad \text{for all } i, \tag{14}$$

where $q_1 = 4\lambda_1/(1 + 4\lambda_1)$.

Let $\bar{\nu}^{(n)}_{\langle q\rangle}$ be the distribution of $(\eta^{(n)}_x, x \in \mathbb{Z}^d)$ defined by [compare with (4)]

$$\begin{aligned}\eta^{(n)}_x := I\{&\exists (y,t) \text{ with } d(x,y) = \lfloor \sqrt{n} \rfloor \text{ or}\\ &t < -\sqrt{n} \text{ s.t. } (y,t) \overset{(q,1-q)}{\to} (x,0)\},\end{aligned} \tag{15}$$

where $(y,t) \overset{(q,1-q)}{\to} (x,0)$ denotes that there is a space–time path from (y,t) to $(x,0)$ in the space–time diagram with Poisson intensity $q/4$ for each of the four types of arrows and Poisson intensity $1-q$ for $*$'s.

It is clear that $\bar{\nu}^{(n)}_{\langle q\rangle}$ dominates $\bar{\nu}_{\langle q\rangle}$; hence, by (14),

$$\bar{\nu}^{(n_i)}_{\langle q_1\rangle}(H(L_{n_i})) > \tilde{\varepsilon} \qquad \text{for all } i. \tag{16}$$

Although $\bar{\nu}^{(n)}_{\lambda}$ is, of course, not the same as $\bar{\nu}^{(n)}_{\langle 4\lambda/(1+4\lambda)\rangle}$, it is straightforward to get analogs of the earlier "approximation lemmas." In particular we get, as an analog of (5),

$$\begin{aligned}&\forall q > 4\lambda_c/(1+4\lambda_c) \exists C_7, C_8 > 0 \text{ s.t. } \forall \Lambda \subset\subset \mathbb{Z}^2\\ &d_{\mathrm{TV}}\big(\bar{\nu}_{\langle q\rangle;\Lambda}, \bar{\nu}^{(n)}_{\langle q\rangle;\Lambda}\big) \le |\Lambda| C_7 \exp(-C_8 n^{1/2}).\end{aligned} \tag{17}$$

Throughout the proof of Theorem 1.1, except at the very end (see Proposition 3.5, where we translate back to parameter λ), we will work with parameter q as described above.

A key step toward application of the results in Section 2.2 is a suitable "time-discretized" version of $\bar{\nu}^{(n)}_{\langle q\rangle}$. A significant obstacle is to obtain an analog of (16) for these discrete variables.

Recall from the beginning of this section that L_n is the box $[n,5n]\times[n,2n]$. To "get ample room for the underlying Poisson points" we also consider the larger box $B_n := [0,6n]\times[0,3n]$. Let $\bar{\nu}^{(n)}_{\langle q\rangle}$ be as before. Note that the collection of random variables $(\eta^{(n)}_x, x \in L_n)$ is completely determined by the (marked) Poisson points in the space–time area $ST(n) := B_n \times [-n, 0]$. (In fact only a subset of that area is involved but for convenience we consider this whole area.) Let as before, $\mathcal{P}_q$ denote the probability measure governing the marked Poisson points.

Let $0 < \alpha < 1$. Later we choose α sufficiently small. Let $\delta = n^{-\alpha}$.

DEFINITION 3.1. We say that an active space–time path π is δ-stable if the following hold:

(i) If s and t are two different jump times of π, then $|t-s| > \delta$.

(ii) If (y,s) is the starting point or endpoint of an arrow of π and there is a $*$ at (y,t), then $|t-s| > \delta$.

The following lemma (and the global structure of its proof) is the analog of Theorem 6.1 for the Voronoi model in [8] and Theorem 8 for the Johnson–Mehl tessellations in [11] (see also [12]). Since the proof is subtle and differs in many details from that in [8] and [11] we give a full proof.

REMARK. In some sense the proof of Lemma 3.2 is easier and shorter than that of the corresponding results in [8] and [11]. This is partly due to the fact that in our model the continuous object that has to be properly discretized (the time axis) is one dimensional. This enables us to "play" with the order (in time) of the Poisson points. On the other hand, our model has some extra complications, for example, there is no natural order on the arrow values assigned to the Poisson points (an arrow to the right is not always better than an arrow to the left). Fortunately these issues can be handled quite smoothly.

LEMMA 3.2 (Stability coupling). *Let $0 < q < q' < 1$. For each n there is a coupling of $\mathcal{P}_q$ and $\mathcal{P}_{q'}$ such that w.h.p. (i.e., with probability tending to 1 as $n \to \infty$) the following holds: For every $x \in L_n$ that has $\eta_x^{(n)} = 1$ in the first copy, there is a $(y,t) \in \mathbb{Z}^2 \times (-\infty, 0)$ with $d(x,y) = \lfloor \sqrt{n} \rfloor$ or $t = -\sqrt{n}$ such that there is a δ-stable space–time path in the second copy from (y,t) to $(x,0)$ and hence $\eta_x^{(n)}$ also equals 1 in the second copy.*

PROOF. Let $\delta_1 = n^{-\alpha/2}$. So $\delta \ll \delta_1$. We partition every "time axis" $\{x\} \times [-\infty, 0]$, $x \in \mathbb{Z}^2$, in intervals $\{x\} \times (-(k+1)\delta_1, -k\delta_1]$, $k = 0, 1, \ldots$, of length δ_1. From now on when we use the word "interval," we will always mean an interval of the above form with $x \in B_n$ and $(k+1)\delta_1 \le n$. Note that the total number of intervals is $M_n := |B_n| \lfloor n/\delta_1 \rfloor$. Let $\mathcal{I}_n$ denote the union of these intervals.

Note that the total number of Poisson points in $\mathcal{I}_n$ is Poisson distributed with mean $\delta_1 M_n$. To construct the coupling first draw a number N according to the above-mentioned Poisson distribution. Now assign N points (called "particles") randomly, uniformly and independently of each other to the above mentioned set $\mathcal{I}_n$. If a particle is assigned to the space–time location (x,t), we say that its time coordinate is t. Call an interval "occupied" if it has at least one particle. Call two different intervals $\{x\} \times (-(k+1)\delta_1, -k\delta_1]$ and $\{y\} \times (-(l+1)\delta_1, -l\delta_1]$ neighbors if $d(x,y) \le 1$ and $|k-l| \le 1$. This gives rise in an obvious way to the notion of "clusters (of occupied intervals)." (This notion of cluster is of course different from that introduced earlier in this paper. Since this "new" notion of cluster is used only in this proof and the other notion is not used here, this should not cause any confusion and we even use the same notation $\mathcal{C}$.)

We have already assigned to each particle a precise location in $\mathcal{I}_n$. However, we "suppress" this precise information and only "keep" the following *partial information*: for each interval the number of particles assigned to it and for each occupied cluster of intervals the relative order (w.r.t. their time coordinates) of all particles

in that cluster. We also assign, with equal probabilities (1/4), a *tentative* $\leftarrow$, $\rightarrow$, $\uparrow$ or $\downarrow$ to each particle (independent of the other particles). The interpretation is that *if* eventually a particle is chosen to represent an arrow, the type of arrow is exactly its above-mentioned tentative one.

REMARK. From now on when we mention a cluster $\mathcal{C}$, we mean not only its corresponding set of intervals but also the above-mentioned partial information about the particle locations as well as the tentative arrows assigned to the particles.

By the size of a cluster we mean the number of particles in the cluster.

CLAIM. There is a constant $D = D(\alpha)$ such that

$$\lim_{n\to\infty} P\big(\exists \text{ an occupied cluster with size } \geq D(\alpha)\big) = 0. \tag{18}$$

PROOF. Let D be a positive integer. If the occupied cluster of a given interval e has size $\geq D$ there is a connected set of D (not necessarily occupied) intervals, such that e is one of these intervals and the number of particles in the union of these intervals is $\geq D$. Note that the number of choices for e is $M_n \leq n^4$ (for n sufficiently large) and that for each choice of e the number of possible connected sets of D intervals is smaller than or equal to some constant $C(D)$ which depends on D only. Further, the number of particles in the union of D given intervals is Poisson distributed with mean $D\delta_1 = Dn^{-\alpha/2}$. So the probability that this number of particles is $\geq D$ is at most $(Dn^{-\alpha/2})^D$. Hence, the probability that there is an occupied cluster of size $\geq D$ is at most

$$n^4 C(D)(Dn^{-\alpha/2})^D.$$

If we take $D = \lceil 9/\alpha \rceil$, this probability goes indeed to 0 as $n \to \infty$. This proves the above claim. □

Let $\mathcal{C}$ be a cluster in the sense given in the remark above. Now consider for both parameter values, q and q', the conditional distribution of the precise configuration for $\mathcal{C}$, that is, the types ($*$, $\leftarrow$, $\rightarrow$, $\uparrow$ or $\downarrow$) and precise locations of all particles in $\mathcal{C}$, given the partial information. The two conditional distributions can be coupled by the following natural procedure which gives two "typical realizations" (one for each of the two parameter values) of the precise configuration.

The first step in this procedure is to assign to each particle i, independent of the other particles, a random variable U_i uniformly distributed on $(0, 1)$. These variables will be used below to decide if a particle corresponds with an arrow or with a $*$.

The next step is to go from relative order of positions to precise positions. Consider the conditional distribution of the precise time coordinates of the particles of $\mathcal{C}$, given their (already known) relative order in time and the intervals they are

located in. Now simply assign the precise locations by drawing from this distribution. Later we will refer to this procedure as the "time assignment procedure."

Note that both steps above are the same for both "realizations," the one for parameter q and the one with parameter q'. However, the next and final step in which the types of the particles are fully determined will take into account the parameter value: for each particle i of $\mathcal{C}$ do the following: If $U_i < q$, the type of i in each of the two copies is equal to the earlier mentioned *tentative* arrow. If $U_i \in (q, q')$, its type is $*$ in copy 1 and equal to the tentative arrow in copy 2. If $U_i > q'$, the type is $*$ in both copies.

Now we have two realizations, say $\omega_{\mathcal{C}}(1)$ and $\omega_{\mathcal{C}}(2)$, and it is easy to see that they are "typical" w.r.t. the two conditional distributions mentioned above (the first for parameter q, the second for parameter q'). So we indeed have a coupling of these two conditional distributions. Also note that $\omega_{\mathcal{C}}(2) \geq \omega_{\mathcal{C}}(1)$ in the sense that the particle locations are exactly the same and each particle in $\omega_{\mathcal{C}}(1)$ that has an arrow-type has the same arrow-type in $\omega_{\mathcal{C}}(2)$. Let this coupling be denoted by $P_{\mathcal{C}}$.

Doing this for each cluster, independently of the other clusters, gives a natural coupling of the two probability measures in the statement of the lemma. However, it is not yet what we want. Although it satisfies the property between brackets at the end of the lemma, it does not necessarily satisfy the stability property in the lemma. The coupling we do want is obtained as follows where we go back to the level of a given cluster $\mathcal{C}$. Recall the two copies $\omega_{\mathcal{C}}(1)$ and $\omega_{\mathcal{C}}(2)$ above and their joint distribution $P_{\mathcal{C}}$. From $P_{\mathcal{C}}$ we will construct a modified distribution $\tilde{P}_{\mathcal{C}}$ of which the two marginal distributions are the same as those of $P_{\mathcal{C}}$. To avoid an abundance of notation we will drop the subscript $\mathcal{C}$ from $\omega_{\mathcal{C}}(1)$ and $\omega_{\mathcal{C}}(2)$.

Recall the time assignment procedure in the second step of the construction of $P_{\mathcal{C}}$. Let B be the event that in $\omega(1)$ [and hence, since the particle locations for $\omega(1)$ and $\omega(2)$ are the same, also in $\omega(2)$] there are two different particles in $\mathcal{C}$ whose time coordinates differ at most δ. The probability of B (or, more precisely, the conditional probability of B given the partial information on $\mathcal{C}$) is maximal if $\mathcal{C}$ consists of one interval only, in which case it is less than or equal to $|\mathcal{C}|^2 \frac{2\delta}{\delta_1} = |\mathcal{C}|^2 2n^{-\alpha/2}$, where $|\mathcal{C}|$ denotes the number of particles in $\mathcal{C}$; so

$$P_{\mathcal{C}}(B) \leq |\mathcal{C}|^2 2n^{-\alpha/2}. \tag{19}$$

Recall the use of the variables U_i in the determination of the types of the points. Let G be the event that each particle in $\omega(1)$ is of type $*$ and each particle in $\omega(2)$ has an arrow type. Note that this event happens if and only if $U_i \in (q, q')$ for all particles i in $\mathcal{C}$ so that we have

$$P_{\mathcal{C}}(G) = (q' - q)^{|\mathcal{C}|}.$$

By this and (19) we have (with $D = D(\alpha)$ as in the claim above)

$$P_{\mathcal{C}}(G) \geq P_{\mathcal{C}}(B) \qquad \text{if } |\mathcal{C}| \leq D \tag{20}$$

and n is sufficiently large. From now on we assume in this proof that n is indeed sufficiently large in this sense.

Now let B' denote $B \setminus G$. If $|\mathcal{C}| \le D$ then by (20) there is a measurable subset $G' \subset G \setminus B$ and a 1–1 map $\psi : B' \to G'$ with the property that ψ and ψ^{-1} are $P_{\mathcal{C}}$-preserving. To each pair $(\omega(1), \omega(2)) \in B'$ this map assigns the pair

$$\psi(\omega(1), \omega(2)) = \big((\psi(\omega(1), \omega(2)))(1), (\psi(\omega(1), \omega(2)))(2)\big).$$

Now a modified coupling called $\tilde{P}_{\mathcal{C}}$ is obtained from $P_{\mathcal{C}}$ by exchange between B' and G' of the second copy, using the map ψ as follows. (Such type of modification is called a "cross-over" in [8].) If $|\mathcal{C}| \ge D$ we simply take $\tilde{P}_{\mathcal{C}} = P_{\mathcal{C}}$. Otherwise, a typical pair $(\tilde{\omega}(1), \tilde{\omega}(2))$ under $\tilde{P}_{\mathcal{C}}$ is drawn as follows. First draw a pair $(\omega(1), \omega(2))$ under $P_{\mathcal{C}}$. If $(\omega(1), \omega(2)) \in (B' \cup G')^c$, take $(\tilde{\omega}(1), \tilde{\omega}(2))$ equal to $(\omega(1), \omega(2))$. If $(\omega(1), \omega(2)) \in B'$, take $\tilde{\omega}(1) = \omega(1)$ and $\tilde{\omega}(2) = (\psi(\omega(1), \omega(2)))(2)$. Finally, if $(\omega(1), \omega(2)) \in G'$ take $\tilde{\omega}(1) = \omega(1)$ and $\tilde{\omega}(2) = (\psi^{-1}(\omega(1), \omega(2)))(2)$. Since in all cases $\tilde{\omega}(1) = \omega(1)$, it is immediate that the first marginal of $\tilde{P}_{\mathcal{C}}$ is equal to that of $P_{\mathcal{C}}$. A short inspection shows that also the second marginal of $\tilde{P}_{\mathcal{C}}$ is equal to that of $P_{\mathcal{C}}$.

Now the "overall" coupling of $\mathcal{P}_q$ and $\mathcal{P}_{q'}$ announced in the statement of the lemma is obtained in a natural and straightforward way by constructing the pair $(\tilde{\omega}_{\mathcal{C}}(1), \tilde{\omega}_{\mathcal{C}}(2)$ for each cluster $\mathcal{C}$ separately, independently of the other clusters.

To check the required properties of this coupling first look again at one single cluster $\mathcal{C}$. Suppose that $|\mathcal{C}| \le D$. Let $(\omega(1), \omega(2))$ and the corresponding pair $(\tilde{\omega}(1), \tilde{\omega}(2))$ be as above. So, in particular, $\tilde{\omega}(1) = \omega(1)$. Suppose that $\tilde{\omega}(1)$ has a certain active space–time path π within $\mathcal{C}$. Note that π is also an active space–time path for $\omega(1)$ and [because $\omega(2) \ge \omega(1)$ in the sense mentioned earlier in this proof] also for $\omega(2)$. For our purpose we may assume that π is part of a path that guarantees for some $x \in L_n$, that $\eta_x^{(n)} = 1$ (see the statement of Lemma 3.2). Therefore, by considering a trajectory of this longer path between entering and leaving the cluster, we may assume that π starts at the bottom of some interval and ends at the top of some interval. We will show that $\tilde{\omega}(2)$ has a δ-stable space–time path $\tilde{\pi}$ that "corresponds" with π. More precisely, although the jump-times of the path $\tilde{\pi}$ may differ a bit from the corresponding jump times of π, it will start and end at the same space–time points as the beginning, respectively end, of π.

First we assume that π makes at least one jump. Since $\omega(1)$ has at least one arrow in $\mathcal{C}$, $(\omega(1), \omega(2))$ is not in G, so we have only the following two possible cases:

(i) If $(\omega(1), \omega(2)) \in B' = B \setminus G$, then its image under the map ψ is in $G \setminus B$. Hence, since the relative order and the tentative arrow types of all the particles are fixed and by the definition of G no particle in $\tilde{\omega}(2)$ has a $*$, there is indeed a natural path $\tilde{\pi}$ in the configuration $\tilde{\omega}(2)$ that corresponds with π. Moreover, by the definition of B^c no two particles in $\tilde{\omega}(2)$ have time coordinates that differ at most δ and hence $\tilde{\pi}$ is δ-stable.

(ii) If $(\omega(1), \omega(2)) \in B^c \cap G^c$, we have $\tilde{\omega}(2) = \omega(2)$. From the definition of B^c it follows that π itself is δ-stable so we can take $\tilde{\pi}$ equal to π.

Now suppose π makes no jump. So π is, in fact, the union of a finite number of consecutive intervals on the time axis of a vertex. Note that by definition of a cluster each of these intervals has at least one particle. Hence, $(\omega(1), \omega(2))$ is not in G because otherwise in the configuration $\omega(1)$ each of these intervals would have a $*$ which contradicts the fact that π is an active path. If it is not in B either, $\tilde{\omega}(2) = \omega(2)$ and we can simply take $\tilde{\pi} = \pi$. Finally, if $(\omega(1), \omega(2))$ is in $B \setminus G$, then its image is in $G \setminus B$ so $\tilde{\omega}(2)$ has no $*$ particles and again the conclusion follows immediately.

Using the above-mentioned δ-stability property of the single-cluster couplings yields a similar property for the "overall" coupling of $\mathcal{P}_q$ and $\mathcal{P}_{q'}$. The only thing that could go "wrong" is if there is a cluster with size $\geq D(\alpha)$. However, by the claim, this has probability going to 0 as $n \to \infty$. The proof of Lemma 3.2 is complete. $\square$

We proceed with the proof of Theorem 1.1. Fix a value $\hat{q}$ in the interval (q_1, q_2), where $q_1 = 4\lambda_1/(4\lambda_1 + 1)$ as before [see below (14)] and $q_2 = 4\lambda_2/(4\lambda_2 + 1)$.

Now we are ready to introduce $0-1$ valued random variables to which we can apply the results in Section 2.2. Let the box B_n and the space–time region $ST(n)$ be as before (see a few lines before Definition 3.1). Now partition every time axis in intervals of length δ, with δ as defined just before Definition 3.1.

As before, we have on each time axis a Poisson point process with density 1 and each Poisson point is, independently of the others, of type $*$ with probability $1-q$ and of each of the types $\rightarrow$, $\leftarrow$, $\uparrow$, $\downarrow$ with probability $q/4$. Let $v \in B_n$ and $k \in \mathbb{N}$, $0 \leq k \leq n/\delta$. By the kth interval of v for the above-mentioned partition, we will mean $\{v\} \times (-k\delta, (-k+1)\delta]$, and we define

$$X_*^{(v,k,\delta)} := I\{\exists \text{ a Poisson point of type } * \text{ in the } k\text{th interval of } v\}.$$

Similarly define

$$X_{\rightarrow}^{(v,k,\delta)} := I\{\exists \text{ a Poisson point of type } \rightarrow \text{ in the } k\text{th interval of } v\}$$

and, analogously, $X_{\leftarrow}^{(v,k,\delta)}$, $X_{\uparrow}^{(v,k,\delta)}$ and $X_{\downarrow}^{(v,k,\delta)}$. Note that this is a collection of independent $0-1$ valued random variables.

Recall the definition of $\eta_v^{(n)}$ and $\bar{\nu}_{\langle q \rangle}^{(n)}$ below equation (14). The X variables defined above give only "crude" information about the space–time diagram; they tell which of the types $*$, $\rightarrow$, etc., occur in each interval but they do not tell their precise locations inside the intervals. Nevertheless, this incomplete information is often enough to conclude that there is a certain space–time path. Let $\eta_v^{(n,\delta)}$ be the indicator of the event that the values of the $X^{(\cdot,\cdot,\delta)}$ variables imply that $\eta_v^{(n)} = 1$.

REMARK. Note that if $\eta_v^{(n)} = 1$ then, after for some Poisson points with mark $*$, this mark is replaced by an arrow still $\eta_v^{(n)} = 1$. The same remark holds for $\eta_v^{(n,\delta)}$ instead of $\eta_v^{(n)}$.

It is easy to see that

$$\eta_v^{(n)} \geq \eta_v^{(n,\delta)} \geq I\big\{\exists(w,t) \text{ with } d(v,w) = \lfloor\sqrt{n}\rfloor \text{ or } t = -\sqrt{n} \text{ s.t.} \\ \exists\delta\text{-stable space–time path from } (w,t) \text{ to } (v,0)\big\}. \tag{21}$$

Hence, with the following notation (where R is a box)

$$H^{(n,\delta)}(R) := \big\{\exists\eta^{(n,\delta)}\text{-occupied horizontal crossing of } R\big\},$$

we get

$$\mathcal{P}_{\hat{q}}\big(H^{(n_i,\delta)}(L_{n_i})\big) \geq \mathcal{P}_{q_1}\big(\exists\eta^{(n_i)}\text{-occupied horizontal crossing of } L_{n_i}\big) - \varepsilon(n_i) \\ = \bar{\nu}_{\langle q_1\rangle}^{(n_i)}\big(H(L_{n_i}) - \varepsilon(n_i)\big),$$

where $\varepsilon(n)$ is a function of n that goes to 0 as $n \to \infty$ and where the inequality comes from the second inequality in (21) and Lemma 3.2 and the equality comes directly from the definitions.

By (16), and obvious monotonicity [see the Remark preceding (21)], this gives the following lemma.

LEMMA 3.3. *For each choice of α the following holds for all sufficiently large i:*

$$\mathcal{P}_q\big(H^{(n_i,\delta)}(L_{n_i})\big) > \frac{\tilde{\varepsilon}}{2}, \qquad q \geq \hat{q}. \tag{22}$$

Now we "wrap around the box B_{n_i} horizontally" by identifying every vertex $(6n_i, y)$ on B_{n_i} with the vertex $(0, y)$ thus turning this box into a cylinder. Define $\eta_v^{(n_i,\delta,C)}$ as the natural analog for the cylinder of $\eta_v^{(n_i,\delta)}$.

REMARK. "By the truncation to distance $\sqrt{n}$ of these variables" and because the left- and right-hand side of L_{n_i} have distance larger than $\sqrt{n}$ to the boundary of B_{n_i}, the event that there is a $\eta_\cdot^{(n_i,\delta,C)}$-occupied horizontal crossing of L_{n_i} and the event that there is a $\eta_\cdot^{(n_i,\delta)}$-occupied horizontal crossing of L_{n_i} are the same.

Let $A^{(n_i,\delta)}$ be the event that at least one of the $(6n_i - 1)$ horizontal translates of L_{n_i} on this cylinder has an $\eta_\cdot^{(n_i,\delta,C)}$-occupied horizontal crossing. Note that the event $A^{(n_i,\delta)}$ is still defined in terms of the random variables $X^{\cdot,k,\delta}$ defined earlier. Moreover, this event is increasing in the X variables corresponding with

arrows and decreasing in those corresponding with $*$'s. For each choice of α the following holds for all sufficiently large i:

$$\mathcal{P}_q\big(A^{(n_i,\delta)}\big) \geq \mathcal{P}_q\big(H^{(n_i,\delta)}(L_{n_i})\big) > \frac{\tilde{\varepsilon}}{2}, \qquad q \geq \hat{q}, \tag{23}$$

where the first inequality is (taking into account the above remark) trivial and the last inequality is exactly Lemma 3.3.

As stated before the X variables are independent $0-1$ valued random variables. Further, for each v and k, $X_*^{(v,k,\delta)}$ has probability $1-\exp(-(1-q)\delta)$ to be 1. Each random variable $X_{\rightarrow}^{(v,k,\delta)}$ has probability $1-\exp(-\delta q/4)$ to be 1. The same holds for the other three arrow types.

Also note that the event $A^{(n_i,\delta)}$ is partially symmetric in the following sense: for fixed value k and fixed $0 \leq l \leq 3n$ all variables $X_{\rightarrow}^{v,k,\delta}$ with $v \in B_n$ with $y-$ coordinate l, "play the same role." In particular, each of them has the same probability to be pivotal for the event $A^{(n_i,\delta)}$. The same statement holds for each of the other three arrow types and for type $*$. Further note that for each k and l the number of such random variables $X_{\rightarrow}^{(v,k,\delta)}$ is of order n. Again, the same statement holds for each of the other types.

We will apply Corollary 2.9 with m equal to our "current" n and with p_1 and p_2 equal to $1-\exp(-\delta q/4)$ and $1-\exp(-(1-q)\delta)$, respectively. For our purpose we should think of n as very large and hence, δ very small. For fixed n (and hence, δ), the p_1 and p_2 above are functions of q and

$$\frac{dp_1}{dq} = \frac{\delta}{4}\exp(-\delta q/4)$$

which is of order δ. More precisely, there are positive constants C' and C'' such that

$$C'\delta \leq \frac{dp_1}{dq} \leq C''\delta \qquad \text{for all } \delta \in (0,1) \text{ and } q \in [\hat{q}, q_2).$$

Similarly, p_1 and p_2 are also of order δ and dp_2/dq is of order $-\delta$. Therefore, when we take the derivative with respect to q of the probability of the event $A^{(n_i,\delta)}$, the factor of order δ that comes from $\max(p_1, p_2)$ in the denominator in the right-hand side of (11) is canceled by a factor of order δ that comes from dp_1/dq and dp_2/dq. Essentially the only "remaining" effect of δ comes from the logarithmic expression in the denominator in the right-hand side of (11). More precisely what we get is

$$\frac{d}{dq}\mathcal{P}_q\big(A^{(n_i,\delta)}\big) \geq \frac{C_9\mathcal{P}_q(A^{(n_i,\delta)})(1-\mathcal{P}_q(A^{(n_i,\delta)})\log n}{\log(2/\delta)}, \qquad q \in [\hat{q}, q_2), \tag{24}$$

where $C_9 > 0$ depends on $\hat{q}$ and q_2 only.

Let $\varepsilon^* > 0$. By (23), (24) and because $\mathcal{P}_q(A^{(n_i,\delta)})$ is clearly nondecreasing in q, it follows that, for every choice of α, the following holds for all sufficiently large i: If $\mathcal{P}_{q_2}(A^{(n_i,\delta)}) < 1 - \varepsilon^*$ then, for all $q \in [\hat{q}, q_2)$,

$$\frac{d}{dq}\mathcal{P}_q\big(A^{(n_i,\delta)}\big) \geq C_9\frac{\tilde{\varepsilon}}{2}\varepsilon^*\frac{\log n}{\log(2/\delta)} \geq \frac{C_{10}\tilde{\varepsilon}\varepsilon^*}{\alpha}$$

(where the last inequality used that $\delta = n^{-\alpha}$) and hence,

$$\mathcal{P}_{q_2}\big(A^{(n_i,\delta)}\big) \geq (q_2 - \hat{q})C_{10}\tilde{\varepsilon}\varepsilon^*/\alpha.$$

By choosing α sufficiently small this gives the following lemma.

LEMMA 3.4. *For every $\varepsilon^* > 0$ there is an $\alpha > 0$ such that for all sufficiently large i*

$$\mathcal{P}_{q_2}\big(A^{(n_i,\delta)}\big) > 1 - \varepsilon^*. \tag{25}$$

Now if there is a horizontal crossing of one of the above-mentioned translates of L_{n_i}, there must be a horizontal crossing in the "hard" direction of at least one of the following (six) translates (on the cylinder) of the rectangle $[0, 3n_i] \times [n_i, 2n_i]$:

$$[jn_i, (j+3)n_i (\mathrm{mod}\, 6n_i)] \times [n_i, 2n_i], \qquad 0 \leq j \leq 5.$$

Hence, by the usual "square root trick,"

$$\mathcal{P}_{q_2}\big(H^{(n_i,\delta)}([0, 3n_i] \times [n_i, 2n_i])\big) \geq 1 - \big(1 - \mathcal{P}_{q_2}\big(A^{(n_i,\delta)}\big)\big)^{1/6}$$

which, combined with Lemma 3.4, immediately gives that for every $\varepsilon^* > 0$ there is an $\alpha > 0$ s.t. for all sufficiently large i

$$\mathcal{P}_{q_2}\big(H^{(n_i,\delta)}([0, 3n_i] \times [0, n_i])\big) > 1 - \varepsilon^*. \tag{26}$$

Finally the following proposition is obtained.

PROPOSITION 3.5.

$$\lim_{i\to\infty} \bar{\nu}_{\lambda_2}(H(3n_i, n_i)) = 1. \tag{27}$$

PROOF. Let $\varepsilon^* > 0$ be given. By (13) (and the definition of q_2), $\bar{\nu}_{\lambda_2}(H(3n_i, n_i)) = \bar{\nu}_{\langle q_2\rangle}(H(3n_i, n_i))$. Hence, by (17), $\liminf_{i\to\infty} \bar{\nu}_{\lambda_2}(H(3n_i, n_i))$ is equal to $\liminf_{i\to\infty} \bar{\nu}^{(n_i)}_{\langle q_2\rangle}(H(3n_i, n_i))$ which by the first inequality in (21) is larger than or equal to

$$\liminf_{i\to\infty} \mathcal{P}_{q_2}\big(H^{(n_i,\delta)}([0, 3n_i] \times [0, n_i])\big).$$

This last expression is, by (the statement ending with) (26) and a suitable choice of α, larger than $1 - \varepsilon^*$. Summarizing, we have that for every $\varepsilon^* > 0$, $\liminf_{i\to\infty} \bar{\nu}_{\lambda_2}(H(3n_i, n_i))$ is larger than $1 - \varepsilon^*$. □

Proposition 3.5, together with the finite-size criterion Lemma 2.3, immediately yields $\bar{\nu}_{\lambda_2}(|\mathcal{C}_O| = \infty) > 0$ which, as observed in the beginning of this section, completes the proof of Theorem 1.1.

Acknowledgments. I thank Oliver Riordan for a very useful discussion concerning influence and sharp-threshold results and in particular for pointing out Corollary 2.6 for general m. The work in this paper was partly done during visits to the Isaac Newton Institute in June 2008, the Institut Henri Poicaré in October 2008 and the Institut Mittag Leffler in April and June 2009. I thank these institutes for their support and hospitality. Finally, I thank Jeff Steif and Geoffrey Grimmett for several useful discussions and an anonymous referee for many valuable detailed comments and suggestions for improvement on an earlier version of this paper.

REFERENCES

[1] AIZENMAN, M. and BARSKY, D. J. (1987). Sharpness of the phase transition in percolation models. *Comm. Math. Phys.* **108** 489–526. MR874906

[2] BÁLINT, A., CAMIA, F. and MEESTER, R. (2009). Sharp phase transition and critical behaviour in 2D divide and colour models. *Stochastic Process. Appl.* **119** 937–965. MR2499865

[3] VAN DEN BERG, J. (2008). Approximate zero-one laws and sharpness of the percolation transition in a class of models including two-dimensional Ising percolation. *Ann. Probab.* **36** 1880–1903. MR2440926

[4] VAN DEN BERG, J., BROUWER, R. and VÁGVÖLGYI, B. (2008). Box-crossings and continuity results for self-destructive percolation in the plane. In *In and Out of Equilibrium* 2 (V. Sidoravičius and M.-E. Vares, eds.). *Progress in Probability* **60** 117–135. Birkhäuser, Basel. MR2477379

[5] VAN DEN BERG, J. and STEIF, J. E. (1999). On the existence and nonexistence of finitary codings for a class of random fields. *Ann. Probab.* **27** 1501–1522. MR1733157

[6] BEZUIDENHOUT, C. and GRIMMETT, G. (1990). The critical contact process dies out. *Ann. Probab.* **18** 1462–1482. MR1071804

[7] BEZUIDENHOUT, C. and GRIMMETT, G. (1991). Exponential decay for subcritical contact and percolation processes. *Ann. Probab.* **19** 984–1009. MR1112404

[8] BOLLOBÁS, B. and RIORDAN, O. (2006). The critical probability for random Voronoi percolation in the plane is 1/2. *Probab. Theory Related Fields* **136** 417–468. MR2257131

[9] BOLLOBÁS, B. and RIORDAN, O. (2006). Sharp thresholds and percolation in the plane. *Random Structures Algorithms* **29** 524–548. MR2268234

[10] BOLLOBÁS, B. and RIORDAN, O. (2006). *Percolation*. Cambridge Univ. Press, New York. MR2283880

[11] BOLLOBÁS, B. and RIORDAN, O. (2008). Percolation on random Johnson–Mehl tessellations and related models. *Probab. Theory Related Fields* **140** 319–343. MR2365477

[12] BOLLOBÁS, B. and RIORDAN, O. (2009). Erratum: Percolation on random Johnson–Mehl tessellations and related models. Available at arXiv:0905.1275v1.

[13] BOURGAIN, J., KAHN, J., KALAI, G., KATZNELSON, Y. and LINIAL, N. (1992). The influence of variables in product spaces. *Israel J. Math.* **77** 55–64. MR1194785

[14] CHAYES, J. T. and CHAYES, L. (1986). Percolation and random media. In *Phénomènes Critiques, Systèmes Aléatoires, Théories de Jauge, Part I, II* (*Les Houches, 1984*) 1001–1142. North-Holland, Amsterdam. MR880545

[15] FRIEDGUT, E. and KALAI, G. (1996). Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.* **124** 2993–3002. MR1371123

[16] GRAHAM, B. T. and GRIMMETT, G. R. (2006). Influence and sharp-threshold theorems for monotonic measures. *Ann. Probab.* **34** 1726–1745. MR2271479

[17] GRIMMETT, G. (1999). *Percolation*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **321**. Springer, Berlin. MR1707339

[18] HIGUCHI, Y. (1993). A sharp transition for the two-dimensional Ising percolation. *Probab. Theory Related Fields* **97** 489–514. MR1246977

[19] KAHN, J., KALAI, G. and LINIAL, N. (1988). The influence of variables on Boolean functions. In *Proc. 29th Annual Symposium on Foundations of Computer Science* 68–80. Computer Society Press.

[20] KÉFI, S., RIETKERK, M., ALADOS, C. L., PUEYO, Y., PAPANASTASIS, V. P., ELAICH, A. and DE RUITER, P. C. (2007). Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. *Nature* **449** 213–217.

[21] KESTEN, H. (1980). The critical probability of bond percolation on the square lattice equals $\frac{1}{2}$. *Comm. Math. Phys.* **74** 41–59. MR575895

[22] KESTEN, H. (1981). Analyticity properties and power law estimates of functions in percolation theory. *J. Stat. Phys.* **25** 717–756. MR633715

[23] LIGGETT, T. M. (1985). *Interacting Particle Systems*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **276**. Springer, New York. MR776231

[24] LIGGETT, T. M. (1999). *Stochastic Interacting Systems*: *Contact*, *Voter and Exclusion Processes*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **324**. Springer, Berlin. MR1717346

[25] LIGGETT, T. M. and STEIF, J. E. (2006). Stochastic domination: The contact process, Ising models and FKG measures. *Ann. Inst. H. Poincaré Probab. Statist.* **42** 223–243. MR2199800

[26] MENSHIKOV, M. V. (1986). Coincidence of critical points in percolation problems. *Dokl. Akad. Nauk SSSR* **288** 1308–1311. MR852458

[27] ROSSIGNOL, R. (2008). Threshold phenomena on product spaces: BKKKL revisited (once more). *Electron. Commun. Probab.* **13** 35–44. MR2372835

[28] RUSSO, L. (1982). An approximate zero-one law. *Z. Wahrsch. Verw. Gebiete* **61** 129–139. MR671248

[29] TALAGRAND, M. (1994). On Russo's approximate zero-one law. *Ann. Probab.* **22** 1576–1587. MR1303654

CWI
SCIENCE PARK 123
1098 XG AMSTERDAM
THE NETHERLANDS
E-MAIL: J.van.den.Berg@cwi.nl