The Annals of Probability
2012, Vol. 40, No. 3, 893–920
DOI: 10.1214/10-AOP641
© Institute of Mathematical Statistics, 2012

LIMIT THEOREMS FOR 2D INVASION PERCOLATION

BY MICHAEL DAMRON[1] AND ARTËM SAPOZHNIKOV[2]

Princeton University and ETH Zürich

We prove limit theorems and variance estimates for quantities related to ponds and outlets for 2D invasion percolation. We first exhibit several properties of a sequence $(\mathbf{O}(n))$ of outlet variables, the nth of which gives the number of outlets in the box centered at the origin of side length 2^n. The most important of these properties describes the sequence's renewal structure and exponentially fast mixing behavior. We use these to prove a central limit theorem and strong law of large numbers for $(\mathbf{O}(n))$. We then show consequences of these limit theorems for the pond radii and outlet weights.

1. Introduction.

1.1. *The model.* Invasion percolation is a stochastic growth model both introduced and numerically studied independently by [2] and [14]. Let $G = (V, E)$ be an infinite connected graph in which a distinguished vertex, the origin, is chosen. Let $(\tau_e)_{e \in E}$ be independent random variables, uniformly distributed on [0, 1]. The *invasion percolation cluster* (IPC) of the origin on G is defined as the limit of an increasing sequence (G_n) of connected subgraphs of G as follows. For an arbitrary subgraph $G' = (V', E')$ of G, we define the outer edge boundary of G' as

$$\Delta G' = \{e = \langle x, y\rangle \in E : e \notin E', \text{but } x \in V' \text{ or } y \in V'\}.$$

We define G_0 to be the origin. Once the graph $G_i = (V_i, E_i)$ is defined, we select the edge e_{i+1} that minimizes τ on ΔG_i. We take $E_{i+1} = E_i \cup \{e_{i+1}\}$ and let G_{i+1} be the graph induced by the edge set E_{i+1}. The graph G_i is called the *invaded region* at time i. Let $E_\infty = \bigcup_{i=0}^\infty E_i$ and $V_\infty = \bigcup_{i=0}^\infty V_i$. Finally, define the IPC

$$\mathcal{S} = (V_\infty, E_\infty).$$

We study invasion percolation on two-dimensional lattices; however, for simplicity *we restrict ourselves hereafter to the square lattice* $\mathbb{Z}^2$ and denote by $\mathbb{E}^2$ the set of nearest-neighbour edges. The results of this paper still hold for lattices

Received September 2010.

[1]Supported by an NSF postdoctoral fellowship.

[2]Supported in part by the Excellence Fund Grant of TU/e of Remco van der Hofstad.

MSC2010 subject classifications. 60K35, 82B43.

Key words and phrases. Invasion percolation, invasion ponds, critical percolation, near critical percolation, correlation length, scaling relations, central limit theorem.

which are invariant under reflection in one of the coordinate axes and under rotation around the origin by some angle. In particular, this includes the triangular and honeycomb lattices.

We define Bernoulli percolation using the random variables τ_e to make a coupling with the invasion immediate. For any $p \in [0, 1]$ we say that an edge $e \in \mathbb{E}^2$ is *p-open* if $\tau_e < p$ and *p-closed* otherwise. It is obvious that the resulting random graph of p-open edges has the same distribution as the one obtained by declaring each edge of $\mathbb{E}^2$ open with probability p and closed with probability $1 - p$, independently of the state of all other edges. The percolation probability $\theta(p)$ is the probability that the origin is in the infinite cluster of p-open edges. There is a critical probability $p_c = \inf\{p : \theta(p) > 0\} \in (0, 1)$. For general background on Bernoulli percolation we refer the reader to [8].

In [3], it is shown that, for any $p > p_c$, the invasion on $(\mathbb{Z}^d, \mathbb{E}^d)$ intersects the infinite p-open cluster with probability one. In the case $d = 2$ this immediately follows from the Russo–Seymour–Welsh theorem (see Section 11.7 in [8]). This result has been extended to much more general graphs in [9]. Furthermore, the definition of the invasion mechanism implies that if the invasion reaches the p-open infinite cluster for some p, it will never leave this cluster. Combining these facts yields that if e_i is the edge added at step i, then $\limsup_{i\to\infty} \tau_{e_i} = p_c$. It is well known that for Bernoulli percolation on $(\mathbb{Z}^2, \mathbb{E}^2)$, the percolation probability at p_c is 0. This implies that, for infinitely many values of i, the weight τ_{e_i} satisfies $\tau_{e_i} > p_c$. The last two results give that $\hat{\tau}_1 = \max\{\tau_e : e \in E_\infty\}$ exists and is greater than p_c. The above maximum is attained at an edge which we shall call $\hat{e}_1$. Suppose that $\hat{e}_1$ is invaded at step i_1, that is, $\hat{e}_1 = e_{i_1}$. Following the terminology of [15], we call the graph G_{i_1-1} the *first pond* of the invasion, denoting it by the symbol $\hat{V}_1$, and we call the edge $\hat{e}_1$ the *first outlet*. The second pond of the invasion is defined similarly. Note that a simple extension of the above argument implies that $\hat{\tau}_2 = \max\{\tau_{e_i} : e_i \in E_\infty, i > i_1\}$ exists and is greater than p_c. If we assume that $\hat{\tau}_2$ is taken on the edge $\hat{e}_2$ at step i_2, we call the graph $G_{i_2-1} \setminus G_{i_1-1}$ the *second pond* of the invasion, and we denote it $\hat{V}_2$. The edge $\hat{e}_2$ is called the *second outlet*. The further ponds $\hat{V}_k$ and outlets $\hat{e}_k$ are defined analogously. For a hydrological interpretation of the ponds we refer the reader to [18].

In this paper, we consider a sequence of outlet variables introduced in [4]. We continue the analysis from that paper, in which almost sure bounds were shown for the sequence's growth rate. Here, we prove limit theorems for the sequence and, as a consequence, we obtain variance estimates for the sequence $(\hat{\tau}_k)$ of outlet weights and for the sequence of pond radii. The current results were inspired by limit theorems for critical percolation obtained by Kesten and Zhang in [13] and later by Zhang in [20]. In those papers, the authors prove central limit theorems for (a) the maximal number of disjoint open circuits around the origin in the box of size n centered at the origin in critical percolation in two dimensions and (b) the number of open clusters in the same box in any dimension in percolation with parameter $p \in [0, 1]$. The martingale methods they use apply to some degree for our questions of invasion percolation, but our techniques, based on mixing properties and moment bounds from [4], seem to reveal more of the underlying structure of the process.

The mixing properties mentioned above are consequences of a more general renewal mechanism that lies inside the invasion process on $\mathbb{Z}^2$. In Section 3, we show that for any $m, k \geq 1$, the invaded regions at distances 2^m and 2^{m+k} from the origin are equal to two statistically independent sets except on an event whose probability decays exponentially in k. Roughly speaking, this means that the invasion has a very weak dependence structure when viewed on exponential length scales.

Last we would like to mention that limit theorems similar to ones we establish in this paper were shown by Goodman [7] for invasion percolation on the regular tree. Those results were also inspiration for the current work. Goodman showed, for example, that the sizes of the ponds grow exponentially, with laws of large numbers, central limit theorems and large deviation results. His analysis is based on representing $\mathcal{S}$ in terms of the outlets weights $\hat{\tau}_n$, as in [1].

1.2. *Notation.* In this section we collect most of the notation and the definitions used in the paper.

For $a \in \mathbb{R}$, we write $|a|$ for the absolute value of a, and, for a site $x = (x_1, x_2) \in \mathbb{Z}^2$, we write $|x|$ for $\max(|x_1|, |x_2|)$. For $n > 0$ and $x \in \mathbb{Z}^2$, let $B(x, n) = \{y \in \mathbb{Z}^2 : |y - x| \leq n\}$ and $\partial B(x, n) = \{y \in \mathbb{Z}^2 : |y - x| = n\}$. We write $B(n)$ for $B(0, n)$ and $\partial B(n)$ for $\partial B(0, n)$. For $m < n$ and $x \in \mathbb{Z}^2$, we define the annulus $\mathrm{Ann}(x; m, n) = B(x, n) \setminus B(x, m)$. We write $\mathrm{Ann}(m, n)$ for $\mathrm{Ann}(0; m, n)$.

We consider the square lattice $(\mathbb{Z}^2, \mathbb{E}^2)$, where $\mathbb{E}^2 = \{\langle x, y\rangle \in \mathbb{Z}^2 \times \mathbb{Z}^2 : |x - y| = 1\}$. Let $(\mathbb{Z}^2)^* = (1/2, 1/2) + \mathbb{Z}^2$ and $(\mathbb{E}^2)^* = (1/2, 1/2) + \mathbb{E}^2$ be the vertices and the edges of the dual lattice. For $x \in \mathbb{Z}^2$, we write x^* for $x + (1/2, 1/2)$. For an edge $e \in \mathbb{E}^2$ we denote its endpoints (left, resp., right or bottom, resp., top) by $e_x, e_y \in \mathbb{Z}^2$. The edge $e^* = \langle e_x + (1/2, 1/2), e_y - (1/2, 1/2)\rangle$ is called the *dual edge* to e. Its endpoints (bottom, resp., top or left, resp., right) are denoted by e_x^* and e_y^*. Note that e_x^* and e_y^* are not the same as $(e_x)^*$ and $(e_y)^*$. For a subset $\mathcal{K} \subset \mathbb{Z}^2$, let $\mathcal{K}^* = (1/2, 1/2) + \mathcal{K}$. We say that an edge $e \in \mathbb{E}^2$ is in $\mathcal{K} \subset \mathbb{Z}^2$ if both its endpoints are in $\mathcal{K}$. For any graph $\mathcal{G}$ we write $|\mathcal{G}|$ for the number of vertices in $\mathcal{G}$.

Let $(\tau_e)_{e \in \mathbb{E}^2}$ be independent random variables, uniformly distributed on [0, 1], indexed by edges. We call τ_e the *weight* of an edge e. We define the weight of an edge e^* as $\tau_{e^*} = \tau_e$. We denote the underlying probability measure by $\mathbb{P}$ and the space of configurations by $([0, 1]^{\mathbb{E}^2}, \mathcal{F})$, where $\mathcal{F}$ is the natural σ-field on $[0, 1]^{\mathbb{E}^2}$. We say that an edge e is *p-open* if $\tau_e < p$ and *p-closed* if $\tau_e > p$. An edge e^* is p-open if e is p-open, and it is p-closed if e is p-closed. The event that two sets of sites $\mathcal{K}_1, \mathcal{K}_2 \subset \mathbb{Z}^2$ are connected by a p-open path is denoted by $\mathcal{K}_1 \overset{p}{\longleftrightarrow} \mathcal{K}_2$.

For any $k \geq 1$, let $\hat{R}_k$ be the radius of the union of the first k ponds. In other words,

$$\hat{R}_k = \max\left\{|x| : x \in \bigcup_{j=1}^{k} \hat{V}_k\right\}.$$

For two functions g and h from a set $\mathcal{X}$ to $\mathbb{R}$, we write $g(z) \asymp h(z)$ to indicate that $g(z)/h(z)$ is bounded away from 0 and ∞, uniformly in $z \in \mathcal{X}$. We will also use the standard notation $g(z) = O(h(z))$ if $g(z)/h(z)$ is bounded away from ∞ uniformly in $z \in \mathcal{X}$, and $g(z) = o(h(z))$ if for each $\varepsilon > 0$, $|g(z)/h(z)| > \varepsilon$ for only a finite number of values of $z \in \mathcal{X}$. For any event A, we write $I(A)$ for the indicator function of A. For any sequence of random variables (X_i) and any $k \geq 0$, we say that the sequence is *k-dependent* if for every $m \geq 1$, the set of variables $\{X_1, \dots, X_m\}$ is independent of the set of variables $\{X_{m+k+1}, \dots\}$. Similarly we say that the sequence of events (A_i) is m-dependent if the sequence of variables $(I[A_i])$ is. Throughout this paper we write log for $\log_2$. All the constants (C_i) in the proofs are strictly positive and finite. Their exact values may be different from proof to proof.

1.3. *Main results.*

1.3.1. *The CLT for outlets.* Let O_k be the number of outlets in the annulus $\mathrm{Ann}(2^{k-1}, 2^k)$ and $a_k = \mathbb{E} O_k$. Let $\mathbf{O}(n) = \sum_{k=1}^{n} O_k$, $a(n) = \mathbb{E}\mathbf{O}(n)$ and $b(n)^2 = \operatorname{Var} \mathbf{O}(n)$.

THEOREM 1. *There exist positive and finite constants c_1 and c_2 such that for all i, $a_i \in [c_1, c_2]$, and the variance of $\mathbf{O}(n)$ satisfies*

$$b(n)^2 \asymp n.$$

Write $N(0, 1)$ for the distribution of a standard normal random variable, and let $\Rightarrow$ denote convergence in distribution.

THEOREM 2. *The sequence $(\mathbf{O}(n))$ satisfies a CLT, that is,*

$$\frac{\mathbf{O}(n) - a(n)}{b(n)} \Rightarrow N(0, 1). \tag{1.1}$$

Furthermore, if $r > 1/2$, then the following convergence is almost sure:

$$\frac{\mathbf{O}(n) - a(n)}{n^r} \to 0. \tag{1.2}$$

1.3.2. *Consequences of the CLT for outlets.* As discussed in Section 1.1, a main intention of this paper is to study the asymptotic behavior of the sequences $(\hat{R}_n)$ (toward infinity) and $(\hat{\tau}_n)$ (toward p_c). In [4] it was proved that these sequences obey the following almost sure bounds. There exist constants $C_1 > 0$ and $C_2 < \infty$ such that with probability one

$$C_1 n \le \log \hat{R}_n \le C_2 n \quad \text{and} \quad C_1 n \le -\log(\hat{\tau}_n - p_c) \le C_2 n$$

for all large n. Motivated by these results, we want study whether or not these sequences converge, after properly shifting and normalizing. Further, we would like know information about rates of convergence. It turns out that from the point of view of these questions, the sequences are closely related to certain sequences (Q_n) and (T_n), which we now define.

Let

$$Q_n = \min\{k : \mathbf{O}(k) \ge n\} \quad \text{and} \quad T_n = \min\{k : a(k) \ge n\}.$$

Note that $\mathbf{O}(Q_n - 1) < n \le \mathbf{O}(Q_n)$ and $a(T_n - 1) < n \le a(T_n)$. We define a sequence of random variables $(a(Q_n))$, where $a(Q_n)$ equals $a(k)$ if and only if $Q_n = k$. By this definition, $a(Q_n)$ takes values in the set $\{a(k) : k \ge 1\}$ with

$$\mathbb{P}\big(a(Q_n) = a(k)\big) = \mathbb{P}(Q_n = k).$$

The CLT for outlets allows us to study the sequence $(a(Q_n))$. Let $\sigma_n^2 = \operatorname{Var} \mathbf{O}(T_n)$.

THEOREM 3.

$$\frac{a(Q_n) - n}{\sigma_n} \Rightarrow N(0, 1).$$

[*Or, equivalently*, $(a(Q_n) - a(T_n))/\sigma_n \Rightarrow N(0, 1)$.] *Moreover*,

$$\mathbb{E}\left(\frac{a(Q_n) - n}{\sigma_n}\right)^2 \to 1 \qquad \textit{as } n \to \infty.$$

REMARK 1. We would like to use Theorem 3 to deduce CLTs for the sequences $(\log \hat{R}_n)$ and $(\log(\hat{\tau}_n - p_c))$, both of which are proved in the case of the regular tree in [7]. It is not difficult to prove these results if one knows that $(Q_n - T_n)/\delta_n$ converges in distribution to a variable with the standard normal distribution for some sequence δ_n. Unfortunately, Theorem 3 does not appear to be strong enough to show this. One possible approach to deduce a CLT for (Q_n) from Theorem 3 is to demonstrate that the sequence (a_n) does not fluctuate too quickly as $n \to \infty$. For instance one could try to prove that there exists $a \in \mathbb{R}$ such that for every sequence (k_n) of natural numbers,

$$(1/n) \sum_{k_n+1}^{k_n+n} a_i \to a \qquad \text{as } n \to \infty.$$

Although we are not able to prove CLTs for $(\log \hat{R}_n)$ and $(\log(\hat{\tau}_n - p_c))$, we show in the next corollaries that the fluctuations are of the correct order of magnitude.

COROLLARY 1.

$$\mathbb{E}(Q_n - T_n)^2 \asymp n, \qquad \mathbb{E}(Q_n - \mathbb{E}Q_n)^2 \asymp n.$$

COROLLARY 2.

$$\mathbb{E}(\log \hat{R}_n - T_n)^2 \asymp n, \qquad \mathbb{E}(\log \hat{R}_n - \mathbb{E}\log \hat{R}_n)^2 \asymp n.$$

In the statements of the next two corollaries, we use the sequence (p_n), defined in Section 2.

COROLLARY 3.

$$\mathbb{E}\left(\log \frac{\hat{\tau}_n - p_c}{p_{2^{T_n}} - p_c}\right)^2 \asymp n, \qquad \mathbb{E}\big(\log(\hat{\tau}_n - p_c) - \mathbb{E}\log(\hat{\tau}_n - p_c)\big)^2 \asymp n.$$

Last we show that the sequences (Q_n), $(\log \hat{R}_n)$ and $(\log(\hat{\tau}_n - p_c))$ satisfy laws of large numbers.

COROLLARY 4. *For any $r > 1/2$, each of the following sequences converges to 0 almost surely*:

$$\left(\frac{Q_n - T_n}{n^r}\right), \qquad \left(\frac{\log \hat{R}_n - T_n}{n^r}\right), \qquad \left(\frac{1}{n^r}\log \frac{\hat{\tau}_n - p_c}{p_{2^{T_n}} - p_c}\right).$$

1.4. *Structure of the paper.* In Section 2 we recall the definition of the correlation length, which is vital to all of our proofs. In Section 3 we describe and prove several properties of the outlet variables (O_k) that will be used in the proofs of Theorems 1 and 2 in Section 4. In Section 5, we prove consequences of the CLT: Theorem 3 and Corollaries 1–4.

2. Correlation length.

2.1. *Definition of correlation length.* For m, n positive integers and $p \in (p_c, 1]$ let

$$\sigma(n, m, p) = \mathbb{P}(\text{there is a } p\text{-open horizontal crossing of } [0, n] \times [0, m]).$$

Given $\varepsilon > 0$, we define

$$L(p, \varepsilon) = \min\{n : \sigma(n, n, p) \geq 1 - \varepsilon\}. \tag{2.1}$$

$L(p,\varepsilon)$ is called the *finite-size scaling correlation length* and it is known that $L(p,\varepsilon)$ scales like the usual correlation length (see [12]). It was also shown in [12] that the scaling of $L(p,\varepsilon)$ is independent of ε given that it is small enough, that is, there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon_1, \varepsilon_2 \le \varepsilon_0$ we have $L(p,\varepsilon_1) \asymp L(p,\varepsilon_2)$. (Here, ε_1 and ε_2 are fixed numbers that do not depend on p.) For simplicity we will write $L(p) = L(p,\varepsilon_0)$ in the entire paper. We also define

$$p_n = \sup\{p : L(p) > n\}.$$

It is easy to see that $L(p) \to \infty$ as $p \to p_c$ and $L(p) \to 0$ as $p \to 1$. In particular, the probability p_n is well defined. It is clear from the definitions of $L(p)$ and p_n and from the RSW theorem that, for positive integers k and l, there exists $\delta_{k,l} > 0$ such that, for any positive integer n and for all $p \in [p_c, p_n]$,

$$\mathbb{P}(\text{there is a } p\text{-open horizontal crossing of } [0,kn]\times[0,ln]) > \delta_{k,l}$$

and

$$\mathbb{P}\big(\text{there is a } p\text{-closed horizontal dual crossing of } ([0,kn]\times[0,ln])^*\big) > \delta_{k,l}.$$

By the FKG inequality and a standard gluing argument [8], Section 11.7, we get that, for positive integers n and $k \ge 2$ and for all $p \in [p_c, p_n]$,

$$\mathbb{P}\big(\mathrm{Ann}(n,kn) \text{ contains a } p\text{-open circuit around the origin}\big) > (\delta_{k,k-2})^4$$

and

$$\begin{aligned}\mathbb{P}\big(&\mathrm{Ann}(n,kn)^* \text{ contains a } p\text{-closed dual circuit around the origin}\big)\\ &> (\delta_{2k,k-1})^4.\end{aligned}$$

2.2. *Properties of correlation length.* We give the following results without proofs.

(1) Reference [12], Theorem 2. There is a constant $D_1 < \infty$ such that, for all $p > p_c$,

$$\theta(p) \le \mathbb{P}\big[0 \stackrel{p}{\longleftrightarrow} \partial B(L(p))\big] \le D_1 \mathbb{P}\big[0 \stackrel{p_c}{\longleftrightarrow} \partial B(L(p))\big], \tag{2.2}$$

where $\theta(p) = \mathbb{P}(0 \stackrel{p}{\longleftrightarrow} \infty)$ is the percolation function for Bernoulli percolation.

(2) Reference [16], Section 4. There is a constant $D_2 > 0$ such that, for all $n \ge 1$,

$$\mathbb{P}\big(B(n) \stackrel{p_n}{\longleftrightarrow} \infty\big) \ge D_2. \tag{2.3}$$

(3) For any $n \ge 1$ and $p \in [0,1]$, let $B_{n,p}$ be the event that there is a p-closed circuit around the origin in the dual lattice with radius at least n. There exist constants $D_3 < \infty$ and $D_4 > 0$ such that for all $p > p_c$,

$$\mathbb{P}(B_{n,p}) \le D_3 \exp\left\{-D_4 \frac{n}{L(p)}\right\}. \tag{2.4}$$

Equation (2.4) follows, for example, from [11], (2.6) and (2.8) (see also [17], Lemma 37 and Remark 38).

(4) There exist constants $D_5 > 0$ and $D_6 < \infty$ such that for all $m, n \geq 1$,

$$D_5\left|\log \frac{m}{n}\right| \leq \left|\log \frac{p_m - p_c}{p_n - p_c}\right| \leq D_6\left|\log \frac{m}{n}\right|. \tag{2.5}$$

This is a consequence of [17], Proposition 34, and a priori bounds on the 4-arm exponent.

3. Properties of the outlet variables. In this section we describe several important properties of the variables (O_k). We first recall the following theorem from [4] that gives k-independent bounds on all of their moments.

THEOREM 4. *There exists $c_1 < \infty$ such that for all $t, k \geq 1$,*

$$\mathbb{E}(O_k^t) \leq (c_1 t)^{3t}. \tag{3.1}$$

One crucial feature of the invasion process that allows us to prove limit theorems is its *renewal structure*. To describe this, we make a couple of definitions. For $k, m \geq 1$ and $1 \leq l \leq \infty$, let $\mathcal{G}(k, l, m)$ be the graph of the invasion process that invades the entire box $B(2^{k-m})$ at step 1 [we take $B(2^{k-m})$ to be the origin if $k < m$], then proceeds with the usual invasion rules and stops when it invades any vertex of $\partial B(2^{k+l+m})$. In the case that $l = \infty$, we allow the invasion to run for all of time. Write $\mathcal{O}$ for the set of all outlets of $\mathcal{S}$, and write $\mathcal{O}(k, l, m)$ for the set of all outlets of $\mathcal{G}(k, l, m)$. In the case of $\mathcal{O}(k, l, m)$, the outlets are defined in the same way as in $\mathcal{O}$; however, note that if the graph $\mathcal{G}(k, l, m)$ is finite (which corresponds to the case of finite l), some of its outlets may have weight below p_c.

For the next theorem, when $l = \infty$, $\mathrm{Ann}(m, l)$ will mean $B(m)^c$.

THEOREM 5 (Renewal structure of the invasion). *There are constants $C < \infty$ and $\delta > 0$ such that for all $k, m \geq 1$ and $1 \leq l \leq \infty$,*

$$\mathbb{P}\big(\mathcal{S} \cap \mathrm{Ann}(2^k, 2^{k+l}) \neq \mathcal{G}(k, l, m) \cap \mathrm{Ann}(2^k, 2^{k+l})\big) < C \exp(-\delta m)$$

and

$$\mathbb{P}\big(\mathcal{O} \cap \mathrm{Ann}(2^k, 2^{k+l}) \neq \mathcal{O}(k, l, m) \cap \mathrm{Ann}(2^k, 2^{k+l})\big) < C \exp(-\delta m).$$

PROOF. Clearly it suffices to prove the theorem for $m > 4$. We first consider the case that $k \geq m$ and $l < \infty$. Observe that $\mathcal{S} \cap \mathrm{Ann}(2^k, 2^{k+l}) = \mathcal{G}(k, l, m) \cap \mathrm{Ann}(2^k, 2^{k+l})$ and $\mathcal{O} \cap \mathrm{Ann}(2^k, 2^{k+l}) = \mathcal{O}(k, l, m) \cap \mathrm{Ann}(2^k, 2^{k+l})$ if (1) there exists a p_c-open circuit around the origin in $\mathrm{Ann}(2^{k-m}, 2^k)$, (2) there exists a $p_{2^{k+l+m/4}}$-closed dual circuit around the origin in the annulus $\mathrm{Ann}(2^{k+l}, 2^{k+l+m/4})^*$, (3) there exists a p_c-open circuit around the origin in $\mathrm{Ann}(2^{k+l+m/2}, 2^{k+l+m})$ and (4) the open circuit from (3) is connected by a

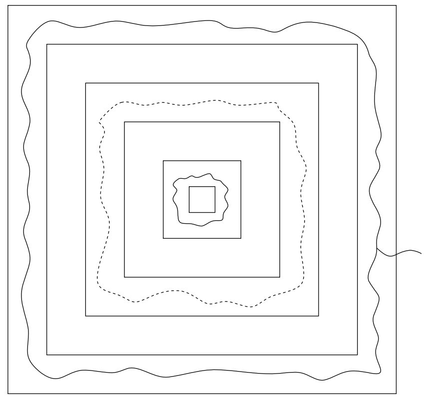

FIG. 1. *The event in the proof of Theorem 5 (in the case $k \geq m$ and $l < \infty$). The boxes, in order from smallest to largest, are $B(2^{k-m})$, $B(2^k)$, $B(2^{k+l})$, $B(2^{k+l+m/4})$, $B(2^{k+l+m/2})$ and $B(2^{k+l+m})$. (Boxes are not drawn to scale.) The dotted path is $p_{2^{k+l+m/4}}$-closed, the path to infinity is $p_{2^{k+l+m/4}}$-open and the other two circuits are p_c-open. If all these paths exist, the sets $\mathcal{S}$ and $\mathcal{G}(k,l,m)$ coincide in* $\mathrm{Ann}(2^k, 2^{k+l})$.

$p_{2^{k+l+m/4}}$-open path to infinity. (See Figure 1 for an illustration of the intersection of these four events.) Indeed, the first condition implies that in the exterior of the p_c-open circuit from (1), $\mathcal{G}(k,l,m)$ is a subset of $\mathcal{S}$. The remaining conditions (2)–(4) imply the existence of an edge e in $\mathrm{Ann}(2^{k+l}, 2^{k+l+m})$, lying in the closure of the exterior of the closed circuit from (2), such that $e \in \mathcal{O} \cap \mathcal{O}(k,l,m)$, and both invasion processes invade e before any vertex of $\partial B(2^{k+l+m})$. Therefore, once this outlet is invaded (by either of the two invasion processes), the set of invaded edges in the interior of the closed circuit from (2) does not change anymore. The RSW theorem and (2.4) imply that the probability that any of (1)–(4) does not hold is bounded from above by $C\exp(-\delta m)$ uniformly in k.

In the case that $k < m$ and $l < \infty$, we exclude condition (1) from the above argument. In the case $k < m$ and $l = \infty$ there is nothing to prove. If $k \geq m$ and $l = \infty$ we argue using only condition (1). $\square$

REMARK 2. Similar ideas were used in the proof of the upper bound in Theorem 1.4 in [4]. Note that there is a typo there in the definition of X_i^n. It should be specified that X_i^n counts only disconnecting edges with weights larger than p_c.

We now present corollaries of Theorem 5 that will help in the proofs of the next section. The first two are about mixing properties of the sequence (X_k). Recall the notation that $a_k = \mathbb{E} O_k$ and let $X_k = O_k - a_k$. For any $m_1 \leq m_2$, let $\Sigma_{m_1}^{m_2}$ be the sigma algebra generated by the variables $X_{m_1}, \dots, X_{m_2}$. Write Σ_{m_1} for $\lim_{m_2 \to \infty} \Sigma_{m_1}^{m_2}$ and Σ^{m_2} for $\Sigma_1^{m_2}$. For $m \geq 0$, define the *strong mixing coefficient*

$$\alpha(m) = \sup_{k \geq 1} \sup_{A,B} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, \tag{3.2}$$

where the supremum is over all $A \in \Sigma^k$ and $B \in \Sigma_{k+m}$.

COROLLARY 5. *There exist constants $C < \infty$ and $\delta > 0$ such that for all m,*

$$\alpha(m) \le C \exp(-\delta m). \tag{3.3}$$

PROOF. Clearly it suffices to prove the corollary for $m > 4$. Fix $k \ge 1$ and let $A \in \Sigma^k$, $B \in \Sigma_{k+m}$. For $j = 1, \dots, k$, let $\tilde{Y}_j$ be the number of outlets in $\mathcal{O}(0, k, \lfloor m/2 \rfloor - 1) \cap \mathrm{Ann}(2^{j-1}, 2^j)$ with weight $> p_c$, and for $j \ge k + m$, let $\tilde{Y}_j$ be the number of outlets in $\mathcal{O}(k + m - 1, \infty, \lfloor m/2 \rfloor - 1) \cap \mathrm{Ann}(2^{j-1}, 2^j)$ with weight $> p_c$. Let $Y_j = \tilde{Y}_j - a_j$. By Theorem 5, there exist constants $C_1 < \infty$ and $\delta_1 > 0$ such that for all $k \ge 1$, $m > 4$,

$$\mathbb{P}(A_{k,m}) \ge 1 - C_1 \exp(-\delta_1 m),$$

where $A_{k,m}$ is the event that $X_j = Y_j$ for all $j \le k$ and for all $j \ge k + m$.

Because $A \in \Sigma^k$, there exists a Borel set $A' \subset \mathbb{R}^k$ such that A is the event that $(X_1, \dots, X_k) \in A'$. Similarly, because $B \in \Sigma_{k+m}$, there exists a Borel set $B' \subset \mathbb{R}^\infty$ (with the product topology) such that B is the event that $(X_{k+m}, \dots) \in B'$. Define A_Y as the event that $(Y_1, \dots, Y_k) \in A'$ and B_Y as the event that $(Y_{k+m}, \dots) \in B'$. Because A_Y and B_Y are independent,

$$|\mathbb{P}(A_Y \cap B_Y) - \mathbb{P}(A_Y)\mathbb{P}(B_Y)| = 0. \tag{3.4}$$

Also, when $A_{k,m}$ occurs, the events A and A_Y (resp., B and B_Y) are identical, so

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A_Y \cap B_Y)| \le \mathbb{P}(A_{k,m}^c) \le C_1 \exp(-\delta_1 m)$$

and

$$\begin{aligned} |\mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(A_Y)\mathbb{P}(B_Y)| &\le \mathbb{P}(A)|\mathbb{P}(B) - \mathbb{P}(B_Y)| \\ &\quad + \mathbb{P}(B_Y)|\mathbb{P}(A) - \mathbb{P}(A_Y)| \\ &\le |\mathbb{P}(B) - \mathbb{P}(B_Y)| + |\mathbb{P}(A) - \mathbb{P}(A_Y)| \\ &\le 2C_1 \exp(-\delta_1 m). \end{aligned}$$

Combining the two above inequalities with (3.4) gives the corollary. □

Now that we have a bound on the decay of the sequence $(\alpha(m))$, we can relate this to the decay of covariances using the following classical result.

COROLLARY 6 ([6], (2.2)). *Let $k, m \ge 1$ and let f and g be functions such that f is Σ^k-measurable and g is Σ_{k+m}-measurable. Suppose that $1/p + 1/q < 1$ and that the moments $\mathbb{E}|f|^p$ and $\mathbb{E}|g|^q$ exist. Then*

$$|\mathbb{E} fg - \mathbb{E} f \mathbb{E} g| \le 12[\mathbb{E}|f|^p]^{1/p}[\mathbb{E}|g|^q]^{1/q}[\alpha(m)]^{1-1/p-1/q}. \tag{3.5}$$

PROOF. For completeness, we will outline the proof in the Appendix. □

Corollaries 5 and 6 tell us that the variables (X_k) are very weakly dependent. This is one main ingredient for proving the CLT and SLLN for this sequence. In the first part of the following corollary, we will bound moments of the sums $(\sum_{k=1}^n X_k)_n$. This is the second main ingredient necessary for proving the CLT. The second part of the corollary will control fluctuations of the sums and will be useful in proving the SLLN.

COROLLARY 7. *The following statements hold.*

(1) *For each* $0 \le t \le 4$, *there exists* $D(t) < \infty$ *such that for all* $k \ge 1$ *and* $m \ge 0$,

$$\mathbb{E}\left|\sum_{j=k}^{k+m} X_j\right|^t \le D(t)m^{t/2}.$$

(2) *There exists* $C < \infty$ *such that for any* $\lambda > 0$ *and* $n \ge 1$,

$$\mathbb{P}\left(\max_{1\le i\le n}\left|\sum_{k=1}^{i} X_k\right| \ge \lambda\right) \le \frac{Cn}{\lambda^2} + \frac{C\sqrt{n}}{\lambda}.$$

PROOF. We will begin with the proof of the first statement. It suffices to consider $t = 4$ because for $t < 4$ we can use Jensen's inequality to reduce to this case. The statement will follow from Proposition 2.2 of [19], which we state below as Lemma 1. For the statement, we need some definitions. For $0 \le k < n$, define

$$c^{(13)}(k,n) = \max_{1\le x_1, x_2 = x_1 + k \le x_3 \le x_4 \le n} \mathbb{E}X_{x_1}X_{x_2}X_{x_3}X_{x_4}$$

and

$$c^{(31)}(k,n) = \max_{1\le x_1 \le x_2 \le x_3, x_4 = x_3 + k \le n} \mathbb{E}X_{x_1}X_{x_2}X_{x_3}X_{x_4}.$$

Also set

$$c(k;1,3) = \sup_{n\ge k}\left[c^{(13)}(k,n) + c^{(31)}(k,n)\right].$$

LEMMA 1. *Suppose that* $\sup_{k\ge 1}\mathbb{E}X_k^4 < \infty$ *and*

$$\sum_{k=0}^{m}(k+1)c(k;1,3) = O(m^\gamma) \qquad \text{as } m \to \infty \text{ for } \gamma \ge 0. \tag{3.6}$$

Then

$$\sup_{b\ge 0}\mathbb{E}(X_b + \cdots + X_{b+a})^4 = O(a^{2+\gamma}) \qquad \text{as } a \to \infty.$$

We make the choice $\gamma = 0$. The condition $\sup_{k\ge 1} \mathbb{E}X_k^4 < \infty$ holds from Theorem 4. As for (3.6), it is not difficult to see that it will hold as long as we show that there exist constants $C_1 < \infty$ and $\delta_1 > 0$ such that for any $m \ge 1$ and for any natural numbers $i_1, \dots, i_4$ such that the distance from i_1 to the set $\{i_2, i_3, i_4\}$ is at least equal to m,

$$|\mathbb{E}X_{i_1} \cdots X_{i_4}| \le C_1 \exp(-\delta_1 m). \tag{3.7}$$

Condition (3.7) holds by Corollary 6. To show this, suppose that $i_1 \le i_2 \le i_3 \le i_4$ (the other cases are handled similarly). We make the choices $f = X_{i_1}$ and $g = X_{i_2} X_{i_3} X_{i_4}$, with $p = 2$ and $q = 4$. From Theorem 4, there exists C_2 such that for all (i_j), both $(\mathbb{E}g^4)^{1/4} \le C_2$ and $(\mathbb{E}f^2)^{1/2} \le C_2$. Since $\mathbb{E}f = 0$, Corollary 6 gives

$$|\mathbb{E}X_{i_1} \cdots X_{i_4}| \le C_2^2 \alpha(m)^{1/4}.$$

Bounding $\alpha(m)$ using Corollary 5 shows (3.7) and completes the proof of the first statement of Corollary 7.

We now prove the second statement. It is the same as the proof of Lemma 2.2 in [6]. Let A_n be the event in the statement, and write $S_n = \sum_{k=1}^n X_k$. Let

$$A_n^i = \{|S_j| < \lambda \text{ for } j = 1, \dots, i-1 \text{ but } |S_i| \ge \lambda\}.$$

Similarly to the proof of Kolmogorov's maximal inequality for independent random variables, one can show that

$$\mathbb{P}(A_n) \le \frac{1}{\lambda^2}\left(\mathbb{E}S_n^2 + 2\sum_{i=1}^n \mathbb{E}\bigl[I[A_n^i]S_i(S_n - S_i)\bigr]\right). \tag{3.8}$$

By the first part of this corollary, $\mathbb{E}S_n^2 \le C_3 n$. Next, write the summand as

$$\mathbb{E}\bigl[I[A_n^i]X_i(S_n - S_i)\bigr] + \sum_{j=i+1}^n \mathbb{E}[I[A_n^i]S_{i-1}X_j].$$

The absolute value of the first term is bounded by

$$\left|\sum_{k=i+1}^n \mathbb{E}[I[A_n^i]X_i X_k]\right| \le \sum_{k=i+1}^n C_4[\alpha(k-i)]^{1/4} \le C_5,$$

where we use Corollary 6 with $f = I[A_n^i]X_i$ and $g = X_k$, with $p = 2$ and $q = 4$ (bounding the moments using Theorem 4) in the first inequality. For the second term we also use Corollary 6 but choose $f = I[A_n^i]S_{i-1}$ and $g = X_j$, with $p = 2$ and $q = 4$. This produces the bound

$$\begin{aligned} C_6 \sum_{j=i+1}^n (\mathbb{E}[S_{i-1}I[A_n^i]]^2)^{1/2}[\alpha(j-i)]^{1/4} &\le C_6 \lambda \sqrt{\mathbb{P}(A_n^i)} \sum_{j=1}^\infty [\alpha(j)]^{1/4} \\ &\le C_7 \lambda \sqrt{\mathbb{P}(A_n^i)}. \end{aligned}$$

Summing over i and using Jensen's inequality with the square root function (recalling that the events A_n^i are disjoint in i), we see that the sum in (3.8) is no bigger than

$$2C_5 n + 2C_7 \lambda n \left(\frac{1}{n} \sum_{i=1}^{n} \sqrt{\mathbb{P}(A_n^i)} \right) \leq C_8 n + C_9 \lambda \sqrt{n}.$$

Putting both this bound and the one on $\mathbb{E}S_n^2$ into (3.8) finishes the proof. □

The following corollary shows a way to construct a sequence of $c \log n$-dependent random variables $(\tilde{O}_k)$ related to (O_k). We will not use this sequence in the rest of the paper; however, the proofs of the CLT and the SLLN given in Section 4 can be replaced by ones that make reference to neither [6] nor [19] but that come from corresponding statements involving independent random variables by using the $\tilde{O}_k$'s. An example of such an approach is the proof of Theorem 1.4 in [4].

COROLLARY 8. *For any $\gamma > 0$, there exists $c < \infty$ such that for all $n \geq 1$, defining $m_n = c \log n$, with probability at least $1 - cn^{-\gamma}$, all random variables $O_{m_n+1}, \dots, O_n$ are equal to some random variables $\tilde{O}_{m_n+1}, \dots, \tilde{O}_n$, which are m_n-dependent and satisfy Theorem 4.*

PROOF. Let c be an integer to be chosen later and let $k \geq c \log n$. We define $\tilde{O}_k$ as the number of outlets in $\mathcal{O}(k-1, 1, \lfloor m_n/2 \rfloor - 1) \cap \mathrm{Ann}(2^{k-1}, 2^k)$ with weight $> p_c$. The reader may verify that exactly the same argument used in [4] for the proof of Theorem 4 applies to each $\tilde{O}_k$. Also, the variables $(\tilde{O}_k)$ are obviously m_n-dependent. By Theorem 5, there exist $C < \infty$ and $\beta > 0$ such that for any $k \geq c \log n$,

$$\mathbb{P}(\tilde{O}_k \neq O_k) \leq C n^{-\beta},$$

where $\beta \to \infty$ as $c \to \infty$. Therefore,

$$\mathbb{P}(\tilde{O}_k \neq O_k \text{ for some } k \in [c \log n, n]) \leq C n^{1-\beta}. \qquad \square$$

4. CLT and SLLN for the outlets.

4.1. *Proof of Theorem 1.* First we will show the statement about the a_k's. Theorem 4 implies the upper bound on a_k, so we need only show the lower bound. The proof is similar to the first part of Theorem 1.4 in [4]. For $k \geq 1$, let A_k be the event that (a) there is a p_{2^k}-closed circuit around the origin in $\mathrm{Ann}(2^{k-1}, 2^k)$, (b) there is a p_{2^k}-open circuit in $\mathrm{Ann}(2^{k-1}, 2^k)$ and (c) the circuit from (b) is connected by a p_{2^k}-open path to infinity. By the RSW theorem and (2.3), there exists $C_1 > 0$ such that for all k,

$$\mathbb{P}(A_k) > C_1.$$

But A_k implies the event $\{O_k \geq 1\}$, so

$$a_k = \mathbb{E}O_k \geq \mathbb{P}(A_k) \geq C_1.$$

We move on to the statement about $b(n)$. The upper bound follows from the case $t = 2$ of the first statement in Corollary 7, so we will focus on the lower bound. Let k be an integer between 1 and n and let $L_n := \log n$. For $i = 1, \ldots, 5$ define $q_k(i) = p_c + i(p_{2^k} - p_c)$. We define $A_{n,k}$ as the event that:

(1) there is an edge e_1 in $\mathrm{Ann}(2^{k+1}, 2^{k+2})$, with weight between $q_k(4)$ and $q_k(5)$, which is connected by a p_c-open path to a p_c-open circuit around the origin that is in $\mathrm{Ann}(2^k, 2^{k+1})$;

(2) the endpoints of e_1^* are connected by a $q_k(5)$-closed dual path in $\mathrm{Ann}(2^{k+1}, 2^{k+2})^*$ such that the union of this path and e_1^* encloses the origin;

(3) there is an edge e_2 in $\mathrm{Ann}(2^{k+2}, 2^{k+3})$, with weight in $[q_k(1), q_k(2)] \cup [q_k(3), q_k(4)]$, which is connected by a p_c-open path to an endpoint of e_1;

(4) the endpoints of e_2^* are connected by a $q_k(5)$-closed dual path in $\mathrm{Ann}(2^{k+2}, 2^{k+3})^*$ such that the union of this path and e_2^* encloses the origin;

(5) there is an edge e_3 in $\mathrm{Ann}(2^{k+3}, 2^{k+4})$ with weight in $[q_k(2), q_k(3)]$, which is connected by a p_c-open path to an endpoint of e_2;

(6) the endpoints of e_3^* are connected by a $q_k(5)$-closed dual path in $\mathrm{Ann}(2^{k+3}, 2^{k+4})^*$ such that the union of this path and e_3^* encloses the origin;

(7) an endpoint of e_3 is connected by a $q_k(1)$-open path to $\partial B(2^{k+L_n})$.

Notice that if $A_{n,k}$ occurs with edges $e_1 - e_3$, it cannot occur with any other edges. It follows from [5], Lemma 6.3, and RSW arguments (similar to the proof of [5], Corollary 6.2) that there exists $C_2 > 0$ such that for any n, k,

$$\mathbb{P}(A_{n,k}) \geq C_2. \tag{4.1}$$

Since, in addition, the events $A_{n,k}$ are L_n-dependent for fixed n, there exists $C_3 > 0$ such that

$$\mathbb{P}(A_{n,3k} \text{ occurs for at least } C_3 n \text{ values of } k \in [1, n/3]) \to 1 \quad \text{as } n \to \infty. \tag{4.2}$$

To see this, we will give the proof in Theorem 1.4 of [4]. Let j be an integer between 1 and L_n, and define $B_i^j = A_{n,3(j+iL_n)}$. Note that the events $(B_i^j)_{i=0}^{\lfloor n/3L_n \rfloor - 1}$ are independent. Therefore we may use Lemma 5.2 from [4]. Its proof is standard, so we omit it.

LEMMA 2. *Let $c > 0$. There exist $\alpha > 0$ and $\beta < 1$ depending on c with the following property. If Y_i are independent 0/1 random variables (not necessarily identically distributed) with $\mathbb{P}(Y_i = 1) > c$ for all i, then for all n,*

$$\mathbb{P}\left(\sum_{i=1}^{n} Y_i < \alpha n\right) < \beta^n.$$

In view of this lemma and (4.1), there exist $\alpha > 0$ and $\beta < 1$ such that for any n and $1 \le j \le L_n$,

$$\mathbb{P}\left(\sum_{i=0}^{\lfloor n/3L_n \rfloor - 1} I[B_i^j] < \frac{\alpha n}{3L_n}\right) < \beta^{n/3L_n}.$$

Therefore,

$$\begin{aligned}
&\mathbb{P}\left(\sum_{j=1}^{L_n} \sum_{i=0}^{\lfloor n/3L_n \rfloor - 1} I[B_i^j] < \alpha n/3\right) \\
&\quad \le \mathbb{P}\left(\sum_{i=0}^{\lfloor n/3L_n \rfloor - 1} I[B_i^j] < \alpha n/3L_n \text{ for some } j \in [1, L_n]\right) \\
&\quad \le L_n \beta^{n/3L_n},
\end{aligned}$$

which converges to 0 as $n \to \infty$. This proves (4.2).

Define $\tilde{A}_{n,k}$ the same way as we defined $A_{n,k}$ except that in item 7, the $q_k(1)$-open path connects e_3 to infinity. (See Figure 2 for an illustration of the event $\tilde{A}_{n,k}$.) Note that if $\tilde{A}_{n,k}$ occurs, then e_1 and e_3 are outlets, and e_2 is an outlet if and only if its weight is in $[q_k(3), q_k(4)]$. If $A_{n,k}$ occurs but $\tilde{A}_{n,k}$ does not, then there exists a $q_k(1)$-closed dual circuit around the origin with radius at least

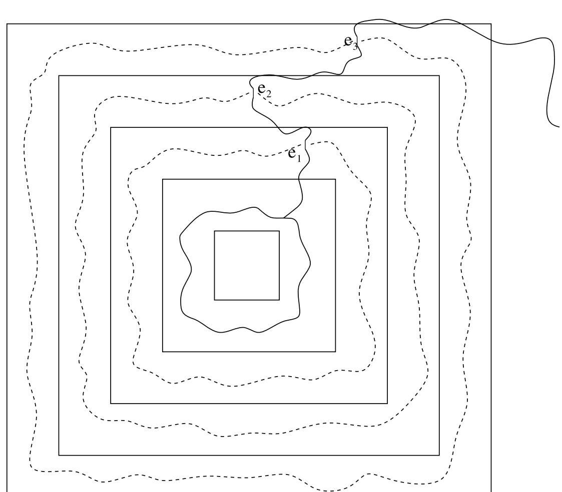

FIG. 2. *The event $\tilde{A}_{n,k}$. The boxes, in order from smallest to largest, are $B(2^{k+i})$ for $i = 0, \dots, 4$. The dotted paths are $q_k(5)$-closed, the path from e_3 to infinity is $q_k(1)$-open and all the other paths are p_c-open. The weight of e_1 is in $[q_k(4), q_k(5)]$, the weight of e_3 is in $[q_k(2), q_k(3)]$ and the weight of e_2 is in $[q_k(1), q_k(2)] \cup [q_k(3), q_k(4)]$. The edges e_1 and e_3 are outlets. The edge e_2 is an outlet if and only if its weight is in $[q_k(3), q_k(4)]$.*

2^{k+L_n}. By (2.4), there exist constants $C_4 < \infty$ and $C_5 > 0$ such that for all n, k, $\mathbb{P}(A_{n,k} \setminus \tilde{A}_{n,k}) \le C_4 \exp(-C_5 2^{k+L_n}/2^k)$, so

$$\mathbb{P}(A_{n,k} \setminus \tilde{A}_{n,k} \text{ occurs for some } k \in [1, n]) \to 0.$$

Therefore we may find $C_6 > 0$ such that for all n,

$$\mathbb{P}(\tilde{A}_{n,3k} \text{ occurs for at least } C_3 n \text{ values of } k \in [1, n/3]) > C_6.$$

Call A the above event whose probability is bounded below by C_6. On the event A, we define the vector $\vec{f} = (f_1, \ldots, f_{\lfloor C_3 n \rfloor})$ whose entries are the first $\lfloor C_3 n \rfloor$ edges e (ordered from distance to the origin) such that there exist edges $\bar{e}_1$ and $\bar{e}_3$ such that $\bar{e}_1$, e and $\bar{e}_3$ satisfy the properties of e_1, e_2 and e_3, respectively, in the definition of $A_{n,3k}$ for some $k \in [1, n/3]$. Write $O_{\vec{f}}(n)$ for the number of outlets that appear in the vector $\vec{f}$, and write $U_{\vec{f}}$ for the number of outlets in $B(2^n)$ that do not appear in $\vec{f}$. At least one of $\{U_{\vec{f}} + (C_2 n)/2 \ge a(n)\}$ or $\{U_{\vec{f}} + (C_2 n)/2 \le a(n)\}$ has probability at least $C_6/3$. Let us assume that it is the first event; if it is the other then the subsequent argument can be easily modified. Write $B = \{U_{\vec{f}} + (C_2 n)/2 \ge a(n)\}$. Since $U_{\vec{f}}$ is defined only on A, we have $B \subset A$.

Associated to each f_k in $\vec{f}$ in the definition of $A_{n,k}$ are two intervals $I_k(1) = [q_k(1), q_k(2)]$ and $I_k(2) = [q_k(3), q_k(4)]$. Let $\eta(\vec{f})$ be the configuration of weights outside of $\vec{f}$. If $\vec{f}$ and $\eta(\vec{f})$ are fixed, then the variable $U_{\vec{f}}$ is a constant function of the weights τ_{f_k}. Also, when these variables are fixed, $O_{\vec{f}}(n)$ is equal to the number of values of $k \in [1, \lfloor C_3 n \rfloor]$ such that $\tau_{f_k} \in I_k(2)$. Since the lengths of $I_k(1)$ and $I_k(2)$ are equal, the distribution of $O_{\vec{f}}(n)$ conditioned on $\vec{f}$, $\eta(\vec{f})$ and B is Binomial$(\lfloor C_3 n \rfloor, 1/2)$. If Y is an independent variable with this distribution, then

$$\begin{aligned}\mathbb{P}(|\mathbf{O}(n) - a_n| \ge \sqrt{n}) &\ge \mathbb{E}[\mathbb{P}(O_{\vec{f}}(n) \ge (C_3 n)/2 + \sqrt{n} \mid B, \vec{f}, \eta(\vec{f}))]\mathbb{P}(B) \\ &\ge (C_6/3)\mathbb{P}(Y \ge (C_3 n)/2 + \sqrt{n}),\end{aligned}$$

which is bounded below uniformly in n. This completes the proof.

4.2. *Proof of Theorem* 2.

PROOF OF THE CLT. We will apply Theorem 2.1 of [19]. To state that theorem, we need to introduce the notion of l-mixing. For $k \ge 0$, $n \ge 1$ and $u \in \mathbb{R}$, set

$$l_n(k, u) = \max_{1 \le j \le n-k} \sup |\mathbb{E}[e^{iuP} e^{-iuF}] - \mathbb{E}e^{iuP}\mathbb{E}e^{-iuF}|, \tag{4.3}$$

where

$$P = b(n)^{-1} \sum_{l=1}^{j} \delta_l X_l, \qquad F = b(n)^{-1} \sum_{l=j+k}^{n} \delta_l X_l,$$

and the supremum in (4.3) is over all $\{\delta_l = 0 \text{ or } 1\}$. Now for $k \geq 0$ and $u \in \mathbb{R}$, set

$$l(k,u) = \sup_{n \geq 1} l_n(k,u).$$

The sequence (X_k) is called *l-mixing* if for all real u, $l(k,u) \to 0$ as $k \to \infty$.

REMARK 3. As mentioned in the discussion below Definition 2.2 in [19], the inequality

$$l(k,u) \leq 16\alpha(k) \tag{4.4}$$

from page 307 in [10] holds for all $k \geq 0$ and $u \in \mathbb{R}$, so since $\alpha(k) \to 0$ as $k \to \infty$, the sequence (X_k) of outlet variables is l-mixing.

For $k \geq 0$, define

$$\tilde{c}(k) = \sup_{j \geq 1} |\mathbb{E} X_j X_{j+k}|.$$

The following is Theorem 2.1 of [19].

LEMMA 3. *The following conditions are sufficient for*

$$\frac{\sum_{k=1}^n X_k}{b(n)} \Rightarrow N(0,1).$$

For some $\varepsilon > 0$ and $\gamma \geq 0$,

$$\sup_{a \geq 1} \mathbb{E}\left|\sum_{k=a}^{a+b} X_k\right|^{2+\varepsilon} = O(b^{1+\varepsilon/2+\gamma}) \qquad \text{as } b \to \infty; \tag{4.5}$$

the sequence (X_k) is l-mixing and for all real u,

$$l(k,u) = o(k^{-\theta}) \qquad \text{as } k \to \infty, \text{ where } \theta = 2\gamma/\varepsilon, \tag{4.6}$$

and

$$b(n) \to \infty \qquad \text{as } n \to \infty \quad \text{and} \quad \sum_{j=0}^{\infty} \tilde{c}(j) < \infty. \tag{4.7}$$

To prove the CLT, we simply need to verify the conditions of Lemma 3. Condition (4.5) holds with $\varepsilon = 2$ and $\gamma = 0$ by the first part of Corollary 7, using $t = 4$. Using (4.4) and Corollary 5, we see that condition (4.6) holds. Also, the first part of Corollary 7 with $t = 2$ shows the first part of condition (4.7). Finally, to verify the second part of (4.7), we appeal to Corollary 6 using $f = X_m$ and $g = X_{m+k}$ (for fixed $m \geq 1$ and $k \geq 0$), with $p = 2$ and $q = 4$. It follows that

$$\tilde{c}(k) \leq C_1 \alpha(k)^{1/4}$$

for some $C_1 < \infty$. In view of Corollary 5, this proves the second part of (4.7) and completes the proof of the CLT. □

PROOF OF THE SLLN. For $i \ge 1$, take $n_i = 2^i$. The second statement of Corollary 7 implies that for any $\varepsilon > 0$,

$$\mathbb{P}\Big(\max_{n_i \le j \le n_{i+1}} |\mathbf{O}(j) - a(j)| \ge \varepsilon n_i^r\Big) \le \frac{C}{\varepsilon^2}\left(\frac{n_{i+1}}{n_i^{2r}} + \frac{\sqrt{n_{i+1}}}{n_i}\right)$$
$$\le C_1\left(\frac{1}{(2^{2r-1})^i} + \frac{1}{(2^{r-1/2})^i}\right).$$

Since $r > 1/2$, this probability is summable in i. Since the function n^r is monotone, it follows that

$$\sum_{i=1}^{\infty} \mathbb{P}\left(\max_{n_i \le j \le n_{i+1}} \frac{|\mathbf{O}(j) - a(j)|}{j^r} \ge \varepsilon\right) < \infty.$$

The Borel–Cantelli lemma finishes the proof. □

5. Further results for invasion percolation.

We begin with a lemma.

LEMMA 4. *There exist constants $C < \infty$ and $\alpha > 0$ such that for all $m, n \ge 1$,*

$$\mathbb{P}\big(\mathbf{O}(n, n+m) \le \alpha m\big) \le C \exp(-m^\alpha),$$

where $\mathbf{O}(n, n+m)$ is the number of outlets in $\mathrm{Ann}(2^n, 2^{n+m})$.

PROOF. The proof of the lower bound in Theorem 1.4 of [4] shows the case $n = 1$. For general n the proof is similar. For $i, m \ge 1$, let $G_{i,m}$ be the event that there is no p_{2^i}-closed dual circuit around the origin with radius larger than $2^{i+\log m}$, and let $K_{i,m}$ be the event that (a) there exists a p_{2^i}-closed dual circuit $\mathcal{C}$ around the origin in $\mathrm{Ann}(2^i, 2^{i+1})^*$, (b) there exists a p_c-open circuit $\mathcal{C}'$ around the origin in $\mathrm{Ann}(2^i, 2^{i+1})$ and (c) the circuit $\mathcal{C}'$ is connected to $\partial B(2^{i+\log m})$ by a p_{2^i}-open path. By the RSW theorem and (2.3), there exists $C_1 > 0$ such that for all $i \ge 0$ and $m \ge 1$,

$$\mathbb{P}(K_{i,m}) \ge C_1.$$

Now let j be an integer between 1 and $\log m$, and define the event $K_{i,m}^j = K_{n+j+i\lfloor \log m\rfloor, m}$. Note that for fixed j, the events $(K_{i,m}^j)_{i=0}^{\lfloor m/\log m\rfloor - 1}$ are independent. Therefore we can apply Lemma 2 to deduce that there exist $\alpha_0 > 0$ and $\beta_0 < 1$ such that for any m, n and $1 \le j \le \log m$,

$$\mathbb{P}\left(\sum_{i=0}^{\lfloor m/\log m\rfloor - 1} I[K_{i,m}^j] < \frac{\alpha_0 m}{\lfloor \log m\rfloor}\right) < \beta_0^{m/\log m}.$$

Therefore,

$$\mathbb{P}\left(\sum_{j=1}^{\lfloor \log m \rfloor} \sum_{i=0}^{\lfloor m/\log m \rfloor - 1} I[K_{i,m}^j] < \alpha_0 m\right)$$
$$\leq \mathbb{P}\left(\sum_{i=0}^{\lfloor m/\log m \rfloor - 1} I[K_{i,m}^j] < \alpha_0 m/\lfloor \log m \rfloor \text{ for some } j \in [1, \log m]\right)$$
$$\leq \log m \beta_0^{m/\log m} \leq C_2 \exp(-m^\alpha)$$

for some $C_2 < \infty$ and $\alpha > 0$. By (2.4), we also have the estimate

$$\sum_{i=0}^{m} \mathbb{P}(G_{n+i,m}^c) \leq C_3 \sum_{i=0}^{m} \exp(-C_4 m) \leq C_3 \exp(-C_5 m).$$

Since the event $K_{i,m} \cap G_{i,m}$ implies $O_{i+1} \geq 1$, we can combine the above estimates to deduce

$$\mathbb{P}\big(\mathbf{O}(n, n+m) \leq \alpha_0 m\big) \leq C_2 \exp(-m^\alpha) + C_3 \exp(-C_5 m),$$

which implies the lemma. $\square$

Recall the definitions of Q_n and T_n from Section 1.3. Since $a(n) \asymp n$, T_n is comparable with n.

PROOF OF THEOREM 3. It follows from the definition of Q_n, the CLT for $\mathbf{O}(n)$ and the fact that for any x, $\sigma_{n+x\sqrt{n}}/\sigma_n \to 1$ as $n \to \infty$ that

$$\mathbb{P}(Q_n < T_{n+x\sigma_n}) \to \Phi(x),$$

where Φ is the standard normal cumulative distribution function. Recall that the a_i's [$a_i = a(i) - a(i-1)$] are uniformly bounded away from 0 and ∞ by Theorem 1. Therefore,

$$\mathbb{P}(Q_n < T_{n+x\sigma_n}) = \mathbb{P}\big(a(Q_n) < a(T_{n+x\sigma_n})\big) = \mathbb{P}\big(a(Q_n) < n + x\sigma_n + r_n\big),$$

where r_n is uniformly bounded in n. It remains to prove the second part of the proposition. The first statement implies that for any $M > 0$,

$$\mathbb{E} \min\left\{\left(\frac{a(Q_n) - n}{\sigma_n}\right)^2, M\right\} \to \mathbb{E} \min\{Z^2, M\},$$

where Z is a standard normal random variable. Therefore, it suffices to show that for any $\varepsilon > 0$ there exists $C_1 > 0$ such that

$$\limsup_{n\to\infty} \mathbb{E}\left(\frac{a(Q_n) - n}{\sigma_n}\right)^2 I\big(|a(Q_n) - n| > C_1 \sigma_n\big) < \varepsilon.$$

This will follow if we show that there exists C_2 such that for all n,

$$\mathbb{E}\left(\frac{a(Q_n)-n}{\sigma_n}\right)^4 < C_2.$$

In other words, we need to show that $\mathbb{E}(a(Q_n)-n)^4 = O(n^2)$. Since the a_i's are uniformly bounded away from 0 and ∞, it suffices to show that $\mathbb{E}(Q_n - T_n)^4 = O(n^2)$. For $c > 0$, consider

$$A_n = \{\mathbf{O}(n, n+k) > ck, \mathbf{O}(n-k, n) > ck \text{ for all } k \geq \sqrt{n}\}.$$

It follows from Lemma 4 that there exists $c > 0$ such that

$$\mathbb{P}(A_n^c) \leq C_3 \exp(-n^c).$$

We write

$$\mathbb{E}(Q_n - T_n)^4 = \mathbb{E}(Q_n - T_n)^4 I(Q_n \leq C_4 n) + \mathbb{E}(Q_n - T_n)^4 I(Q_n > C_4 n).$$

If C_4 is large enough, $\mathbb{E}(Q_n - T_n)^4 I(Q_n > C_4 n) = o(n^2)$ (One can write $I(Q_n > C_4 n)$ as $\sum_{k=1}^{\infty} I(Q_n \in (kC_4 n, (k+1)C_4 n])$ and use Lemma 4). We now bound the first expectation.

$$\begin{aligned}
&\mathbb{E}(Q_n - T_n)^4 I(Q_n \leq C_4 n) \\
&\quad \leq (C_4 n)^4 \mathbb{P}(A_{T_n}^c) + T_n^2 + \mathbb{E}(Q_n - T_n)^4 I\big(|Q_n - T_n| > \sqrt{T_n}, A_{T_n}\big).
\end{aligned}$$

The first two summands are bounded by $C_5 n^2$. It remains to bound the last summand

$$\begin{aligned}
&\mathbb{E}(Q_n - T_n)^4 I\big(Q_n - T_n > \sqrt{T_n}, A_{T_n}\big) \\
&\quad \leq \mathbb{E}(Q_n - T_n)^4 I\big(Q_n > T_n, \mathbf{O}(T_n, Q_n - 1) > c(Q_n - 1 - T_n)\big) \\
&\quad \leq \frac{8}{c^4}\mathbb{E}\big(\mathbf{O}(Q_n - 1) - \mathbf{O}(T_n)\big)^4 I(Q_n > T_n) + 8 \\
&\quad \leq \frac{8}{c^4}\mathbb{E}\big(\mathbf{O}(T_n) - n\big)^4 + 8 \\
&\quad \leq C_6 n^2,
\end{aligned}$$

where the last inequality follows from Corollary 7. Similarly, one can show that $\mathbb{E}(Q_n - T_n)^4 I(Q_n - T_n < -\sqrt{T_n}, A_{T_n}) \leq C_7 n^2$. $\square$

PROOF OF COROLLARY 1. It follows from Theorem 1 that $a(T_n) = n + O(1)$, $\sigma_n \asymp \sqrt{n}$ and $|a(m) - a(n)| \asymp |m - n|$ independently of m, n. Therefore, the first statement of Corollary 1 follows directly from Theorem 3. The upper bound in the second statement follows immediately from the upper bound in the first statement.

For the lower bound, we may apply the CLT for $(a(Q_n))$ to deduce that there exists $C > 0$ such that for all n,

$$\mathbb{P}(Q_n \geq T_n + \sqrt{n}) > C \quad \text{and} \quad \mathbb{P}(Q_n \leq T_n - \sqrt{n}) > C.$$

The lower bound follows from these two estimates. Indeed, if $\mathbb{E}Q_n \geq T_n$, then $Q_n \leq T_n - \sqrt{n}$ implies that $Q_n \leq \mathbb{E}Q_n - \sqrt{n}$ and so

$$\mathbb{E}(Q_n - \mathbb{E}Q_n)^2 \geq n\mathbb{P}(Q_n \leq T_n - \sqrt{n}) > Cn.$$

If $\mathbb{E}Q_n \leq T_n$, then the argument is similar. $\square$

PROOF OF COROLLARY 2. The proofs of both statements are similar so we only show the proof of the first. We first prove the lower bound. The CLT for $(a(Q_n))$ implies that there exists C_1 such $\mathbb{P}(Q_n > T_n + \sqrt{n}) > C_1$. It is obvious that $\hat{R}_n \geq 2^{Q_n - 1}$. Therefore, $\mathbb{P}(\hat{R}_n \geq 2^{T_n + \sqrt{n} - 1}) > C_1$, which implies that $\mathbb{E}(\log \hat{R}_n - T_n)^2 \geq (\sqrt{n} - 1)^2 C_1$.

We now prove the upper bound. We first observe that by Theorem 4, using $t = 4$,

$$\begin{aligned} \mathbb{P}(\hat{R}_n < 2^{\sqrt{n}}) &\leq \mathbb{P}(Q_n \leq \sqrt{n}) \leq \mathbb{P}(\mathbf{O}(\sqrt{n}) \geq n) \\ &\leq \mathbb{E}\mathbf{O}(\sqrt{n})^4 / n^4 = O(n^{-2}). \end{aligned} \tag{5.1}$$

Therefore, $\mathbb{E}(\log \hat{R}_n - T_n)^2 I(\hat{R}_n < 2^{\sqrt{n}}) = o(n)$. We next rule out the case when $\hat{R}_n > 2^{C_2 n}$ for large enough C_2.

$$\begin{aligned} \mathbb{E}(\log \hat{R}_n - T_n)^2 I(\hat{R}_n > 2^{C_2 n}) &\leq \mathbb{E}(\log \hat{R}_n)^2 I(\hat{R}_n > 2^{C_2 n}) \\ &\leq \sum_{k=1}^{\infty} (C_2 n(k+1))^2 \mathbb{P}(\hat{R}_n > 2^{C_2 nk}). \end{aligned}$$

Note that $\mathbb{P}(\hat{R}_n > 2^{C_2 nk})$ is bounded above by

$$\begin{aligned} &\mathbb{P}(\text{there is no } p_c\text{-open circuit around the origin in } \operatorname{Ann}(2^{C_2 nk - \sqrt{C_2 nk}}, 2^{C_2 nk})) \\ &\quad + \mathbb{P}(Q_n > C_2 nk - \sqrt{C_2 nk}). \end{aligned}$$

Using the RSW theorem and Lemma 4 if C_2 is large enough, this gives the bound

$$\mathbb{P}(\hat{R}_n > 2^{C_2 nk}) \leq C_3 \exp(-(nk)^{C_4}). \tag{5.2}$$

Therefore,

$$\mathbb{E}(\log \hat{R}_n - T_n)^2 I(\hat{R}_n > 2^{C_2 n}) = o(n).$$

Let $\tilde{A}_n$ be the event that there exists a p_c-open circuit around the origin in $\operatorname{Ann}(2^{k - \sqrt{k}}, 2^k)$ for all $k \geq \sqrt{n}$. It follows from the RSW theorem that

$$\mathbb{P}(\tilde{A}_n^c) \leq C_5 \exp(-n^{C_6}). \tag{5.3}$$

Therefore,

$$\mathbb{E}(\log \hat{R}_n - T_n)^2 I(\hat{R}_n \le 2^{C_2 n}, \tilde{A}_n^c) = o(n).$$

Moreover, if $\hat{R}_n > 2^{\sqrt{n}}$ and $\tilde{A}_n$ occurs, then $Q_n \ge \log \hat{R}_n - \sqrt{\log \hat{R}_n} - 1$. Hence

$$\begin{aligned}
&\mathbb{E}(\log \hat{R}_n - T_n)^2 I\bigl(2^{\sqrt{n}} \le \hat{R}_n \le 2^{C_2 n}, \tilde{A}_n\bigr) \\
&\quad \le 2\mathbb{E}(Q_n - T_n)^2 + 2\mathbb{E}(\log \hat{R}_n - Q_n)^2 I\bigl(Q_n > \log \hat{R}_n - \sqrt{C_2 n} - 1\bigr) \\
&\quad \le C_7 n.
\end{aligned}$$

The last inequality follows from Corollary 1 and from the fact that $Q_n \le \log \hat{R}_n + 1$. The upper bound is proved. □

PROOF OF COROLLARY 3. For any n, let

$$f(n) = \max\{\tau_e : e \text{ is an outlet in } B(n)^c\}$$

and $g(n) = \min\{\tau_e : e \text{ is an outlet in } B(n)\}$ if there is an outlet in $B(n)$ and $g(n) = 0$ otherwise.

LEMMA 5. *There exists $C < \infty$ such that for any $t \ge 1$ and $n \ge 1$,*

$$\mathbb{E}\biggl(\biggl|\log \frac{f(n) - p_c}{p_n - p_c}\biggr|^t\biggr) \le (Ct)^{Ct} \quad \textit{and} \quad \mathbb{E}\biggl(\biggl|\log \frac{g(n) - p_c}{p_n - p_c}\biggr|^t\biggr) \le (Ct)^{Ct}.$$

PROOF. Using the RSW theorem and (2.4), respectively, we see that there exist constants $C_1 < \infty$ and $C_2 > 0$ such that for any $n \ge 1$ and $p \in (0,1)$,

$$\mathbb{P}\bigl(f(n) < p\bigr) \le \mathbb{P}\bigl(B(n) \overset{p}{\leftrightarrow} \infty\bigr) \le C_1 \biggl(\frac{n}{L(p)}\biggr)^{C_2}$$

and

$$\mathbb{P}\bigl(f(n) \ge p\bigr) \le \mathbb{P}(B_{n,p}) \le C_1 \exp\biggl(-C_2 \frac{n}{L(p)}\biggr),$$

where $B_{n,p}$ is defined directly above (2.4). For $k \in \mathbb{Z}$, let $q_k = p_{2^k n}$

$$\begin{aligned}
&\mathbb{E}\biggl(\biggl|\log \frac{f(n) - p_c}{p_n - p_c}\biggr|^t\biggr) \\
&\quad = \sum_k \mathbb{E}\biggl(\biggl|\log \frac{f(n) - p_c}{p_n - p_c}\biggr|^t I\bigl(f(n) \in [q_{k+1}, q_k)\bigr)\biggr) \\
&\quad \le \sum_{k \ge 0} \biggl|\log \frac{q_{k+1} - p_c}{p_n - p_c}\biggr|^t \mathbb{P}\bigl(f(n) < q_k\bigr) \\
&\qquad + \sum_{k < 0} \biggl|\log \frac{q_k - p_c}{p_n - p_c}\biggr|^t \mathbb{P}\bigl(f(n) \ge q_{k+1}\bigr).
\end{aligned}$$

The first result of the lemma follows from (2.5) and the above estimates. It remains to prove the second statement. Note that for any $n \geq 1$ and $p \in (0, 1)$,

$$\mathbb{P}\big(g(n) < p\big) \leq \mathbb{P}\big(B(n) \overset{p}{\leftrightarrow} \infty\big) \leq C_1 \left(\frac{n}{L(p)} \right)^{C_2}.$$

To bound $\mathbb{P}(g(n) \geq p)$, note that if $g(n) \geq p$, then there is an outlet in $B(n)$. For $1 \leq m \leq \lfloor \log n \rfloor + 1$, consider the event $A_{m,n}$ that $\mathrm{Ann}(\lfloor n/2^m \rfloor, n)$ contains an outlet. (For the case $m = \lfloor \log n \rfloor + 1$, we use the convention that $\mathrm{Ann}(\lfloor n/2^m \rfloor, n) = B(n)$.) Note that for $m = 0$, $A_{m,n}$ is equal to the null event and that for fixed n, the events $A_{m,n}$ are increasing in m. By Lemma 4, there exists $C_3 < \infty$ and $C_4 > 0$ such that for all m, n,

$$\mathbb{P}(A_{m,n}^c) \leq C_3 \exp(-m^{C_4}).$$

Using this estimate, we get

$$\begin{aligned}
\mathbb{P}\big(g(n) \geq p\big) &= \sum_{m=0}^{\lfloor \log n \rfloor} \mathbb{P}\big(g(n) \geq p, A_{m,n}^c, A_{m+1,n}\big) \\
&\leq \sum_{m=0}^{\lfloor \log n \rfloor} \mathbb{P}\big(B_{\lfloor n/2^{m+1} \rfloor, p}, A_{m,n}^c\big) \\
&\leq \sum_{m=0}^{\lfloor \log n \rfloor} \mathbb{P}\big(B_{\lfloor n/2^{m+1} \rfloor, p}\big)^{1/2} \mathbb{P}(A_{m,n}^c)^{1/2} \\
&\leq C_5 \sum_{m=0}^{\lfloor \log n \rfloor} \left[\exp\left(-C_2 \frac{\lfloor n/2^{m+1} \rfloor}{L(p)} \right) \exp(-m^{C_4}) \right]^{1/2}
\end{aligned}$$

for some $C_5 < \infty$. In particular, for $k < 0$,

$$\mathbb{P}\big(g(n) \geq q_k\big) \leq C_6 \exp(-|k|^{C_7}).$$

The remainder of the proof of the lemma is similar to the proof of the first statement. $\square$

We proceed with the proof of the corollary. We will only prove the first statement; the proof of the second is similar. Inequality (2.5) and Corollary 1 imply that

$$\mathbb{E}\left(\log \frac{p_{2^{Q_n}} - p_c}{p_{2^{T_n}} - p_c} \right)^2 \asymp \mathbb{E}(Q_n - T_n)^2 \asymp n.$$

Note that

$$g(2^{Q_n}) \leq \hat{\tau}_n \leq f(2^{Q_n - 1}).$$

Therefore the corollary will follow if we show that there exists C_8 such that for all n,

$$\mathbb{E}\biggl(\log \frac{g(2^{Q_n}) - p_c}{p_{2^{Q_n}} - p_c}\biggr)^2 \le C_8\sqrt{n} \quad \text{and} \quad \mathbb{E}\biggl(\log \frac{f(2^{Q_n-1}) - p_c}{p_{2^{Q_n}} - p_c}\biggr)^2 \le C_8\sqrt{n}.$$

Let D_n be the event that (a) there exists a p_c-open circuit in the annulus $\mathrm{Ann}(2^{n-n^{1/4}}, 2^{n-1})$, (b) this circuit is connected to infinity by a $p_{2^{n-2n^{1/4}}}$-open path and (c) there exists a $p_{2^{n+n^{1/4}}}$-closed dual circuit around $B(2^n)^*$. The RSW theorem and (2.4) imply that there exist constants C_9 and C_{10} such that for all n,

$$\mathbb{P}(D_n^c) \le C_9 e^{-n^{C_{10}}}. \tag{5.4}$$

Recall that for all n,

$$\mathbb{E}\biggl(\log \frac{g(2^n) - p_c}{p_{2^n} - p_c}\biggr)^4 \le C_{11} \quad \text{and} \quad \mathbb{E}\biggl(\log \frac{f(2^{n-1}) - p_c}{p_{2^n} - p_c}\biggr)^4 \le C_{11}.$$

Therefore,

$$\mathbb{E}\biggl(\log \frac{g(2^{Q_n}) - p_c}{p_{2^{Q_n}} - p_c}\biggr)^2 I(D_{Q_n}^c) \le \sum_{k=1}^{\infty} \mathbb{E}\biggl(\log \frac{g(2^k) - p_c}{p_{2^k} - p_c}\biggr)^2 I(D_k^c) \le C_{12},$$

where D_{Q_n} is the event $\bigcup_k (D_k \cap \{Q_n = k\})$. Similarly,

$$\mathbb{E}\biggl(\log \frac{f(2^{Q_n-1}) - p_c}{p_{2^{Q_n}} - p_c}\biggr)^2 I(D_{Q_n}^c) \le C_{12}.$$

On the other hand, if D_n occurs and, moreover, there is an outlet in the annulus $\mathrm{Ann}(2^{n-1}, 2^n)$, then

$$g(2^n) \text{ and } f(2^{n-1}) \text{ are both in } [p_{2^{n+n^{1/4}}}, p_{2^{n-2n^{1/4}}}].$$

This observation and inequality (2.5) imply [note that $\mathrm{Ann}(2^{Q_n-1}, 2^{Q_n})$ always contains an outlet]

$$\mathbb{E}\biggl(\log \frac{g(2^{Q_n}) - p_c}{p_{2^{Q_n}} - p_c}\biggr)^2 I(D_{Q_n}) \le C_{13}\mathbb{E}Q_n^{1/2} \le C_{14}\sqrt{n}$$

and, similarly,

$$\mathbb{E}\biggl(\log \frac{f(2^{Q_n-1}) - p_c}{p_{2^{Q_n}} - p_c}\biggr)^2 I(D_{Q_n}) \le C_{14}\sqrt{n}.$$

This completes the proof of the corollary. $\square$

PROOF OF COROLLARY 4. We start with the proof of the first statement. Take $r > 1/2$. The SLLN for outlets gives that $(\mathbf{O}(n) - a(n))/n^r \to 0$ a.s. as $n \to \infty$.

Theorem 1.4 in [4] states that there are constants $C_1 > 0$ and $C_2 < \infty$ such that with probability 1, for all large n,

$$C_1 n < \mathbf{O}(n) < C_2 n.$$

This implies that there exist constants $C_3 > 0$ and $C_4 < \infty$ such that with probability 1, for all large n,

$$C_3 n < Q_n < C_4 n.$$

Therefore,

$$\frac{\mathbf{O}(Q_n) - a(Q_n)}{n^r} \to 0 \qquad \text{a.s. as } n \to \infty.$$

Because $a(Q_n) - n \asymp Q_n - T_n$, the first statement of the corollary will follow if we show that $(\mathbf{O}(Q_n) - n)/n^r \to 0$ a.s. Note that $n \le \mathbf{O}(Q_n) \le n + O_{Q_n}$ by the definition of Q_n. Since there exists a finite constant C_5 such that (a) $Q_n < C_4 n$ a.s. for all large n and (b) $\mathbb{P}(O_i > n^{r/2}$ for some $i = 1, \dots, C_4 n) \le C_5/n^2$ (this second statement is a consequence of Theorem 4), it follows that, a.s. for all large n, $O_{Q_n} \le n^{r/2}$. The desired convergence follows.

The second and third statements follow easily from the first and from estimates developed in the proofs of Corollaries 2 and 3. Indeed, since $(Q_n - T_n)/n^r \to 0$ a.s., the statements about $\log \hat{R}_n$ and $\hat{\tau}_n$ will follow if we show that

$$\frac{\log \hat{R}_n - Q_n}{n^r} \to 0 \quad \text{and} \quad \frac{1}{n^r} \log \frac{\hat{\tau}_n - p_c}{p_{2^{Q_n}} - p_c} \to 0 \qquad \text{a.s.} \tag{5.5}$$

It follows from the proof of Corollary 2 and the Borel–Cantelli lemma that there exists $C_6 < \infty$ such that, a.s., for all large n,

$$\log \hat{R}_n - \sqrt{C_6 n} - 1 \le Q_n \le \log \hat{R}_n + 1. \tag{5.6}$$

To see this, note first that by (5.1), with probability one, $\log \hat{R}_n \ge \sqrt{n}$ for all large n. Next, let $\tilde{A}_n$ be the event that for all $k \ge \sqrt{n}$, there is a p_c-open circuit around the origin in the annulus $\mathrm{Ann}(2^{k-\sqrt{k}}, 2^k)$. By (5.3), with probability one the events $(\tilde{A}_n)$ occur for all large n. Last, by (5.2) (setting $k = 1$ there), there exists $C_6 < \infty$ such that with probability one, for all large n, $\log \hat{R}_n \le C_6 n$. Since $\tilde{A}_n \cap \{\sqrt{n} < \log \hat{R}_n < C_6 n\}$ implies (5.6), it in fact occurs a.s. for all large n. This implies the desired SLLN for $(\log \hat{R}_n)$.

Similarly, one may use arguments from the proof of Corollary 3 and the Borel–Cantelli lemma to show that a.s., for all large n,

$$p_{2^{Q_n + (C_4 n)^{1/4}}} \le \hat{\tau}_n \le p_{2^{Q_n - 2(C_4 n)^{1/4}}}. \tag{5.7}$$

To prove this, define D_n as in the proof of that corollary: it is the event that (a) there exists a p_c-open circuit in $\mathrm{Ann}(2^{n-n^{1/4}}, 2^{n-1})$, (b) this circuit is connected to infinity by a $p_{2^{n-2n^{1/4}}}$-open path and (c) there exists a $p_{2^{n+n^{1/4}}}$-closed dual circuit around $B(2^n)^*$. By (5.4), a.s. D_n occurs for all large n. The fact that if D_n occurs and there is an outlet in $\mathrm{Ann}(2^{n-1}, 2^n)$, then $g(2^n)$ and $f(2^{n-1})$ are in the interval $[p_{2^{n+n^{1/4}}}, p_{2^{n-2n^{1/4}}}]$, combined with the fact that $\mathrm{Ann}(2^{Q_n-1}, 2^{Q_n})$ always contains an outlet, shows (5.7) a.s. for all large n. Along with (2.5), this implies the second part of (5.5) and completes the proof of Corollary 4. $\square$

APPENDIX: COVARIANCE ESTIMATES

Here we give the proof of Corollary 6. The proof we present is directly from [6]. We begin with a lemma, which is (17.2.2) from [10].

LEMMA 6. *Suppose that f is Σ^k-measurable, and g is Σ_{k+m}-measurable, and there are constants $C_1, C_2 < \infty$ such that $|f| \le C_1$ and $|g| \le C_2$ a.s. Then*

$$|\mathbb{E}[fg] - \mathbb{E}f\mathbb{E}g| \le 4C_1C_2\alpha(m), \tag{A.1}$$

where $\alpha(m)$ was defined in (3.2).

PROOF. We write the left-hand side of (A.1) as

$$\big|\mathbb{E}\big[f\mathbb{E}[g - \mathbb{E}g \mid \Sigma^k]\big]\big| \le C_1\mathbb{E}\big[\big|\mathbb{E}[g - \mathbb{E}g \mid \Sigma^k]\big|\big] = C_1\mathbb{E}\big[f_1\mathbb{E}[g - \mathbb{E}g \mid \Sigma^k]\big],$$

where $f_1 = \mathrm{sgn}(\mathbb{E}[g - \mathbb{E}g \mid \Sigma^k])$. Since f_1 is Σ^k-measurable,

$$|\mathbb{E}[fg] - \mathbb{E}f\mathbb{E}g| \le C_1|\mathbb{E}[f_1g] - \mathbb{E}f_1\mathbb{E}g|.$$

Similarly comparing g to $g_1 = \mathrm{sgn}(\mathbb{E}[g - \mathbb{E}g \mid \Sigma_{k+m}])$,

$$|\mathbb{E}[fg] - \mathbb{E}f\mathbb{E}g| \le C_1C_2|\mathbb{E}[f_1g_1] - \mathbb{E}f_1\mathbb{E}g_1|.$$

Define $A = \{f_1 = 1\}$ and $B = \{g_1 = 1\}$. Then the right-hand side of the above inequality is bounded above by

$$\begin{aligned} C_1C_2|&\mathbb{P}(A, B) + \mathbb{P}(A^c, B^c) - \mathbb{P}(A^c, B) - \mathbb{P}(A, B^c) \\ &- \mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(A^c)\mathbb{P}(B^c) + \mathbb{P}(A^c)\mathbb{P}(B) + \mathbb{P}(A)\mathbb{P}(B^c)|, \end{aligned}$$

which is bounded above by $4C_1C_2\alpha(m)$. $\square$

Now we will suppose that one function is bounded and the other is in L^p for $p > 1$. The following is Lemma 2.1 from [6].

LEMMA 7. *Suppose that f is Σ^k-measurable, and g is Σ_{k+m}-measurable and that there exists $C < \infty$ such that $|g| \le C$ a.s. Further, suppose that there is $p > 1$ such that the moment $\mathbb{E}|f|^p < \infty$ exists. Then*

$$|\mathbb{E}[fg] - \mathbb{E}f\mathbb{E}g| \le 6C[\mathbb{E}|f|^p]^{1/p}\alpha(m)^{1/q}, \tag{A.2}$$

where $1/p + 1/q = 1$.

PROOF. Let N be a positive number to be chosen later and set $f_N = fI[|f| < N]$. Applying the previous lemma to f_N and g, we get

$$|\mathbb{E}[f_N g] - \mathbb{E} f_N \mathbb{E} g| \le 4CN\alpha(m). \tag{A.3}$$

To estimate the difference between this and the quantity in this lemma, note that the left-hand side of (A.2) is bounded above by

$$|\mathbb{E}[f_N g] - \mathbb{E} f_N \mathbb{E} g| + |\mathbb{E}[\tilde{f}_N g] - \mathbb{E}\tilde{f}_N \mathbb{E} g|,$$

where $\tilde{f}_N = f - f_N$. Since $|g| \le C$, we find that the second term is no bigger than $2C\mathbb{E}|\tilde{f}_N|$, and

$$\mathbb{E}|\tilde{f}_N| = \mathbb{E}\big[|f|^p |f|^{1-p} I[|f| \ge N]\big] \le N^{1-p}\mathbb{E}|f|^p.$$

Combining this with (A.3) gives

$$|\mathbb{E}[fg] - \mathbb{E} f \mathbb{E} g| \le 4CN\alpha(m) + 2CN^{1-p}\mathbb{E}|f|^p. \tag{A.4}$$

Choosing $N = [\mathbb{E}|f|^p]^{1/p}\alpha(m)^{-1/p}$ yields (A.2). □

For the proof of Corollary 6 we use a similar method to the one given above. We let C be a positive number to be chosen later and set $g_C = gI[|g| < C]$. By Lemma 7,

$$|\mathbb{E}[fg_C] - \mathbb{E} f \mathbb{E} g_C| \le 6C[\mathbb{E}|f|^p]^{1/p}\alpha(m)^{1/p'},$$

where $1/p + 1/p' = 1$. To estimate the difference, we write $\tilde{g}_C = g - g_C$ and again see that

$$|\mathbb{E}[fg] - \mathbb{E} f \mathbb{E} g| \le |\mathbb{E}[fg_C] - \mathbb{E} f \mathbb{E} g_C| + |\mathbb{E}[f\tilde{g}_C] - \mathbb{E} f \mathbb{E}\tilde{g}_C|.$$

We bound the last term using Hölder's inequality by

$$[\mathbb{E}|f|^p]^{1/p}[\mathbb{E}|\tilde{g}_C - \mathbb{E}\tilde{g}_C|^{p'}]^{1/p'}$$

and then use

$$[\mathbb{E}|\tilde{g}_C - \mathbb{E}\tilde{g}_C|^{p'}]^{1/p'} \le 2[\mathbb{E}|\tilde{g}_C|^{p'}]^{1/p'},$$

which we can bound above by $2(C^{p'-q}\mathbb{E}|g|^q)^{1/p'}$ as in (A.4). Choosing $C = [\mathbb{E}|g|^q]^{1/q}\alpha(m)^{-1/q}$ and combining the estimates as before completes the proof.

Acknowledgments. We would like to thank A. Hammond for discussions related to the law of large numbers. We also thank C. Newman and R. van der Hofstad for valuable comments and advice. Last we would like to acknowledge an anonymous referee for many helpful suggestions to make the paper more readable.

REFERENCES

[1] ANGEL, O., GOODMAN, J., DEN HOLLANDER, F. and SLADE, G. (2008). Invasion percolation on regular trees. *Ann. Probab.* **36** 420–466. MR2393988

[2] CHANDLER, R., KOPLICK, J., LERMAN, K. and WILLEMSEN, J. F. (1982). Capillary displacement and percolation in porous media. *J. Fluid Mech.* **119** 249–267.

[3] CHAYES, J. T., CHAYES, L. and NEWMAN, C. M. (1985). The stochastic geometry of invasion percolation. *Comm. Math. Phys.* **101** 383–407. MR0815191

[4] DAMRON, M. and SAPOZHNIKOV, A. (2011). Outlets of 2D invasion percolation and multiple-armed incipient infinite clusters. *Probab. Theory Related Fields* **150** 257–294. MR2800910

[5] DAMRON, M., SAPOZHNIKOV, A. and VÁGVÖLGYI, B. (2009). Relations between invasion percolation and critical percolation in two dimensions. *Ann. Probab.* **37** 2297–2331. MR2573559

[6] DAVYDOV, Y. (1968). Convergence of distributions generated by stationary stochastic processes. *Theory Probab. Appl.* **13** 691–696.

[7] GOODMAN, J. (2009). Exponential growth of ponds for invasion percolation on regular trees. Available at arXiv:0912.5205.

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

[9] HÄGGSTRÖM, O., PERES, Y. and SCHONMANN, R. H. (1999). Percolation on transitive graphs as a coalescent process: Relentless merging followed by simultaneous uniqueness. In *Perplexing Problems in Probability*. *Progress in Probability* **44** 69–90. Birkhäuser, Boston, MA. MR1703125

[10] IBRAGIMOV, I. A. and LINNIK, Y. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen. MR0322926

[11] JÁRAI, A. A. (2003). Invasion percolation and the incipient infinite cluster in 2D. *Comm. Math. Phys.* **236** 311–334. MR1981994

[12] KESTEN, H. (1987). Scaling relations for 2D-percolation. *Comm. Math. Phys.* **109** 109–156. MR0879034

[13] KESTEN, H. and ZHANG, Y. (1997). A central limit theorem for "critical" first-passage percolation in two dimensions. *Probab. Theory Related Fields* **107** 137–160. MR1431216

[14] LENORMAND, R. and BORIES, S. (1980). Description d'un mecanisme de connexion de liaision destine a l'etude du drainage avec piegeage en milieu poreux. *C. R. Acad. Sci.* **291** 279–282.

[15] NEWMAN, C. and STEIN, D. L. (1995). Broken ergodicity and the geometry of rugged landscapes. *Phys. Rev. E* **51** 5228–5238.

[16] NGUYEN, B. G. (1985). Correlation lengths for percolation processes. Ph.D. thesis, Univ. California, Los Angeles.

[17] NOLIN, P. (2008). Near-critical percolation in two dimensions. *Electron. J. Probab.* **13** 1562–1623. MR2438816

[18] VAN DEN BERG, J., JÁRAI, A. A. and VÁGVÖLGYI, B. (2007). The size of a pond in 2D invasion percolation. *Electron. Commun. Probab.* **12** 411–420. MR2350578

[19] WITHERS, C. S. (1981). Central limit theorems for dependent variables. I. *Z. Wahrsch. Verw. Gebiete* **57** 509–534. MR0631374

[20] ZHANG, Y. (2001). A martingale approach in the study of percolation clusters on the $\mathbf{Z}^d$ lattice. *J. Theoret. Probab.* **14** 165–187. MR1822899

PRINCETON UNIVERSITY
FINE HALL, WASHINGTON RD.
PRINCETON, NEW JERSEY 08544
USA
E-MAIL: mdamron@princeton.edu

ETH ZÜRICH
RÄMISTRASSE 101
8092 ZÜRICH
SWITZERLAND
E-MAIL: artem.sapozhnikov@math.ethz.ch