The Annals of Probability
2001, Vol. 29, No. 4, 1405–1425

HOW TO FIND AN EXTRA HEAD: OPTIMAL RANDOM SHIFTS OF BERNOULLI AND POISSON RANDOM FIELDS

BY ALEXANDER E. HOLROYD[1] AND THOMAS M. LIGGETT[2]

University of California, Los Angeles

We consider the following problem: given an i.i.d. family of Bernoulli random variables indexed by $\mathbb{Z}^d$, find a random occupied site $X \in \mathbb{Z}^d$ such that relative to X, the other random variables are still i.i.d. Bernoulli. Results of Thorisson imply that such an X exists for all d. Liggett proved that for $d = 1$, there exists an X with tails $P(|X| \geq t)$ of order $ct^{-1/2}$, but none with finite 1/2th moment. We prove that for general d there exists a solution with tails of order $ct^{-d/2}$, while for $d = 2$ there is none with finite first moment. We also prove analogous results for a continuum version of the same problem. Finally we prove a result which strongly suggests that the tail behavior mentioned above is the best possible for all d.

1. Introduction. The following problems were considered in [7]. They were originally motivated by some problems involving tagged particles in the exclusion and zero-range processes. We refer to [7] for more on those connections.

For $0 < \rho < 1$ and a positive integer d, let $\nu_\rho = \nu_\rho(d)$ denote the product measure with parameter ρ on $\{0, 1\}^{\mathbb{Z}^d}$ (with the product σ-algebra). Let η have distribution ν_ρ (so that $\{\eta(k) : k \in \mathbb{Z}^d\}$ are i.i.d. Bernoulli with parameter ρ), and define the measure ν_ρ^* by

$$\nu_\rho^*(\cdot) = \nu_\rho(\eta \in \cdot \mid \eta(0) = 1).$$

PROBLEM A. If η has distribution ν_ρ, find a $\mathbb{Z}^d$-valued random variable X (possibly using additional randomization) such that $\eta(X + \cdot)$ has distribution ν_ρ^*.

In the following continuum version of the above problem, Π is a spatial Poisson process regarded as a random non-negative integer-valued Borel measure on $\mathbb{R}^d$.

PROBLEM B. If Π is a rate-1 spatial Poisson process on $\mathbb{R}^d$, find an $\mathbb{R}^d$-valued random variable Y (possibly using additional randomization) such that $\Pi(\{Y\}) = 1$ a.s., and $\Pi(Y + \cdot)$ is a rate-1 spatial Poisson process on $\mathbb{R}^d$ with an added point at 0.

Received November 2000; revised January 2001.
[1]Supported in part by NSF Grants DMS-00-72398.
[2]Supoorted in part by NSF Grants DMS-97-03830 and DMS-00-70465.
AMS 2000 *subject classifications*. Primary 60G60; secondary 60G55, 60K35
Key words and phrases. product measure, random shift, Poisson process, tagged particle, shift coupling.

It is by no means obvious that either problem has a solution, but in fact solutions to both exist for all d, ρ. This follows from much more general results in [9]; for more information see [7], [10].

Writing $|\cdot|$ for the Euclidean norm on $\mathbb{R}^d$, it is natural to ask how large $|X|, |Y|$ must be for solutions to Problems A, B. In particular, what can be said about their moments and tail behavior? These questions were essentially fully answered for $d = 1$ in [7]. In particular the following two results were proved there.

THEOREM 1. *Let* $d = 1$.

(A) *For each* $0 < \rho < 1$ *there exists a solution* X *to Problem* A *satisfying*

$$P(|X| \geq t) \leq c_1 \rho^{-1/2} t^{-1/2},$$

where $c_1 < \infty$.

(B) *There exists a solution* Y *to Problem* B *satisfying*

$$P(|Y| \geq t) \leq c_2 t^{-1/2},$$

where $c_2 < \infty$.

THEOREM 2. *Let* $d = 1$.

(A) *Any solution to Problem* A *satisfies* $E|X|^{1/2} = \infty$.

(B) *Any solution to Problem* B *satisfies* $E|Y|^{1/2} = \infty$.

Actually, even tighter bounds involving $E|X \wedge t|$ were obtained in [7]. We remark that the proof of Theorem 1 in [7] is constructive, in the sense that an algorithm is given for choosing X (respectively Y) given η (respectively Π). Additional randomization is not required for Problem B, or for Problem A when ρ is the reciprocal of an integer. The explicit dependence on ρ in the bound in Theorem 1(A) is of interest because it allows Theorem 1(B) to be deduced via an alternative limiting argument.

We will prove the following.

THEOREM 3. *Let* $d \geq 1$.

(A) *For each* $0 < \rho < 1$ *there exists a solution* X *to Problem* A *satisfying*

$$P(|X| \geq t) \leq c_1 \rho^{-1/2} t^{-d/2},$$

where $c_1 = c_1(d) < \infty$.

(B) *There exists a solution* Y *to Problem* B *satisfying*

$$P(|Y| \geq t) \leq c_2 t^{-d/2},$$

where $c_2 = c_2(d) < \infty$.

THEOREM 4. *Let $d = 2$.*

(A) *Any solution to Problem* A *satisfies* $E|X| = \infty$.
(B) *Any solution to Problem* B *satisfies* $E|Y| = \infty$.

Our proof of Theorem 3(A) is also constructive, and makes use of the one-dimensional construction in [7]. The appearance of $\rho^{-1/2}$ in Theorem 3(A) will be important because it will allow (B) to be deduced from it via a limiting argument. As remarked in [7], the technique used to prove Theorem 2 gives no information in the case $d \geq 2$. It does not even rule out the possibility that X or Y is bounded. Our proof of Theorem 4 uses an entirely different approach.

Our results provide an almost complete answer to the question posed above for $d = 2$. It is natural to ask what can be said for $d \geq 3$. In particular, is it the case that $E|X|^{d/2}$, $E|Y|^{d/2}$ must be infinite for all d? For $d \geq 3$ we have been unable to improve on the following obvious lower bounds (see the remarks in the next paragraph, however). For Problem A, $|X|$ is stochastically greater than the distance from the origin to the closest occupied site of η. [A site $k \in \mathbb{Z}^d$ is said to be *occupied* if $\eta(k) = 1$.] For Problem B, $|Y|$ is stochastically greater than the distance from the origin to the closest point of Π.

Finally we will prove Theorem 8, a result which strongly suggests that the tail behavior of $|X|$ in Theorem 3(A) is the best that can be achieved for all d. Loosely speaking, Theorem 8 can be expressed as follows. If X is constructed from η by sequentially examining sites according to some algorithm, then the distance from the origin to the furthest site to be examined has tails of order at least $ct^{-d/2}$. In the course of proving this we will make use of another result, Proposition 9. This latter proposition relates to a natural generalization of Problem A to permutations of $\mathbb{Z}^d$.

The following question is natural. Suppose Y is a solution to Problem B. Can one use it to construct a solution to Problem A? The answer is yes; this follows from an argument used in the proof of Theorem 4(B).

The paper is organized as follows. Theorems 3(A) and 4(A) are proved in Sections 2 and 3 respectively. The (B) parts of both theorems are proved in Section 4. Theorem 8 and Proposition 9 mentioned above are stated and proved in Section 5.

2. Construction in *d* dimensions. In this section we prove Theorem 3(A). The following is an appealing idea for solving Problem A for general d using the solution for $d = 1$ [Theorem 1(A)]. Let s be an injective mapping from $\mathbb{Z}^1$ to $\mathbb{Z}^d$. Now for $\eta \in \{0,1\}^{\mathbb{Z}^d}$, define $\eta^1 \in \{0,1\}^{\mathbb{Z}^1}$ by $\eta^1(k) = \eta(s(k))$, apply the one-dimensional construction to η^1 to obtain X^1, and let $X = s(X^1)$. It seems natural to guess that X solves Problem A for η, but this is in general false, as was shown in [7]. An exception is the case when $s(k) = (k, 0, \dots, 0)$, but clearly this s will only give an X with the same tail behavior as X^1. Our approach here will be based on the above idea, but s will be a suitably chosen *random* mapping.

Suppose s is a bijection from $\mathbb{Z}^1$ to $\mathbb{Z}^d$ satisfying $s(0) = 0$. For $y \in \mathbb{Z}^d$, we define $\theta_y s$ to be 's viewed from y' thus:

$$\theta_y s(k) = s(s^{-1}(y) + k) - y.$$

Note that $\theta_y s$ is also a bijection satisfying $\theta_y s(0) = 0$.

PROPOSITION 5. *For any $d \geq 2$ there exists a random bijection S satisfying $S(0) = 0$ such that*:

(i) *For every $y \in \mathbb{Z}^d$, $\theta_y S$ has the same distribution as S.*

(ii) *$|S(k)| \leq C|k|^{1/d}$ for all $k \in \mathbb{Z}^1$ a.s., where $C = C(d) < \infty$ is a non-random constant.*

We note that Proposition 5 is not obvious, even if restriction (ii) is dropped. It may be proved using a construction based on space-filling curves. It appears that this construction has been known for some time; a version in the case $d = 2$ appears in [2]. We give a proof of the full result at the end of this section. We are grateful to Yuval Peres and the anonymous referee for advice on this point.

We now describe the construction for Theorem 3(A) (followed by the proof that it works). Let η have distribution $\nu_\rho(d)$. Choose S as in Proposition 5, independent of η. Define $\eta^1 \in \{0, 1\}^{\mathbb{Z}^1}$ by

$$\eta^1(k) = \eta(S(k)).$$

We will use the superscript ^{1} to denote "1-dimensional" objects throughout. It is clear that η^1 has distribution $\nu_\rho(1)$ (we will check this below). Therefore we may use Theorem 1(A) to find a $\mathbb{Z}^1$-valued random variable X^1 which solves Problem A for η^1. Indeed, choose X^1 to be conditionally independent of (η, S) given η^1 [in other words, any additional randomization in the construction of X^1 is taken to be independent of (η, S)]. Now let $X = S(X^1)$.

We will make extensive use of the following simple lemma.

LEMMA 6. *Let U, V be independent random variables (taking values in arbitrary spaces) and suppose $W = f(U, V)$, where W is a random variable (also taking values in an arbitrary space) and f is a deterministic function. Suppose that $f(U, v)$ has the same distribution α for every deterministic v in the support of V. Then V and W are independent, and W has distribution α.*

Lemma 6 is of course trivial in the case when V is a discrete random variable. The general version may be proved by a straightforward application of Fubini's theorem. We omit the details.

PROOF OF THEOREM 3(A). For $d = 1$ the result is exactly Theorem 1(A). For $d \geq 2$ we claim that X as defined above solves Problem A and has the stated tail behavior.

First we claim that

$$\eta^1 \text{ and } S \text{ are independent.} \tag{1}$$

This may be proved by applying Lemma 6 to the independent random variables η and S. For every deterministic bijection s, the composition $\eta(s(\cdot))$ clearly has distribution ν_ρ, so we deduce that $\eta^1 = \eta(S(\cdot))$ and S are independent, and also that η^1 has distribution ν_ρ.

Now by the conditional independence assumption on X^1, it follows from (1) that

$$(\eta^1, X^1) \text{ and } S \text{ are independent.} \tag{2}$$

Define $R = \theta_X S$. Recall that $X = S(X^1)$, so that $R = \theta_{S(X^1)}S$, and $R(k) = S(X^1 + k) - S(X^1)$. We claim

$$(\eta^1, X^1) \text{ and } R \text{ are independent.} \tag{3}$$

We will prove this using Lemma 6 and (2). By definition, R is a function of (η^1, X^1) and S, and does not depend on η^1. Therefore, defining $R_{x^1} = \theta_{S(x^1)}S$, it suffices to prove that R_{x^1} has the same distribution for each x^1. Consider the event

$$A = A(m, x^1, r) = \{R_{x^1}(k) = r(k) \text{ for all } -m \le k \le m\},$$

for m a positive integer and r a bijection. We will prove that for fixed m and r, $P(A)$ is constant in x^1 provided $-m \le x^1 \le m$, and the above claim then follows. To check this, observe that if $-m \le x^1 \le m$ then A equals the event

$$\left\{\theta_{-r(-x^1)}S(k) = r(k) \text{ for all } -m \le k \le m\right\}.$$

[The key point here is that, taking $k = -x^1$ in either expression for the event and using the fact that $R_{x^1}(k) = S(x^1+k) - S(x^1)$, we have that on the event, $S(x^1) = -r(-x^1)$ and therefore $R_{x^1} = \theta_{-r(-x^1)}S$]. Hence by Proposition 5 (i) we have

$$P(A) = P(S(k) = r(k) \text{ for all } -m \le k \le m),$$

which does not depend on x^1. Therefore (3) is proved.

Now define $\gamma^1 \in \{0, 1\}^{\mathbb{Z}^1}$ by $\gamma^1(k) = \eta^1(X^1 + k)$. Since X^1 solves Problem A, γ^1 has distribution $\nu_\rho^*(1)$. Since γ^1 is a function of (η^1, X^1), (3) implies that

$$\gamma^1 \text{ and } R \text{ are independent.} \tag{4}$$

We are now ready to check that X has the property required by Problem A. Define $\gamma \in \{0, 1\}^{\mathbb{Z}^d}$ by $\gamma(k) = \eta(X + k)$, so that it is required to check that γ has distribution $\nu_\rho^*(d)$. First observe that

$$\gamma(k) = \eta(X + k) = \eta^1(S^{-1}(X + k)) = \gamma^1(S^{-1}(X + k) - X^1) = \gamma^1(R^{-1}(k)).$$

Now for any deterministic bijection r with $r(0) = 0$, it is clear that $\gamma^1(r^{-1}(\cdot))$ has distribution $\nu_\rho^*(d)$, so the required fact follows from Lemma 6 and (4).

Finally, we check that X has the claimed tail behavior. Since $X = S(X^1)$, by Proposition 5 (ii) and Theorem 1(A) we have

$$\begin{aligned} P(|X| \geq t) &\leq P(C|X^1|^{1/d} \geq t) \\ &= P(|X^1| \geq C^{-d}t^d) \\ &\leq c_1(1)C^{d/2}\rho^{-1/2}t^{-d/2}. \end{aligned}$$ □

It remains to prove Proposition 5. The bijection S which we will construct will have the additional property that

$$|S(k) - S(l)| = 1 \text{ whenever } |k - l| = 1. \tag{5}$$

We will construct S from a doubly-infinite directed path which visits all the elements of $\mathbb{Z}^d$. Here is some notation. Elements of $\mathbb{Z}^d$ are called *vertices*. A (directed) *edge* is an ordered pair, written $[u, v\rangle$, of vertices $u, v \in \mathbb{Z}^d$ such that $|u-v| = 1$. The vertices u, v are said to be *incident* to the edge $[u, v\rangle$. We write $\mathbb{E}$ for the set of all edges. A *graph* is a subset of $\mathbb{E}$. The *vertex set* of a graph is the set of all vertices incident to its edges. For a graph E and a set of vertices V, $E(V)$ is the set of all edges in E having both incident vertices in V. A *path* is a graph of the form $\{[v_1, v_2\rangle, [v_2, v_3\rangle, \dots, [v_{n-1}, v_n\rangle\}$, where $v_1, \dots, v_n$ are distinct. A *tour* is a graph of the form $\{\dots, [v_{-1}, v_0\rangle, [v_0, v_1\rangle, [v_1, v_2\rangle, \dots\}$, where the v_i are distinct, whose vertex set is $\mathbb{Z}^d$.

We will construct a random tour T, and then use it to define S as follows. $S(0) = 0$. For $n > 0$, $S(n)$ is defined inductively to be the unique vertex such that $[S(n-1), S(n)\rangle \in T$, and similarly $S(-n)$ is defined so that $[S(-n), S(-n+1)\rangle \in T$. It is clear that S defined in this way is a bijection from $\mathbb{Z}^1$ to $\mathbb{Z}^d$ satisfying (5).

We will use the following lemma. For an integer $n \geq 0$, and for $v \in \mathbb{Z}^d$, let $C_n(v)$ be the set of vertices in the cube of side $2^n - 1$ with its minimum corner at v:

$$C_n(v) = v + \{0, \dots, 2^n - 1\}^d \subseteq \mathbb{Z}^d.$$

For $m \leq n$ we say that $C_m(u)$ is a *descendant* of $C_n(v)$ if $C_m(u) \subseteq C_n(v)$ and all the coordinates of $u - v$ are multiples of 2^m. If E is a graph, by a *copy* of E we mean an image of E under an isometry of $\mathbb{Z}^d$.

LEMMA 7. *Let $d \geq 2$. There exists a (deterministic) sequence of graphs H_n for $n \geq 0$ such that*:

(i) *H_n is a path with vertex set $C_n(0)$.*
(ii) *If $C_m(v)$ is a descendant of $C_n(0)$, then $H_n(C_m(v))$ is a copy of H_m.*

We omit the proof of Lemma 7. The required graphs H_n are the polygonal approximations used in the construction of the Hilbert space-filling curve. For $d = 2$ the construction is well known, appearing originally in [5]. Details of the construction for general d may be found in [3]. For further information about space-filling curves see [8].

PROOF OF PROPOSITION 5. Our aim is to construct a random tour T as described above. We write Ω for the set of all graphs, which we associate in the usual way with $\{0, 1\}^{\mathbb{E}}$. We will construct probability measures μ_n, μ on the corresponding product σ-algebra. When describing events of Ω, we will sometimes write G for a typical graph in Ω. For each $n \geq 1$ we will define a random graph T_n with distribution μ_n. The graph T with distribution μ will be the weak limit of this sequence.

The graph T_n is constructed as follows. We choose a cube of side $2^n - 1$ uniformly among those which contain the origin, and fill it with a copy of H_n with its orientation chosen uniformly at random. Then we deterministically fill the remainder of space with disjoint translated copies of this graph. See Figure 1 for an illustration. Formally, let σ to be an isometry of $\mathbb{Z}^d$ chosen uniformly at random from the $d!2^d$ which preserve $C_n(0)$, and independently choose $a \in C_n(0)$ uniformly at random. Then define

$$T_n = \bigcup_{u \in \mathbb{Z}^d} \left(\sigma(H_n) + a + 2^n u \right). \tag{6}$$

Let μ_n be the distribution of T_n.

Given a graph $G \in \Omega$ (which we think of as chosen according to μ_n), we say that $C_n(v)$ is an *n-box* (of G) if $G(C_n(v))$ is a *maximal* path in G. [For $G = T_n$ arising as in (6), this corresponds to v being of the form $a + 2^n u$.] For $m < n$, we also say that $C_m(w)$ is an *m-box* if it is a descendant of an n-box. See Figure 1. Thus, in a graph chosen according to μ_n, for $m \leq n$ every vertex

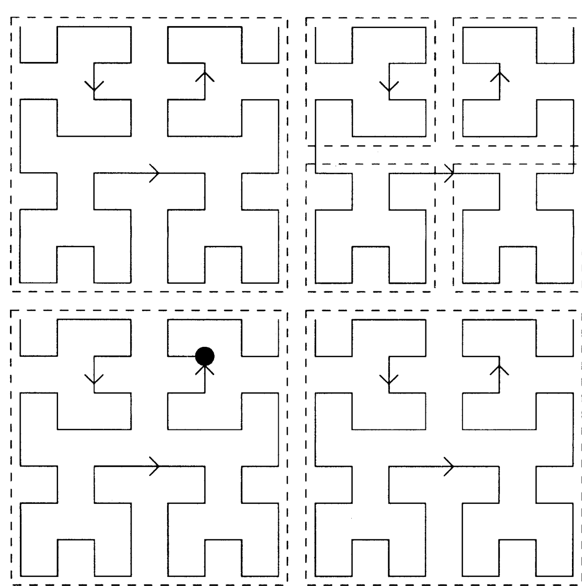

FIG. 1. *Part of a typical realization of T_3 for $d = 2$. The origin 0 is marked with a blob. Some 3-boxes and 2-boxes are marked with broken lines.*

of $\mathbb{Z}^d$ lies in exactly one m-box almost surely, while for $m > n$ there are no m-boxes. Furthermore, Lemma 7(ii) implies that if C is any m-box of T_n, then $T_n(C)$ is a copy of H_m.

Next we show that the sequence μ_n has a weak limit. Let A be a cylinder event of Ω, depending only on a finite set of edges K. Let $W_m = W_m(A)$ be the event that the vertex set of K lies entirely within a single m-box. Note that since K is finite, if m is sufficiently large and $n \geq m$ we have $0 < \mu_n(W_m) < 1$. Thus for some m_0 (depending on K), if $m_0 \leq m \leq n$ we have

$$\mu_n(A) = \mu_n(A \mid W_m)\mu_n(W_m) + \mu_n(A \mid W_m^c)(1 - \mu_n(W_m)). \tag{7}$$

We claim that for $m_0 \leq m \leq n$:

(i) $\mu_n(W_m)$ is constant in n, and $\mu_m(W_m) \to 1$ as $m \to \infty$.
(ii) $\mu_n(A \mid W_m)$ is constant in n.

To see (i), note that the set of m-boxes has the same distribution under each μ_n with $n \geq m$, and when 2^m is large compared with the diameter of K, $\mu_m(W_m)$ is close to 1. For (ii), note that under $\mu_n(\cdot \mid W_m)$, the m-box, C say, which contains K is equally likely to be any cube $C_m(v)$ containing K. Furthermore, by Proposition 7(ii) and (6), $T_n(C)$ is a copy of H_m, with all possible orientations being equally likely.

Now letting $n \to \infty$ and then $m \to \infty$ in (7) and using (i), (ii) yields

$$\limsup_{m\to\infty} \mu_m(A \mid W_m) \leq \liminf_{n\to\infty} \mu_n(A) \leq \limsup_{n\to\infty} \mu_n(A) \leq \liminf_{m\to\infty} \mu_m(A \mid W_m).$$

Hence the limits $\lim_{m\to\infty} \mu_m(A \mid W_m)$ and $\lim_{n\to\infty} \mu_n(A)$ exist and are equal. We let μ be the weak limit of the sequence μ_n, and let T be a graph with distribution μ. Thus for A a cylinder event,

$$\mu(A) = \lim_{m\to\infty} \mu_m(A \mid W_m(A)). \tag{8}$$

Next we must check that T is a tour almost surely. For a vertex v, let $A = A(v)$ be the event that T contains edges $[u, v\rangle$ and $[v, w\rangle$ for some $u \neq w$, and no other edges incident to v. For vertices u, v and $l > 0$, let $B = B(u, v, l)$ be the event that T contains a path of length at most l with u and v in its vertex set. It is sufficient to check that $\mu(A(v)) = 1$ for all v, and $\mu(B(u, v, l)) = 1$ for all u, v and some $l = l(u, v)$. But it is easily seen that $A(v)$ and $B(u, v, l)$ are cylinder events, and that for suitable $l = l(u, v)$, $\mu_m(A \mid W_m(A)) = \mu_m(B \mid W_m(B)) = 1$ for m sufficiently large. Hence the required conclusion follows from (8).

Now construct S in terms of T as described earlier. To complete the proof we must check properties (i) and (ii) in Proposition 5. Property (i) is equivalent to the assertion that the distribution of T is invariant under translations of $\mathbb{Z}^d$, and this follows immediately since T_n has the same property.

We now check (ii). Let G be a graph (chosen according to μ or μ_n). For a vertex $v \in \mathbb{Z}^d$, define $\Delta(v)$ to be the minimum size of a path in G with 0 and v in its vertex set (that is, the 'distance from 0 to v along G'). For a positive integer j, let $F(j)$ be the event that every vertex v with $\Delta(v) \leq j$

satisfies $|v| \le C(d)j^{1/d}$, where $C(d) = 2\sqrt{d+3}$. It is sufficient to check that $\mu(F(j)) = 1$ for all j. Clearly $F(j)$ is a cylinder event, and we will show that $\mu_n(F(j)) = 1$ for all n sufficiently large, implying the required result.

Given j, choose m so that $2^{d(m-1)} \le j < 2^{dm}$, and let $G = T_n$, where $n \ge m$. Suppose v satisfies $\Delta(v) \le j < 2^{dm}$. By the properties of H_n (Lemma 7), 0 and v must lie in the same m-box or in adjacent m-boxes (two disjoint m-boxes C, C' are said to be adjacent if there exist $w \in C$ and $w' \in C'$ with $|w-w'| = 1$). It follows (by Pythagoras' theorem) that

$$|v| \le \sqrt{(d-1)(2^m)^2 + (2 \cdot 2^m)^2} \le C(d)j^{1/d}.$$

Hence we have proved that $\mu_n(F(j)) = 1$ as required. □

3. Necessary condition in two dimensions. In this section we prove Theorem 4(A). We begin with a few comments about the idea behind the proof. Suppose X is a solution to Problem A. Then the construction that led to X from η can be applied to the random field relative to site X to find a second site known to be occupied. This construction can be repeated to obtain a sequence of occupied sites. If X has finite first moment, then the ergodic theorem implies that this sequence of sites can move away from the origin with at most a linear rate. In a box of size n there are n^2 sites in total, and at least εn of them are known to be occupied by this argument. Since X solves Problem A, the *other* sites should each be occupied with probability ρ. But this should contradict the central limit theorem for the number of occupied sites in the box viewed in terms of the original random field.

PROOF OF THEOREM 4(A). The random choice of X necessarily depends on η, but may also depend on other random choices. In the proof, it will be important to separate explicitly these two types of dependence. We will therefore begin by writing X as a function g of η and a random variable V that is independent of η and uniform on $[0, 1]$. (The choice of the uniform distribution is made for specificity only–any continuous distribution would work just as well.) To do so, let $k_1, k_2, \dots$ be any ordering of $\mathbb{Z}^2$. Define

$$g : \{0, 1\}^{\mathbb{Z}^2} \times [0, 1] \to \mathbb{Z}^2$$

as follows: Using the joint distribution of X and η, put

$$\Gamma_n(\eta) = P(X \in \{k_1, \dots, k_n\} \mid \eta).$$

This is defined for almost every η, and is monotone in n, so that $\Gamma_n(\eta)$ can be extended to all η so as to be monotone in n. Now put $g(\eta, v) = k_n$ if $\Gamma_{n-1}(\eta) < v \le \Gamma_n(\eta)$. Taking V to be independent of η and uniform on $[0, 1]$, we see that

$$\begin{aligned} P(g(\eta, V) = k_n \mid \eta) &= P(\Gamma_{n-1}(\eta) < V \le \Gamma_n(\eta) \mid \eta) \\ &= \Gamma_n(\eta) - \Gamma_{n-1}(\eta) \\ &= P(X = k_n \mid \eta) \qquad \text{a.s.} \end{aligned}$$

for each $n \geq 1$. Therefore, without loss of generality, we may take $X = g(\eta, V)$ in the proof.

Next, we make a construction similar to that used in the proof of Proposition 2.2 of [7]. Let the random variables $\eta_0 = \eta$, $\{\sigma_k : k \geq 0\}$, and $\{V_k : k \geq 0\}$ be all independent, with the following distributions: η_0 has distribution ν_ρ, σ_k is Bernoulli with parameter ρ, and V_k is uniform on $[0, 1]$. Let $X_0 = g(\eta_0, V_0)$, and then define η_1 by

$$\eta_1(k) = \begin{cases} \eta_0(X_0 + k), & \text{if } k \neq 0, \\ \sigma_0, & \text{if } k = 0. \end{cases}$$

Since X_0 solves Problem A, η_1 has distribution ν_ρ. Therefore, it is natural to define $X_1 = g(\eta_1, V_1)$. More generally, we define successively X_n and η_n by $X_n = g(\eta_n, V_n)$ and

$$\eta_{n+1}(k) = \begin{cases} \eta_n(X_n + k), & \text{if } k \neq 0, \\ \sigma_n, & \text{if } k = 0. \end{cases}$$

Writing this recursion in the functional form,

$$\eta_{n+1} = F(X_n, \eta_n, \sigma_n) = F(g(\eta_n, V_n), \eta_n, \sigma_n)$$

exhibits η_n as a Markov process with stationary distribution ν_ρ. Since η_0 has this distribution, η_n is a stationary process, and hence so is X_n. Let $S_n = X_0 + \cdots + X_{n-1}$ be the partial sums of this sequence. We will now check that the following hold:

(i) $\eta(S_{n+1}) = 1$ a.s. for each $n \geq 0$,
(ii) $\eta_n(X_n + \cdot)$ has distribution ν_ρ^* for each $n \geq 0$,
(iii) $\eta_n(X_n + k)$

$$= \begin{cases} \eta(S_{n+1} + k), & \text{if } S_{n+1} + k \neq S_i \text{ for all } 1 \leq i \leq n, \\ \sigma_{j-1}, & \text{if } S_{n+1} + k = S_j, \text{ but } S_{n+1} + k \neq S_i \\ & \quad \text{for all } j < i \leq n. \end{cases}$$

Property (ii) for general n follows from the case $n = 0$ (which holds by the defining property of X_0) and the fact that (η_n, X_n) is a stationary process.

The key to checking properties (i) and (iii) is the following fact. If for some $0 \leq j < n$ and integer k

$$S_{n+1} + k \neq S_i \qquad \text{for all } j < i \leq n,$$

then $\eta_l(S_{n+1} + k - S_l)$ is constant in l for $j \leq l \leq n$, and in particular takes the same values for $l = n$ and for $l = j$:

$$\eta_n(X_n + k) = \eta_n(S_{n+1} + k - S_n) = \eta_j(S_{n+1} + k - S_j).$$

To check this fact, note that for $j \leq l < n$, the defining recursion gives

$$\eta_{l+1}(S_{n+1} + k - S_{l+1}) = \eta_l(X_l + S_{n+1} + k - S_{l+1}) = \eta_l(S_{n+1} + k - S_l),$$

since $S_{n+1} + k - S_{l+1} \neq 0$. If in addition $S_{n+1} + k = S_j$, it follows that

$$\eta_n(X_n + k) = \eta_j(S_{n+1} + k - S_j) = \eta_j(0) = \sigma_{j-1}.$$

This gives the second part of (iii). If $S_{n+1} + k \neq S_i$ for all $0 < i \leq n$, we can use the above fact with $j = 0$ to get

$$\eta_n(X_n + k) = \eta_0(S_{n+1} + k),$$

which is the first part of (iii). For (i), we argue as follows. If $S_{n+1} \neq S_i$ for all $1 \leq i \leq n$ then the above fact with $k = j = 0$ gives

$$\eta_0(S_{n+1}) = \eta_n(X_n) = 1.$$

On the other hand, if $S_{n+1} = S_{i+1}$ and $S_{n+1} \notin \{S_1, \dots, S_i\}$, then

$$\eta_0(S_{n+1}) = \eta_0(S_{i+1}) = 1$$

by the previous case.

Now take $I = (0, a]^2$ for some integer $a > 0$ to be chosen later, and let

$$A_n = \sum_{k \in nI} \eta_n(X_n + k).$$

By choices to be made later, we will be able to assume that most of the sites $S_1, \dots, S_n$ lie in nI with high probability. By (ii) above, A_n is binomially distributed with parameters $|nI \cap \mathbb{Z}^2|$ and ρ. On the other hand, by (i) and (iii),

$$\begin{aligned} A_n &= \sum_{\substack{l \in S_{n+1}+nI: \\ l \neq S_i \ \forall\ 1 \leq i \leq n}} \eta(l) + \sum_{\substack{j: 1 \leq j \leq n, \\ S_j \in S_{n+1}+nI, \\ S_j \neq S_i \ \forall\ j < i \leq n}} \sigma_{j-1} \\ &= \sum_{l \in S_{n+1}+nI} \eta(l) - \sum_{\substack{j: 1 \leq j \leq n, \\ S_j \in S_{n+1}+nI, \\ S_j \neq S_i \ \forall\ j < i \leq n}} (1 - \sigma_{j-1}) \qquad (9) \\ &= B_n - C_n, \end{aligned}$$

where B_n and C_n are defined as in the last equality in (9). Our aim is to show that each of A_n and B_n is approximately normal with mean $n^2 a^2 \rho$ and variance $n^2 a^2 \rho(1 - \rho)$, while C_n is of order n, giving a contradiction.

Assume now for purposes of getting a contradiction that $E|X_0| < \infty$. The ergodic theorem then implies that

$$\lim_{n \to \infty} \frac{S_n}{n} = Z \qquad (10)$$

a.s. and in L_1 for some random $Z \in \mathbb{R}^2$. We may assume without loss of generality that the distribution of X_0, and hence of Z, has been modified by adding any given deterministic $h \in \mathbb{Z}^2$. To see this, suppose $g(\eta, V)$ is a solution to Problem A, and define g' by

$$g'(\eta, v) = g(T_h \eta, v) + h,$$

where T_h is the shift on $\{0,1\}^{\mathbb{Z}^2}$ defined by $(T_h\eta)(k) = \eta(h+k)$. Define $X' = g'(\eta, V)$ and $X = g(T_h\eta, V)$. Then $X' = X + h$, and

$$\begin{aligned}\eta(X'+l) &= \eta(g'(\eta, V)+l)\\ &= \eta(g(T_h\eta, V)+h+l)\\ &= (T_h\eta)(g(T_h\eta, V)+l)\\ &= (T_h\eta)(X+l)\end{aligned}$$

so that $\eta(X'+\cdot)$ has distribution ν_ρ^*.

The point of the previous observation is that given $\varepsilon > 0$, we may assume without loss of generality that

$$P(Z \le (-1,-1)) \ge 1-\varepsilon \tag{11}$$

and

$$P(S_n = 0 \text{ for some } n \ge 1) \le \varepsilon. \tag{12}$$

[The first inequality in (11) is understood to be coordinate-wise.] That (11) can be achieved by adding a constant to g is immediate, since adding a constant to g has the effect of adding the same constant to Z. To check that (12) also can be so achieved, note that (10) implies that

$$\sup_n \frac{|S_n|}{n} < \infty \qquad \text{a.s.},$$

so that for $h \in \mathbb{Z}^2$ with $|h|$ sufficiently large,

$$P(S_n - nh = 0 \text{ for some } n \ge 1) = P\left(\frac{S_n}{n} = h \text{ for some } n \ge 1\right) \le \varepsilon.$$

By (11), a can be taken sufficiently large so that

$$P\Big((-a-1,-a-1) \le Z \le (-1,-1)\Big) \ge 1-2\varepsilon. \tag{13}$$

Since

$$P(-S_j \in jI) \le P(-S_j \in nI)$$

for $j \le n$, it follows from (10) and (13) that

$$\sup_{n/2 \le j \le n} P(-S_j \notin nI) \le 3\varepsilon \tag{14}$$

for sufficiently large n.

Now, we have

$$
\begin{aligned}
EC_n &= \sum_{j=1}^{n} P(\sigma_{j-1} = 0, S_j \in S_{n+1} + nI, S_j \neq S_i \ \forall\ j < i \leq n) \\
&\geq \sum_{j=1}^{\lfloor n/2 \rfloor} \Big[P(\sigma_{j-1} = 0) - P(S_j \notin S_{n+1} + nI) \\
&\qquad\qquad - P(S_j = S_i \text{ for some } j < i \leq n) \Big] \\
&= \sum_{j=1}^{\lfloor n/2 \rfloor} \Big[1 - \rho - P(-S_{n+1-j} \notin nI) \\
&\qquad\qquad - P(S_k = 0 \text{ for some } 1 \leq k \leq n - j) \Big].
\end{aligned}
$$

(Here $\lfloor \cdot \rfloor$ denotes the integer part.) Applying (12) and (14), it follows that

$$
\liminf_{n\to\infty} \frac{EC_n}{n} \geq \frac{1 - \rho - 4\varepsilon}{2} > 0, \tag{15}
$$

provided ε is sufficiently small.

By the central limit theorem,

$$
\frac{A_n - n^2 a^2 \rho}{n} \Rightarrow N(0, a^2 \rho(1 - \rho)). \tag{16}
$$

Here $\Rightarrow$ denotes convergence in distribution, and $N(0, \sigma^2)$ is the normal distribution with mean 0 and variance σ^2. We want to prove that also

$$
\frac{B_n - n^2 a^2 \rho}{n} \Rightarrow N(0, a^2 \rho(1 - \rho)). \tag{17}
$$

Before proving (17), we note that (9), (15), (16) and (17) are incompatible, so we will have a contradiction to the assumption that $E|X_0| < \infty$. To see this, suppose that these four statements are correct. By the definition of C_n in (9), $0 \leq C_n \leq n$, so the distributions of C_n/n are tight. Passing to a subsequence, we may assume that

$$
\left(\frac{B_n - n^2 a^2 \rho}{n}, \frac{C_n}{n} \right) \Rightarrow (B, C),
$$

where B is $N(0, a^2\rho(1 - \rho))$ by (17) and $EC > 0$ by (15) and the bounded convergence theorem for convergence in distribution. But by (9),

$$
\frac{A_n - n^2 a^2 \rho}{n} = \frac{B_n - n^2 a^2 \rho}{n} - \frac{C_n}{n}.
$$

The limiting distribution of the right side above is that of $B - C$, which has a nonzero mean, while the limit of the left side is normal with mean zero by (16). This is a contradiction.

So it remains to prove (17). Let $\mathscr{R}$ be the set of rectangles

$$\mathscr{R} = \{[a, b] \times [c, d] : a, b, c, d \in \mathbb{R}\}.$$

For $A \in \mathscr{R}$, let

$$W_n(A) = \frac{1}{n\sqrt{\rho(1-\rho)}} \sum_{l \in nA} [\eta(l) - \rho],$$

and let W be the standard Brownian sheet. That is, for $A, B \in \mathscr{R}$, $W(A)$ has a normal distribution with mean zero, $E[W(A)W(B)]$ equals the area of $A \cap B$, and $W(A)$ is a.s. continuous in A. The following invariance principle is proved in [6]:

$$W_n \quad \Rightarrow \quad W, \tag{18}$$

where now $\Rightarrow$ denotes weak convergence in the space D, which we define next. For an integer $n \geq 1$, let D_n be the set of all functions on $\mathscr{R}$ which, as a function of A, change only when an edge of A crosses a point in $n^{-1}\mathbb{Z}^2$. Then D is the closure of $\cup_{n=1}^{\infty} D_n$ in the topology of uniform convergence on compact sets. Note that D is a complete separable metric space with this topology, and that it contains the continuous functions. Furthermore, $W_n \in D$. (Actually, in [6], (18) is proved for a smoothed version of W_n. However, since we are considering only rectangles A, the smoothed version implies the one we want; see, e.g., Corollary 4.6 of [1] for more on this point.)

By (10) and (18), the sequence $(W_n, S_n/n)$ is tight in $D \times \mathbb{R}^2$, and therefore relatively compact. We will show that

$$(W_n, S_n/n) \Rightarrow (W, Z), \tag{19}$$

where W and Z are taken to be independent. By the relative compactness, it is enough to show that any subsequential limit of $(W_n, S_n/n)$ has independent components. To simplify notation, we will not distinguish between the full sequence and a convergent subsequence.

So, let Z be the limit in (10) and let L_n be the following increasing sequence of squares in $\mathbb{Z}^2$:

$$L_n = [-n^{1/3}, n^{1/3}]^2 \cap \mathbb{Z}^2.$$

Let $\mathscr{F}_n$ be the σ-algebra generated by the random variables $\{\eta(k) : k \in L_n\}$, $\{\sigma_l : l \geq 0\}$ and $\{V_l : l \geq 0\}$, and put $Z_n = E(Z \mid \mathscr{F}_n)$. Then Z_n is independent of $\{\eta(k) : k \notin L_n\}$, and $E|Z_n - Z| \to 0$ by the martingale convergence theorem. It follows from (10) that

$$E\left|\frac{S_n}{n} - Z_n\right| \to 0. \tag{20}$$

Now let

$$W'_n(A) = \frac{1}{n\sqrt{\rho(1-\rho)}} \sum_{l \in nA, l \notin L_n} [\eta(l) - \rho].$$

Then

$$|W_n(A) - W'_n(A)| \le \frac{|L_n|}{n\sqrt{\rho(1-\rho)}} \to 0 \tag{21}$$

as $n \to \infty$. Now it follows from (20) and (21) that whenever $(W_n, S_n/n)$ has a weak limit, so does (W'_n, Z_n), and the limits are the same. But (W'_n, Z_n) does have independent coordinates, since W'_n is a function of $\{\eta(l) : l \notin L_n\}$, and Z_n is independent of these random variables. This completes the verification of (19).

Finally, note that

$$\frac{B_n - n^2a^2\rho}{n\sqrt{\rho(1-\rho)}} = W_n\biggl(\frac{S_{n+1}}{n} + I\biggr) = \Phi\biggl(W_n, \frac{S_{n+1}}{n}\biggr),$$

where $\Phi : D \times \mathbb{R}^2 \to \mathbb{R}$ is given by

$$\Phi(w, x) = w(x + I).$$

Since Φ is continuous at all (w, x) for which w is continuous on $\mathscr{R}$, and W is continuous on $\mathscr{R}$, it follows from (19) that

$$\frac{B_n - n^2a^2\rho}{n\sqrt{\rho(1-\rho)}} \Rightarrow \Phi(W, Z),$$

where W and Z are independent. But

$$\Phi(W, Z) = W(Z + I).$$

Since $W(z + I)$ has distribution $N(0, a^2)$ for each deterministic z, it follows by Lemma 6 that $W(Z + I)$ has distribution $N(0, a^2)$ as well. This completes the proof of (17) as required. □

4. Poisson case. In this section we deduce the (B) parts of Theorems 3 and 4 from the (A) parts.

PROOF OF THEOREM 3(B). For $n \ge 1$ and $k \in \mathbb{Z}^d$ let $K_n(k)$ be the (open) cube of side n^{-1} centered at $n^{-1}k$:

$$K_n(k) = n^{-1}k + (-(2n)^{-1}, (2n)^{-1})^d$$

(here $(\cdot, \cdot)$ denotes an interval of $\mathbb{R}$). Given the Poisson process Π, define $\eta_n \in \{0, 1\}^{\mathbb{Z}^d}$ by

$$\eta_n(k) = 1 \wedge \Pi(K_n(k)),$$

where $\wedge$ denotes minimum; thus $\eta_n(k) = 1$ if and only if $K_n(k)$ contains at least one Poisson point. Then η_n has distribution ν_{ρ_n} where

$$\rho_n = P\Bigl(\Pi(K_n(k)) \ge 1\Bigr) = 1 - e^{-n^{-d}} \sim n^{-d} \tag{22}$$

as $n \to \infty$. Using Theorem 3(A), choose X_n solving Problem A for η_n. Then we have

$$
\begin{aligned}
P(n^{-1}|X_n| \geq t) &= P(|X_n| \geq nt) \\
&\leq c_1 \rho_n^{-1/2} n^{-d/2} t^{-d/2} \\
&\leq c_2 t^{-d/2}
\end{aligned} \tag{23}
$$

for some constant c_2, by (22).

Let Π_n be the point process that has a point at $n^{-1}k$ if $\eta_n(k) = 1$, and no others. Let Π^* be a rate-1 Poisson process with an extra point at the origin. We claim that

$$\Pi_n \Rightarrow \Pi \tag{24}$$

and

$$\Pi_n(n^{-1}X_n + \cdot) \Rightarrow \Pi^*(\cdot), \tag{25}$$

where $\Rightarrow$ denotes weak convergence in the space $\mathscr{M}$ of locally finite measures on $\mathbb{R}^d$. Both statements are consequences of the Poisson approximation to the Binomial, since Π_n and $\Pi_n(n^{-1}X_n + \cdot)$ are Bernoulli random fields on the grid of points $n^{-1}\mathbb{Z}^d$.

By (23) and (24), the sequence $(\Pi_n, n^{-1}X_n)$ is tight on $\mathscr{M} \times \mathbb{R}^d$, and therefore relatively compact by Prohorov's Theorem ([4], Chapter 18, Theorem 17). Let (Π_0, Y_0) be a weak limit point of this sequence. Then (24) and (25) imply that Π_0 and $\Pi(Y_0 + \cdot)$ have the distributions required by Problem B. Note that these are simply properties of the joint distribution of the random vector (Π_0, Y_0), not of any particular realization of it.

The remaining issue is to see how to use the joint distribution of (Π_0, Y_0) constructed above to choose an Y with the right properties for a given rate-1 Poisson process Π. The solution is simply to use the conditional distribution of Y_0 given Π_0, and apply this to the the given Π: by enlarging the probability space, one can choose Y via the prescription

$$P(Y \in \cdot \mid \Pi) = P(Y_0 \in \cdot \mid \Pi_0). \qquad \square$$

PROOF OF THEOREM 4(B). Let Π be a rate-1 spatial Poisson process on $\mathbb{R}^2$, and suppose Y is a solution to Problem B for Π with $E|Y| < \infty$. We will deduce a contradiction to Theorem 4(A).

Let Q be the unit square $[0, 1)^2 \subseteq \mathbb{R}^2$, and let A have uniform distribution on Q, independent of (Π, Y). For $u \in \mathbb{R}$ we write $\lfloor u \rfloor$ for the greatest integer less than or equal to u, and $\{u\} = u - \lfloor u \rfloor$. For $v = (v_1, v_2) \in \mathbb{R}^2$ we write $\lfloor v \rfloor = (\lfloor v_1 \rfloor, \lfloor v_2 \rfloor)$ and $\{v\} = (\{v_1\}, \{v_2\})$. Now define $X = \lfloor Y + A \rfloor$ and $B = \{Y + A\}$. (We imagine a grid of unit squares with its origin at $-A$. Square 0 is the square containing 0, and square X is the square containing Y. The location of Y within square X is B.) Note that

$$X + B = Y + A. \tag{26}$$

We claim that B has uniform distribution on Q, and that it is independent of (Π, Y). This is proved by an application of Lemma 6 as follows. Since A and (Π, Y) are independent, and $B = \{Y + A\}$, it suffices to check that $\{y + A\}$ is uniform on Q for each fixed y; but this is an elementary property of the uniform distribution on Q.

Now define $\eta \in \{0, 1\}^{\mathbb{Z}^2}$ by

$$\eta(k) = 1 \wedge \Pi(Q - A + k). \tag{27}$$

[Thus $\eta(k)$ is 1 if and only if square k contains at least one Poisson point.] For any fixed a, it is easily seen that η_a defined by $\eta_a(k) = 1 \wedge \Pi(Q - a + k)$ has distribution $\nu_{1-e^{-1}}$; hence Lemma 6 implies that η has distribution $\nu_{1-e^{-1}}$.

We claim that X is a solution to Problem A for η. To prove this, define γ by $\gamma(k) = \eta(X + k)$, and Λ by $\Lambda(\cdot) = \Pi(Y + \cdot)$. Note that by (26) and (27),

$$\gamma(k) = 1 \wedge \Pi(Q - A + X + k) = 1 \wedge \Pi(Q - B + Y + k) = 1 \wedge \Lambda(Q - B + k)$$

[compare (27)]. Since B and (Π, Y) are independent, B and Λ are independent, and since Y solves Problem B, Λ is a Poisson process with an added point at 0. For any $b \in Q$ it is easy to check that γ_b defined by $\gamma_b(k) = 1 \wedge \Lambda(Q - b + k)$ has distribution $\nu^*_{1-e^{-1}}$ (note in particular that $0 \in Q - b$ so $\gamma_b(0) = 1$ a.s.), so a final application of Lemma 6 shows that γ has distribution $\nu^*_{1-e^{-1}}$ as claimed.

Finally we have $E|X| \leq E|Y| + \sqrt{2} < \infty$, contradicting Theorem 4(A). $\square$

5. The set of examined sites. Let X be a solution to Problem A. Imagine that X is constructed from η by some algorithm which sequentially examines the values $\eta(k)$ for sites $k \in \mathbb{Z}^d$ until some stopping time, finally choosing an X in the set of examined sites. The algorithm may use additional randomization if desired. We formalize this idea as follows. Define the box $B_n \subseteq \mathbb{Z}^d$ by

$$B_n = \{-n, \ldots, n\}^d.$$

Let η have distribution ν_ρ, suppose X is a $\mathbb{Z}^d$-valued random variable and suppose N is a non-negative integer-valued random variable. We say that (X, N) is a *stopping solution* to Problem A if the following hold:

(i) X is a solution to Problem A.

(ii) For any x, n, the σ-algebra generated by the event $\{X = x, N = n\}$ and the random variables $\{\eta(k) : k \in B_n\}$ is independent of that generated by $\{\eta(k) : k \notin B_n\}$.

The random set B_N should be thought of as the set of examined sites. As a consequence of (i), (ii), any stopping solution has the additional property

$$X \in B_N \qquad \text{a.s.} \tag{28}$$

To see this, note that by (i),

$$1 = P(\eta(X) = 1) = P(\eta(X) = 1 \mid X \in B_N)P(X \in B_N)$$
$$+P(\eta(X) = 1 \mid X \notin B_N)(1 - P(X \in B_N)),$$

but we have $P(\eta(X) = 1 \mid X \notin B_N) = \rho < 1$ by (ii).

From [7] and the proof of Theorem 3(A), it may be seen that for the X constructed in Theorem 3(A), (X, N) is indeed a stopping solution for some N. Furthermore this N can be chosen such that $P(N \geq n) \leq c_1' n^{-d/2}$ for some $c_1' = c_1'(\rho, d) < \infty$. For $d = 1$, this is achieved by taking $N = X$ in Theorem 1(A); that this is possible follows from the proof in [7]. For general d, the following extension of Proposition 5 is required. In part (ii) of that result, all of the sites $S(0), S(1), \dots, S(k)$ lie within distance $C|k|^{1/d}$ of 0. This follows directly from the proof of Proposition 5.

THEOREM 8. *Let* $d \geq 1$ *and* $0 < \rho < 1$. *If* (X, N) *is any stopping solution to Problem* A *then*

$$P(N \geq n) \geq c_3 n^{-d/2},$$

where $c_3 = c_3(\rho, d) > 0$.

As remarked in the Introduction, Theorem 8 makes it plausible that the tail behavior of $|X|$ in Theorem 3(A) is essentially the best possible for all d. This is because it would be very surprising if the construction of a solution X with optimal tail behavior required the examination of sites "much further" from the origin than the final choice of X.

In the proof of Theorem 8 we will make use of Proposition 9 below, which is of interest in its own right. Let π be a permutation of $\mathbb{Z}^d$. We define

$$|\pi| = \min\{n \geq 0 : \pi(k) = k \text{ for all } k \notin B_n\},$$

where the minimum of the empty set is taken to be ∞. Thus, $|\pi| \leq n$ if π disturbs only sites in B_n. A permutation π acts on elements η of $\{0, 1\}^{\mathbb{Z}^d}$ via $(\pi\eta)(k) = \eta(\pi(k))$, and on measures μ on $\{0, 1\}^{\mathbb{Z}^d}$ via $(\pi\mu)(A) = \mu(\pi^{-1}(A))$. The following result relates to a natural generalization of Problem A to permutations.

PROPOSITION 9. *Let* η *have distribution* ν_ρ, *and suppose* π *is a random permutation of* $\mathbb{Z}^d$ *with the property that* $\pi\eta$ *has distribution* ν_ρ^*. *Then*

$$P(|\pi| > n) \geq c_4 n^{-d/2},$$

where $c_4 = c_4(\rho, d) > 0$.

We remark that the bound in Proposition 9 is essentially optimal; that is, there exists such a π with $P(|\pi| > n) \leq c_5 n^{-d/2}$, where $c_5 = c_5(\rho, d) < \infty$. To prove this, let (X, N) be a stopping solution satisfying the bound of Theorem 3(A), as described above. It will follow from the proof of Theorem 8 that we can construct a π having the required properties from (X, N).

PROOF OF PROPOSITION 9. Our proof is similar to that of Theorem 3.1 in [7], which in turn is based on the shift-coupling inequality; see [10], Chapter 7. The total variation norm $\|\cdot\|$ on signed measures is defined as usual by $\|\mu\| = \sup_{f:|f|\le 1}\int f\,d\mu$. We claim that

$$\left\| \frac{1}{|B_n|!}\sum_{\alpha:|\alpha|\le n} \alpha\nu_\rho - \frac{1}{|B_n|!}\sum_{\alpha:|\alpha|\le n}\alpha\nu_\rho^*\right\| \le 2P(|\pi| > n), \tag{29}$$

where each sum is over all $|B_n|!$ permutations α for which $|\alpha|\le n$. Equation (29) is a variant of the shift-coupling inequality.

To prove (29), note that by the assumptions of the proposition, the left side equals the supremum over $|f|\le 1$ of

$$\frac{1}{|B_n|!}\sum_{\alpha:|\alpha|\le n} Ef(\alpha\eta) - \frac{1}{|B_n|!}\sum_{\alpha:|\alpha|\le n} Ef((\pi\alpha)\eta), \tag{30}$$

where $\pi\alpha$ is the composition of permutations defined by $(\pi\alpha)(k) = \pi(\alpha(k))$. By conditioning on π, (30) equals

$$E\left(\frac{1}{|B_n|!}\sum_{\alpha:|\alpha|\le n} E[f(\alpha\eta)\mid\pi] - \frac{1}{|B_n|!}\sum_{\alpha:|\alpha|\le n} E[f((\pi\alpha)\eta)\mid\pi]\right),$$

and we may bound this expression as follows. For every fixed π with $|\pi|\le n$, we have $\{\pi\alpha : |\alpha|\le n\} = \{\alpha : |\alpha|\le n\}$, so the two sums are equal in this case. On the other hand, each of the two terms is bounded above in absolute value by 1, for every fixed π. So (30) is at most

$$0P(|\pi|\le n) + 2P(|\pi| > n).$$

Now taking the supremum over f gives (29).

Next we compute the left side of (29) as follows. Evidently, $\alpha\nu_\rho = \nu_\rho$. Furthermore, we have $\alpha\nu_\rho^* = \nu_\rho^{*,\alpha^{-1}(0)}$ where $\nu_\rho^{*,k}$ is defined by

$$\nu_\rho^{*,k}(\cdot) = \nu_\rho(\eta\in\cdot\mid\eta(k)=1).$$

Hence the left side of (29) equals

$$\left\|\nu_\rho - \frac{1}{|B_n|}\sum_{k\in B_n}\nu_\rho^{*,k}\right\|. \tag{31}$$

We now proceed as in [7]. The expression above can be written as

$$\int\left|1 - \frac{1}{|B_n|}\sum_{k\in B_n}\frac{d\nu_\rho^{*,k}}{d\nu_\rho}\right| d\nu_\rho.$$

But under the measure ν_ρ, the sum in this expression is equal to $1/\rho$ times a binomial random variable with parameters $|B_n|$, ρ. It follows by the central limit theorem that (31) is asymptotic to

$$c'|B_n|^{-1/2} = c'(2n+1)^{-d/2},$$

for some $c' = c'(\rho) > 0$. Combining this with (29) yields the required result. □

PROOF OF THEOREM 8. Let η have distribution ν_ρ and let (X, N) be a stopping solution. A *cuboid* is a subset of $\mathbb{Z}^d$ which is a direct product of sets of the form $\{a, a+1, \dots, b\}$. Define $R = R(X, N)$ to be the smallest cuboid which contains $B_N \cup (B_N + X)$ as a subset. (See Figure 2 for an illustration in the case $d = 2$). Now define $\pi : \mathbb{Z}^d \to \mathbb{Z}^d$ by

$$\pi(k) = \begin{cases} X - k, & \text{if } k \in R, \\ k, & \text{otherwise.} \end{cases}$$

It is easy to check that π is a permutation, and that $|\pi| \le 2N$ [by (28); also see Figure 2]. We will prove that $\pi\eta$ has distribution ν_ρ^*. The required result will then follow immediately from Proposition 9.

Write $\lambda = \pi\eta$, so that

$$\lambda(k) = \begin{cases} \eta(X - k), & \text{if } k \in R, \\ \eta(k), & \text{otherwise.} \end{cases}$$

Define $\xi \in \{0, 1\}^{\mathbb{Z}^d}$ by $\xi(k) = \eta(X - k)$. Since $\eta(X + \cdot)$ has distribution ν_ρ^*, ξ has this distribution also. We need to prove that λ and ξ have the same distribution. In fact we will prove the stronger statement

$$P(\lambda \in \cdot, X = x, N = n) = P(\xi \in \cdot, X = x, N = n) \tag{32}$$

for all x, n; this gives the required fact on summing over x, n. It is sufficient to check (32) for increasing cylinder events, and we therefore proceed as follows. Fix x and n, and write T for the event $\{X = x, N = n\}$. Let K be any finite subset of $\mathbb{Z}^d$, and write $K' = K \cap B_n$, $K'' = K \cap (R(x, n) \setminus B_n)$ and $K''' = K \cap R(x, n)^c$. Then we have the chain of equalities

$$\begin{aligned} &P\big(\lambda(k) = 1 \ \forall\ k \in K, T\big) \\ &\quad = P\Big(\eta(x - k) = 1 \ \forall\ k \in K', \eta(x - k) = 1 \ \forall\ k \in K'', \\ &\qquad\qquad \eta(k) = 1 \ \forall\ k \in K''', T\Big) \end{aligned}$$

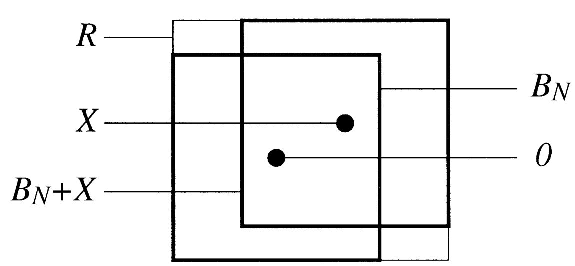

FIG. 2. *The cuboid R.*

$$
\begin{aligned}
&= P(\eta(x-k) = 1\ \forall\ k \in K', T)\ P(\eta(x-k) = 1\ \forall\ k \in K'') \\
&\quad \times P(\eta(k) = 1\ \forall\ k \in K''') \\
&= P(\eta(x-k) = 1\ \forall\ k \in K', T)\ P(\eta(x-k) = 1\ \forall\ k \in K'') \\
&\quad \times P(\eta(x-k) = 1\ \forall\ k \in K''') \\
&= P\Big(\eta(x-k) = 1\ \forall\ k \in K', \eta(x-k) = 1\ \forall\ k \in K'', \\
&\qquad\qquad\qquad \eta(x-k) = 1\ \forall\ k \in K''', T\Big) \\
&= P(\xi(k) = 1\ \forall\ k \in K, T).
\end{aligned}
$$

In the second and fourth equalities we have used property (ii) of a stopping solution, the independence of the random variables $\{\eta(k) : k \in \mathbb{Z}^d\}$, and the fact that if $k \in K'''$ then $k, x-k \notin R(x,n) \supseteq B_n$. We have thus established (32), as required. □

REFERENCES

[1] ALEXANDER, K. (1987). Central limit theorems for stochastic processes under random entropy conditions. *Probability Theory Related Fields* **75** 351–378.

[2] BURTON, R. M. and KEANE, M. (1991). Topological and metric properties of infinite clusters in stationary two-dimensional site percolation. *Israel J. Math.* **76** 299–316.

[3] BUTZ, A. R. (1969). Convergence with Hilbert's space-filling curve. *J. Comput. System Sci.* **3** 128–146.

[4] FRISTEDT, B. and GRAY, L. (1997). *A Modern Approach to Probability Theory*. Birkhäuser, Boston.

[5] HILBERT, D. (1891). Über die stetige abbildung einer linie auf ein flächenstück. *Math. Ann.* **38** 459–460.

[6] KUELBS, J. (1968). The invariance principle for a lattice of random variables. *Ann. Math. Statist.* **39** 382–389.

[7] LIGGETT, T. M. (2001). Tagged particle distributions, or how to choose a head at random. In *Proceedings of the 2000 Brazilian Summer School in Probability*. To appear. Birkhäuser, Boston.

[8] SAGAN, H. (1994). *Space-Filling Curves*. Springer, Berlin.

[9] THORISSON, H. (1996). Transforming random elements and shifting random fields. *Ann. Probab.* **24** 2057–2064.

[10] THORISSON, H. (2000). *Coupling, Stationarity and Regeneration*. Springer, Berlin.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
405 HILGARD AVENUE
LOS ANGELES, CALIFORNIA 90095–1555
E-MAIL: holroyd@math.ucla.edu
tml@math.ucla.edu