Random nearest neighbor graphs: the translation invariant case

Bounghun Bock[1,a], Michael Damron[2,b] and Jack Hanson[3,c]

[1]*School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta, GA 30332, USA,* [a]*bbock1106@gmail.com*
[2]*School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta, GA 30332, USA,* [b]*mdamron6@protonmail.com*
[3]*Department of Mathematics, City University of New York, City College, 160 Convent Ave, NAC 6/292, New York, NY 10031, USA,*
[c]*jhanson@ccny.cuny.edu*

Abstract. If $(\omega(e))$ is a family of random variables (weights) assigned to the edges of $\mathbb{Z}^d$, the nearest neighbor graph is the directed graph induced by all edges $\langle x, y \rangle$ such that $\omega(\{x, y\})$ is minimal among all neighbors y of x. That is, each vertex points to its closest neighbor, if the weights are viewed as edge-lengths. Nanda-Newman introduced nearest neighbor graphs when the weights are i.i.d. and continuously distributed and proved that a.s., all components of the undirected version of the graph are finite. We study the case of translation invariant, distinct weights, and prove that nearest neighbor graphs do not contain doubly-infinite directed paths. In contrast to the i.i.d. case, we show that in this stationary case, the graphs can contain either one or two infinite components (but not more) in dimension two, and k infinite components for any $k \in [1, \infty]$ in dimension ≥ 3. The latter constructions use a general procedure to exhibit a certain class of directed graphs as nearest neighbor graphs with distinct weights, and thereby characterize all translation invariant nearest neighbor graphs. We also discuss relations to geodesic graphs from first-passage percolation and implications for the coalescing walk model of Chaika-Krishnan.

Résumé. ???.

1. Introduction

Random nearest neighbor graphs were introduced by Nanda-Newman [17] in the context of the cubic lattice $\mathbb{Z}^d$, but we will define them on general graphs. The directed nearest neighbor graph N_D is defined on a given graph $G = (V, E)$ using real-valued edge-weights $(\omega(e))_{e \in E}$. We will assume that G does not have self-loops (it has no edges of the form $\{v, v\}$ for $v \in V$) and does not have multiple edges between any two vertices (so that each edge is uniquely identified by its endpoints). The vertex set of N_D is V and the edge set is the set

$$\{\langle x, y \rangle : \{x, y\} \in E \text{ and } \omega(\{x, y\}) \leq \omega(\{x, z\}) \text{ for all } z \text{ with } \{x, z\} \in E\}.$$

That is, each vertex points to its neighbors y which minimize the weight $\omega(\{x, y\})$. If some vertex x has infinite degree, and no neighbor minimizes ω, then x does not point to any neighbor. The undirected nearest neighbor graph N is the undirected version of N_D, with vertex set V and edge set $\{\{x, y\} : \langle x, y \rangle$ an edge of $N_D\}$. Note that if a vertex has finite degree in G, it has out-degree at least one in N_D and if the weights $(\omega(e))$ are all distinct, any vertex has out-degree at most one in N_D.

1.1. Background

In [17], Nanda-Newman studied nearest neighbor graphs on $\mathbb{Z}^d$ in the case that the weights are i.i.d. with a common uniform $(0, 1)$ distribution. (Here, a.s. each vertex has out-degree exactly one in N_D.) One of their main results was that

a.s., N has only finite components[1] and, further, that connection probabilities decay rapidly:

$$P\left(0 \text{ and } x \text{ are in the same component of } N\right) \le \frac{C^{\|x-y\|_1}}{\|x-y\|_1!}.$$

[17, Lemma 2.3] also contains a complete description of all finite clusters of N: if C is one of the connected components of N, let C_D be the directed subgraph of N_D induced by the vertices of C. Then a.s. for all C,

1. C is a tree,
2. C_D contains exactly one "miniloop" between some vertices x and y, and
3. every edge in C_D (besides $\langle x, y\rangle$ and $\langle y, x\rangle$) is directed toward both x and y.

Here, a "miniloop" is a directed circuit of length two (see the definition of a directed circuit in the next section). These results give a more-or-less complete description of i.i.d. nearest neighbor graphs on $\mathbb{Z}^d$, and were used by Nanda-Newman to study "influence graphs" arising from energy minimization procedures in disordered Ising models.

In this paper, we study the structure of N and N_D under weaker assumptions on the weights: that they are translation invariant and a.s. distinct. We first show in Theorem 1.1 that under these general conditions, N_D cannot contain doubly-infinite directed paths. Furthermore, although these graphs in the i.i.d. case contain only finite components, in Corollary 1.1, we prove that for dimension $d = 2$ N can contain exactly one or exactly two infinite components, and for dimensions $d \ge 3$ and any $k \in [1, \infty]$ N can contain exactly k infinite components. (Theorem 1.1 immediately implies that for $d = 1$ there are only finite components; see Remark 1.) These constructions follow from a general result, Theorem 1.3, which shows that directed graphs with certain properties (all vertices have out-degree one, there are no directed cycles of length at least three, and there are no doubly-infinite directed paths) can be realized as nearest neighbor graphs with distinct weights. In contrast to the situation in dimension $d \ge 3$ for dimension two, we show in Theorem 1.2 that N cannot have more than two infinite components. This result follows from a detailed analysis of the topological structure of infinite components in the plane.

Although random nearest neighbor graphs on $\mathbb{Z}^d$ appear only to have been studied by Nanda-Newman, similar graphs have appeared in the literature. For example, [13] studies percolation properties of some neighbor graphs, and in the context of Poisson models [2, 12, 15] and bipartite graphs [18], several authors have studied nearest neighbor-type models. Other work on Poisson models in [8, 14] establishes absence of percolation under various conditions by way of the mass transport principle. Furthermore, Chaika-Krishnan [6, 7] have introduced models of stationary coalescing walks, and the related directed graphs share some of the features of our graphs. In Section 1.3, we will explain the implications our results have for their models.

One motivation for studying nearest neighbor models with translation invariant weights comes from geodesic graphs constructed in first-passage percolation [1, 4, 9]. (See in particular above [4, Prop. 2.4] for the definition of geodesic graphs.) These are distributional limits of directed graphs whose edge sets are unions of point-to-hyperplane geodesics. It is known [4, Thm. 1.1] that geodesic graphs in any dimension $d \ge 2$ do not contain doubly-infinite paths, but each of their vertices has out-degree one and there are no directed cycles. Therefore they satisfy the conditions of Theorem 1.3 and can be realized as nearest neighbor graphs. It is an important question to determine the number of infinite components of geodesic graphs in general dimensions, It is known that there is only one component in two dimensions, but for $d \ge 3$ the number of infinite components is not even known to be 1 or infinity. The results of this paper show that there are nearest neighbor graphs in any dimension $d \ge 3$ with any number of infinite components, so any work on these questions for geodesic graphs must use more detailed properties of the percolation model.

1.2. Main Results

Our probabilistic results will concern the cubic lattice $\mathbb{Z}^d, E^d$ with translation invariant weights. (General graphs are considered below.) For this reason, our probability space will be the product space $\Omega = \mathbb{R}^{E^d}$ for some $d \ge 1$ with the product Borel sigma-algebra. Our probability measure P will be assumed to satisfy the following conditions:

---

[1]A reviewer of this paper has suggested the following short argument that in i.i.d. nearest neighbor graphs, N has only finite components. Let $b \in \mathbb{R}$ such that $p := P(\omega(\cdot) > b) > 0$. Let ℓ be a positive integer. Consider the event $A(0, b, \ell)$ corresponding to "there exists a directed path of length ℓ starting at 0 and whose weight of the first edge is smaller than b." The i.i.d. hypothesis should allow to prove that $P(A(0, b, \ell + 1)) \le (1 - p^{3^{d-1}})P(A(0, b, \ell))$ for any ℓ and p as defined as before. It is then not difficult to conclude from this inequality.

Assumption A:

A1. P is translation invariant. That is, for any $z \in \mathbb{Z}^d$, $P = P \circ T_z^{-1}$, where T_z is the translation by z: for $\omega \in \Omega$

(1.1) $$T_z \omega \in \Omega \text{ is given by } (T_z \omega)(e) = \omega(e + z),$$

where $e + z = \{x + z, y + z\}$ if $e = \{x, y\}$

A2. For any distinct $e, f \in E$, $P(\omega(e) = \omega(f)) = 0$

Note that assumption **A** holds if the weights are i.i.d. with a continuous common distribution. Furthermore, under item **A2**, a.s. every vertex has out-degree exactly one in N_D.

Our first result states that under assumption **A**, N_D has no infinite backward paths. For its statement, if x, y are vertices of a directed graph, we write $x \to y$ if there is a directed path from x to y. We use the convention that $x \to x$ for any x, so the graph C_x defined below always has at least one vertex.

Theorem 1.1. *Let $d \geq 1$ and P be a measure satisfying assumption* **A**. *For any $x \in \mathbb{Z}^d$, write C_x for the subgraph of N_D induced by the vertices y such that $y \to x$ in N_D. Then*

$$a.s., C_x \text{ is finite for all } x \in \mathbb{Z}^d.$$

We will prove Theorem 1.1 in Section 2.1. Because it relies on the mass transport principle, the argument can be extended to more general graphs satisfying the unimodular condition (see [16, Sec. 8.2]).

Remark 1. By Theorem 1.1, in the case $d = 1$ all components of N must be finite. Indeed, by translation invariance, a.s. one of the following three must occur: (a) all edges in N_D point left, (b) all edges point right, or (c) there are infinitely many left-pointing and right-pointing edges (and each left-pointing edge has a right-pointing edge somewhere to its left and somewhere to its right). Cases (a) and (b) cannot occur by the theorem. In case (c), all components of N are finite.

Theorem 1.1 states that nearest neighbor graphs cannot have infinite backward paths in any dimension. It is natural then to ask whether they can have infinite components at all and, if so, then how many there can be. In the next result we show that in two dimensions, there can be at most two infinite components.

Theorem 1.2. *Let $d = 2$ and P be a measure satisfying assumption* **A**. *A.s., N has at most two infinite components.*

We prove Theorem 1.2 in Section 2.2. The argument gives more information than what is stated in the theorem. It shows that if there are two infinite components, their closures in $\mathbb{Z}^2$ must be topological half-planes possibly separated by infinitely many finite components. In Remark 3, we show that such finite separating components need not exist, but if they do, they can be isolated from each other.

In the third result, we show that certain directed graphs (and therefore certain random graph models) can be realized as nearest neighbor graphs. In its statement, a directed cycle of length ℓ in a directed graph is a sequence of directed edges $\langle x_0, x_1 \rangle, \langle x_1, x_2 \rangle, \ldots, \langle x_{\ell-1}, x_\ell \rangle$ such that $x_0, \ldots, x_{\ell-1}$ are all distinct and $x_\ell = x_0$.

Theorem 1.3. *Let $G = (V, E)$ be a graph such that E is countable and let $\vec{G} = (V, \vec{E})$ be a directed graph with the same vertex set V. Assume that*

1. *if $\langle x, y \rangle \in \vec{E}$ then $\{x, y\} \in E$,*
2. *each $x \in V$ has out-degree one in $\vec{G}$,*
3. *$\vec{G}$ has no directed cycles of length at least three, and*
4. *for each vertex $x \in V$, writing C_x for the subgraph of $\vec{G}$ induced by $y \in V$ such that $y \to x$ in $\vec{G}$, C_x is finite.*

There exists a collection $(\omega(e))_{e \in E}$ of distinct weights such that the nearest neighbor graph N_D corresponding to these weights is $\vec{G}$.

The proof of Theorem 1.3 will be given in Section 2.3.

Remark 2. Theorem 1.3 states that items 1-4 are sufficient for a given graph $\vec{G}$ to be a nearest neighbor graph. If the graph $\vec{G}$ is a random directed graph sampled from a translation invariant distribution, then the definition of the weights in (2.11) ensures that the corresponding P satisfies assumption **A**. Conversely, the statement and proof of Theorem 1.1 (see the enumerated properties of N in Section 2.1) show that in the case where $G = \mathbb{Z}^d$ (or more generally, a graph satisfying the unimodular condition), these properties listed in items 1-4 are also necessary. In other words, these results characterize all nearest neighbor graphs under assumption **A**.

As a consequence, we can construct various different random nearest neighbor graphs on $\mathbb{Z}^d$ with $d \geq 2$ for measures satisfying **A**.

Corollary 1.1.   *1. Let $d = 2$ and $k \in \{0, 1, 2\}$. There is a measure $\mathbf{P}$ satisfying $\mathbf{A}$ such that a.s., N has exactly k infinite components.*

   2. Let $d \geq 3$ and $k \in \{0, 1, 2, \ldots\} \cup \{\infty\}$. There is a measure $\mathbf{P}$ satisfying $\mathbf{A}$ such that a.s., N has exactly k infinite components.

The proof of Corollary 1.1 will be given in Section 2.4. Because nearest neighbor graphs with i.i.d., continuously distributed weights provide examples with $k = 0$ we need only focus on the cases $k \geq 1$.

1.3. Implications for coalescing walks

Our results, especially those of Theorem 1.2 in the Z^2 setting, relate to some of those of Chaika-Krishnan [6, 7]. The "stationary coalescing walk" model studied there amounts to a directed random graph G (whose distribution is ergodic under lattice shifts) on the undirected graph (Z^d, E^d), satisfying conditions similar to the items in our Theorem 1.3. The main changes to these conditions are a) there are no cycles (i.e. "miniloops" are disallowed), b) C_x is allowed to be infinite, and c) paths are assumed to pass hyperplanes: given any infinite directed path $(x_0, x_1, \ldots)$ in G, we have for each k, $x_i \cdot e_1 > k$ for all large i. (The model of [6] is defined more generally, but these are the assumed conditions for their theorems about component structure on Z^2.)

   In this setting, in the case $d = 2$, Chaika-Krishnan show a dichotomy: either C_x is a.s. finite for each x and also the undirected version of G has one component, or a.s. each infinite directed path in G contains a site x with $\#C_x = \infty$ (and the undirected version of G must have infinitely many components). In their language, one says that G either exhibits coalescence without bi-infinite trajectories, or each component contains a bi-infinite trajectory. While assumption c) is natural in some coalescing walk models (notably first-passage percolation), there are many examples for which it fails, and for these, the dichotomy can be false. See the example in Section 2.4.1 below, which exhibits neither bi-infinite trajectories nor coalescence, its undirected version having two infinite components a.s.

   Our Theorem 1.2 shows that this example demonstrates the most extreme failure of coalescence allowed: if G does not exhibit bi-infinite trajectories, its undirected version has at most two components a.s. In other words, if the undirected version of G has at least three infinite components, then G must exhibit bi-infinite trajectories. It is perhaps worth noting that the dichotomy breaks down in other ways without assumption c): for instance, see [3] for a model which exhibits a bi-infinite trajectory and also exhibits coalescence. Perhaps the techniques used to prove Theorem 1.2 can be used to completely classify the allowed behavior of stationary coalescing walks which do not necessarily pass hyperplanes.

2. Proofs

In this section, we prove the main results, starting with Theorem 1.1 in Section 2.1, moving to Theorem 1.2 in Section 2.2, and finishing with Theorem 1.3 in Section 2.3 and Corollary 1.1 in Section 2.4.

2.1. Proof of Theorem 1.1

Throughout the proof we make assumption $\mathbf{A}$. We first note that for $x, y, z \in Z^d$ with $x \neq z$, from $\mathbf{A2}$,

(2.1)   a.s. if $\langle x, y \rangle$ and $\langle y, z \rangle$ are edges in N_D then $\omega(\{x, y\}) > \omega(\{y, z\})$.

As a consequence, we have the following facts:

1. N_D a.s. has no directed cycles of length at least 3. To show this, consider an outcome ω for which $\omega(e) \neq \omega(f)$ for all $e \neq f$ and a directed cycle with vertices $x_0, x_1, \ldots, x_k$ such that $\langle x_i, x_{i+1} \rangle \in N_D$ for $i = 0, \ldots, k-1$, with $x_0, \ldots, x_{k-1}$ all distinct and $x_0 = x_k$ (so that the cycle has length k). Writing x_{k+1} also for x_1, note that if $k \geq 3$ then for each $i = 0, \ldots, k-1$, the edges $\{x_i, x_{i+1}\}$ and $\{x_{i+1}, x_{i+2}\}$ are distinct and share an endpoint x_{i+1}, so $\omega(\{x_i, x_{i+1}\}) > \omega(\{x_{i+1}, x_{i+2}\})$. Iterating this bound, we obtain

$$\omega(\{x_k, x_{k+1}\}) = \omega(\{x_0, x_1\}) > \cdots > \omega(\{x_{k-1}, x_k\}) > \omega(\{x_k, x_{k+1}\}),$$

   a contradiction.
2. For $x \in Z^d$, write Γ_x for the subgraph of N_D induced by the vertices y such that $x \to y$ in N_D. Then a.s. there are two possibilities:

   (a) Γ_x is finite and ends in a cycle of length two.

   (b) Γ_x is an infinite vertex self-avoiding directed path.

To see why, since each vertex has out-degree one, we can follow each out-edge starting at x and label the vertices in order as $x = x_0, x_1, \ldots$. There are two possibilities: either all x_i's are distinct, or there is a first $i \geq 2$ such that x_i is an element of $\{x_0, \ldots, x_{i-1}\}$. In the first case, Γ_x is an infinite vertex self-avoiding directed path. In the second, item 1 implies that $x_i = x_{i-2}$, and so Γ_x ends in a cycle of length two.

3. If Γ_x is infinite, write its vertices in order as $x = x_0, x_1, \ldots$. Then

$$\omega(\{x_i, x_{i+1}\}) > \omega(\{x_{i+1}, x_{i+2}\}) \text{ for all } i \geq 0.$$

In other words, Γ_x is a monotone decreasing path. This follows from (2.1) and item 2: Γ_x is an infinite vertex self-avoiding directed path and we can apply (2.1) to each pair of adjacent edges. Similarly, if Γ_x is finite and not a cycle of length two, we write its vertices in order as $x = x_0, x_1, \ldots, x_N, x_{N+1}$, where x_N is the first vertex in a cycle of length two. Then

$$\omega(\{x_i, x_{i+1}\}) > \omega(\{x_{i+1}, x_{i+2}\}) \text{ if } 0 \leq i \leq N - 1.$$

Given the three properties above, we continue with the proof of Theorem 1.1. For a contradiction, we assume that

(2.2) $$P(C_0 \text{ is infinite}) > 0.$$

By **A1**, if we show that (2.2) is false, then Theorem 1.1 will follow. Under (2.2), we also have

$$P(C_0 \text{ is infinite and } \Gamma_0 \text{ contains at least three vertices}) > 0.$$

This, along with assumption **A2**, implies that if we define for $x \in \mathbb{Z}^d$

$$I_x = \inf\{\omega(\{u, v\}) : \{u, v\} \text{ is an edge of } \Gamma_x\} \text{ and}$$
$$S_x = \sup\{\omega(\{u, v\}) : \{u, v\} \text{ is an edge of } \Gamma_x\},$$

then $P(C_0 \text{ is infinite and } I_0 < S_0) > 0$ and we can therefore find (a deterministic) r such that

(2.3) $$P(C_0 \text{ is infinite and } I_0 < r < S_0) > 0.$$

Following (2.3), for $x, y \in \mathbb{Z}^d$, we say that y is the "r-descendant of x" if

1. y is a vertex of Γ_x and, writing z for the a.s. unique vertex such that $\{y, z\}$ is an edge of N_D, one has $\omega(\{y, z\}) \geq r$ and
2. for any edge $\{u, v\}$ of Γ_z, one has $\omega(\{u, v\}) < r$

In other words, y is the last vertex in Γ_x whose out-edge has weight $\geq r$.

Define the following random variable ("mass transport function") $m(x, y)$ for $x, y \in \mathbb{Z}^d$:

$$m(x, y) = \begin{cases} 1 & \text{if } y \text{ is the } r\text{-descendant of } x \\ 0 & \text{otherwise.} \end{cases}$$

Because m satisfies $m(x, y)(\omega) = m(x - z, y - z)(T^z \omega)$ for $x, y, z \in \mathbb{Z}^d$ where T^z was defined in (1.1), **A1** implies that we can apply the mass transport principle (see [11, 16] for an introduction) to obtain

(2.4) $$E \sum_{x \in \mathbb{Z}^d} m(x, 0) = E \sum_{x \in \mathbb{Z}^d} m(0, x).$$

Because $\sum_{x \in \mathbb{Z}^d} m(0, x) \leq 1$ a.s., we obtain

(2.5) $$E \sum_x m(x, 0) = E \sum_{x \in \mathbb{Z}^d} m(0, x) \leq 1.$$

However a.s. on the event in (2.3), there is a vertex y such that $\sum_x m(x, y) = \infty$. Indeed, if $I_0 < r < S_0$, then 0 cannot be in a cycle of length two, and item 3 gives that the weights along Γ_0 are decreasing, so 0 has an r-descendant. Furthermore, for each vertex w of C_0 (of which there are infinitely many) a.s. the graph Γ_w is equal to Γ_0 with finitely many directed

edges appended in sequence to the beginning. All of these edges by item 3 have weight $> r$ and so any such w has the same r-descendant as does 0. Writing y for this r-descendant, we obtain

$$\text{for an outcome as above,} \quad \sum_x m(x, y) \geq \sum_{w \text{ a vertex of } C_0} m(w, y) = \infty.$$

Therefore from (2.3) we deduce that

$$P\left(\text{for some } y \in \mathbb{Z}^d, \ \sum_x m(x, y) = \infty \right) > 0.$$

By **A1**, we obtain $E \sum_x m(x, 0) = \infty$ and this contradicts (2.5). We find then that (2.3) must have been false, and therefore so was (2.2). This completes the proof of Theorem 1.1.

2.2. Proof of Theorem 1.2

The proof will be split over two subsections. In Section 2.2.1, we derive some basic results about vertex sets and their boundaries. In Section 2.2.2, we analyze the structure of infinite components in N and give the proof of Theorem 1.2.

To do this, we begin with some simple definitions. If $x, y \in \mathbb{Z}^2$, then x and y are site-neighbors if $\|x - y\|_1 = 1$ (this is just a redefinition of "nearest-neighbors" in $\mathbb{Z}^2$, made to distinguish from neighbors in N). A set $V \subset \mathbb{Z}^2$ of vertices is site-connected if for each $x, y \in V$ there is a path (here a sequence of vertices $x = x_0, x_1, \ldots, x_n = y$ such that x_i and x_{i+1} are site-neighbors for all i) from x to y remaining in V. A site-component of V is a maximal site-connected subset of V. As usual, in addition to the (primal) lattice $\mathbb{Z}^2$, E^2 with edge set E^2 consisting of those edges between neighboring vertices, we use the dual lattice $\mathbb{Z}^{2\,*}$, $E^{2\,*}$, with vertex set

$$\mathbb{Z}^{2\,*} = \mathbb{Z}^2 + \left(\frac{1}{2}, \frac{1}{2} \right)$$

and edge set $E^{2\,*}$ consisting of edges between neighbors. Each dual edge e^* bisects a unique edge e.

2.2.1. Basic topological properties of vertex sets
In this section, we derive some simple properties of vertex sets. These will be used in the main proof in the following section. In the following definition, the set $\overline{V}$ is intended to be V "with its holes filled in," so its complement should only have infinite site-components.

Definition 1. Let $V \subset \mathbb{Z}^2$.

1. The closure of V, written $\overline{V}$, is the union of V with all finite site-components of V^c.
2. The dual edge boundary of V, written $B(V)$, is the subgraph of $\mathbb{Z}^{2\,*}$, $E^{2\,*}$ induced by the set of dual edges whose unique bisecting edge $\{x, y\}$ has $x \in \overline{V}$ and $y \notin \overline{V}$.

We note some simple properties of the definitions.

Lemma 2.1. *Let $V \subset \mathbb{Z}^2$. Then $\overline{V}^c$ has only infinite site-components.*

Proof. Assume for a contradiction that $\overline{V}^c$ had a finite site-component containing a vertex x. Since $\overline{V}^c \subset V^c$ and the finite site-components of V^c are in $\overline{V}$, x must be in an infinite site-component of V^c. Then pick a vertex self-avoiding path π, starting from x and remaining in V^c, which has infinitely many vertices. This path π must leave the site-component of $\overline{V}^c$ containing x, so it contains a vertex $q \in \overline{V}$. But $q \in \pi \subset V^c$, so q must be in $\overline{V} \setminus V$ and therefore is in a finite site-component $\tilde{C}_q$ of V^c. However then $\pi \subset \tilde{C}_q$, giving a contradiction since π is infinite. □

Next we show that the vertices on the site-boundary of $\overline{V}$ are actually in V. (Otherwise, they would be in finite holes in the complement of V.) As a consequence, if V is site-connected, so is $\overline{V}$.

Lemma 2.2. *If $x \in \overline{V}$ has a neighbor in $\overline{V}^c$, then $x \in V$.*

Proof. The neighbor y of x that is in $\overline{V}^c$ is not in V, and is therefore in an infinite site-component $\tilde{C}$ of V^c (otherwise it would be in $\overline{V}$). If x were not in V, then it would be in $\tilde{C}$, as it is adjacent to an element of this site-component. But this means x would be in an infinite site-component of V^c and therefore would not be in $\overline{V}$, a contradiction. □

Lemma 2.3. *Let* $V \subset \mathbb{Z}^2$ *be site-connected. Then each vertex in the graph* $B(V)$ *has degree two. Therefore* $B(V)$ *is a vertex-disjoint union of vertex self-avoiding circuits and vertex self-avoiding doubly-infinite paths. If* V *is also infinite, then* $B(V)$ *contains no vertex self-avoiding circuits.*

Proof. Because $B(V)$ is induced by a set of edges, each vertex has at least degree one. Because edges of $B(V)$ separate $\overline{V}$ from $\overline{V}^c$, its vertices must have degree 2 or 4, so we rule out degree 4. Assume for a contradiction that some dual vertex x has degree 4 in $B(V)$. Then we can enumerate the (primal) vertices at distance $\overline{\sqrt{2}}/2$ of x by $x_1, \ldots, x_4$ in clockwise order so that $x_i \in \overline{V}$ if and only if i is odd. Because x_1 and x_3 are site-neighbors of $\overline{V}^c$, Lemma 2.2 implies that they are in V. We can then choose a vertex self-avoiding path π from x_1 to x_3 which remains in V, and then a plane curve P which starts at x_1 and proceeds as follows. First, P connects the vertices of π in order by straight line segments. At x_3, P connects to x_1 by a diagonal line segment (going through x). P is a Jordan curve, and therefore its complement (in $\mathbb{R}^2$) has two components: one bounded (its interior) and one unbounded (its exterior). Proceeding in a straight line from x_2 to x_4, we cross P exactly once. Since neither x_2 nor x_4 is on P, one must be in each component. By symmetry, let's say that x_2 is in the interior. Then $x_2 \in V^c$ and must be in a bounded site-component of V^c, since any infinite vertex self-avoiding path starting at x_2 must leave the interior of P and therefore touch π (it cannot touch the interior of the other segment composing P). This is a contradiction, since $x_2 \in \overline{V}^c$, and $\overline{V}$ contains the bounded site-components of V^c. We conclude that x must have degree 2 in $B(V)$.

Now suppose that V is infinite and, for a contradiction, assume that $B(V)$ contains a vertex self-avoiding circuit. Again, form a Jordan curve P by proceeding along the circuit, using straight line segments to connect its vertices. Let $x \in V$ and $y \in \overline{V}^c$ be such that the edge $\{x, y\}$ bisects a dual edge in P. Then x and y are in different components of the complement of P (the edge $\{x, y\}$ crosses P exactly once), so one of them is in the bounded component. First suppose that it is x; then because V is infinite and connected, there is an infinite vertex self-avoiding path π_x started at x which remains in V. But π_x must then exit the bounded component of the complement of P and cross P to a vertex of $\overline{V}^c$, which is a contradiction. If instead y is in the bounded component, then by Lemma 2.1, we can choose an infinite vertex self-avoiding path π_y started at y which remains in $\overline{V}^c$. By the same reasoning, we obtain another contradiction. Therefore $B(V)$ contains no vertex self-avoiding circuit. $\qquad\square$

2.2.2. *Topology of nearest neighbor components*

Our aim is to show that for $d = 2$ if the number of infinite components of N is ≥ 2 then it must be 2. Essentially we will show that in this case, the two components are both topologically half-planes which can come within distance 1 of each other, and they may be separated by finite components of N. The first step is to show that if C is an infinite component of N, then the closure of its vertex set is topologically either a full-plane or half-plane. We do this in the following proposition.

Proposition 2.1. *For any component* C *of* N*, write* $V(C)$ *for its vertex set and* $B(C)$ *for* $B(V(C))$*. Then under assumption* **A***, a.s., for each infinite component* C *of* N*,* $B(C)$ *is either empty or is a vertex self-avoiding doubly-infinite path.*

Proof. Assume for a contradiction that

$$(2.6) \qquad \mathbb{P}\left(\begin{array}{l} B(C) \text{ contains at least two distinct vertex self-avoiding} \\ \text{doubly-infinite paths for some infinite component } C \text{ of } N \end{array}\right) > 0.$$

We will show that for almost every outcome in this event, there is a vertex x with $\#C_x = \infty$. This will be a contradiction, as we have shown in Theorem 1.1 that this has zero probability.

By Lemma 2.3, the two doubly-infinite paths in (2.6) can be assumed to be vertex disjoint. So, for such an outcome, let C be an infinite component of N and let π_1 and π_2 be vertex disjoint, vertex self-avoiding doubly-infinite paths contained in $B(C)$. Choose a dual edge e_0 in π_1 and enumerate the edges in either direction along π_1 as $\ldots, e_{-1}, e_0, e_1, \ldots$. For $n \in \mathbb{Z}$ write $\{x_n, x_n^0\}$ for the edge which bisects e_n, so that $x_n \in V(C)$ and $x_n^0 \in \overline{V(C)}^c$. For $n \geq 1$ since x_n, x_{-n} are in $V(C)$, there is a vertex self-avoiding path P_n in N connecting x_n to x_{-n}. Fixing any $y \in V(C)$ such that some edge $\{y, y^0\}$ bisects an edge of π_2, we can also find a path π_3 in N connecting x_0 to y. See Figure 1.

We will now argue by the Jordan curve theorem that P_n must intersect π_3. To do this, define the plane curve P_n as follows. It proceeds from x_n to x_{-n} along the edges of N that connect vertices of P_n, then it connects x_{-n} to π_1 halfway through the edge $\{x_{-n}, x_{-n}^0\}$, then it proceeds along π_1 until it meets the edge $\{x_n, x_n^0\}$, and last moves to x_n halfway through this edge. The curve P_n is a Jordan curve, and so its complement in $\mathbb{R}^2$ has a bounded component (the interior) and an unbounded component (the exterior). We first note that

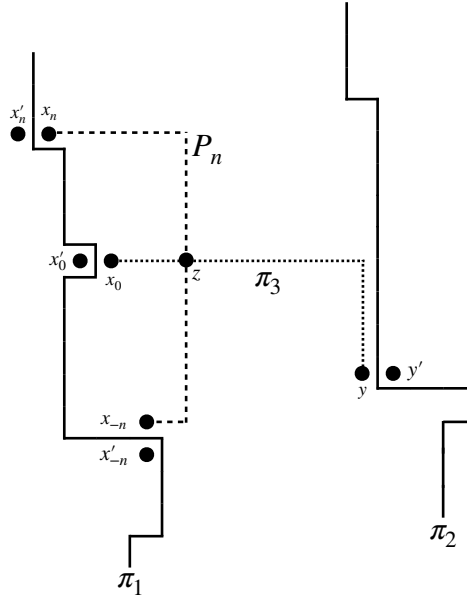$$(2.7) \qquad \text{for } n \geq 1, \; \overline{V(C)}^c \subset \text{ext } P_n.$$

FIG 1. Illustration of the argument for Proposition 2.1. The closure $\overline{V(C)}$ of the component C lies between the two paths π_1, π_2, which are part of the boundary $B(C)$. The x_i's are the endpoints of edges dual to those of π_1 that lie in $V(C)$, and P_n is a path in N connecting x_n to x_{-n}. The path π_3 is in N and connects x_0 to y, an endpoint of an edge dual to one in π_2. By the Jordan curve theorem, P_n must intersect π_3 at some vertex z. Because this is true for all n, one can argue that some such z has $\#C_z = \infty$.

Indeed, if $u \in \overline{V(C)}^c$, then (by Lemma 2.1) u is in an unbounded site-component of $\overline{V(C)}^c$, and so there is an infinite vertex self-avoiding path π starting from u and remaining in this component. However π cannot touch P_n because it never leaves this component and therefore never comes in contact with P_n or π_1. Therefore u must be in the unbounded component of the complement of P_n, and this shows (2.7).

Due to (2.7), we can now argue that x_0 is either on P_n or in its interior. So suppose that x_0 is not on P_n; this means it is not on P_n. Then the edge $\{x_0, x_0^0\}$ only touches P_n at π_1, and it only touches it once. However x_0^0 is in the exterior of P_n, so x_0 must be in the interior. By similar reasoning, we can argue that y (the other endpoint of π_3) is either on P_n or in its exterior. Indeed, if it is not on P_n, then it is not on P_n, and then the edge $\{y, y^0\}$ does not touch P_n but ends at the vertex y^0, which is in the exterior. Therefore y is also in the exterior. Because x_0 is not in the exterior of P_n, and y is not in the interior, the curve formed by following straight line segments between the vertices of π_3 must touch P_n. Since π_3 doesn't leave C, it must touch a vertex of P_n. Therefore we find that for all n, P_n shares a vertex with π_3, as desired.

Because π_3 is finite, there is a $z \in \pi_3$ that is in P_{n_k} for some subsequence (n_k) of integers with $n_k \to \infty$. We claim that either $x_{n_k} \to z$ in N_D or $x_{-n_k} \to z$ in N_D. This is clear if one of x_{n_k} or x_{-n_k} equals z. Otherwise, z has degree two on P_{n_k}. Since a.s., all vertices have out-degree one in N_D (and therefore so does z), one of its neighbors on P_{n_k}, say w_0, is such that $\langle w_0, z \rangle$ is an edge of N_D. Enumerating the subsequent vertices (beyond w_0) of P_{n_k} as $w_1, \ldots, w_r$, each w_i has out-degree one in N_D, so $\langle w_{i+1}, w_i \rangle$ is an edge of N_D. Thus $w_r \to z$ in N_D. Since $w_r = x_{n_k}$ or x_{-n_k}, this proves the claim and furthermore establishes that (2.6) implies

$$P\left(\exists z \in \mathbb{Z}^2 \text{ such that } \#C_z = \infty\right) > 0.$$

This is a contradiction and completes the proof. □

Now that each infinite component of N is topologically a half-space, we must show that the complement cannot contain more than one other infinite component. This could happen, for example, if three infinite components were separated by an infinite union of finite components with at least three topological ends, or if three infinite components come within distance 1 of each other. To rule out these and other possibilities, we will make use of the results of Burton-Keane in [5]. To begin, we make a few definitions, and state a structural lemma which follows from arguments of [5].

Definition 2. Write X for the subset of $\mathbb{Z}^2$ defined by

$$X = \mathbb{Z}^2 \setminus \bigcup_C \overline{V(C)},$$

where the union is over all infinite components of N. Then $\mathbb{Z}^2$ is the disjoint union of sets of the following three types:

(a) $\overline{V(C)}$ for an infinite component C of N,

(b) an infinite site-component of X, and

(c) a finite site-component of X.

Note that because the complement of X contains only infinite site-components, each set of type (b) or (c) is equal to its closure. For the next lemma, we say that vertices x and y are $*$neighbors (or $*$adjacent) if $\|x - y\|_\infty = 1$. We extend this notion to sets in the usual way.

Lemma 2.4. *Under assumption* **A**, *a.s., none of the following occur.*

1. *There is a set of type (b) whose complement has at least 3 site-components.*
2. *There is a dual vertex within Euclidean distance $\sqrt{2}/2$ of three different sets of types (a) or (b).*
3. *There is a set of type (c) which is $*$adjacent to at least three different sets of types (a) or (b).*

Proof. Item 1 is a direct application of [5, Thm. 2], which states that in stationary site percolation on $\mathbb{Z}^2$, a.s., there is no "ribbon" whose complement contains at least 3 site-components. To apply this result, we define variables $(x_v)_{v \in \mathbb{Z}^2}$ as $x_v = \mathbf{1}_{\{v \in X\}}$. The x_v form a stationary site percolation, and any infinite site-component of X (a type (b) set) is a "1-ribbon" (the closure of an infinite 1-cluster). The result follows.

Items 2 and 3 have similar proofs, and follow that of [5, Thm. 1] (see also [10]), which states that in stationary site percolation on $\mathbb{Z}^2$, a.s., no "rock" has at least 3 ribbons as $*$neighbors. Because the details are the same, we only sketch the proofs. To any dual vertex v as described in item 2, we associate three infinite vertex self-avoiding site-paths $\gamma_i(v)$ in different sets of types (a) or (b), starting from vertices within Euclidean distance $\sqrt{2}/2$ of v. For any such $v \in [0, n]^2$, consider the first intersection $x_i(v)$ of $\gamma_i(v)$ with $\partial[0, n]^2$, and ordering them so that $x_1(v)$, $x_2(v)$, $x_3(v)$ are in counterclockwise order, we say that $x_2(v)$ is the "central point" associated with v. (It appears that even in the original argument of [5], a bit more care needs to be taken to define the central point: the $x_i(v)$ should be chosen as functions of the intersection of their corresponding (a) or (b) set with $[0, n]^2$.) One then uses a Jordan curve argument to prove that central points corresponding to different v's are distinct, and therefore there can be at most $4n$ such v's in $[0, n]^2$. (In this part, it is important that for given v, v', the starting points of the $\gamma_i(v')$'s must all lie in the closure of one of the regions between the $\gamma_i(v)$'s.) However, if the event described in item 2 has positive probability, translation invariance implies that the expected number of v in $[0, n]^2$ is of order n^2, a contradiction for large n.

The argument for item 3 is similar, defining three paths corresponding to each type (c) set as described, paths γ_i, and corresponding central points. Again, central points associated to distinct such type (c) sets are distinct, and we conclude as above. □

Given these preparations, we now prove that N has at most 2 infinite components.

Proof of Theorem 1.2. We start by defining the $*$boundary for an infinite component C of N, using our dual path $B(C)$. Writing D for the event in Proposition 2.1, consider a configuration in D in which N has at least 2 infinite components. Let C be any one of them, and note that $B(C)$ is a (nonempty) doubly-infinite, vertex self-avoiding dual path. Enumerate the dual edges of $B(C)$ as $\ldots, e_{-1}, e_0, e_1, \ldots$, and write x_i for the vertex of $\overline{V(C)}^c$ which is an endpoint of the edge whose dual is e_i. The sequence (x_n) is $*$connected (it is in fact a $*$path), but it not necessarily site-connected. To remedy this, we define a doubly-infinite sequence $P = P(C)$ of vertices by following the x_i's, but inserting between any x_i and x_{i+1} which are $*$neighbors but are not site-neighbors their unique common site-neighbor which is in $\overline{V(C)}^c$. (For example, if $x_i = (0, 0)$ and $x_{i+1} = (1, 1)$ with $e_i = \{(-1/2, 1/2), (1/2, 1/2)\}$ and $e_{i+1} = \{(1/2, 1/2), (1/2, 3/2)\}$ then this neighbor is the vertex $(1, 0)$. In fact, this is the only possible case up to translation, reflection, and rotation by multiples of $\pi/2$.) Then P is a natural enumeration of the $*$boundary of $\overline{V(C)}$ and is clearly site-connected (it is a path which might not be vertex self-avoiding). For our given C and P, we will consider the different possible sets of types (a)-(c) which can intersect P.

First, we argue that no infinite component can be a $*$neighbor to two other sets of type (a) or (b):

(2.8) $$\mathbf{P}\left(\begin{array}{c} \exists x, y \in P = P(C) \text{ in different sets of type (a)} \\ \text{or (b) for some infinite component } C \text{ of } N \end{array} \right) = 0.$$

For a contradiction, assume that this probability is positive. We will show that in this case, we can find either a dual vertex as in item 2 of Lemma 2.4 or a set of type (c) as in item 3 of Lemma 2.4. To do this, consider an outcome in the intersection of D with the event in (2.8), with C, P, x, y as described, and write B_1 for the set of type (a) or (b) containing x. Following P from x to y, we must exit B_1.
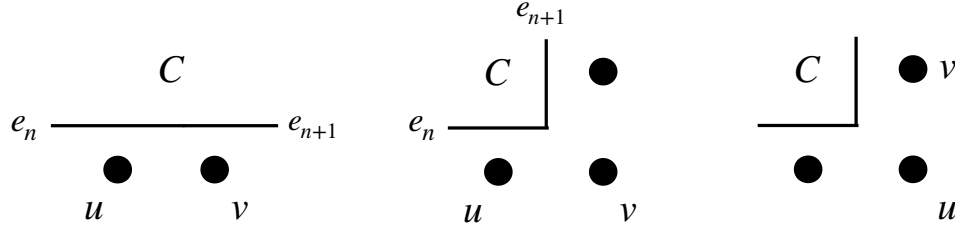
FIG 2. The three possibilities in the argument of case 1 for equation (2.8). In the first, $u = x_n$ for some n and v is as well. In the second, $u = x_n$ for some n, but v is not, although it is in P. In the last, $u \in P$ is not of the form x_n but v is. In all possibilities, the central dual vertex (taken to be $(1/2, 1/2)$) satisfies the condition of item 2 of Lemma 2.4.

Case 1. If we enter a different type (a) or (b) set, say B_2, then write u for the last point in P of B_1 before entering B_2, and v for the first point of B_2. If $u = x_n$ for some n, then after translating, rotating, and reflecting, we may assume that $u = (0, 0)$ and $(0, 1) \in V(C)$, so that $e_n = \{(-1/2, 1/2), (1/2, 1/2)\} \in B(C)$, $e_{n+1} = \{(1/2, 1/2), (3/2, 1/2)\}$. Then $v = x_{n+1} = (1, 0)$, then the dual vertex $(1/2, 1/2)$ satisfies item 2 of Lemma 2.4 with sets B_1, B_2, $\overline{V(C)}$ (See Figure 2.) If instead $e_{n+1} = \{(1/2, 1/2), (1/2, 3/2)\}$, then $v = (1, 0)$ is in P but not in the sequence (x_n); however, $(1/2, 1/2)$ still satisfies item 2 with sets B_1, B_2, $\overline{V(C)}$. We cannot have $e_{n+1} = \{(1/2, 1/2), (1/2, -1/2)\}$ because then $v = x_n = u$, but u and v are in different sets of type (a) or (b). The last possibility is that u is not any of the x_n's, in which case after translating, rotating, and reflecting, we may assume that $u = (1, 0)$ and $v = (1, 1)$ with $(0, 1) \in V(C)$. Here, again the vertex $(1/2, 1/2)$ satisfies item 2 with the same sets B_1, B_2, $\overline{V(C)}$.

Case 2. The other possibility is that, after we leave B_1, we enter a set Y of type (c). In this case, B_1 must be of type (a). If, as we proceed along P, we next re-enter B_1, we simply wait until we leave B_1 once again. Otherwise, once we leave Y, we enter another set B_2 of type (a) which is not B_1 and also not $\overline{V(C)}$. In this case, the set Y is site-adjacent to both B_1 and B_2 but also $*$adjacent to $\overline{V(C)}$ (as every vertex of P is $*$adjacent to $\overline{V(C)}$). Therefore Y satisfies item 3 of Lemma 2.4.

In either case, assuming that (2.8) fails implies that at least one of the events described in Lemma 2.4 has positive probability, a contradiction. Therefore (2.8) holds.

Next, we argue that if any infinite component of N is a $*$neighbor of another one (by the above, the component can have at most one such $*$neighbor), then there are exactly two infinite components of N:

$$(2.9) \qquad P\left(\begin{array}{c} \exists x \in P = P(C) \text{ in a set of type (a) for some infinite} \\ \text{component } C \text{ of } N \text{ and } N \text{ has at least } 3 \text{ infinite components} \end{array}\right) = 0.$$

As before, we argue by contradiction and assume that this probability is positive. Choose any outcome in the intersection of D and this event such that the event in (2.8) does not occur, and pick C, x as described in (2.9). Note that if $\overline{V(C')}$ is the set of type (a) containing x, then infinitely many vertices of P (in both directions) are in $\overline{V(C')}$. (This implies that P consists of vertices of $\overline{V(C')}$ separated by finite segments of vertices in (c) components.) Indeed, if this were not true, then in some direction along P, all vertices from some point on would be in X (since they could not be in another type (a) set due to the event in (2.8) not occurring). But these vertices are site-connected, so they would be part of a set of type (b), and this would also lead us back to the event in (2.8).

Because the event in (2.9) occurs, we can choose, in addition to the components C and C', yet another infinite component C'' of N. Pick a site-connected path π from C'' to C and follow it until its first vertex in $\overline{V(C)} \cup \overline{V(C')}$ (this cannot be the initial vertex of π). By symmetry, we may assume that this vertex is in $\overline{V(C)}$. Let w be the vertex of π directly before it. Since $w \in P$ (as it is site-adjacent to $\overline{V(C)}$), but not in $\overline{V(C')}$, and vertices of P are either in $\overline{V(C')}$ or in type (c) sets, w must be in a set Y of type (c). Note that then Y is site-adjacent to $\overline{V(C)}$, but it is also site-adjacent to $\overline{V(C')}$, as we can follow P through Y directly to $\overline{V(C')}$. Following π backward from w until we exit Y, we must enter another type (a) set, but this set cannot be $\overline{V(C)}$ or $\overline{V(C')}$. We conclude that Y is site-adjacent to 3 sets of type (a); that is, Y satisfies the condition of item 3 of Lemma 2.4. Just as in case 1, we see that our assumption that (2.9) fails implies the event described in item 3 of Lemma 2.4 has positive probability, a contradiction. Therefore (2.9) holds.

Last, we deal with the remaining possibility: that all vertices of $P = P(C)$ are in one set of type (b), and that this holds for all infinite components C of N. In this case, we will show that the number of infinite components of N is at most two:

$$(2.10) \qquad P\left(\begin{array}{c} \text{all } x \in P = P(C) \text{ are in a set of type (b) for all infinite} \\ \text{components } C \text{ of } N \text{ and } N \text{ has at least } 3 \text{ infinite components} \end{array}\right) = 0.$$

For a contradiction, assume this probability is positive, and consider any outcome in the intersection of D and this event. Pick infinite components C_1, C_2, C_3 of N and write P_1, P_2, P_3 for their corresponding $*$boundary paths. Then P_i is contained entirely in some set Y_i of type (b). In this case, we will show that all Y_i's are equal and that their complement has at least 3 site-components, as in item 1 of Lemma 2.4.

Choose a site-connected path π from C_i to C_j for some $i \neq j$. Let w be the vertex directly before entering $\overline{V(C_j)}$ for the first time and note that since w is site-adjacent to $\overline{V(C_j)}$, it must be in P_j, and therefore in Y_j. Following π backward from w until it last leaves Y_j at some vertex z, we see that z must be in $\overline{V(C_i)}$. The reason is that otherwise z must be in another set of type (a), and this set would have $*$boundary path fully contained in Y_j (as it is site-adjacent to Y_j), so to reach C_i, we would need to re-enter Y_j, contradicting the definition of z. We find, therefore, that Y_j is site-adjacent to $\overline{V(C_i)}$ as well, and so $Y_i = Y_j$.

Because $Y_1 = Y_2 = Y_3$, each of $\overline{V(C_1)}$, $\overline{V(C_2)}$, $\overline{V(C_3)}$ are $*$adjacent only to Y_1, This implies that the sets $\overline{V(C_i)}$ are contained in different site-components of Y_1^c (otherwise we could move from one to the other without touching Y_1). Thus Y_1 satisfies the condition of item 1 of Lemma 2.4. This means that if we assume that (2.10) fails, then the event described in item 1 has positive probability, a contradiction. Therefore (2.10) holds. This completes the proof. □

2.3. Proof of Theorem 1.3

Suppose that $G = (V, E)$ is a graph such that E is countable and $\mathsf{G} = (V, \mathsf{E})$ is a directed graph with the properties stated in the theorem: if $\langle x, y \rangle \in \mathsf{E}$ then $\{x, y\} \in E$, each $x \in V$ has out-degree one in G, G has no directed cycles of length at least three, and for each vertex $x \in V$, C_x is finite. We define weights $\omega(e)$ for edges $e \in E$ as follows. Let $(U(e))_{e \in E}$ be a collection of i.i.d. uniform $[0, 1]$ random variables and let E be the set of edges $\{x, y\} \in E$ such that $\langle x, y \rangle \in \mathsf{E}$. If $e \in \mathsf{E}$ with $e = \{x, y\}$ and $\langle x, y \rangle \in \mathsf{E}$, write $V(e)$ for the number of vertices in C_x. Note that V is well-defined: if $\langle x, y \rangle$ and $\langle y, x \rangle$ are both in E, then the graphs C_x and C_y are the same, so they have the same number of vertices. Our definition of $\omega(e)$ is

$$(2.11) \qquad \omega(e) = \begin{cases} \dfrac{1}{V(e) + U(e)} & \text{if } e \in \mathsf{E} \\ 1 + U(e) & \text{otherwise.} \end{cases}$$

Note that a.s., the weights $\omega(e)$ are distinct. Therefore to prove Theorem 1.3, we will show that a.s., the nearest neighbor graph N_D constructed from the weights is equal to G. First suppose that $\langle x, y \rangle$ is an edge of G; we will prove that it is in N_D. To do this, we will show that for $z \neq y$ with $\{z, x\} \in E$ we have $\omega(\{z, x\}) > \omega(\{x, y\})$. There are two cases. If $\{z, x\} \notin \mathsf{E}$, then a.s.

$$\omega(\{z, x\}) = 1 + U(\{z, x\}) > 1 > \frac{1}{V(\{x, y\}) + U(\{x, y\})} = \omega(\{x, y\}),$$

since $V(\{x, y\}) \geq 1$. Alternatively, if $\{z, x\} \in \mathsf{E}$, then because x has out-degree one in G, $\langle z, x \rangle \in \mathsf{E}$ and therefore $V(\{z, x\})$ is the number of vertices in C_z. We claim that C_z is strictly contained in C_x: each vertex of C_z is in C_x, but $x \notin C_z$. Assuming this for the moment, we obtain $V(\{x, z\}) \leq V(\{x, y\}) - 1$ and so a.s.

$$\omega(\{x, y\}) = \frac{1}{V(\{x, y\}) + U(\{x, y\})} < \frac{1}{V(\{z, x\}) + 1} \leq \frac{1}{V(\{z, x\}) + U(\{z, x\})} = \omega(\{z, x\}).$$

This implies that a.s. if $\langle x, y \rangle \in \mathsf{E}$ then $\langle x, y \rangle$ is an edge of N_D.

To prove the claim, observe that if w is a vertex of C_z then it is clearly a vertex of C_x: any directed path from w to z can be extended to x by simply appending the edge $\langle z, x \rangle$ to the end. So we assume for a contradiction that x is a vertex of C_z. Then there is a directed path π from x to z in G (which we may assume is vertex self-avoiding). This path cannot contain an edge from x to z since $\langle x, z \rangle \notin \mathsf{E}$, so it must first visit some u which is not equal to x or z. But then appending the edge $\langle z, x \rangle$ to the end of π produces a directed cycle of length at least three in G, a contradiction. We conclude that x is not a vertex of C_z and therefore the claim holds.

To complete the proof of Theorem 1.3, suppose that $\langle x, y \rangle \notin \mathsf{E}$ but $\{x, y\} \in E$. Since x has out-degree one in G, there is some $z \neq y$ such that $\langle x, z \rangle \in \mathsf{E}$. As we saw above, this implies a.s. that $\langle x, z \rangle$ is an edge of N_D. Since x a.s. has out-degree at most one in N_D (as the weights are distinct), we find that $\langle x, y \rangle$ cannot be an edge of N_D. This shows that the edges of G and N_D are the same, and finishes the proof.

2.4. Proof of Corollary 1.1

To prove the corollary, we will build various random directed graphs and use Theorem 1.3 to exhibit them as nearest neighbor graphs for some choices of weights. Because i.i.d., continuously distributed weights produce nearest neighbor graphs with all finite components, it suffices to take $k \geq 1$. Our underlying graph will be $\mathbb{Z}^d$ for some $d \geq 2$, and we will identify any directed graph G with vertex set $\mathbb{Z}^d$ with a point in the space $\{0, 1\}^{E^d}$, where E^d is the set $\{\langle x, y \rangle : x, y \in \mathbb{Z}^d, \|x - y\| = 1\}$ of directed edges of the lattice. (We give this space the usual product Borel sigma-algebra.) Any translation T^z acts on this space just as it did on the edge-weight space Ω: for $\eta \in \{0, 1\}^{E^d}$, we set $(T^z \eta)(\langle x, y \rangle) = \eta(\langle x + z, y + z \rangle)$. Last, to a directed graph $G = (\mathbb{Z}^d, E)$ we naturally associate the point $\eta = \eta(G)$:

$$\eta(\langle x, y \rangle) = \begin{cases} 1 & \text{if } \langle x, y \rangle \in E \\ 0 & \text{otherwise.} \end{cases}$$

By Theorem 1.3, to show that there is a measure P satisfying **A** such that a.s. the graph N_D has, say, property P, it suffices to show that there is a random directed graph G (measure on $\{0, 1\}^{E^d}$) which is invariant under all translations T^z such that a.s., G has property P and the properties stated in the theorem: if $\langle x, y \rangle \in E$ then $\{x, y\} \in E^d$, each $x \in \mathbb{Z}^d$ has out-degree one in G, G has no directed cycles of length at least three, and for each $x \in \mathbb{Z}^d$, C_x is finite. This is because if the distribution of G is invariant under translations, then the weights $(\omega(e))$ defined in (2.11) will be as well, so **A1** will hold, and the weights are all distinct, so **A2** will hold as well.

Given these preliminaries, we move to the constructions.

2.4.1. Case $d = 2$ and $k = 2$

To construct the measure P for the case $d = 2$ and $k = 2$, we use the graph from [19, Sec. 3]. It is a.s. a union of two directed trees, each built from coalescing random walks: one tree moves up-right and the other moves down-left, both trees being dual to each other. Let $(B_x)_{x \in \mathbb{Z}^2}$ be a family of i.i.d. Bernoulli$(1/2)$ random variables and define a directed graph G_0 using these variables as follows. The vertex set of G_0 is $\mathbb{Z}^2$ and the edge set is

$$\bigcup_{x: B_x = 1} \{\langle 2x, 2x + e_2 \rangle, \langle 2x + e_2, 2x + 2e_2 \rangle, \langle 2x + e_1, 2x + e_1 - e_2 \rangle, \langle 2x + e_1 + e_2, 2x + e_2 \rangle\}$$

$$\bigcup \bigcup_{x: B_x = 0} \{\langle 2x, 2x + e_1 \rangle, \langle 2x + e_1, 2x + 2e_1 \rangle, \langle 2x + e_2, 2x + e_2 - e_1 \rangle, \langle 2x + e_1 + e_2, 2x + e_1 \rangle\} .$$

Here e_1 and e_2 are the standard basis vectors of $\mathbb{R}^2$. (See [19, Figs. 2,3] for illustrations of the structure of G_0.) Although G_0 is not translation invariant, we can remedy this by letting U be an independent uniform vector on the set $\{0, e_1, e_2, e_1 + e_2\}$ and setting G to be the translation of G_0 by U. That is, G has vertex set $\mathbb{Z}^2$ but edge set

$$\{\langle x + U, y + U \rangle : \langle x, y \rangle \text{ an edge of } G_0\} .$$

In [19, p. 1730] it is shown that the distribution of G is invariant (and even ergodic) under lattice translations.

By construction, a.s. each vertex has out-degree one in G and the graph Γ_x obtained by starting with a vertex x and following each out-edge is a symmetric random walk that, once it reaches a vertex z, steps either (a) up twice or right twice if $z \cdot (1, 1)$ is even or (b) down twice or left twice if $z \cdot (1, 1)$ is odd. Because all paths of type (a) intersect, as do all paths of type (b), but paths of type (a) do not intersect those of type (b), it follows that the undirected version of G (the graph with the same vertex set but edge set equal to $\{\{x, y\} : \langle x, y \rangle$ an edge of $G\}$) has exactly two components a.s. (These are the two directed trees mentioned above.) In [19, p. 1730] it is shown that "for any x the subtree for which x is a.s. finite." In our notation, this means that for each $x \in \mathbb{Z}^2$, the number of vertices in C_x is finite. Because this G a.s. satisfies the conditions of Theorem 1.3, is invariant under translations, and its undirected version has two components, this completes the case $d = 2$ and $k = 2$.

Remark 3. In the language of the proof of Theorem 1.2, the above example exhibits the lattice $\mathbb{Z}^2$ as a union of two disjoint type-(a) sets, V_1 and V_2, corresponding to the up-right tree and the down-left tree. The dual edge boundaries $B(V_1)$ and $B(V_2)$ coincide and, furthermore, each vertex of $\mathbb{Z}^2$ is an endpoint of an edge dual to one on this boundary. One can modify this example to produce a model consisting of two type-(a) sets separated by type-(c) sets as follows. For any z such that all vertices x with $\langle x, z \rangle$ an edge of G satisfy $\# C_x = 1$, we remove the out-edge of z from G and add a new out-edge from z to any such x (choosing one in a deterministic manner). The resulting directed graph is then seen to be a nearest neighbor graph for weights satisfying assumption **A**, and every z listed above becomes part of a type-(c) set separating the two type-(a) sets.
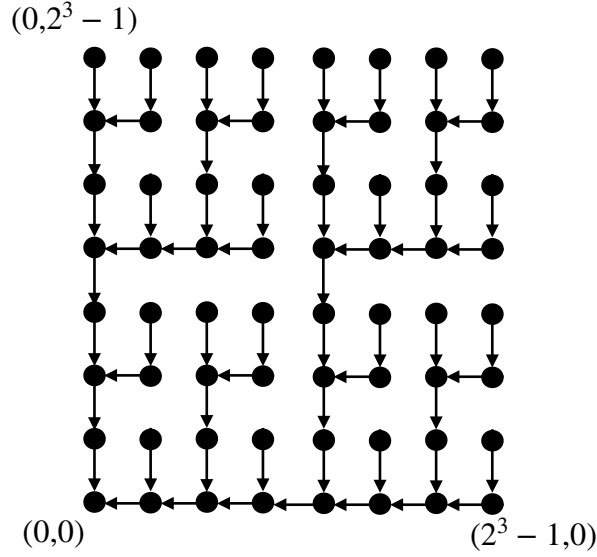
FIG 3. Illustration of the dyadic construction of η_0 for $d = 2$ restricted to the box C_3 of side-length $8 = 2^3$. In each block of side-length 2, arrows point down from the top level, and to the left from the lower-right vertex. In each block of side-length 4, we repeat this construction, viewing the side-length 2 blocks as vertices: the top-level blocks point down and the lower-right block points left. In blocks of side-length 8 (the whole figure), the top-level side-length 4 blocks point down and the lower-right one points left.

2.4.2. Case $d \geq 3$ and $k = \infty$

To prove item 2 in the case $d \geq 3$ and $k = \infty$ we use a layered construction. We produce a translation-invariant random graph G with vertex set $\mathbb{Z}^d$ which satisfies the conditions of Theorem 1.3 and whose undirected version has infinitely many infinite components. We do this by induction, so suppose there is such a random graph G_d for a given dimension d with at least two infinite components (by the above argument, we know this is true for $d = 2$); we will show one exists in dimension $d + 1$ with infinitely many infinite components. Write η_d for the (random) point in $\{0, 1\}^{\mathbb{E}^d}$ corresponding to G_d in d-dimensions and define an element $\eta \in \{0, 1\}^{\mathbb{E}^{d+1}}$ by

$$
\eta\left(\langle x, y\rangle\right) = \begin{cases} \eta_d\left(\langle \pi_d(x), \pi_d(y)\rangle\right) & \text{if } \pi_d(x) \neq_d \pi_d(y) \\ 0 & \text{otherwise.} \end{cases}
$$

Here $\pi_d : \mathbb{Z}^{d+1} \to \mathbb{Z}^d$ is the projection $\pi_d(x) = \sum_{i=1}^d (x \cdot e_i) e_i$. If G is the graph corresponding to η, then the intersection of G with each hyperplane $\{x \cdot e_{d+1} = n\}$ (for $n \in \mathbb{Z}$) is a copy of G_d, and there are no edges in G connecting these hyperplanes. Therefore the distribution of G is invariant under translations. Furthermore, each $x \in \mathbb{Z}^{d+1}$ has out-degree one, it has no directed cycles of length at least three, and each C_x is finite. Last, since the undirected version of G_d has multiple infinite components, so does the undirected version of G (in fact it has infinitely many). This proves the case $d \geq 3$ and $k = \infty$.

2.4.3. Case $d \geq 2$ and $k = 1$

For the case $d \geq 2$ and $k = 1$, we give a dyadic construction. Define the orthant $O = \{x \in \mathbb{Z}^d : x \cdot e_i \geq 0 \text{ for all } i\}$. For any nonzero $x \in O$ let

$$
k(x) = \min\left\{ k \geq 1 : x/2^k \notin \mathbb{Z}^d \right\}.
$$

Because $x/2^{k(x)-1} \in \mathbb{Z}^d$ but $x/2^{k(x)} \notin \mathbb{Z}^d$, at least one coordinate of $x/2^{k(x)-1}$ is odd. Let $i(x)$ be the largest such index. (For example, if $x = (4, 8, 16)$ then $k(x) = 1$ and $i(x) = 3$, and if $x = (0, 8, 16)$ then $k(x) = 4$ and $i(x) = 2$.) Now define $\eta_0 \in \{0, 1\}^{\mathbb{E}^d}$ by

$$
\eta_0\left(\langle x, x - 2^{k(x)} e_{i(x)} \rangle\right) = 1 \text{ for all nonzero } x \in O,
$$

and $\eta_0(\langle x, y\rangle) = 0$ for all other directed edges $\langle x, y\rangle$. Write G_0 for the directed graph corresponding to η_0. See Figure 3.

In G_0, the vertex 0 has out-degree 0, as does every vertex that is not in O. Each non-zero $x \in O$ has out-degree one. Therefore the directed subgraph Γ_x^0 of G_0 induced by the vertices y such that $x \to y$ is a directed path. Starting from any

$x \in O$ and moving along Γ_x^0, the vertices are obtained from x as follows. We decrement the largest odd coordinate of x by 1, then the next largest odd coordinate by 1, and so on, until all coordinates are even. Then we decrement the largest coordinate that is not divisible by 4 repeatedly until it becomes divisible by 4, then the next largest coordinate that is not divisible by 4, and so on, until all coordinates are divisible by 4. We then iterate the steps with 8 in place of 4, then 16, and all powers of 2, until we reach the origin. We observe that in this procedure,

(2.12) $\qquad$ after all coordinates are divisible by 2^k, at most one is not at any time.

For any $k \geq 1$, O is a disjoint union of "2^k-boxes" of the form $2^k z + C_k$, where $C_k = \{0, \ldots, 2^k - 1\}^d$ and $z \in O$. Here,

(2.13) $\qquad$ if $x \in 2^k z + C_k$, then $z = ((x/2^k) \cdot e_1, \ldots, (x/2^k) \cdot e_d)$.

Note that for $k \geq 1$ and $z \in O$,

(2.14) $\qquad$ if $x \in 2^k z + C_k$ then Γ_x^0 touches $2^k z$ within $d 2^k$ steps without leaving $2^k z + C_k$.

This follows from the above description of Γ_x^0: as we traverse Γ_x^0, we start at x and decrement coordinates in the order described above until all coordinates are divisible by 2^k.

Before we translate and average to build a measure, we note the following properties of G_0:

(A) If (hx, yi) is an edge of G_0, then $\{x, y\} \in E^d$. This is clear by the construction.
(B) Each $x \in O$ that is nonzero has out-degree one in G_0. This was stated above.
(C) G_0 has no directed cycles. This is because each directed edge points in the direction of decreasing i-th coordinate for some i.
(D) If $x \in O$ is in a 2^k-box $2^k z + C_k$ for $z \in O$ and $k \geq 1$, and has at least two coordinates which are not multiples of 2^k, then C_x^0 (the subgraph of G_0 induced by y such that $y \to x$ in G_0) contains no vertices outside of $2^k z + C_k$. To prove this, suppose that y is a vertex of C_x^0. If y is in a different 2^k-box from x, say $2^k z^0 + C_k$ for $z^0 \neq z$, then by (2.14), traversing Γ_y^0 leads us to $2^k z^0$ without leaving $2^k z^0 + C_k$ (in particular not touching x). Because of (2.12) and the assumed properties of x, Γ_y^0 never touches x.
(E) For $k \geq 1$, if x, y are elements of the same k-box, then while traversing Γ_x^0 starting from x, we intersect Γ_y^0 within $d 2^k$ steps. Indeed, we note that such x and y can be written for some $z \in O$ as

$$x = 2^k z + x^0, \quad y = 2^k z + y^0 \text{ for } x^0, y^0 \in C_k.$$

So by (2.14), after at most $d 2^k$ steps on Γ_x^0, we reach $2^k z$ (and similarly for Γ_y^0).

The next step is to define a sequence of variables $(Z_n)_{n \geq 1}$ such that Z_n is uniform on $\{0, \ldots, 2^n - 1\}^d$, and set $\eta_n = T^{Z_n} \eta_0$ to be the translation of η_0 by Z_n. Because the η_n form a tight sequence (the space $\{0, 1\}^{E^d}$ is compact), there is a subsequence (n_i) such that η_{n_i} converges in distribution to some η. (In fact, a subsequence is not necessary.) Letting G be the (random) directed graph corresponding to η, it is standard that G is invariant under translations. We are then left to prove that a.s., G satisfies the conditions of Theorem 1.3, and that a.s., the undirected version of G has only one component.

To show the desired properties of G, we start with item 1 of Theorem 1.3, and this is the most obvious. For any $x, y \in \mathbb{Z}^d$, the event $\tau \in \{0, 1\}^{E^d} : \tau(hx, yi) = 1$ is a cylinder event, so its indicator is a (bounded) continuous function. Therefore $P(\eta(hx, yi) = 1) = \lim_{k \to \infty} P(\eta_{n_k}(hx, yi) = 1)$. If x and y are not neighbors (that is, $kx - yk_1 > 1$), this probability is zero by item (A) above. This means G satisfies item 1 of Theorem 1.3 a.s.

For item 2, note that the event that the origin has out-degree one is a cylinder event. Again this implies that

$$P(0 \text{ has out-degree one in } G) = \lim_{i \to \infty} P(Z_{n_i} \text{ has out-degree one in } G_0).$$

By item (B), the right side equals $\lim_{i \to \infty} (1 - 1/2^{d n_i}) = 1$. By translation invariance, we conclude item 2. By a similar argument, we can show item 3: a.s. G has no directed cycles of length at least 3. (In fact, it has no directed cycles.) Letting C be a deterministic (finite) directed cycle, the event that all directed edges in C are present in the graph (the event $\{\tau : \tau(hx, yi) = 1 \text{ for all edges } hx, yi \text{ in } C\}$) is again a cylinder event. By item (C) above, the probability that η is in this event is zero. Taking a union over all finite cycles shows item 3.

We now show item 4: a.s. for each $x \in \mathbb{Z}^d$, the graph C_x is finite. By translation invariance, it suffices to consider $x = 0$. Writing $\#C_0$ for the number of vertices in C_0, note that because $\{\#C_0 > \lambda\}$ is a cylinder event, we have

$$P(\#C_0 > \lambda) = \lim_{i \to \infty} P\left(\#C^0_{Z_{n_i}} > \lambda\right).$$

For $\lambda \geq 2^d$ and $k = b\log_2 \lambda c$, write z_{n_i} for the unique point of O such that $Z_{n_i} \in 2^k z_{n_i} + C_k$. Then by item (D),

$$\lim_{i \to \infty} P\left(\#C^0_{Z_{n_i}} > \lambda\right) \leq \lim_{i \to \infty} P\left(C_{Z_{n_i}} \text{ contains a vertex outside } 2^k z_{n_i} + C_k\right)$$

$$\leq \lim_{i \to \infty} P\left(\text{at most one coordinate of } Z_{n_i} \text{ is not divisible by } 2^k\right)$$

$$= 2^{-dk} + d2^{-k(d-1)}(1 - 2^{-k})$$

$$\leq \frac{d + 1}{2^{k(d-1)}}$$

$$\leq \frac{2^{d-1}(d + 1)}{\lambda^{\frac{d-1}{d}}}.$$

Letting $\lambda \to \infty$, we obtain $\#C_0 < \infty$ a.s. and this proves item 4.

Finally we prove that the undirected version of G has one infinite component a.s. For $x \in \mathbb{Z}^d$ and an integer $\lambda > 0$, let $E_x = E_x(\lambda)$ be the event that there are directed paths π_0 in Γ_0 starting from 0 and π_x in Γ_x starting from x, both with λ many edges, which intersect. Note that this event is defined even when some vertices have out-degree greater than one, but in our graphs Γ_0 and Γ_x are directed paths. (As usual, Γ_x is the subgraph induced by the vertices y such that $x \to y$.) Because E_x is also a cylinder event,

$$(2.15) \qquad P(\eta \in E_x) = \lim_{i \to \infty} P(\eta_{n_i} \in E_x).$$

The event on the right is that $\Gamma^0_{Z_{n_i}}$ and $\Gamma^0_{Z_{n_i} + x}$ (recall these are the paths in the graph G_0) both have at least λ many steps, and, following either from its starting point, we intersect the other within λ many steps. Note that if every coordinate of Z_{n_i} is at least $\lambda + kx k_\infty$, then both paths will have at least λ many steps. Furthermore, if both of these points are in the same 2^k-box $2^k z_{n_i} + C_k$, then by item (E), the paths will intersect within $d2^k$ steps. Therefore if we put $k = b\log_2(\lambda/d)c$ (for $\lambda \geq 2d$), we see that if Z_{n_i} and $Z_{n_i} + x$ are in the same 2^k-box, then $\Gamma^0_{Z_{n_i}}$ and $\Gamma^0_{Z_{n_i} + x}$ will intersect within λ many steps. This means the right side in (2.15) is at least

$$\lim_{i \to \infty} P\left(Z_{n_i} \text{ and } Z_{n_i} + x \text{ are in the same } 2^k\text{-box and have all coordinates} \geq \lambda\right)$$

$$(2.16) \qquad \geq \lim_{i \to \infty} \left(1 - \frac{d(\lambda + kx k_\infty)}{2^{n_i}} - P\left(Z_{n_i} \text{ and } Z_{n_i} + x \text{ are in different } 2^k\text{-boxes}\right)\right).$$

If Z_{n_i} and $Z_{n_i} + x$ are in different 2^k-boxes, then by (2.13),

$$(b(Z_{n_i}/2^k) \cdot e c, \ldots, b_N(Z/2^k) \cdot e c) \neq (b((Z_{n_i} + x)/2^k) \cdot e c, \ldots, b_N((Z + x)/2^k) \cdot e c).$$

The probability of this is at most $2dkx k_\infty / 2^k$. Putting this in (2.16), and then combining with (2.15), we find

$$P(\eta \in E_x) \geq \lim_{i \to \infty} \left(1 - \frac{d(\lambda + kx k_\infty)}{2^{n_i}} - \frac{2dkx k_\infty}{2^k}\right) \geq 1 - \frac{4d^2 kx k_\infty}{\lambda}.$$

Recalling the definition of E_x, if we take $\lambda \to \infty$, we see that a.s. both Γ_0 and Γ_x are infinite and intersect. Since this is true for all x, a.s. the undirected version of G has one infinite component. This completes the proof of the case $d \geq 2$ and $k = 1$.

2.4.4. *Case $d \geq 3$ and $k \in [2, \infty)$*

We have seen in Section 2.4.2 that when $d \geq 3$, it is possible to construct nearest-neighbor graphs whose undirected versions have infinitely many infinite components, contrary to the situation when $d = 2$. This could lead one to ask about the possibility of some arbitrary finite number $k \geq 2$ of components. Fix some such integer $k \in [2, \infty)$ for the remainder of this section. We will explicitly construct a translation-invariant measure P such that, a.s., N has exactly k infinite

components. We do this by describing how to generate a translation-invariant random directed graph G which satisfies the hypotheses of Theorem 1.3 and whose unoriented version has exactly k infinite components. We write $G = (\mathbb{Z}^d, \vec{E})$; we construct the (directed) edge set $\vec{E}$ in stages, writing $\vec{E} = \vec{E}^{(1)} \cup \cdots \cup \vec{E}^{(k+1)}$.

Recall the nearest-neighbor measure constructed in Section 2.4.3. This was a translation-invariant nearest-neighbor model with the property that, a.s. N has exactly one infinite component. Let $G^{(1)}, \ldots, G^{(k)}$ denote k independent samples from this measure. We begin by using $G^{(1)}$ to define $\vec{E}^{(1)}$. Given an edge $\{x, x + e_i\}$ of E^d, we include in $\vec{E}^{(1)}$ an appropriate orientation of each of the edges $\{4kx, 4kx + e_i\}$, $\{4kx + e_i, 4kx + 2e_i\}$, $\ldots$, $\{4kx + (4k - 1)e_i, 4k(x + e_i)\}$; exactly one orientation of each of these edges will be chosen to appear in $\vec{E}^{(1)}$. Which orientations of each of these edges appears in $\vec{E}^{(1)}$ depends on whether a) $\langle x, x + e_i\rangle$ is an edge of $G^{(1)}$, b) $\langle x + e_i, x\rangle$ is an edge of $G^{(1)}$, or c) neither $\langle x, x + e_i\rangle$ nor $\langle x + e_i, x\rangle$ is an edge of $G^{(1)}$ (note that, by the construction of $G^{(1)}$, these are a.s. the only possibilities).

- In case a), the edges $\langle 4kx + \ell e_i, 4kx + (\ell + 1)e_i\rangle \in \vec{E}^{(1)}$ for each $0 \le \ell \le 4k - 1$ (and the other orientation is omitted: $\langle 4kx + (\ell + 1)e_i, 4kx + \ell e_i\rangle \notin \vec{E}^{(1)}$).
- Case b) is identical but reflected: each edge $\langle 4kx + (\ell + 1)e_i, 4kx + \ell e_i\rangle \in \vec{E}^{(1)}$ for each $0 \le \ell \le 4k - 1$, and the other orientation is again omitted.
- Lastly, in case c), we orient edges toward the closer of $4kx$ and $4k(x + e_i)$: the edges $\langle 4kx + e_i, 4kx\rangle$, $\langle 4kx + 2e_i, 4kx + e_i\rangle$, $\ldots$, $\langle 4kx + 2ke_i, 4kx + (2k - 1)e_i\rangle \in \vec{E}^{(1)}$, and also $\langle 4kx + (2k + 1)e_i, 4kx + (2k + 2)e_i\rangle$, $\ldots$, $\langle 4kx + (4k - 1)e_i, 4k(x + e_i)\rangle \in \vec{E}^{(1)}$, with none of the reversed orientations of these edges appearing in $\vec{E}^{(1)}$.

We can think of this construction as in a sense "stretching out the lattice $\mathbb{Z}^d$" by a factor $4k$: each edge of the lattice is turned into a segment of $4k$ edges. The above definition guarantees that these segments are traversed by an oriented path in $\vec{E}^{(1)}$ exactly when the corresponding "un-stretched" edges of E^d appear in $G^{(1)}$.

Let $V^{(1)}$ denote the set of endpoints of the edges considered above — in other words,

$$V^{(1)} = \{4kx + \ell e_i : x \in \mathbb{Z}^d, 0 \le \ell \le 4k - 1, 1 \le i \le d\}.$$

We can consider $\vec{E}^{(1)}$ as inducing a random directed graph with vertex set $V^{(1)}$. We note several properties of this graph which follow directly from the definition and from properties of $G^{(1)}$. First,

(2.17) $\qquad\qquad\qquad\qquad$ a.s., each $y \in V^{(1)}$ has out-degree one in $\vec{E}^{(1)}$;

Next, for each $x \in \mathbb{Z}^d$,

(2.18) $\qquad\quad \#\{y \in V^{(1)} : y \to 4kx \text{ by a path in } \vec{E}^{(1)}\} \le 8kd\#\{z \in \mathbb{Z}^d : z \to x \text{ by a path in } G^{(1)}\}.$

From (2.18), we immediately see that

(2.19) $\qquad\qquad$ a.s., for each $y \in V^{(1)}$, $\quad \#\{z \in V^{(1)} : z \to y \text{ by a path in } \vec{E}^{(1)}\} < \infty.$

No edge of $\vec{E}^{(j)}$, $j \ge 2$ will be incident to any vertex of $V^{(1)}$, so the above properties will be preserved throughout the remainder of the construction. Statements (2.17) and (2.19) (and their analogues for the endpoints of edges in $\vec{E}^{(j)}$, $j \ge 2$) will guarantee that the graph G satisfies the hypotheses of Theorem 1.3 and hence can be represented as a nearest neighbor graph. Let $E^{(1)}$ denote the set of undirected versions of edges in $\vec{E}^{(1)}$; a final important property of the above is that

$$(V^{(1)}, E^{(1)}) \text{ a.s. has exactly one infinite component.}$$

This is again easy to see from the definition. Indeed, the construction above preserves the component structure of vertices of the form $4kx$: there is a path of edges of $E^{(1)}$ from $4kx_1$ to $4kx_2$ if and only if there is a path of (undirected versions of) edges of $G^{(1)}$ from x_1 to x_2, and other vertices of $V^{(1)}$ can only connect up via vertices of the form $4kx$.

The construction of $\vec{E}^{(2)}, \ldots, \vec{E}^{(k)}$ proceeds analogously, but on shifted sublattices. The vertex set $V^{(2)} = V^{(1)} + (4, 4, \ldots, 4)$. Edges are included in $\vec{E}^{(2)}$ via an analogous version of the above procedure, but using the realization $G^{(2)}$ and with the entire construction shifted by $(4, \ldots, 4)$: we use the status of edges of the form $\langle x, x + e_i\rangle$ to determine the status of directed versions of edges of the form $\{4kx + (4, 4, \ldots, 4), 4kx + (4, 4, \ldots, 4) + (\ell + 1)e_i\}$. To construct $\vec{E}^{(3)}$, we proceed analogously but with $V^{(3)} = V^{(1)} + (8, 8, \ldots, 8)$, and so on. A vertex $x \in V^{(j)}$ has the property that

(2.20) $\qquad\qquad\qquad\qquad\qquad x \cdot e_i \equiv 4(j - 1) \mod 4k$

for at least $d-1$ values of i; hence, $V^{(j)} \cap V^{(k)} = \varnothing$ when $j \neq k$. In other words, we have constructed k "noninteracting" and independent directed graphs on distinct sublattices of $\mathbb{Z}^d$, each of which obeys the properties (2.17) and (2.18), and whose undirected version a.s. has exactly one infinite component.

To complete the construction, we must choose $\vec{E}^{(k+1)}$ in a way which guarantees the hypotheses of Theorem 1.3 are satisfied. Once this is done, we will have constructed a random graph which is realizable as a nearest-neighbor graph, though (as in the construction in Section 2.4.3) this nearest-neighbor model will not be translation-invariant. To finish the construction and recover translation invariance, we conclude by shifting the entire graph by an independent random integer vector in the cube $[0, 4k-1)$.

It remains to choose $\vec{E}^{(k+1)}$. Suppose we chose $\vec{E}^{(k+1)} = \varnothing$, or in other words had the (oriented) edge set of $\vec{G}$ be $\vec{E}^{(1)} \cup \cdots \cup \vec{E}^{(k)}$. Then the hypotheses of Theorem 1.3 would not be satisfied: the set $V^{(k+1)} := \mathbb{Z}^d \setminus (V^{(1)} \cup \cdots \cup V^{(k)})$ is nonempty, and each vertex of $V^{(k+1)}$ would have both out- and in-degree zero under this oriented edge set. We complete the construction by choosing $\vec{E}^{(k+1)}$ in a way such that

- Each vertex of $V^{(k+1)}$ has out-degree one in $(V^{(k+1)}, \vec{E}^{(k+1)})$;
- No edge of $\vec{E}^{(k+1)}$ connects a vertex of $V^{(k+1)}$ to a vertex of $\mathbb{Z}^d \setminus V^{(k+1)}$ (or vice-versa);
- The undirected version $E^{(k+1)}$ of $\vec{E}^{(k+1)}$ has components of ℓ^∞ diameter no larger than $4k$.

These properties guarantee that the directed graph $(V^{(k+1)}, \vec{E}^{(k+1)})$ again "does not interact with" the graphs $(V^{(i)}, \vec{E}^{(i)})$ for $i \leq k$, and that the graph $\vec{G}$ satisfies the hypotheses of Theorem 1.3. Moreover, since there are no infinite components in $(V^{(k+1)}, \vec{E}^{(k+1)})$, the graph $\vec{G}$ has exactly k infinite components.

We choose the edges of $\vec{E}^{(k+1)}$ to have both endpoints in a common cube of the form $4kx + [-2k, 2k)$; this will guarantee the diameter condition above holds. To do this, for each $x \in \mathbb{Z}^d$, consider the site-components of $R_x := 4kx + [-2k, 2k) \setminus (V^{(1)} \cup \ldots \cup V^{(k)})$. For each site-component C, choose a deterministic spanning tree of the vertices of C oriented toward some deterministic root, then insert a single additional oriented edge from this root toward one of its neighbors (note that this choice can be done in a non-random way, identically for each x; all of the R_x's are translates of one another).

This construction guarantees that the above bulleted properties hold: the latter two are obvious, and the first holds by construction as long as each site-component C has at least two vertices. To see why each site-component has at least two vertices, consider a vertex $y \in R_0$, and let C be the corresponding site component of y. By symmetry, we may assume $y \cdot e_1 \geq 0$. If $y + e_1 \in C$, then we are done. Otherwise, if $y + e_1 \notin C$, then either i) $y + e_1 \notin [-2k, 2k)$ or ii) $y + e_1 \in V^{(j)}$ for some $1 \leq j \leq k$. In case i), we have $y \cdot e_1 = 2k$ and then $(y - e_1) \cdot e_1 = 2k - 1$, whence $y - e_1$ cannot be in $V^{(m)}$ for any $1 \leq m \leq k$. In case ii), $y - e_1 \in [-2k, 2k)$ since $y \cdot e_1 \geq 0$. Moreover, $y - e_1 \notin V^{(j)}$, since otherwise we would also have $y \in V^{(j)}$. Lastly, $y - e_1$ cannot be in $V^{(m)}$ for any $m \neq j$, because then $V^{(j)}$ and $V^{(m)}$ would be at Euclidean distance two from each other. We conclude that in either case, $y - e_1$ is also an element of y's site component C. Thus, the site components of R_x are not singletons, and so the first bulleted property above holds.

Acknowledgments

Funding

References

[1] Ahlberg, D.; Hoffman, C. Random coalescing geodesics in first-passage percolation. preprint, 2016.

[2] Ballister, P.; Bollobás, B. Percolation in the k-nearest neighbor graph. In *Recent Results in Designs and Graphs: a Tribute to Lucia Gionfriddo*, Quaderni di Matematica, Volume 28, edited by Marco Buratti, Curt Lindner, Francesco Mazzocca, and Nicola Melone. (2013), 83–100.

[3] Boivin, D.; Derrien, J.-M. Geodesics and recurrence of random walks in disordered systems. *Ann. Probab.* **7** (1991), 101–115.

[4] Brito, G.; Damron, M.; Hanson, J. Absence of backward infinite paths for first-passage percolation in arbitrary dimension. preprint, 2020.

[5] Burton, R.; Keane, M. Topological and metric properties of infinite clusters in stationary two-dimensional site-percolation. *Isr. J. Math.* **76** (1991), 299–316.

[6] Chaika, J.; Krishnan, A. Stationary coalescing walks on the lattice. *Probab. Theory Relat. Fields.* **175** (2018), 655–675.

[7] Chaika, J.; Krishnan, A. Stationary coalescing walks on the lattice II: Entropy. preprint, 2019.

[8] Coupier, D.; Dereudre, D.; Le Stum, Simon. Existence and percolation results for stopped germ-grain models with unbounded velocities. *Stoch. Proc. Appl.* bf 142 (2021), 549–579.

[9] Damron, M.; Hanson, J. Busemann functions and infinite geodesics in two-dimensional first-passage percolation. *Commun. Math. Phys.* **325** (2014), 917–963.

[10] Guiol, H. A note about Burton Keane's theorem. preprint, 1997.

[11] Häggström, O. Invariant percolation on trees and the mass-transport method.   In *Bulletin of the International Statistical Institute.*   52nd session proceedings, Tome LVIII, Book 1, Helsinki, 1999, pp. 363–366.

[12] Häggström, O.; Meester, R. Nearest neighbor and hard sphere models in continuum percolation. *Random Structures Algorithms.* **9** (1996), 295–315.

[13] Harris, M.; Meester, R. Nontrivial phase transition in a dependent parametric bond percolation model. *Markov Process. Related Fields.* **2** (1996), 513–528.

[14] Hirsch, C. On the absence of percolation in a line-segment based lilypond model. *Ann. Inst. H. Poincaré Probab. Statist.* **52** (2016), 127–145.

[15] Kozakova, I.; Meester, R.; Nanda, S. The size of components in continuum nearest-neighbor graphs. *Ann. Probab.* **34** (2006), 528–538.

[16] Lyons, R.; Peres, Y. *Probability on Trees and Networks.* Vol. 42. Cambridge University Press, 2017.

[17] Nanda, S.; Newman, C. M. Random nearest neighbor and influence graphs on Z^d. *Random Structures Algorithms.* **15** (1999), 262–278.

[18] Pittel, B.; Weishaar, R. The random bipartite nearest neighbor graphs. *Random Structures Algorithms.* **15** (1999), 279–310.

[19] Zerner, M.; Merkl, F. A zero-one law for planar random walks in random environment. *Ann. Probab.* **29** (2001), 1716–1732.