Electronic Journal of Probability

Electron. J. Probab. **17** (2012), no. 57, 1–21.
ISSN: 1083-6489 DOI: 10.1214/EJP.v17-1977

Long-range percolation on the hierarchical lattice

Vyacheslav Koval* Ronald Meester† Pieter Trapman‡

Abstract

We study long-range percolation on the hierarchical lattice of order N, where any edge of length k is present with probability $p_k = 1 - \exp(-\beta^{-k}\alpha)$, independently of all other edges. For fixed β, we show that $\alpha_c(\beta)$ (the infimum of those α for which an infinite cluster exists a.s.) is non-trivial if and only if $N < \beta < N^2$. Furthermore, we show uniqueness of the infinite component and continuity of the percolation probability and of $\alpha_c(\beta)$ as a function of β. This means that the phase diagram of this model is well understood

Keywords: long-range percolation; renormalisation; ergodicity.
AMS MSC 2010: Primary 60K35; 37F25, Secondary 47A35.
Submitted to EJP on April 25, 2012, final version accepted on June 29, 2012.
Supersedes arXiv:1004.1251.

1 Introduction and main results

The study of long-range percolation on $\mathbb{Z}^d$ goes back to [23] and led to a series of interesting problems and results [2, 22, 4, 5, 24]; see [6, Section 2] for an extensive overview. In [11] asymptotic long-range percolation is studied on the hierarchical lattice Ω_N (to be defined below) for $N \to \infty$. The contact process on Ω_N for fixed N has been studied in [3].

In this paper we study the case of finite N. Long range percolation on the hierarchical lattice is quite different from the usual lattice: classical methods break down and results are different.

We note that independently of the present paper, Dawson and Gorostiza [10] also studied long range percolation on Ω_N for fixed N and obtained partly overlapping results, using different methods.

The research of this paper is inspired by questions from epidemiology. For references to the use of (long-range) percolation theory in epidemiology see [9, 24]. In the most basic epidemiological models, all individuals are interacting in the same way, and every pair of individuals make contacts, which may lead to transmission of infectious

*Utrecht University, The Netherlands. E-mail: slakov@gmail.com
†VU University Amsterdam, The Netherlands. E-mail: r.w.j.meester@vu.nl
‡Stockholm University, Sweden. E-mail: ptrapman@math.su.se

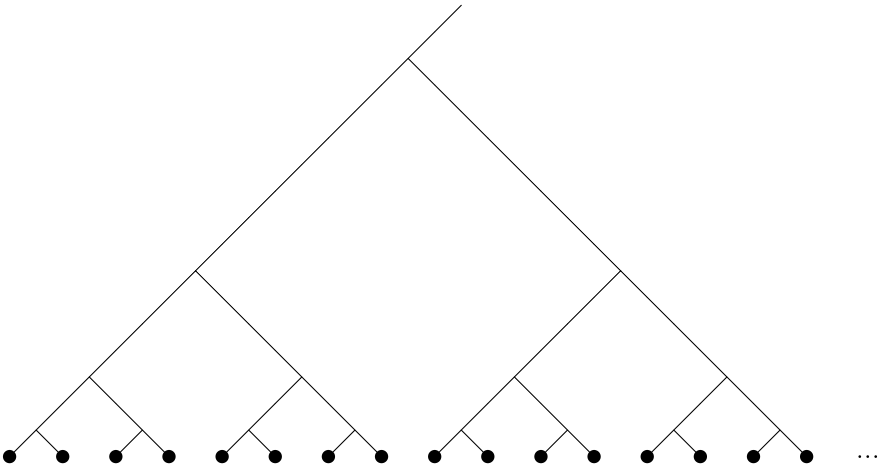


Figure 1: Hierarchical lattice (the "leaves") of order 2 with the metric generating tree attached.

diseases, at the same intensity [13]. To analyse those models, branching process approximations are often used [20].

Clearly the assumptions for this model are unrealistic. We want to investigate how a hierarchical structure in the population, governing the contacts between individuals, influences the spread of an epidemic. In particular we focus on the phase diagram of the epidemic through considering the corresponding phase diagram in a percolation model, and note that the probability that a given vertex is in an infinite percolation cluster, corresponds with the probability of a large outbreak in the population under consideration. Obviously this model is also unrealistic as the most basic epidemic model. However, studying this "other extreme" will provide insight in the behaviour of epidemics in more realistic populations. Furthermore, the model is very interesting in its own right from a probabilistic perspective.

The hierarchical lattice is defined as follows. For an integer $N \geq 2$, we define the set

$$\Omega_N := \{\mathbf{x} = (x_1, x_2, \dots) : x_i \in \{0, 1, \dots, N-1\}, \sum_i x_i < \infty\},$$

and define a metric on it by

$$d(\mathbf{x}, \mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x} = \mathbf{y}, \\ \max\{i : x_i \neq y_i\} & \text{if } \mathbf{x} \neq \mathbf{y}. \end{cases}$$

The pair (Ω_N, d) is called the *hierarchical lattice of order* N.

One can think of the vertices in the hierarchical lattice as the leaves of a regular tree without a root, see Figure 1. The metric d can then be interpreted as the number of generations (levels) till the "most recent common ancestor" of two vertices. Let $\mathbb{N}$ be the non-negative integers, including 0 and $\mathbb{N}_+ := \mathbb{N} \setminus \{0\}$. The set Ω_N is countable, and we can introduce a natural labeling of its vertices via the map $f : \Omega_N \to \mathbb{N}$ given by

$$f : (x_1, x_2, \dots) = \sum_{i=1}^{\infty} x_i N^{i-1}. \tag{1.1}$$

We will sometimes abuse notation and write n for $f^{-1}(n) \in \Omega_N$.

The metric space (Ω_N, d) satisfies the strengthened version of the triangle inequality

$$d(x, y) \le \max(d(x, z), d(z, y))$$

for any triple $x, y, z \in \Omega_N$. Such spaces are called ultrametric (or non-Archimedean) [25]. For $x \in \Omega_N$, define $\mathcal{B}_r(x)$ to be the ball of radius r around x. Several important geometrical properties follow from the definition of the space (Ω_N, d) and its ultrametricity:

1. $\mathcal{B}_r(x)$ contains N^r vertices for any x;
2. for every $x \in \Omega_N$ there are $(N-1)N^{k-1}$ vertices at distance k from it;
3. if $y \in \mathcal{B}_r(x)$ then $\mathcal{B}_r(x) = \mathcal{B}_r(y)$;
4. as a consequence of the previous property, for all $x, y \in \Omega_N$ and for all $r \ge 0$ we either have $\mathcal{B}_r(x) = \mathcal{B}_r(y)$ or $\mathcal{B}_r(x) \cap \mathcal{B}_r(y) = \emptyset$.

Now consider long-range percolation on Ω_N. Every pair of vertices $(x, y) \in \Omega_N \times \Omega_N$ is (independently of all other edges) connected by a single edge with probability

$$p_k = 1 - \exp\left(-\frac{\alpha}{\beta^k}\right),$$

where $k = d(x, y)$ and where $0 \le \alpha < \infty$ and $0 < \beta < \infty$ are the parameters of the model. The edges are not directed. The vertices $x \in \Omega_N$ and $y \in \Omega_N$ are in the same connected component if there exists a *path* from x to y, that is, if there exists a finite sequence $x = x_0, x_1, \ldots, x_n = y$ of vertices such that every pair (x_{i-1}, x_i) of points with $1 \le i \le n$ shares an edge.

We denote the size of a set S of vertices by $|S|$. The connected component (also called "cluster") containing the vertex x is denoted by $\mathcal{C}(x)$. Since there is, for every vertex $x \in \Omega_N$, an automorphism on Ω_N which projects x to 0, the $|\mathcal{C}(x)|$ have the same distribution for every $x \in \Omega_N$ we may study $|\mathcal{C}(0)|$ instead of $|\mathcal{C}(x)|$.

Let $\mathbb{P}_{\alpha,\beta}$ be the probability measure governing this percolation process (on the appropriate probability space and sigma-algebra) and $\mathbb{E}_{\alpha,\beta}$ the corresponding expectation operator. When no confusion is possible, we omit the subscripts α and β. Denote

$$\theta(\alpha, \beta) := \mathbb{P}_{\alpha,\beta}\left(|\mathcal{C}(0)| = \infty\right). \tag{1.2}$$

It follows from a standard coupling argument that $\theta(\alpha, \beta)$ is non-decreasing in α for any given β. Therefore, it is reasonable to define

$$\alpha_c(\beta) := \inf\{\alpha \ge 0 : \theta(\alpha, \beta) > 0\}.$$

Note that also $\theta(\alpha, \beta)$ is non-increasing in β, for any given α.

Throughout the paper we use the following notation. For a set S of vertices, let $\overline{S} := \Omega_N \setminus S$ denote its complement. The set $\mathcal{C}_n(x)$ is the cluster of vertices that are connected to the origin by a path that uses only vertices inside $\mathcal{B}_n(x)$. For disjoint sets $S_1, S_2 \subset \Omega_N$, the event that at least one edge connects a vertex in S_1 with a vertex in S_2 is denoted by $S_1 \leftrightarrow S_2$. The notation $S_1 \not\leftrightarrow S_2$ denotes the event that such an edge does not exist. Let $\mathcal{C}_n^m(x)$ be the largest cluster in $\mathcal{B}_n(x)$; if more such clusters exist, $\mathcal{C}_n^m(x)$ is defined to be one of them, chosen uniformly among all possible candidates. In any case,

$$|\mathcal{C}_n^m(x)| = \max_{y \in \mathcal{B}_n(x)} |\mathcal{C}_n(y)|.$$

Theorem 1.1. *((non)-triviality of the phase transition)*

(a) $\alpha_c(\beta) = 0$ *for* $\beta \leq N$*,*

(b) $0 < \alpha_c(\beta) < \infty$ *for* $N < \beta < N^2$*,*

(c) $\alpha_c(\beta) = \infty$ *for* $\beta \geq N^2$*.*

Theorem 1.2. *(uniqueness of the infinite cluster) There is a.s. at most one infinite cluster, for any value of* α *and* β*.*

Theorem 1.3. *(continuity of* θ*) The percolation function* $\theta(\alpha, \beta)$ *is continuous whenever* $\alpha > 0$.

Theorem 1.4. *(continuity of* $\alpha_c(\beta)$*) The critical value* $\alpha_c(\beta)$ *is continuous for* $\beta \in (0, N^2)$ *and strictly increasing for* $\beta \in [N, N^2)$*. Finally,* $\alpha_c(\beta) \nearrow \infty$ *for* $\beta \nearrow N^2$*.*

In order to prove Theorems 1.3 and 1.4, we need the following result, which is interesting in its own right.

Theorem 1.5. *(size of the large components) If* α *and* β *are such that* $\theta := \theta(\alpha, \beta) > 0$*, then for every* $\varepsilon > 0$*,*

$$\lim_{k\to\infty} \mathbb{P}_{\alpha,\beta}\left(|\mathcal{C}_k^m(0)| > (\theta - \varepsilon)N^k\right) = 1. \tag{1.3}$$

In the next two sections we prove Theorem 1.1 and Theorem 1.2 respectively. After that we prove the remaining results, and we end with a discussion about possible generalisations.

While we were working on this paper, we learned that Dawson and Gorostiza [10] also studied long-range percolation on Ω_N and focused on whether or not an infinite cluster exists with positive probability for given edge probabilities. They provide an independent proof of our Theorem 1. Furthermore they study some deviations of the model around $\beta = N^2$, where α is replaced by a function of k. In particular, they analyze the model for which p_k is decreasing in k and

$$p_{k_n} = \min\left(1, \frac{C + a \log n N^{b \log n}}{N^{2k_n}}\right),$$

where $k_n = \lfloor Kn \log n \rfloor$ and $C \geq 0$, $a > 0$, $b \geq 0$ and $K \geq 1$. (Here and throughout the paper $\lceil x \rceil := \inf\{n \in \mathbb{Z}; n \geq x\}$ is the ceiling of x and $\lfloor x \rfloor := \sup\{n \in \mathbb{Z}; n \leq x\}$ is the floor of x.) They prove that it is possible to chose C, a and b such that percolation occurs.

2 Proof of Theorem 1.1

Proof of (a). Denote by E_k the event that the origin shares an edge with at least one vertex at distance k. Then

$$\mathbb{P}(E_k) = 1 - \exp\left(-\frac{\alpha}{\beta^k}(N-1)N^{k-1}\right)$$

and the events $(E_k)_{k\geq 1}$ are independent. It is easy to see that if $\beta \leq N$ then $\sum_{k=1}^{\infty} \mathbb{P}(E_k)$ diverges for any $\alpha > 0$. Therefore, by the second Borel-Cantelli lemma, infinitely many of the events E_k occur with probability 1, so the origin has infinite degree with probability 1 and $\theta(\alpha, \beta) = 1$, for any $\alpha > 0$ and $0 < \beta \leq N$. This implies that $\alpha_c(\beta) = 0$ for $0 < \beta \leq N$. □

Proof of (c). By monotonicity it suffices to prove that $\alpha_c(N^2) = \infty$, so we now take $\beta = N^2$. Let $n_0 = 0$ and $n_{i+1} = \inf\{n \geq n_i : \mathcal{B}_{n_i}(0) \not\leftrightarrow \overline{\mathcal{B}_n(0)}\}$. Since

$$\{\mathcal{C}(0) = \infty\} \subset \{\mathcal{B}_0(0) \leftrightarrow \overline{\mathcal{B}_0(0)}\} \cap \bigcap_{i=1}^{\infty} \{\left(\mathcal{B}_{n_i}(0) \setminus \mathcal{B}_{n_{i-1}}(0)\right) \leftrightarrow \overline{\mathcal{B}_{n_i}(0)}\},$$

it is enough to prove that there a.s. exists i such that $\left(\mathcal{B}_{n_i}(0) \setminus \mathcal{B}_{n_{i-1}}(0)\right) \not\leftrightarrow \overline{\mathcal{B}_{n_i}(0)}$, that is, $n_{i+1} = n_i$.

Writing $j = n_i$, we compute

$$\begin{aligned} \mathbb{P}(n_{i+1} \neq n_i | n_i \neq n_{i-1}) &= \mathbb{P}(\left(\mathcal{B}_{n_i}(0) \setminus \mathcal{B}_{n_{i-1}}(0)\right) \leftrightarrow \overline{\mathcal{B}_{n_i}(0)}) \\ &\leq \mathbb{P}(\mathcal{B}_j(0) \leftrightarrow \overline{\mathcal{B}_j(0)}) \\ &= 1 - \exp\left(-\alpha N^j \frac{(N-1)}{N^2} \sum_{k=1}^{\infty} \frac{N^{j+k-1}}{N^{2(j+k-1)}}\right) = 1 - \exp\left(-\frac{\alpha}{N}\right). \end{aligned}$$

This upper bound is strictly less than 1, and independent of n_i and n_{i-1}. The result now follows by the second Borel-Cantelli lemma. □

Proof of (b). The strict positivity of $\alpha_c(\beta)$ follows from the fact that

$$\begin{aligned} \sum_{k=1}^{\infty}(N-1)N^{k-1}p_k &= \sum_{k=1}^{\infty}(N-1)N^{k-1}\left(1 - \exp(-\frac{\alpha}{\beta^k})\right) \leq \\ &\leq \frac{\alpha(N-1)}{N} \sum_{k=1}^{\infty} \left(\frac{N}{\beta}\right)^k = \alpha(N-1)\frac{1}{\beta - N}, \end{aligned}$$

which can be made strictly smaller than 1 by choosing α small enough. Hence the expected number of edges from a given vertex is strictly smaller than 1, and by coupling with a subcritical branching process (cf. [24]), the almost sure finiteness of the percolation cluster follows.

The second inequality is more involved, and we start by explaining the idea of the proof.

Fix an integer K and a large enough number η, such that

$$\sqrt{\beta} < \eta \leq \left(N^K - 1\right)^{1/K}; \tag{2.1}$$

this is possible since $\sqrt{\beta} < N$. (The reason for precisely this conditions will become clear in the proof.) We define

$$\varepsilon_n := \exp\left(-\frac{\alpha}{\beta^K}\left(\frac{\eta^2}{\beta}\right)^{nK}\right) \tag{2.2}$$

and

$$s_n = \mathbb{P}\left(|\mathcal{C}^m_{nK}(0)| \geq \eta^{nK}\right). \tag{2.3}$$

Note that $s_0 = 1$. We use renormalisation techniques to deduce that

$$s_{n+1} \geq \mathbb{P}\Big(Bin\left(N^K, s_n(1-\varepsilon_n)\right) \geq N^K - 1\Big), \tag{2.4}$$

where $Bin(n, p)$ denotes a random variable with a binomial distribution with parameters n and p.

Observe that equation (2.4) is close to the usual iteration formula in fractal percolation [12]. In fractal percolation, one studies, for some given m to be fixed, the map

$$\pi(p) = \mathbb{P}(Bin(m, p) \geq m - 1).$$

Define $u_0 = 1$ and, for $n \in \mathbb{N}$

$$u_{n+1} = \pi(pu_n).$$

Writing $G_p(\cdot)$ for $\pi(p\cdot)$ we then obtain

$$u_{n+1} = G_p(u_n). \tag{2.5}$$

In [12] it is shown that the limit $u = \lim_{n\to\infty} u_n$ always exists and is positive if and only if p is so large that the equation $G_p(x) = x$ has a positive solution.

Now observe that (2.4) can be rewritten as

$$s_{n+1} \geq G_{1-\varepsilon_n}(s_n). \tag{2.6}$$

This is very similar to (2.5), the only difference being that the subscript of the iteration function depends on n now. However, we show that for α large enough, ε_n goes down exponentially fast at any given rate. From this we derive below that for every $\gamma \in (0, 1/2)$ and $\alpha = \alpha(\gamma)$ large enough, the probability that the size of the largest cluster restricted to a ball of radius nK is at least η^{nK} with probability bounded below by $1 - \gamma^{n+1}$.

Using α and γ as above, we show that conditioned on the event that the cluster of the origin restricted to the ball of radius nK has at least size η^{nK}, the probability that the cluster of the origin restricted to the ball of radius $(n+1)K$ has at least size $\eta^{(n+1)K}$ is bounded below by $1 - \gamma^{n+2}$. This implies that for all $n \in \mathbb{N}$

$$t_n := \mathbb{P}\big(|\mathcal{C}_{nK}(0)| \geq \eta^{nK}\big) > 1 - \sum_{k=0}^{\infty} \gamma^{k+1} > 0$$

and therefore $\lim_{n\to\infty} t_n > 0$, which proves the theorem.

Next we turn to the formal proof. Choose an integer K and a real number η such that (2.1) is satisfied. We say that a ball of radius nK, $\mathcal{B}_{nK}(x)$ is *good* if its largest connected component has size at least η^{nK}, i.e., if $|\mathcal{C}^m_{nK}(x)| \geq \eta^{nK}$. That is s_n as defined in (2.3) is the probability that a ball of radius nK is good. By convention, we set $s_0 = 1$.

Consider the ball of radius $(n+1)K$, $\mathcal{B}_{(n+1)K}(y)$ and the set of largest clusters restricted to balls of radius nK within $\mathcal{B}_{(n+1)K}(y)$, i.e., we consider the N^K clusters $\{\mathcal{C}^m_{nK}(x)\}_{x\in\mathcal{B}_{(n+1)K}(y)}$ and denote these clusters by $\mathcal{C}(n, y; 1)$, $\mathcal{C}(n, y; 2), \cdots, \mathcal{C}(n, y; N^K)$. The order of the clusters is not important.

We say that a ball of radius nK, $\mathcal{B}_{nK}(x) \subset \mathcal{B}_{(n+1)K}(y)$ is *very good* if

1. it is good,
2. $\mathcal{C}^m_{nK}(x) = \mathcal{C}(n, y; J)$ or there is an edge connecting a vertex in $\mathcal{C}^m_{nK}(x)$ with a vertex in $\mathcal{C}(n, y; J)$, where $J = \min\{i; |\mathcal{C}(n, y; i)| \geq \eta^{nK}\}$.

Note that according to this definition, the first good sub-ball of diameter nK in a ball of radius $(n+1)K$ is automatically very good.

Since $(N^K - 1) \geq \eta^K$, $\mathcal{B}_{(n+1)K}(y)$ will certainly be good if

(a) it contains $N^K - 1$ good sub-balls of radius nK, and
(b) all these good sub-balls are very good.

We next estimate the probability of the events in (a) and (b).

Clearly, the number of good balls of radius nK in $\mathcal{B}_{(n+1)K}(y)$ has a binomial distribution with parameters N^K and s_n. Furthermore, given the collection of good balls of radius nK, the probability that the first such good ball is very good is equal to 1 by definition, and the probability for any of the other good balls of radius nK to be very good is at least $1 - \varepsilon_n$, where ε_n is defined as in (2.2), since the distance between two vertices in a ball of radius $(n+1)K$ is at most $(n+1)K$ and the largest component of a good ball of radius nK contains at least η^{nK} vertices. We conclude that the number of very good balls of radius nK is stochastically larger than a random variable having a binomial distribution with parameters N^K and $s_n(1-\varepsilon_n)$. Inequality (2.4): $s_{n+1} \geq \mathbb{P}\Big(Bin\left(N^K, s_n(1-\varepsilon_n)\right) \geq N^K - 1\Big)$, now follows.

We also have

$$\mathbb{P}(Bin(k,p) \geq k-1) \geq 1 - \binom{k}{2}(1-p)^2,$$

and writing $C = \binom{N^K}{2}$ we arrive at the inequality

$$s_{n+1} \geq 1 - C(1 - s_n + s_n\varepsilon_n)^2 \geq 1 - C(1 - s_n + \varepsilon_n)^2. \tag{2.7}$$

Writing $\xi_n = 1 - s_n$ this gives

$$\xi_{n+1} \leq C(\xi_n + \varepsilon_n)^2. \tag{2.8}$$

Choose first γ so small that $4C \leq \gamma^{-1}$ and then α so large that both $\varepsilon_n \leq \gamma^n$ and $\xi_1 = \mathbb{P}\left(|\mathcal{C}_K^m(0)| < \eta^K\right) \leq \gamma^2$. This is possible because $1/x > e^{-x}$ for all $x > 0$, and therefore we have

$$\varepsilon_n = \left(\exp\left(-\left(\frac{\eta^2}{\beta}\right)^{nK}\right)\right)^{\frac{\alpha}{\beta^K}} \leq \left(\left(\frac{\beta}{\eta^2}\right)^{\alpha K \beta^{-K}}\right)^n.$$

In an inductive fashion, if $\xi_n \leq \gamma^{n+1}$ then

$$\xi_{n+1} \leq C(\xi_n + \varepsilon_n)^2 \leq 4C(\gamma^{n+1})^2 \leq \gamma^{2n+1} \leq \gamma^{n+2}, \tag{2.9}$$

which implies that $\xi_n \leq \gamma^{n+1}$ for all n. Hence, for α large enough, s_n converges to 1 exponentially fast.

As written in the outline of this proof, the exponential convergence of s_n to 1 is not quite enough for our purposes. Indeed, s_n represents the probability that a ball of radius nK contains a component of size at least η^{nK}, but this component does not necessarily contain the origin. Therefore, we have to make one extra step. Let

$$t_n := \mathbb{P}\big(|\mathcal{C}_{nK}(0)| \geq \eta^{nK}\big).$$

We claim that

$$t_{n+1} \geq t_n \times \mathbb{P}(Bin(N^K - 1, s_n(1-\varepsilon_n)) \geq N^K - 2). \tag{2.10}$$

To see this, we argue as before. If $|\mathcal{C}_{nK}(0)| \geq \eta^{nK}$, then $\mathcal{B}_{nK}(0)$ will be the first good sub-ball in the derivation above. If this component is connected to at least $N^K - 2$ other large components in $\mathcal{B}_{(n+1)K}(0)$ as above, then the component of the origin in $\mathcal{B}_{(n+1)K}$ is large enough, that is, has size at least $\eta^{(n+1)K}$. From this, (2.10) follows. Since a simple coupling gives that

$$\mathbb{P}(Bin(N^K-1, s_n(1-\varepsilon_n)) \geq N^K - 2) \geq \mathbb{P}(Bin(N^K, s_n(1-\varepsilon_n)) \geq N^K - 1),$$

and since the derivation above actually gives that the right hand side of this inequality is bounded below by $1 - \gamma^{n+1}$, it follows that

$$\theta(\alpha,\beta) := \lim_{n\to\infty} t_n > 1 - \sum_{i=1} \gamma^i > 0, \tag{2.11}$$

which is enough to prove the result. □

Remarks (I) By the proof of the strict positivity of $\alpha_c(\beta)$ for $\beta > N$, we also may deduce that $\alpha_c(\beta)$ is not differentiable at $\beta = N$. Indeed, $\alpha_c(\beta) = 0$ for $\beta \leq N$ and since

$$\sum_{k=1}^{\infty}(N-1)N^{k-1}p_k \leq \alpha(N-1)\frac{1}{\beta - N}$$

for $\beta > N$, we have $\alpha_c(\beta) \geq \frac{\beta - N}{N-1}$ for all $\beta > N$. This in turn implies that for all $\beta > N$ we have

$$\frac{\alpha_c(\beta) - \alpha_c(N)}{\beta - N} \geq \frac{1}{N-1} > 0.$$

(II) Since we may choose γ arbitrary small in equation (2.9), we in fact have that for every $\varepsilon > 0$, we can choose α so large, that $t_n > 1 - \varepsilon$ for all n. This implies that for every $\beta \in (N, N^2)$, we can choose α so large such that by 2.11 we have $\theta(\alpha, \beta) > 1 - \varepsilon$.

3 Proof of Theorem 1.2

We will use Theorem 0 from [16]:

Theorem 3.1. *(Gandolfi et al. [16]) Consider long range percolation on $\mathbb{Z}^d$ with the properties*

1. *the model is translation-invariant, and*
2. *the model satisfies the positive finite energy condition.*

Then there can be a.s. at most one infinite component.

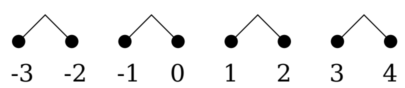
Here the positive finite energy condition is that for every pair of vertices $\{v_1, v_2\}$, with $v_1, v_2 \in \mathbb{Z}^d$, the probability that $v_1 \leftrightarrow v_2$ given the configuration of all other possible edges in $\mathbb{Z}^d \times \mathbb{Z}^d$ is almost surely positive.

In order to be able to use this result, we will first embed the metric generating tree into $\mathbb{Z}$ in a stationary (and ergodic) way. The embedding will be such that for each r, we have

a. any ball of radius r will be represented by N^r consecutive integers,
b. the collection of balls of radius r partitions $\mathbb{Z}$.

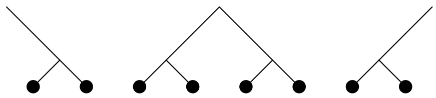
We first describe the construction rather loosely, and after that provide an explicit construction. For ease of description, a collection of m consecutive integers is called an *interval of length* m.

The ball of radius 1 containing 0, that is, $\mathcal{B}_1(0)$ is chosen uniformly at random among all N possible intervals of length N containing the origin of $\mathbb{Z}$. Once we have chosen this ball, all other balls of radius 1 are determined by requirements (a) and (b) above, although it is not yet clear at this point to exactly which balls in Ω_N they correspond. To get an idea of this first step of the procedure, note that for $N = 2$ there are only two possibilities, one of which is depicted here:



The other possibility is obtained by translating the edges over one unit to the right (or to the left, for that matter).

Next, we determine $\mathcal{B}_2(0)$. The ball $\mathcal{B}_2(0)$ is a union of N balls of radius 1 and contains $\mathcal{B}_1(0)$. There are N possible ways to achieve this, keeping in mind that any ball of radius 2 must - according to (a) above - be an interval of length N^2. We now simply choose one of the N possible ways to do this, with probability $1/N$ each. Once we have chosen $\mathcal{B}_2(0)$, all other balls of radius 2 are determined for the same reason as before. The following picture illustrates a possible choice for $\mathcal{B}_2(0)$ given the choice of $\mathcal{B}_1(0)$ made before.



We can continue this procedure as long as we wish, and in doing so we obtain a metric generating tree which is isomorphic to the tree depicted in Figure 1. This last statement perhaps requires some reflection: one can see that this holds by first identifying the two 0's in both graphs, and then build up the balls $\mathcal{B}_r(0)$, $r = 1, 2, \ldots$, in that order.

It is intuitively clear that this construction yields a stationary metric generating tree in the sense that the distribution of the stochastic process which assigns to each pair $\{z, z'\}$ of points in $\mathbb{Z}$ the distance between them, is invariant under integer translations. However, we would like to explicitly construct the tree on an appropriate probability space in such a way that not only stationarity follows as an easy corollary, but we also obtain that the embedding of the metric generating tree is in fact ergodic with respect to translations.

A possible explicit construction is the following. Our probability space is the unit interval $[0, 1]$ endowed with Lebesgue measure on its Borel sigma field. For $\gamma \in [0, 1]$, let $\gamma = 0.\gamma_1\gamma_2\cdots$ be its N-adic expansion, that is,

$$\gamma = \sum_{n=1}^{\infty} \gamma_n N^{-n},$$

where $\gamma_i \in \{0, 1, \ldots, N-1\}$ and we ignore those γ for which the expansion is not unique - this is a set of Lebesgue measure zero anyway. In the construction above, we saw that for each r, $\mathcal{B}_{r-1}(0)$ can be seen as one of the balls of radius $r-1$ among the balls making up $\mathcal{B}_r(0)$. The metric generating tree corresponding to $\gamma \in [0, 1]$ is obtained as follows. We let $\mathcal{B}_r(0)$ be such that $\mathcal{B}_{r-1}(0)$ is the $(\gamma_r + 1)$-st ball in $\mathcal{B}_r(0)$, counted from left to right. For instance, in the preceding two figures with $N = 2$, we have that $\gamma_1 = 1$ and $\gamma_2 = 0$. The map which assigns to each (apart from the exceptional null set discussed before) γ a metric generating tree is denoted by ϕ. This map $\phi : [0, 1] \to \mathcal{T}$, where $\mathcal{T}$ is the set of metric generating trees, is invertible on a set of full Lebesgue measure.

One can write down explicitly the transformation $S : [0, 1] \to [0, 1]$ which corresponds to the left-shift T on the space of metric generating trees in the sense that $\phi \circ S = T \circ \phi$, hence $T = \phi S \phi^{-1}$. Indeed, a little reflection shows that S can be described as follows: if $Y(\gamma) = \min\{n; \gamma_n \neq N-1\}$ then $S(\gamma)_k$ (that is, the k-th digit in $S(\gamma)$) is given by

$$S(\gamma)_k = \begin{cases} 0 & \text{if } k < Y(\gamma), \\ \gamma_k + 1 & \text{if } k = Y(\gamma), \\ \gamma_k & \text{if } k > Y(\gamma). \end{cases}$$

This transformation has been studied in the literature and goes by the name Kakutani - Von Neumann transformation, see e.g. [14] or [15]. It is easy to check that Lebesgue measure is invariant under the action of S, and this immediately proves that the construction of our random metric generating tree is stationary on $\mathbb{Z}$.

Proof of Theorem 1.2. The construction above shows that the metric generating tree can be embedded into $\mathbb{Z}$ in a stationary way. We claim that this implies that the whole long range percolation process on the hierarchical lattice can be realised as a stationary percolation process on $\mathbb{Z}$. To see this, we assign a uniformly-$[0,1]$ distributed random variable U_e to each edge e in such a way that the collection is independent.

Given a realisation of the metric generating tree, we declare edge e to be open if $U_e \leq 1 - \exp(-\alpha/\beta^{|e|})$, where $|e|$ denotes the length of e with respect to the metric generating tree. This gives a realisation of the percolation process with the correct distribution, and shows that we have embedded the full percolation process on the hierarchical tree in a stationary way. Since every pair of vertices shares an edge, with positive probability, irrespective of the presence or absence of other edges, the positive finite energy condition is met and the result follows. □

With a little more work one can also see that the construction is in fact ergodic, that is, any event which is invariant under the shift on $\mathbb{Z}$ has probability 0 or 1. To show this, we first show that Lebesgue measure on $[0,1]$ is ergodic with respect to S. This result is known (see e.g. [15] Theorem 1), but we give a simple (and new) proof for the convenience of the reader.

Lemma 3.2. *Lebesgue measure on* $[0,1]$ *is ergodic with respect to* S*.*

Proof. Consider the first digit in each of γ, $S(\gamma)$, $S^2(\gamma), \ldots$. From the construction we immediately conclude that the first digit follows the periodic pattern $0, 1, \ldots, N-1, 0, 1, 2, \ldots, N-1, \ldots$ starting at any number. Hence the first digit is just adding 1 modulo N. The second digit can only change when the first digit is an $N-1$, and then it also changes according to adding 1 modulo N. In general the k-th digit can only change when the $(k-1)$-st digit is an $N-1$, and then the change consists of adding 1 modulo N. It follows from these observations that the orbit of γ under the action of S visits any N-adic interval $I_{m,k} = [kN^{-m}, (k+1)N^{-m}]$ with frequency N^{-m}, that is,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}_{S(\gamma)\in I_{m,k}} = N^{-m}, \tag{3.1}$$

where $\mathbf{1}_A$ denotes the indicator function of A.

Now consider the collection $\mathcal{M}$ of invariant probability measures for the transformation S. From the fact that Lebesgue measure preserves measure under S we see that $\mathcal{M}$ is not empty. It is well known and easy to see that the set $\mathcal{M}$ is convex, and that the ergodic measures are precisely the extremal points of $\mathcal{M}$. Since $\mathcal{M} \neq \emptyset$, this implies that there is at least one ergodic measure with respect to S.

Let ν be any ergodic measure with respect to S. It follows from the ergodic theorem and (3.1) that $\nu(I_{m,k}) = N^{-m}$ for any m and $k = 0, \ldots, N^m - 1$. However, there is only one measure that satisfies this condition, namely Lebesgue measure on $[0,1]$. Hence ν must be Lebesgue measure, which we already know is indeed invariant. □

Theorem 3.3. *The embedding of our long range percolation process on the hierarchical lattice into* $\mathbb{Z}$ *is ergodic.*

Proof. From Lemma 3.2 it follows that the metric generating tree is embedded ergodically. Adding the i.i.d. random variables U_e as before does not destroy ergodicity, and the final configuration is a factor of this ergodic process and hence ergodic itself. □

4 Proof of Theorem 1.5

We provide the proof for $\beta \geq N$. By monotonicity and the observation made in the proof of Theorem 1.1(a) that $\theta(\alpha, N) = 1$ for $\alpha > 0$, it follows that the result holds for all $\beta > 0$. The proof consists of three steps:

1. For every constant $K > 0$ the indicator function of the event that both $|\mathcal{C}(0)| = \infty$ and $|\mathcal{C}_n(0)| < K(\beta/N)^n$ converges a.s. to 0 as $n \to \infty$.
2. The fraction of the vertices in $\mathcal{B}_n(0)$ which are in a cluster of size at least $K(\beta/N)^n$, converges a.s. to θ as $n \to \infty$.
3. Use the previous two steps to prove the theorem.

Step 1. Observe that if $\beta = N$, then it follows from the a.s. infinite degree of the origin that $\mathbb{P}(|\mathcal{C}_n(0)| < K) \to 0$ as $n \to \infty$ and therefore the claim follows. For $\beta > N$, we compute

$$\mathbb{P}\left(|\mathcal{C}(0)| = \infty \middle| |\{n \in \mathbb{N}; |\mathcal{C}_n(0)| \leq K\left(\frac{\beta}{N}\right)^n\}| = \infty\right).$$

Note that the event conditioned on has positive probability, since $|\mathcal{C}_n(0)|$ might be finite. Let n_1 be the smallest n for which $|\mathcal{C}_n(0)| \leq K(\beta/N)^n$.If $\mathcal{C}_{n_i}(0) \leftrightarrow \overline{\mathcal{B}_{n_i}(0)}$, then n_{i+1} is the smallest $n > n_i$ for which $\mathcal{C}_{n_i}(0) \not\leftrightarrow \overline{\mathcal{B}_n(0)}$ and for which $|\mathcal{C}_n(0)| \leq K(\beta/N)^n$.

Since $|\mathcal{C}_{n_i}(0)| \leq K(\beta/N)^{n_i}$, we have

$$\begin{aligned}
\mathbb{P}\left(\mathcal{C}_{n_i}(0) \leftrightarrow \overline{\mathcal{B}_{n_i}(0)}\right) &\leq \mathbb{P}\left(\mathcal{C}_{n_i}(0) \leftrightarrow \overline{\mathcal{B}_{n_i}(0)} \middle| |\mathcal{C}_{n_i}(0)| = \left\lfloor K\left(\tfrac{\beta}{N}\right)^{n_i}\right\rfloor\right) \\
&\leq 1 - \exp\left(-\alpha K\left(\tfrac{\beta}{N}\right)^{n_i} \sum_{j=n_i+1}^{\infty} (N-1)\frac{N^{j-1}}{\beta^j}\right) \\
&= 1 - \exp\left(-\alpha K \tfrac{N-1}{\beta - N}\right).
\end{aligned} \tag{4.1}$$

The right hand side is strictly less than 1 and is independent of n_i. So there will be an n_i for which $\{\mathcal{C}_{n_i}(0) \not\leftrightarrow \overline{\mathcal{B}_{n_i}(0)}\}$, and it follows that

$$\mathbb{P}\left(|\mathcal{C}(0)| = \infty \middle| |\{n \in \mathbb{N}; |\mathcal{C}_n(0)| \leq K\left(\frac{\beta}{N}\right)^n\}| = \infty\right) = 0,$$

which implies

$$\mathbb{P}\left(|\mathcal{C}(0)| = \infty \cap |\{n \in \mathbb{N}; |\mathcal{C}_n(0)| \leq K\left(\frac{\beta}{N}\right)^n\}| = \infty\right) = 0.$$

Step 2. We use the random embedding of the hierarchical lattice in $\mathbb{Z}$ introduced in the previous section. We start by considering all vertices between $-N^n$ and N^n. After that, we show how we can use that to study the vertices in $\mathcal{B}_n(0)$.

Consider

$$A(n) := \frac{1}{2N^n + 1} \sum_{x=-N^n}^{N^n} \mathbf{1}\left(|\mathcal{C}_n(x)| > K\left(\frac{\beta}{N}\right)^n\right).$$

First observe that for every $k \in \mathbb{N} \cap [0, n]$

$$\begin{aligned}
A(n) &\leq \frac{1}{2N^n + 1} \sum_{x=-N^n}^{N^n} \mathbf{1}\left(|\mathcal{C}(x)| > K\left(\frac{\beta}{N}\right)^n\right) \\
&\leq \frac{1}{2N^n + 1} \sum_{x=-N^n}^{N^n} \mathbf{1}\left(|\mathcal{C}(x)| > K\left(\frac{\beta}{N}\right)^k\right).
\end{aligned}$$

By Theorem 3.3 and the ergodic theorem, we have that for every $k \in \mathbb{N}$,

$$\frac{1}{2N^n+1}\sum_{x=-N^n}^{N^n}\mathbf{1}\left(|\mathcal{C}(x)|>K\left(\frac{\beta}{N}\right)^k\right)\xrightarrow[n\to\infty]{a.s.}\mathbb{P}\left(|\mathcal{C}(0)|>K\left(\frac{\beta}{N}\right)^k\right),$$

which decreases to θ, as $\beta > N$ and $k\to\infty$, while for$\beta = N$ the right-hand side is equal to 1. In particular, we have that for every $\epsilon > 0$

$$\mathbf{1}(A(n)<\theta+\epsilon)\xrightarrow[n\to\infty]{a.s.}1. \tag{4.2}$$

Then, use that for every $k \in \mathbb{N}\cap[0,n]$,

$$A(n)\geq\frac{1}{2N^n+1}\sum_{x=-N^n}^{N^n}\mathbf{1}\left(\bigcap_{j=k}^{\infty}\left\{|\mathcal{C}_j(x)|>K\left(\frac{\beta}{N}\right)^j\right\}\right).$$

By Theorem 3.3 and the ergodic theorem we have that for every $k \in \mathbb{N}$

$$\frac{1}{2N^n+1}\sum_{x=-N^n}^{N^n}\mathbf{1}\left(\bigcap_{j=k}^{\infty}\left\{|\mathcal{C}_j(x)|>K\left(\frac{\beta}{N}\right)^j\right\}\right)\xrightarrow[n\to\infty]{a.s.}\mathbb{P}\left(\bigcap_{j=k}^{\infty}\left\{|\mathcal{C}_j(0)|>K\left(\frac{\beta}{N}\right)^j\right\}\right).$$

Since $\bigcap_{j=k}^{\infty}\left\{|\mathcal{C}_j(0)|>K\left(\frac{\beta}{N}\right)^j\right\}\subset\{|\mathcal{C}(0)|=\infty\}$, we have

$$\mathbb{P}\left(\bigcap_{j=k}^{\infty}\left\{|\mathcal{C}_j(0)|>K\left(\frac{\beta}{N}\right)^j\right\}\right)\leq\theta.$$

Combined with Step 1 this implies that $\mathbb{P}\left(\bigcap_{j=k}^{\infty}\left\{|\mathcal{C}_j(0)|>K\left(\frac{\beta}{N}\right)^j\right\}\right)$ increases to θ, as $\beta\geq N$ and $k\to\infty$. In particular, we have that for every $\epsilon>0$

$$\mathbf{1}(A(n)>\theta-\epsilon)\xrightarrow[n\to\infty]{a.s.}1.$$

Together with (4.2) this implies $A(n)\xrightarrow[n\to\infty]{a.s.}\theta$.

Note that the collection of vertices $\{-N^n,-N^n+1,-N^n+2,\dots,N^n\}$ contains the image under the embedding of the ball $\mathcal{B}_n(0)$ and this image contains a fraction $\frac{N^n}{2N^n+1}$ of those vertices. Furthermore, the events $|\mathcal{C}_n(x)|>K(\beta/N)^n$ are independent for vertices x in different n-balls, so

$$A_1(n):=\frac{1}{2N^n+1}\sum_{x\in\mathcal{B}_n(0)}\mathbf{1}\left(|\mathcal{C}_n(x)|>K\left(\frac{\beta}{N}\right)^n\right)$$

and $A_2(n):=A(n)-A_1(n)$ are independent. We have that

$$A_1(n)+A_2(n)=A(n)\xrightarrow[n\to\infty]{a.s.}\theta,$$

and we want to conclude that in fact $A_1(n)\xrightarrow[n\to\infty]{a.s.}\theta/2$. Since $A_1(n)$ and $A_2(n)$ are bounded above by 1 and have asymptotically the same mean $\theta/2$, it is enough to show that $A_1(n)$ converges a.s. to some constant. By independence, the joint distribution of $A_1(n)$ and $A_2(n)$ is a product measure μ_n say. Since the only product measures on $[0,1]$ which concentrates on the line $x+y=\theta$ are point masses at a point on this line, it is easy to see that the marginal distributions of μ_n must also converge to a point mass, which proves that $A_1(n)$ also converges a.s. to a constant.

Step 3. The strategy is to split those components in $\mathcal{B}_{n+1}(0)$ which are at least of size $K(\beta/N)^n$ into clusters roughly of size $K(\beta/N)^n$. Then we use those clusters as "meta-vertices" for an N-partite graph, in which meta-vertices in different n-balls are connected if the clusters they represent are connected by an edge of length $n+1$. Meta-vertices in the same n-ball never share an edge. We show that if we choose K and n large enough, then the largest component of the graph of meta-vertices contains a fraction of the meta-vertices close to 1, which shows that for large n, the fraction of vertices in the largest cluster of $\mathcal{B}_{n+1}(0)$ is close to θ. We will be more precise now.

By step 2 we know that for every $K>0$, every $\varepsilon>0$ and all large enough n, it holds that

$$\mathbb{P}\left(\left|\left\{x\in\mathcal{B}_n(0); |\mathcal{C}_n(x)|>K\left(\frac{\beta}{N}\right)^n\right\}\right|>(\theta-\varepsilon)N^n\right)>1-\varepsilon.$$

We now fix ε. The ball $\mathcal{B}_n(y)$ is said to be *good* if

$$\left|\left\{x\in\mathcal{B}_n(y); |\mathcal{C}_n(x)|>K\left(\frac{\beta}{N}\right)^n\right\}\right|>(\theta-\varepsilon)N^n.$$

We condition on the event that all n-balls in $\mathcal{B}_{n+1}(0)$ are good. The probability of this event is bounded below by $(1-\varepsilon)^N>1-N\varepsilon$.

Now, for every good ball $\mathcal{B}_n(y)$, $y\in\Omega_N$, we make a partition of the set

$$\left\{x\in\mathcal{B}_n(y); |\mathcal{C}_n(x)|>K\left(\frac{\beta}{N}\right)^n\right\}$$

in "meta-vertices". For the moment we denote this set by $\mathcal{B}'_n(y)$. For $x\in\mathcal{B}'_n(y)$ we make a partition of $\mathcal{C}_n(x)$ in $\lfloor|\mathcal{C}_n(x)|/(\lceil K(\beta/N)^n\rceil)\rfloor$ sets, which all have size at least $\lceil K(\beta/N)^n\rceil$. The vertices that are not in such a cluster are ignored for the moment. Denote the collection of meta-vertices that contain vertices in $\mathcal{B}_{n+1}(0)$ by $\mathcal{V}_n$. We note that if $\mathcal{B}_n(y)$ is good and K is large enough, then it contains at least

$$(\theta-\varepsilon)N^n/\lceil 2K(\beta/N)^n\rceil\geq(\theta-\varepsilon)N^n/(3K(\beta/N)^n)$$

meta-vertices.

We construct a new N-partite graph on $\mathcal{V}_n$ as follows. Let $\mathcal{V}_n$ be the vertex set and let $\mathcal{E}_n$ be the set of edges between those vertices. This edge set is obtained as follows. Choose $\lceil K(\beta/N)^n\rceil$ original vertices from every meta-vertex in $\mathcal{V}_n$. Choosing those vertices may be done in any way that is independent of the presence of edges of length $n+1$ or larger, e.g. by choosing the first $\lceil K(\beta/N)^n\rceil$ vertices in the vertex sets in $\mathcal{V}_n$ according to the labeling generated by the function f as defined in (1.1). Denote these new sets of meta-vertices by $\mathcal{A}_n$. The meta-vertices $x,y\in\mathcal{V}_n$ share an edge in $\mathcal{E}_n$, if there is at least 1 edge in the original graph that is shared by vertices that make up the sets in $\mathcal{A}_n$ corresponding to x and y, and if the original vertices that make up x and y are at distance $n+1$ of each other. Otherwise there is no edge between the meta-vertices.

As observed before, the number of meta-vertices in $\mathcal{V}_n$ that consist of vertices from a good ball $\mathcal{B}_n(x)$, is at least $(\theta-\varepsilon)N^n/(3K(\beta/N)^n)$. Since $\beta<N^2$, this quantity grows to ∞ as $n\to\infty$. The expected degree of a vertex in $\mathcal{V}_n$ exceeds

$$\frac{(N-1)(\theta-\varepsilon)N^n}{3K(\beta/N)^n}\left(1-\exp\left(-\alpha\beta^{-(n+1)}(K)^2\left(\frac{\beta}{N}\right)^{2n}\right)\right),$$

which is larger than $\lambda := (N-1)(\theta-\varepsilon)\alpha K/(6\beta)$, for all large enough n. This holds for every $K > 0$, and therefore the expected degree can be chosen to be arbitrary large.

This N-partite graph falls within the class of inhomogeneous random graphs of Bollobás, Janson and Riordan [7]. The degree of every meta-vertex is asymptotically Poisson distributed, with mean bounded below by λ and we know [7, Thm. 3.1] that the (unique) largest component of such an N-partite graph contains with high probability (in the limit for $n \to \infty$) a fraction ρ of the meta-vertices, where ρ is the largest solution of

$$1 - \rho = e^{-\lambda\rho}.$$

By tuning K, λ can be chosen arbitrary large and ρ can be taken such that $\rho > 1 - \varepsilon$. So, for every $\varepsilon > 0$ and large enough n the graph $(\mathcal{V}_n, \mathcal{E}_n)$ contains a unique giant component, which contains a fraction $1 - \varepsilon$ of the vertices in $\mathcal{V}_n$, with probability at least $1 - \varepsilon$.

Since we have conditioned on the event that all n-balls in $\mathcal{B}_{n+1}(0)$ are good, the fraction of vertices in $\mathcal{B}_{n+1}(0)$, that are part of vertices in $\mathcal{V}_n$ is bounded below by $\theta - 2\varepsilon$. (The factor 2, is due to the fact that the sizes of different meta-vertices differ at most a factor 2). Therefore, with the same conditioning, the largest cluster in $\mathcal{B}_{n+1}(0)$ is at least of size

$$(\rho - \varepsilon)(\theta - 2\varepsilon)N^n > (1 - 2\varepsilon)(\theta - 2\varepsilon)N^n$$

with probability exceeding $1 - \varepsilon$. Now multiplying by the probability that all n-balls in $\mathcal{B}_{n+1}(0)$ are good, gives that the probability that the largest cluster in $\mathcal{B}_{n+1}(0)$ is at least of size $(1 - 2\varepsilon)(\theta - 2\varepsilon)N^n$ is bounded below by $(1 - \varepsilon)(1 - N\varepsilon)$. By choosing $\varepsilon' \in (0, \varepsilon/\max(4, N+1))$, we obtain that $\mathbb{P}(|\mathcal{C}_n(0)| > (\theta - \varepsilon')N^n)$ is at least $1 - \varepsilon'$ and this finishes the proof. □

Remark We realize that it is possible to prove the statement of Step 2 by using the strong law of large numbers. If we do this, then it is only a small step from the proof of Theorem 1.5 to a proof of Theorem 1.2. However, we think that the proof presented in the previous section contains some valuable ideas and therefore should be included in this paper.

5 Proof of Theorem 1.3

Continuity proofs of percolation functions typically split into separate proofs for left and right continuity, one of which typically follows from standard arguments [17]. In this case, continuity from the right in α and continuity from the left in β are the easy parts:

Lemma 5.1. *$\theta(\alpha, \beta)$ is continuous from the right in $\alpha > 0$ and continuous from the left in $\beta > 0$.*

Proof. We use that a decreasing limit of increasing (resp. decreasing) functions, which are continuous from the right (resp. left) is continuous from the right (resp. left), and apply this to the sequence $\mathbb{P}_{\alpha,\beta}\left(\mathcal{C}_i(0) \leftrightarrow \overline{\mathcal{B}_i(0)}\right)$, viewed as functions of α and β. Note that

$$\{\mathcal{C}_i(0) \leftrightarrow \overline{\mathcal{B}_i(0)}\} \subset \{\mathcal{C}_{i-1}(0) \leftrightarrow \overline{\mathcal{B}_{i-1}(0)}\}$$

and

$$\{|\mathcal{C}(0)| = \infty\} = \bigcap_{i=0}^{\infty} \{\mathcal{C}_i(0) \leftrightarrow \overline{\mathcal{B}_i(0)}\}. \tag{5.1}$$

Therefore, the sequence of probabilities has the appropriate limit. Furthermore, the probability $\mathbb{P}_{\alpha,\beta}\left(\mathcal{C}_i(0) \leftrightarrow \overline{\mathcal{B}_i(0)}\right)$ is increasing in α and decreasing in β.

The only thing left to prove is that $\mathbb{P}_{\alpha,\beta}\left(\mathcal{C}_i(0) \leftrightarrow \overline{\mathcal{B}_i(0)}\right)$ is continuous in α and β, which is not entirely trivial since the event depends on infinitely many pairs of vertices.

For $\alpha > 0$ and $\beta \le N$, $\theta(\alpha,\beta) = 1$, so the statement of the lemma holds in that domain. A straightforward computation yields that for $\beta > N$,

$$\mathbb{P}\left(\mathcal{C}_i(0) \leftrightarrow \overline{\mathcal{B}_i(0)}\right) = \mathbb{E}\left(1 - \exp\left(-\alpha|\mathcal{C}_i(0)|\frac{N-1}{\beta - N}\left(\frac{N}{\beta}\right)^i\right)\right). \tag{5.2}$$

Since $|\mathcal{C}_i(0)|$ depends on the state of only finitely many edges, this expectation is continuous in α and β for $\alpha > 0$ and $\beta > N$. Furthermore, $\mathbb{P}\left(\mathcal{C}_i(0) \leftrightarrow \overline{\mathcal{B}_i(0)}\right) \to 1$ as $\beta \searrow N$, so continuity holds in the whole domain and the statement of the lemma follows. □

In order to prove that $\theta(\alpha,\beta)$ is continuous from the left in $\alpha > 0$ and continuous from the right in $\beta > 0$, we use a renormalisation argument. Fix $\alpha > 0$ and $N \le \beta < N^2$. To get insight in the argument we first (falsely) assume that for given $\varepsilon > 0$ and large enough finite k, there is a $\delta > 0$ such that,

$$\mathbb{P}_{\alpha-\delta,\beta+\delta}(|\mathcal{C}_k^m(0)| > (\theta(\alpha,\beta) - \varepsilon)N^k) = 1.$$

(Although the assumption is false, we can get this probability arbitrary close to 1, by choosing k large enough, δ small enough and using Theorem 1.5.)

Now we use renormalisation. The balls of radius k are considered as vertices of Ω_N which we call "meta-vertices". If two vertices in the original model have distance $k + l$, then the meta-vertices in which they are contained are at distance l. Vertices in the new model are connected if and only if the largest clusters in the original k-balls, represented by these vertices, are connected by an edge. The new model is again a percolation model on Ω_N.

Let x and y be meta-vertices, at distance l of each other. Define, for $\delta > 0$ small,

$$\alpha' := (\alpha - \delta)((\theta(\alpha,\beta) - \varepsilon))^2 \frac{N^{2k}}{(\beta+\delta)^k}.$$

Given the states of all other edges, the (conditional) probability that x and y are connected to each other is always bounded below by

$$1 - \exp(-(\alpha - \delta)((\theta(\alpha,\beta) - \varepsilon)N^k)^2(\beta+\delta)^{-(k+l)})$$

and by the choice of α', this is just

$$1 - \exp(-\alpha'(\beta+\delta)^{-l}).$$

Hence, the renormalized model stochastically dominates the percolation model with parameters α' and $\beta + \delta$.

Since $N^2/(\beta+\delta) > 1$, α' can be chosen arbitrary large by choosing k large. In particular it can be chosen such that $\theta(\alpha', \beta+\delta) > 1 - \varepsilon$, (by the second remark after the proof of Theorem 1.1). It follows that for large enough k

$$\begin{aligned}
\mathbb{P}_{\alpha-\delta,\beta+\delta}\left(|\mathcal{C}| = \infty\right) &\ge \theta(\alpha',\beta+\delta)\mathbb{P}_{\alpha-\delta,\beta+\delta}(0 \in \mathcal{C}_k^m(0)) \\
&\ge (1-\varepsilon)(\theta(\alpha,\beta) - \varepsilon) \\
&\ge \theta(\alpha,\beta) - 2\varepsilon.
\end{aligned}$$

The only problem is that we have incorrectly assumed that

$$\mathbb{P}_{\alpha-\delta,\beta+\delta}\left(|\mathcal{C}_k^m(0)| > (\theta(\alpha,\beta) - \varepsilon)N^k\right) = 1,$$

and we will deal with this problem now. We need the notion of mixed percolation (cf. [8]). Mixed percolation involves independently removing vertices, together with all of its adjacent edges. Formally, the measure $\mathbb{P}^{mixed}_{\alpha,\beta,\gamma}$ is constructed as follows. A vertex in Ω_N is open with probability $1-\gamma$, independently of the states (open or closed) of the other vertices in Ω_N. If x and y are both open vertices then they share an edge with probability $1-\exp(-\alpha/\beta^{d(x,y)})$. Conditioned on the states of the vertices the presence or absence of edges are independent.

Lemma 5.2. *Let $\beta > N$. For every $\varepsilon > 0$, there exists $\gamma > 0$, such that*

$$\mathbb{P}_{\alpha,\beta}(|\mathcal{C}(0)| = \infty) \le \mathbb{P}^{mixed}_{\alpha(1+\varepsilon),\beta,\gamma}(|\mathcal{C}(0)| = \infty).$$

Before giving the proof of this result, we show how it can be used to prove Theorem 1.3. The following lemma suffices.

Lemma 5.3. *If $\theta(\alpha,\beta) > 0$, then for every $\epsilon \in (0,\theta(\alpha,\beta))$, there exists $\delta > 0$ such that $\theta(\alpha-\delta,\beta+\delta) > \theta(\alpha,\beta)-\epsilon$.*

Proof. Fix α, β and $\epsilon > 0$. Let α' be such that

$$\theta(\alpha', (N^2+\beta)/2) > 1-\varepsilon/3,$$

which is possible by Theorem 1.1(b) and the second remark after its proof. Furthermore, let $\gamma \in (0,\epsilon/3)$ be such that

$$\theta(\alpha', (N^2+\beta)/2) \le \mathbb{P}^{mixed}_{2\alpha',(N^2+\beta)/2,\gamma}(|\mathcal{C}(0)| = \infty),$$

which is possible by Lemma 5.2 and the monotonicity of the right-hand side in γ. Let K be such that the following conditions are satisfied:

1. $\alpha(\theta(\alpha,\beta)-\epsilon/2)2(N^2/\beta)^K > 3\alpha'$,
2. $\mathbb{P}_{\alpha,\beta}(|\mathcal{C}^m_K(0)| > (\theta(\alpha,\beta)-\epsilon/3)N^K) > 1-\gamma/2$,

which are possible by respectively $N^2 > \beta$ and Theorem 1.5. Finally, let $\delta > 0$ be such that $\delta < \min(\alpha/3, (N^2-\beta)/2)$ and

$$\mathbb{P}_{\alpha-\delta,\beta+\delta}(|\mathcal{C}^m_K(0)| > (\theta(\alpha,\beta)-\epsilon/3)N^K) > 1-\gamma, \tag{5.3}$$

which is possible by the continuity of the probability in α and β for finite K.

We say that the ball $\mathcal{B}_K(x)$ is *good* if the size of $\mathcal{C}^m_K(x)$ is at least $(\theta(\alpha,\beta)-\epsilon/3)N^K$. Delete all vertices that are in a ball of diameter K which is not good and also all vertices that are not in the largest cluster of good balls. As above, we interpret the remaining components as the vertices of the hierarchical lattice of order N in which vertices are independently deleted with probability at most γ, by (5.3). Remaining clusters in the original graph, of which the vertices are at distance $K+l$, are connected by at least one edge with probability at least

$$1-\exp(-(\alpha-\delta)(\theta(\alpha,\beta)-\epsilon/3)2N^{2K}\beta^{-(K+l)}) > 1-\exp(-2\alpha'\beta^{-l}),$$

irrespective of the existence or absence of other connections. Here we have used that $\alpha-\delta > 2\alpha/3$. Hence the rescaled process stochastically dominates a mixed percolation process with parameters $2\alpha'$, β and γ.

Now note that by exchangeable

$$\begin{aligned} & \mathbb{P}_{\alpha-\delta,\beta+\delta}(|\mathcal{C}_K(0)| \ge (\theta(\alpha,\beta)-\epsilon/3)N^K) \\ \ge\ & (\theta(\alpha,\beta)-\epsilon/3)\mathbb{P}_{\alpha-\delta,\beta+\delta}(|\mathcal{C}^m_K(0)| > (\theta(\alpha,\beta)-\epsilon/3)N^K) \\ \ge\ & (\theta(\alpha,\beta)-\epsilon/3)(1-\gamma). \end{aligned}$$

Furthermore, conditioned on 0 being in the largest cluster of a good ball, the probability that 0 is in an infinite cluster if the parameters are $\alpha-\delta$ and $\beta+\delta$ is larger than $1-\epsilon/3$. Combining these observations and $\gamma<\epsilon/3$ gives

$$\theta(\alpha-\delta,\beta+\delta)>(\theta(\alpha,\beta)-\epsilon/3)(1-\epsilon/3)^2>\theta(\alpha,\beta)-\epsilon.$$

□

It remains to prove Lemma 5.2. Before giving a proof of this lemma we define a directed long-range percolation model and relate this to the undirected model. In the directed version, vertices in Ω_N are open with probability $1-\gamma$. If vertex x is open, then a directed edge from x to y is present with probability $1-\exp\left(-\alpha\beta^{-d(x,y)}\right)$. Conditioned on the states of the vertices (open or closed) the presence or absence of an edge is independent of the presence or absence of other edges. The corresponding measure we denote by $\hat{\mathbb{P}}^{mixed}_{\alpha,\beta,\gamma}$. The set of vertices which can be reached by a path from vertex x is denoted by $\hat{\mathcal{C}}(x)$. Note that in the directed model, the presence of a path from x to y does not necessarily imply that there exists a path from y to x. We define the directed version of the original (not mixed) measure, $\hat{\mathbb{P}}_{\alpha,\beta}$, in a similar way and note that $\hat{\mathbb{P}}_{\alpha,\beta}=\hat{\mathbb{P}}^{mixed}_{\alpha,\beta,0}$. Standard arguments (see e.g. [9, 21]) can be used to show that

$$\mathbb{P}^{mixed}_{\alpha,\beta,\gamma}(|\mathcal{C}(0)|=\infty)=\hat{\mathbb{P}}^{mixed}_{\alpha,\beta,\gamma}(|\hat{\mathcal{C}}(0)|=\infty).$$

Proof of Lemma 5.2. The directed mixed percolation graph with parameters α, β and γ can be obtained as follows (the ordinary model can be obtained by taking $\gamma=0$). We assign i.i.d. random variables X_x to the vertices $x\in\Omega_N$, all Poisson distributed with parameter $\alpha(N-1)/(\beta-N)$. We construct a directed multi-graph (a graph in which multiple edges between two vertices in the same direction are allowed). Vertices are open with probability $1-\gamma$, independently of each other. If x is open, then X_x directed edges start at x. The endpoints of these edges are independently chosen from $\Omega_N\setminus x$, and a vertex at distance r of x is chosen with probability $(\beta-N)(N-1)^{-1}\beta^{-r}$. If x is closed, then no edges start at x. We obtain the original directed graph by replacing the collection of all edges from x to y (if there is at least one) by a single edge from x to y, for all $x,y\in\Omega_N$.

Let Z_1 be a Poisson distributed random variable with mean $\alpha\frac{N-1}{\beta-N}$. Furthermore, let $Z_2=Y_1Y_2$, where Y_1 is equal to 1 with probability $1-\gamma$ and equal to 0 with probability γ, and where the random variable Y_2 is independent of Y_1 and Poisson distributed with parameter $\alpha(1+\varepsilon)(N-1)/(\beta-N)$. For the ordinary percolation model the number of edges starting at x in the multigraph is distributed as Z_1, while for the mixed percolation model, the number of edges starting at x is distributed as Z_2. It is now easy to check that for $\varepsilon>0$ there is a $\gamma>0$, such that $\mathbb{P}(Z_1=0)=\mathbb{P}(Z_2=0)$ and for this γ and all $k>0$ we have,

$$\mathbb{P}(Z_2>k|Z_2>0)=\mathbb{P}(Y_2>k|Y_2>0)>\mathbb{P}(Z_1>k|Z_1>0).$$

The statement of Lemma 5.2 now follows by a straightforward coupling argument. □

Remark In [4, Theorem 1.5] percolation on the Euclidean lattice $\mathbb{Z}^d$, $d\geq 1$ is studied for $p_r=1-e^{-\alpha r^{-s}}$ where $\alpha>0$ and $d<s<2d$. Here r is the Euclidean distance between two vertices. Berger shows that for given s, the percolation probability is continuous in α. (For ease of exposition we have formulated this result slightly less general than Berger did.) This result is strongly related to Theorem 1.3, but the proof of Berger relies on translation invariance and independence of the presence or absence of edges between different pairs of vertices. In our model we cannot have both of these properties simultaneously, therefore we provided a detailed model specific proof of Theorem 1.3 above.

6 Proof of Theorem 1.4

The proof of Theorem 1.4 is split into separate proofs of continuity from the right and from the left of $\alpha_c(\beta)$.

Lemma 6.1. $\alpha_c(\beta)$ *is strictly increasing on* $\beta \in (N, N^2)$ *and continuous from the right on* $\beta \in (0, N^2)$.

Proof Theorem 1.3 implies that $\theta(\alpha_c(\beta), \beta) = 0$ for $\beta > N$. By observing that for every $\varepsilon > 0$, $\mathbb{P}_{(1+\varepsilon)\alpha,(1+\varepsilon)\beta}$ is stochastically dominated by $\mathbb{P}_{\alpha,\beta}$, we deduce that

$$\alpha_c(\beta(1+\varepsilon)) \geq (1+\varepsilon)\alpha_c(\beta).$$

Since by Theorem 1.1, $\alpha_c(\beta) > 0$ for $\beta \in (N, N^2)$, this gives that $\alpha_c(\beta)$ is strictly increasing on (N, N^2).

In order to prove continuity from the right, we note that for all $\delta > 0$, $\theta(\alpha_c(\beta)+\delta, \beta)$ is strictly positive by definition. By the continuity of $\theta(\alpha, \beta)$ we obtain that there exist $\varepsilon > 0$, such that $\theta(\alpha_c(\beta)+\delta, \beta+\varepsilon) > 0$ and therefore $\alpha_c(\beta+\varepsilon) < \alpha_c(\beta)+\delta$. This together with $\alpha_c(\beta+\varepsilon) > \alpha_c(\beta)$ completes the proof. □

Lemma 6.2. $\alpha_c(\beta)$ *is continuous from the left for* $\beta \in (0, N^2)$.

To prove this, we use [1]. In that paper it is shown that for long-range percolation on $\mathbb{Z}^d$,

$$\inf\{\alpha : \theta(\alpha,\beta) > 0\} = \sup\{\alpha : \mathbb{E}_{\alpha,\beta}(|\mathcal{C}(0)|) < \infty\}. \tag{6.1}$$

Inspection of the proof of this result yields that this proof also works on the hierarchical lattice. Now use the following lemma.

Lemma 6.3. *Let* $\alpha > 0$ *and* $\beta > N$. *If* $\mathbb{E}_{\alpha,\beta}(|\mathcal{C}(0)|) < \infty$, *then there exist* $\varepsilon > 0$ *such that* $\mathbb{E}_{\alpha,\beta(1-\varepsilon)}(|\mathcal{C}(0)|) < \infty$.

Proof of Lemma 6.2 (given Lemma 6.3). For every $\varepsilon > 0$, we have that

$$\mathbb{P}_{\alpha_c(\beta),\beta(1-\varepsilon)}(|\mathcal{C}(0)| = \infty) > 0,$$

by the strict increase of $\alpha_c(\beta)$. This implies that $\mathbb{E}_{\alpha_c(\beta),\beta(1-\varepsilon)}(|\mathcal{C}(0)|) = \infty$, for every $\varepsilon > 0$ and therefore $\mathbb{E}_{\alpha_c(\beta),\beta}(|\mathcal{C}(0)|) = \infty$ by Lemma 6.3.

Furthermore, by equation (6.1) we know that $\mathbb{E}_{\alpha_c(\beta)-\delta,\beta}(|\mathcal{C}(0)|) < \infty$ for every $\delta > 0$. Therefore, there exist $\varepsilon > 0$ such that

$$\mathbb{E}_{(\alpha_c(\beta)-\delta),\beta(1-\varepsilon)}(|\mathcal{C}(0)|) < \infty,$$

which in turn implies that for all $\delta > 0$, there exists a strictly positive constant ε such that $\alpha_c(\beta(1-\varepsilon)) > \alpha_c(\beta) - \delta$. This together with $\alpha_c(\beta(1-\varepsilon)) < \alpha_c(\beta)$ gives continuity from the left of $\alpha_c(\beta)$. □

Proof of Lemma 6.3. Assign independent uniform $(0,1)$ random variables to all pairs of vertices in Ω_N. The random variable assigned to the pair (x,y) is denoted by $U(x,y)$. We say that the vertices x and y share an edge for the parameters α and β if $U(x,y) < 1 - \exp(-\alpha\beta^{-d(x,y)})$. This construction provides a coupling for long-range percolation models with different values of α and β. Define $\mathcal{C}(x;\alpha,\beta)$ as the cluster of vertices that can be reached by paths from vertex x if the parameters are α and β.

Assume that $a := \mathbb{E}_{\alpha,\beta}(|\mathcal{C}(0)|) < \infty$ and take $\varepsilon > 0$ small enough (we will see later exactly how small). Define $\mathcal{A}_0(0) := \mathcal{C}(0;\alpha,\beta)$. In an inductive fashion, let $\mathcal{A}'_{i+1}(0)$ be the set of vertices not in $\cup_{j=0}^{i}\mathcal{A}_j(0)$ that can be reached from $\mathcal{A}_i(0)$ by crossing an edge present for the parameters α and $\beta(1-\varepsilon)$. $\mathcal{A}_i(0)$ is defined by

$$\mathcal{A}_i(0) := \left(\cup_{x\in\mathcal{A}'_i(0)}\mathcal{C}(x;\alpha,\beta)\right) \setminus \cup_{j=0}^{i}\mathcal{A}_j(0)$$

Note that $\mathcal{A}'_i(0) \subset \mathcal{A}_i(0)$ by definition and that by construction we may conclude that $\mathcal{C}(0;\alpha,\beta(1-\varepsilon)) = \cup_{i=1}^{\infty}\mathcal{A}_i(0)$. The next step in the proof is to bound $\mathbb{E}(|\mathcal{A}_i(0)|)$. By definition $\mathbb{E}(|\mathcal{A}_0(0)|) = a$. Since the graph is transitive, for every $x \in \mathcal{A}_i(0)$ the expected size of the set

$$\{y \in \Omega_N \setminus \cup_{j=0}^{i}\mathcal{A}_j(0); U(x,y) < 1-\exp(-\alpha((1-\varepsilon)\beta)^{-d(x,y)})\}$$

is bounded above by

$$\begin{aligned}
b &:= \sum_{y\in\Omega_N}\mathbb{P}\Big(U(0,y) < 1-\exp(-\alpha((1-\varepsilon)\beta)^{-d(0,y)})\Big| U(0,y) > 1-\exp[-\alpha\beta^{-d(0,y)}]\Big)\\
&= \sum_{y\in\Omega_N}\frac{\mathbb{P}(1-\exp(-\alpha\beta^{-d(0,y)}) < U(0,y) < 1-\exp(-\alpha((1-\varepsilon)\beta)^{-d(0,y)})}{\mathbb{P}(1-\exp(-\alpha\beta^{-d(0,y)}) < U(0,y))}\\
&= \sum_{i=1}^{\infty}(N-1)N^{i-1}\frac{\exp(-\alpha\beta^{-i}) - \exp(-\alpha\beta^{-i}(1-\varepsilon)^{-i})}{\exp(-\alpha\beta^{-i})}\\
&= \sum_{i=1}^{\infty}(N-1)N^{i-1}(1-\exp(-\alpha\beta^{-i}((1-\varepsilon)^{-i}-1)))\\
&\leq \sum_{i=1}^{\infty}(N-1)N^{i-1}\alpha\beta^{-i}[(1-\varepsilon)^{-i}-1]\\
&= (N-1)\alpha\frac{1}{(1-\varepsilon)\beta-N} - (N-1)\alpha\frac{1}{\beta-N}\\
&= \frac{\alpha\beta\varepsilon(N-1)}{(\beta(1-\varepsilon)-N)(\beta-N)},
\end{aligned}$$

which converges to 0, if $\varepsilon \searrow 0$. Therefore, we can choose $\varepsilon > 0$ sufficiently small, such that $b < a^{-1}$. Note that $\mathbb{E}(|\mathcal{A}'_{i+1}(0)|) \leq b\mathbb{E}(|\mathcal{A}_i(0)|)$, and because of the transitivity of the graph, $\mathbb{E}(|\mathcal{A}_{i+1}(0)|) \leq a\mathbb{E}(|\mathcal{A}'_{i+1}(0)|)$. So, we have

$$\mathbb{E}(|\mathcal{A}_i(0)|) \leq (ab)^i\mathbb{E}(|\mathcal{A}_0(0)|).$$

Since $\mathcal{C}(0;\alpha,\beta(1-\varepsilon)) = \cup_{i=1}^{\infty}\mathcal{A}_i(0)$ and $ab < 1$, we have

$$\mathbb{E}(\mathcal{C}(0;\alpha,\beta(1-\varepsilon))) \leq \sum_{i=0}^{\infty}a(ab)^i = \frac{a}{1-ab} < \infty.$$

□

Proof of Theorem 1.4. The only things left to prove is that $\alpha_c(\beta) \nearrow \infty$ for $\beta \nearrow N^2$, but this follows immediately from Lemma 6.3, equality (6.1) and observing that $\alpha_c(N^2) = \infty$. □

7 Possible generalizations

Possible generalizations of the model considered in this paper include:

1. *Randomness in the hierarchical lattice.* The hierarchical lattice Ω_N is generated by a N-regular tree. An interesting question is how randomness in the underlying

tree (or induced random metric) affects the percolation process on the resulting lattice. One possibility is that the metric generating tree, is a Galton-Watson tree. Analysis of long-range percolation on such random hierarchical structures is not a trivial extension of the analysis in this paper, since in renormalisation schemes, one has to take care of all kinds of dependencies of the sizes of balls of given diameter.

2. *More general connection function* $p(k)$. In this paper we focused on the connection function $p_k = 1 - \exp\left(-\frac{\alpha}{\beta^k}\right)$. What are necessary and sufficient conditions on $g(k)$ so that when $p_k = 1 - \exp\left(-\alpha g(k)\right)$ we have $0 < \alpha_c < \infty$?

3. *Random cluster models.* We only consider independent percolation on the hierarchical lattice. We did not try to incorporate Random cluster (or Fortuin-Kasteleyn) model [18] yet. Some work has already been done for the Ising model on the hierarchical lattice [19].

Acknowledgments. We thank Odo Diekmann for suggesting to study this percolation model. VK was supported by NWO grant NWO-CLS Project 635.100.002. PT was supported by a grant from Riksbankens jubileumsfond.

References

[1] M. Aizenman and D.J. Barsky. Sharpness of the phase transition in percolation models. *Comm. Math. Phys.*, 108(3):489–526, 1987.

[2] M. Aizenman and C.M. Newman. Discontinuity of the percolation density in one-dimensional $1/|x-y|^2$ percolation models. *Comm. Math. Phys.*, 107(4):611–647, 1986.

[3] S.R. Athreya and J.M. Swart. Survival of contact processes on the hierarchical group. *Probab. Theory Related Fields*, 147(3-4):529–563, 2010.

[4] N. Berger. Transience, recurrence and critical behavior for long-range percolation. *Comm. Math. Phys.*, 226(3):531–558, 2002.

[5] M. Biskup. On the scaling of the chemical distance in long-range percolation models. *Ann. Probab.*, 32(4):2938–2977, 2004.

[6] M. Biskup. Graph diameter in long-range percolation. *Random Structures Algorithms*, 39(2):210–227, 2011.

[7] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms*, 31(1-2):3–122, 2007.

[8] L. Chayes and R.H. Schonmann. Mixed percolation as a bridge between site and bond percolation. *Ann. Appl. Probab.*, 10(4):1182–1196, 2000.

[9] J.T. Cox and R. Durrett. Limit theorems for the spread of epidemics and forest fires. *Stochastic processes and their applications*, 30(2):171–191, 1988.

[10] D.A. Dawson and L.G. Gorostiza. Percolation in an ultrametric space. *ArXiv e-prints*.

[11] D.A. Dawson and L.G. Gorostiza. Percolation in a hierarchical random graph. *Commun. Stoch. Anal.*, 1(1):29–47, 2007.

[12] F.M. Dekking and R.W.J. Meester. On the structure of Mandelbrot's percolation process and other random Cantor sets. *J. Stat. Phys.*, 58:1109–1126, 1990.

[13] O. Diekmann and J.A.P. Heesterbeek. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation.* Wiley, 2000.

[14] N.A. Friedman. *Introduction to ergodic theory.* Van Nostrand Reinhold, New York, 1970.

[15] Nathaniel A. Friedman. Replication and stacking in ergodic theory. *Amer. Math. Monthly*, 99(1):31–41, 1992.

[16] A. Gandolfi, M.S. Keane, and C.M. Newman. Uniqueness of the infinite component in a random graph with applications to percolation and spin glasses. *Probab. Theory Related Fields*, 92(4):511–527, 1992.

[17] G.R. Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1999.

[18] G.R. Grimmett. *The random cluster model*, volume 333 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

[19] T. Hara, T. Hattori, and H. Watanabe. Triviality of hierarchical Ising model in four dimensions. *Commun. Math. Phys.*, 220(1):13–40, 2001.

[20] P. Jagers. *Branching processes with biological applications*. Wiley-Interscience [John Wiley & Sons], London, 1975. Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics.

[21] R. Meester and P. Trapman. Bounding basic characteristics of spatial epidemics with a new percolation model. *Adv. in Appl. Probab.*, 43(2):335–347, 2011.

[22] C.M. Newman and L.S. Schulman. One dimensional $1/|j-i|^s$ percolation models: The existence of a transition for $s \leq 2$. *Communications in Mathematical Physics*, 104(4):547–571, 1986.

[23] L.S. Schulman. Long range percolation in one dimension. *Journal of Physics A: Mathematical and General*, 16:L639–L641, 1983.

[24] P. Trapman. The growth of the infinite long-range percolation cluster. *Ann. Probab.*, 38(4):1583–1608, 2010.

[25] A.C.M. Van Rooij. Non-Archimedean functional analysis, vol. 51 of Monographs and Textbooks in Pure and Applied Math, 1978.