Combinatorics, Probability and Computing (2020), **29**, pp. 587–615
doi:10.1017/S096354832000005X

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Percolation on an infinitely generated group

Agelos Georgakopoulos*,† and John Haslegrave†

Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK
*Corresponding author.

(Received 28 June 2018; revised 28 October 2019; first published online 20 February 2020)

Abstract

We give an example of a long range Bernoulli percolation process on a group non-quasi-isometric with $\mathbb{Z}$, in which clusters are almost surely finite for all values of the parameter. This random graph admits diverse equivalent definitions, and we study their ramifications. We also study its expected size and point out certain phase transitions.

2010 MSC Codes: Primary 60K35; Secondary 05C80, 05C81

1. Introduction

We consider an instance of (long range) Bernoulli percolation on the group $\bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$, providing the first example of Bernoulli percolation that is subcritical for every value of the parameter on a group non-quasi-isometric with $\mathbb{Z}$. We observe that this random graph arises in other contexts, and point out further interesting properties.

Most random graph models studied enjoy some form of invariance. For example, the distribution of the Erdős–Rényi random graph on the vertex set $[n]$ is invariant under permutations of $[n]$, and this is also true in much more general models; see *e.g.* [14]. Percolation theory provides further examples where a random graph is invariant under the action of some group, *e.g.* $\mathbb{Z}^d$, on the set of vertices. In a random geometric graph in the sense of [16], the group does not act on the vertex set directly but on an ambient space in which the vertices live. These models display a 'spatial invariance', but there are also examples of 'temporal invariance': any random (regular) rooted graph arising as a limit of a sequence of finite graphs in the sense of Benjamini and Schramm [4] is invariant under taking a step of simple random walk from the root and then declaring the destination to be the root. Dynamic percolation [18] serves as an example of a model with both spatial and temporal invariance. The random graphs introduced in this paper also enjoy both spatial and temporal invariance, but we need to use two seemingly unrelated – and *a posteriori* equivalent – definitions to see this.

The 'spatial' definition is via percolation on $\Gamma = \bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$, *i.e.* the direct sum of infinitely many copies of the cyclic group with b elements: we join each pair $x, y \in \Gamma$ to a random number of (parallel) edges with distribution $\mathrm{Po}(\lambda b^{1-2h(x,y)})$, where $\lambda \in \mathbb{R}_+$ is the parameter of the model

†Supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 639046).

© Cambridge University Press 2020

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

(proportional to the average degree of a vertex), $h(x, y)$ is the first coordinate at which x, y differ, and $\mathrm{Po}(\mu)$ denotes the Poisson distribution with mean μ. The reader is not yet expected to appreciate why $\lambda b^{1-2h(x,y)}$ was the right choice; for the time being we just note that this random (multi-)graph[1] is invariant under the natural action of Γ on itself, and apart from that it is hard to say anything about it.

The 'temporal' definition of this model is given by the following proposition.

Proposition 1.1. *For every $\lambda \in \mathbb{R}_+$, there is a unique rooted connected random multigraph $(G(\lambda), o)$ with the root having finite expected degree which is invariant under the following operation.*

Replace each vertex v of $G(\lambda)$ (including the root o) with b vertices $v_0, \dots, v_{b-1}$, and for each $i \neq j \in \mathbb{Z}_b$ join v_i to v_j with a random number of edges with distribution $\mathrm{Po}(\lambda/b)$.
Moreover, replace each edge uv of $G(\lambda)$ with one of the b^2 edges $v_i u_j : i, j \in \mathbb{Z}_2$, chosen uniformly at random. All these random experiments are made independently from each other. (1.1)
Choose the root of the resulting graph to be each of $o_0, \dots, o_{b-1}$ with probability $1/b$.
Finally, if the graph is now disconnected, discard all components not containing the root.

It turns out that $(G(\lambda), o)$ has the same distribution as the component (a.k.a. cluster) of the origin in the aforementioned percolation model on Γ. (In fact it is possible to obtain a statement similar to Proposition 1.1 when all components are retained, and the corresponding disconnected random graph has the same distribution as our percolation on all of Γ.) The fact that these two random multigraphs coincide is far from clear at first sight; we prove this by showing (in Section 3) that they both coincide with a third random graph. The vertices of the latter random graph are the leaves of an infinite tree T_∞ (the *canopy tree*, defined in Section 2.1). Following the general construction of *group walk random graphs* (GWRGs) [7], the choice of which pairs to connect with an edge is made using an experiment involving random walks on T_∞. Thus the choice of the coefficients $\lambda b^{1-2h(x,y)}$ above was dictated by the behaviour of random walk. This choice is also unique in that it makes the two aforementioned models coincide, and 'critical' in a sense explained in Section 1.1.

In fact this 'third' definition (given in detail in Section 2) was the starting point of our work. The general motivation is that GWRGs link groups to geometric random graphs, and studying the interactions could be fruitful; we refer the interested reader to [7] for more details on the background of this construction.

Despite having several equivalent definitions, it is very hard to say anything about the structure of $G(\lambda)$. It is not even obvious whether it is finite or infinite, but our first main result (proved in Section 5) implies that it is almost surely finite.

Theorem 1.2. *The expected number of vertices $\chi(\lambda)$ of $G(\lambda)$ satisfies*

$$e^{c\lambda} < \chi(\lambda) < e^{e^{C\lambda}}$$

for some constants $c, C > 0$.

These bounds leave an enormous gap, but it seems to be very hard to improve them significantly. Computer simulations we performed for $b = 2$ and $\lambda \leqslant 12$ suggest that $\chi(\lambda)$ might be of order $\lambda^{c\lambda}$. Conjecturing that $\chi(\lambda) \sim \lambda^{c\lambda}$ led us to wonder whether χ is a continuous/smooth function of λ. By adapting a well-known technique of Kesten [13], the first author and C. Panagiotis [8] proved that $\chi(\lambda)$ is an analytic function at every $\lambda \in \mathbb{R}_+$, and this statement holds in the full generality of all Bernoulli long or short range percolation models on groups.

[1] A *multigraph* is a graph in which we can have several 'parallel' edges between the same pair of vertices.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Thus $G(\lambda)$ displays no phase transitions, at least as far as χ is concerned. Still, we observed some rougher phase transition phenomena. We consider finite versions of our percolation model on $\bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$ obtained, roughly speaking, by restriction to finite subgroups, and determine the threshold value of λ for obtaining a connected graph. Alternatively, these finite versions of our model can be described by replacing the canopy tree in our other definition with the full binary tree of depth $n \in \mathbb{N}$. We prove that there is a phase transition for connectedness, occurring at a sharp threshold λ_{conn} logarithmic in the size of the graph, while the transition occurs in a window of width proportional to the logarithm of λ_{conn} (Section 7). We remark that this restriction on finite subgroups of $\bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$ is somewhat related to percolation on Hamming hypercubes (which are Cayley graphs of such subgroups), which has attracted a lot of interest recently; see [10, 11] and references therein.

1.1 Percolation on groups

A well-known conjecture of Benjamini and Schramm [3], recently proved by Duminil-Copin, Goswami, Raoufi, Severo and Yadin [5], states that $p_c < 1$ holds for Bernoulli percolation on every Cayley graph of a group which is not a finite extension of $\mathbb{Z}$. Our result that $G(\lambda)$ is almost surely finite for every λ means that the analogue of this statement for long range percolation on infinitely generated groups is false. To explain what we mean by long range percolation, let μ be a probability measure on a (countable) group Γ. We say that μ is a *generating measure* of Γ if the support of μ generates Γ, and μ is symmetric, *i.e.* $\mu(g) = \mu(g^{-1})$ for every $g \in \Gamma$.

Every generating measure μ naturally defines a percolation process on Γ as follows. Given $\lambda \in \mathbb{R}_+$, we define a random (multi-)graph $\Gamma_\mu(\lambda)$ with vertex set Γ, by letting the number of (parallel) edges between two elements $g, h \in \Gamma$ be an independent Poisson random variable with mean $\lambda\mu(g^{-1}h)$ (we may as well remove any parallel edges to obtain a simple graph). Such models were already considered in [1], for example.

Note that $\Gamma_\mu(\lambda)$ is a Γ-invariant percolation model, that is, the natural action of Γ on $\Gamma_\mu(\lambda)$ defined by multiplication from the left preserves the probability distribution of $\Gamma_\mu(\lambda)$.

Similarly to the standard percolation threshold p_c, we define

$$\lambda_c = \lambda_c(\mu) := \sup\{\lambda \mid \mathbb{P}(\Gamma_\mu(\lambda) \text{ has an infinite component}) = 0\}.$$

We remark that λ_c may be infinite, as is the case with our $G(\lambda)$. Another result of this paper implies, however, that $\lambda_c < \infty$ for other choices of μ on the same group $\Gamma = \bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$: suppose $\mu(g)$ is proportional to $\alpha^{-\ell(g)}$, where $\ell(g)$ denotes the index of the last non-zero coordinate of $g \in \bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$. Then for $\alpha = b^2$ we obtain $G(\lambda)$ as the component of the origin by the definitions. Moreover, we prove that this α is 'critical' in the sense that percolation does occur – for large enough λ – for any $\alpha \in (b, b^2)$, but not for $\alpha \geqslant b^2$ (Section 4).

It is interesting to compare this fact with long range percolation on $\Gamma := \mathbb{Z}$, which is the most studied instance of this model. Let $\mu(i) = c|i|^{-s}$, with $s \in (1, \infty)$, and c a suitable normalizing constant that matters little. It has been proved that for $s > 2$ we have no percolation, *i.e.* $\lambda_c = \infty$ [17], while for $s \in (1, 2]$ we have $\lambda_c < \infty$ [15]. The case $s = 2$ is of particular interest, and is considered to be the 'critical' case. Indeed, when $s = 2$, the percolation density is discontinuous at λ_c [2].

Interestingly, this case is related to our critical case $\alpha = b^2$ in Example 3 above. Indeed, if we enumerate the elements of $\bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$ appropriately, namely by thinking of the elements of $\bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$ as natural numbers expressed in base b, then for pairs of 'numbers' x, y far apart, the probability of joining x to y with an edge decays like $|x - y|^{-2}$ in both models. (But when $|x - y|$ is small, then in our example this probability can be much smaller than the corresponding probability for $\mathbb{Z}$.)

But perhaps a more interesting connection is that, as mentioned in [7], the critical ($s = 2$) case for $\Gamma := \mathbb{Z}$ can be obtained as a special case of GWRG, just like the critical ($\alpha = b^2$) case for $\Gamma := \bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$. This raises the question of whether there is a general method for finding critical generating measures for other groups, which we hope to explore in future work.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

1.2 Outline of the paper

This paper is structured as follows. In Section 2 we make precise the definition of our model via the canopy tree, and the closely related model arising from GWRGs. In Section 3 we prove Proposition 1.1 by showing that the cluster of the origin obtained by iterating (1.1) converges in distribution to that of the model defined in Section 2.

In Section 4 we analyse the model of Section 1.1 more generally, showing that the exponent $\alpha = b^2$ is critical for percolation to occur (Theorem 4.1); we also give more precise bounds on the critical window (Theorem 4.2). In Section 5 we give the lower and upper bounds required for Theorem 1.2.

In Section 7 we establish sharp thresholds for connectedness in the finite versions of the models of Section 2 (Theorems 7.5 and 7.7). In one version the threshold for connectedness coincides with the threshold at which no isolated vertices remain, but, perhaps surprisingly, this is not the case for the model arising from a GWRG.

2. Random graph models based on random walks on trees

In this section we provide the third alternative definition of the random graph $G(\lambda)$ from the Introduction. This definition, called the 'Poisson edge model', will be useful in our proof of Proposition 1.1 in Section 3. This Poisson edge model is closely related to an instance of 'group-walk random graphs' as introduced in [7], which we define below as the 'Poisson particle model'.

2.1 Some notation for trees

Fix an integer $b \geqslant 2$. We inductively define the b-ary tree T_h of height h as follows. Let T_0 be the one-vertex tree with single vertex v_0. For each $h \geqslant 0$, T_{h+1} is the graph obtained from T_h by adding $b-1$ additional copies of T_h and a new vertex v_{h+1}, and adding edges between v_{h+1} and the vertices of degree b (or 0) in our b copies of T_h. Note that T_h has b^h leaves and $(b^{h+1}-1)/(b-1)$ vertices. Define the (b-ary) *canopy tree* T_∞, as $T_\infty = \bigcup_h T_h$, where we think of T_h as a subtree of T_{h+1}. For $h \in \mathbb{N} \cup \{\infty\}$, write L_h for the set of leaves of T_h.

Note that, for each $h > 0$, removing the vertex v_h divides T_∞ into b finite components and one infinite component, which contains v_{h+1}. Therefore T_∞ contains a unique infinite path starting at any vertex. Moreover, the group $\Gamma = \bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$ from the Introduction can be realized as a subgroup of the automorphism group of T_∞ acting transitively and faithfully on the set of leaves L_∞ of T_∞. Indeed, if we label the edges of T_∞ with $0, 1, \ldots, b-1$, in such a way that for each non-leaf v, each label appears exactly once among the offspring of v, and all edges along the unique infinite path from v_0 are labelled 0, then every element g of Γ can be identified with the unique leaf v_g in L_∞ such that the sequence of labels along the unique infinite path in T_∞ starting at v_g coincides with the sequence g. With this identification, multiplication by an element h of Γ defines an automorphism of T_∞.

We define the *height* $h(v)$ of a vertex v in T_∞ (or T_h) as the distance to the nearest leaf. The *apex* of T_h is the unique vertex of height h. Each vertex in T_∞ has a unique higher neighbour, its *parent*. We say that two distinct vertices are *siblings* if they have the same parent. x is a *descendant* of y if there is a path $x \cdots y$ in which each vertex except the first is the parent of the previous one; we include the possibility of a path of length 0, so that x is a descendant of itself. We say that x is an *ancestor* of y if y is a descendant of x. Write $D(v)$ for the descendants of v.

2.2 The two models and their relationship

We define a random multigraph $G_n(\lambda)$ for every $n \in \mathbb{N} \cup \{\infty\}$ and $\lambda \in \mathbb{R}_+$ as follows. The vertex set of $G_n(\lambda)$ is the set L_n of leaves of T_n. For every pair $x, y \in L_n$, the number of x–y edges in $G_n(\lambda)$

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

is a random variable with distribution $\mathrm{Po}(\lambda b^{1-d(x,y)})$, where $d(x, y)$ is the distance between x, y in T_n. Note that in the case $n = \infty$ we have $\sum_{y \neq x} b^{1-d(x,y)} = 1$. Since L_∞ can be identified with the group $\bigoplus_{i \in \mathbb{N}} \mathbb{Z}_b$ (see the remark above), $G_\infty(\lambda)$ can be obtained as a special case of our long range percolation model from Section 1.1. We call $G_n(\lambda)$ the *Poisson edge model*. If no n is specified in the context, then the term Poisson edge model will refer to $G_\infty(\lambda)$.

Next, we describe an instance of the random graph model of [7], which will turn out to be very similar to the Poisson edge model.

Define a random multigraph $G_n^\bullet(\lambda)$ for every $n \in \mathbb{N} \cup \{\infty\}$ and $\lambda \in \mathbb{R}_+$ as follows. The vertex set of $G_n^\bullet(\lambda)$ is again the set L_n of leaves of T_n. The edge set of $G_n^\bullet(\lambda)$ is determined by the following random experiment. At each vertex $v \in L_n$, we start a number of particles, and these numbers are independent and identically distributed random variables with distribution $\mathrm{Po}(\lambda/2)$. These particles perform simple random walks on T_n, and are stopped upon their first return to L_n. For each of these particles p, we put an edge e_p in $G_n^\bullet(\lambda)$ connecting the vertex at which p was started to the last vertex of its random walk. We call this random multigraph $G_n^\bullet(\lambda)$ the *Poisson particle model*. (We choose $\lambda/2$ as the mean of the number of particles started at v so that the number of edges at v, *i.e.* the number of particles starting or ending at v, has mean λ.)

It is possible to show that $G_n(\lambda)$ converges in distribution to $G_\infty(\lambda)$ and likewise for $G_n^\bullet(\lambda)$ and $G_\infty^\bullet(\lambda)$;[2] we will not use or prove this fact directly, but a lot of the intuition underlying this paper was based on it.

It is not too hard to see that the expected number of x–y edges in the Poisson particle model decays like $\lambda b^{1-d(x,y)}$; in Section 6 we provide some precise calculations. This implies that there are constants $0 < c < C$ such that the Poisson edge model with parameter λ lies between $G_\infty^\bullet(c\lambda)$ and $G_\infty^\bullet(C\lambda)$ for every λ (in the sense that the three processes may be coupled so that their edge sets are nested as $E(G_\infty^\bullet(c\lambda)) \subseteq E(G_\infty(\lambda)) \subseteq E(G_\infty^\bullet(C\lambda))$); see Lemma 6.3 and the remark after it.

3. Proof of Proposition 1.1

For $n \in \mathbb{N} \cup \{\infty\}$, let $C_n(\lambda)$ denote the component of v_0 – the unique vertex of T_0 – in the Poisson edge model $G_n(\lambda)$. Since the automorphism group of T_n acts transitively on its leaves and this action preserves the probability distributions of edges between pairs of vertices, the component of any fixed vertex has the same distribution, up to isomorphism, as $C_n(\lambda)$.

First we show that $C_\infty(\lambda)$ is almost surely finite,[3] by showing that there is almost surely some $n < \infty$ such that $V(C_\infty(\lambda)) \subseteq L_n$, where as in Section 2.1 L_n denotes the set of leaves of T_n, which we think of as a subset of the leaves of T_{n+1}, and therefore of T_∞.

Lemma 3.1. *For any $m > n$ (including the case $m = \infty$) we have $\mathbb{P}(V(C_m(\lambda)) \subseteq L_n) > 1 - (1 - \mathrm{e}^{-\lambda}/b)^n$.*

Proof. It is sufficient to prove the case $m = \infty$, since we may obtain $G_m(\lambda)$ as an induced subgraph of $G_\infty(\lambda)$.

For each $i \in \mathbb{N}$, let e_i be the edge from T_i to $T_\infty \setminus T_i$. For each i write E_i for the event that at least one edge of $G_\infty(\lambda)$ starts and finishes on different sides of e_i, *i.e.* starts in L_i and finishes outside it or *vice versa*. It suffices to show that

$$\mathbb{P}\left(\bigcap_{i=1}^{n} E_i\right) < (1 - \mathrm{e}^{-\lambda}/b)^n,$$

since whenever $V(C_m(\lambda)) \not\subseteq L_n$, the events $E_1, \ldots, E_n$ must all hold.

[2] We thank Gourab Ray for this observation.
[3] We thank Omer Angel for this observation.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Note that $\mathbb{P}(E_i) = (1 - \mathrm{e}^{-\lambda})$. We claim that, given $\bigcup_{j \leqslant i} E_j$, the probability that every edge which crosses e_i also crosses e_{i+1} is less than $1/b$. The number of edges crossing e_i but not e_{i+1} is given by a Poisson random variable of mean $\lambda(b-1)/b$, and the number of edges crossing both e_i and e_{i+1} by an independent Poisson random variable of mean λ/b. Thus, conditional on there being k edges which cross e_i, the number of edges crossing both e_i and e_{i+1} has distribution Bin $(k, 1/b)$. The probability that every edge which crosses e_i also crosses e_{i+1} is therefore b^{-k}; conditional on $k \geqslant 1$ this is at most $1/b$.

Now we proceed as follows. Set $k_0 = 1$. For each i, we look sequentially, first for edges which cross e_{k_i} but not e_{k_i+1}, then for edges which cross e_{k_i}, e_{k_i+1} but not e_{k_i+2}, and so on. If we eventually find such an edge, let $e_{k_{i+1}}$ be the first edge it does not cross and move on to $i+1$; otherwise stop the process and set $k = k_i$. Note that the event $E_{k_{i+1}}$ is independent of the information we have after finding an edge which crosses e_{k_i}, since edges between different pairs of vertices occur independently. Now E_k is the first event which does not occur, and k is dominated by the variable $X = \sum_{i=1}^{Y} Z_i$, where $Y \sim \mathrm{Geo}\,(\mathrm{e}^{-\lambda})$ and $Z_i \sim \mathrm{Geo}\,(1/b)$ are independent.[4] Writing $p = \mathrm{e}^{-\lambda}$ and $q = 1/b$, the generating function of Y is $f_Y(s) = ps/(1-(1-p)s)$, and that of each Z_i is $f_Z(s) = qs/(1-(1-q)s)$. Then

$$\begin{aligned} f_X(s) &= f_Y(f_Z(s)) \\ &= p\frac{qs}{1-(1-q)s} \cdot \frac{1}{1-(1-p)\frac{qs}{1-(1-q)s}} \\ &= \frac{pqs}{1-(1-q)s-(1-p)qs} \\ &= \frac{pqs}{1-(1-pq)s}, \end{aligned}$$

so $X \sim \mathrm{Geo}\,(pq)$, giving the required result. □

Remark. It follows that $C_n(\lambda)$ converges in distribution to $C_\infty(\lambda)$, since there is a natural coupling for which $C_n(\lambda)$ is always a subgraph of $C_\infty(\lambda)$, and for which Lemma 3.1 gives almost sure convergence.

We now prove Proposition 1.1 by showing that $C_\infty(\lambda)$ is the unique random multigraph with finite average degree invariant under (1.1).

For this, note that if we add an extra layer L'_∞ to T_∞ by attaching b new leaves $u_0, \dots, u_{b-1}$ to each leaf u of T_∞, then the resulting tree T'_∞ is isomorphic to T_∞. Therefore, if we repeat the definition of $G_\infty(\lambda)$ using T'_∞ instead of T_∞, we obtain a random multigraph $G'_\infty(\lambda)$ which is identically distributed with $G_\infty(\lambda)$. Moreover, it is straightforward to check that if we apply operation (1.1) to $G_\infty(\lambda)$ we obtain a realization of $G'_\infty(\lambda)$, because of the choice of the rates $b^{1-2h}\lambda$ at which edges appear. It follows that if we let $G(\lambda)$ denote $C_\infty(\lambda)$, then $G(\lambda)$ is indeed invariant under (1.1), as (1.1) can be thought of as choosing the component of o_0 in $G'_\infty(\lambda)$, which is identically distributed with the component of o in $G_\infty(\lambda)$.

To prove the uniqueness of $G(\lambda)$, let (X, o) be another random rooted multigraph – possibly depending on λ – with these properties. Let δ denote the expected degree of o in X, and recall that we are assuming that δ is finite. This means that with positive probability q, the root will become isolated if we perform operation (1.1) on (X, o); indeed, the root becomes isolated whenever all edges of o are inherited by $b-1$ of its offspring, no new edges are formed between these offspring and the remaining one o_i, and we choose o_i as the new root. Moreover, if we perform

[4] We adopt the convention that a geometric random variable represents the number of trials up to and including the first success.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

(1.1) once more on the resulting graph, then the probability of obtaining an isolated root is still q. As the choices we make each time we apply (1.1) are independent from what happened in earlier applications, it follows that for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that if we perform (1.1) N times on (X, o), then the probability of obtaining an isolated root at least once is at least $1 - \varepsilon$. Note that if the root is an isolated vertex, then performing (1.1) M times on it yields a random graph with the law of the component $C_M(\lambda)$. But as $C_M(\lambda)$ converges in distribution to $C_\infty(\lambda)$, and the distribution of (X, o) is preserved after performing (1.1) any number of times, it follows that this distribution coincides with $C_\infty(\lambda)$. This proves Proposition 1.1.

We have just proved that $G(\lambda)$ coincides with $G_\infty(\lambda)$. Since the leaves L_∞ of T_∞ can be identified with the elements of the group $\bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$ as remarked in Section 2.1, it follows that the Poisson edge model $G_\infty(\lambda)$ coincides with the percolation model on $\bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$ defined in the Introduction, and hence also with $G(\lambda)$ as claimed there.

We next show that the assumption that the root has finite expected degree is necessary for Proposition 1.1 to hold: without this restriction $G(\lambda)$ is not unique.

Proposition 3.2. *There is a random connected rooted multigraph with infinite degrees which is invariant under* (1.1).

Proof. Let Γ^* be the group whose elements are the two-way infinite sequences $(x^{(i)})_{i\in\mathbb{Z}}$ over $\mathbb{Z}_b$ such that $\{i\colon x^{(i)} = 1\}$ is finite and $x^{(-2i)} = 0$ for every $i > 0$; the group operation is componentwise addition in $\mathbb{Z}_b$. Let $\mu^*(x) = b^{1-\ell'(x)}$, where $\ell'(x) = \min\{i\colon x^{(i)} = 1\}$. Let G be the random graph obtained by applying (1.1) to $\Gamma^*_{\mu^*}(\lambda)$, where the vertices of G are $\{x_0, \dots, x_{b-1} \mid x \in \Gamma^*\}$; note that G is almost surely connected. We may define a bijection between the vertices of G and of $\Gamma^*_{\mu^*}(\lambda)$ as follows: $x_i \mapsto y$ where $y^{(0)} = i$ and $y^{(j)} = x^{j-2}$ for $j \neq 0$. It is easy to check that this bijection preserves the rates of all edges, and hence $\Gamma^*_{\mu^*}(\lambda)$ is also invariant under (1.1). □

4. Percolation on $\bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$

The group $G := \bigoplus_{i\in\mathbb{N}} \mathbb{Z}_b$ consists of all sequences of elements of $\mathbb{Z}_b$ which have only finitely many non-zero terms, endowed with the operation of componentwise addition, where $\mathbb{Z}_b = \mathbb{Z}/b\mathbb{Z}$ is the cyclic group of order b (the reader will lose nothing by assuming that $b = 2$ throughout this section).

In this section we study the question of whether percolation occurs in $G_\mu(\lambda)$ for large enough λ for various generating measures μ on G. More precisely, the aim of this section is to determine the critical asymptotic decay for μ that separates the $\lambda_c = \infty$ from the $\lambda_c < \infty$ regime. We shall sometimes neglect to normalize μ to be a probability measure; this does not affect the results of this section since normalizing μ is equivalent to rescaling λ. We use standard asymptotic notation: if $f, g\colon [0,\infty) \to [0,\infty)$ we write

$$f(x) = \begin{cases} o(g(x)) & \text{if } \lim_{x\to\infty} f(x)/g(x) = 0, \\ O(g(x)) & \text{if } \limsup_{x\to\infty} f(x)/g(x) < \infty, \\ \Omega(g(x)) & \text{if } \liminf_{x\to\infty} f(x)/g(x) > 0, \\ \Theta(g(x)) & \text{if } f(x) = O(g(x)) \text{ and } f(x) = \Omega(g(x)). \end{cases}$$

For $y \in G$, write $\ell(y)$ for the position of the last non-zero term of y (we take $\ell(\mathbf{0}) = 0$). We will concentrate on generating measures $\mu(y)$ that depend on $\ell(y)$ only, and are monotone decreasing in $\ell(y)$: given a real number $\alpha > b$, we define a generating measure μ_α by letting $\mu(y) = \mu_\alpha(y) := \alpha^{-\ell(y)}$. The reason why we do not consider $\alpha \leqslant b$ is that μ_α fails to be a finite measure in that case, since G has b^{k-1} elements y with $\ell(y) = k$.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

We consider the percolation process $G_{\mu_\alpha}(\lambda)$ as defined in Section 1.1. In this case percolation does not occur for overly small λ, since an exploration of the component of the identity is dominated by a subcritical Galton–Watson tree.

The Poisson edge model of Section 2.2 is obtained by taking $\alpha = b^2$, and by Lemma 3.1 there is no percolation, *i.e.* $\lambda_c = \infty$, in this case. Easily, the same arguments imply that $\lambda_c = \infty$ when $\alpha > b^2$ as well. Our next result shows that this value $\alpha = b^2$ is in a sense critical.

Theorem 4.1. *For $\alpha \in (b, \infty)$, the percolation model $G_{\mu_\alpha}(\lambda)$ satisfies $\lambda_c < \infty$ if $\alpha < b^2$ and $\lambda_c = \infty$ if $\alpha \geqslant b^2$.*

Proof. The second statement follows from Lemma 3.1, and so we may assume $\alpha < b^2$. Write V_k for $\{x \in G \mid \ell(x) \leqslant k\}$, and X_k for the component of the identity in the subgraph of $G_{\mu_\alpha}(\lambda)$ spanned by V_k. Fix a real β with $\alpha/b < \beta < b$. We aim to show that for large enough λ we have $\mathbb{P}(|X_k| \geqslant \beta^k \text{ for every } k) > 0$. In fact, we shall show that $\mathbb{P}(|X_{k_i}| \geqslant k_i\beta^{k_i} \text{ for every } i) > p$ for some $p > 0$ and a strictly increasing sequence $(k_i)_{i\geqslant 0}$ chosen so that

$$\beta^{k_{i+1}} \leqslant k_i\beta^{k_i}. \tag{4.1}$$

This is sufficient since if $k_i < k < k_{i+1}$ then $|X_k| \geqslant |X_{k_i}| \geqslant k_i\beta^{k_i} \geqslant \beta^{k_{i+1}} > \beta^k$. We defer the choice of k_0 until later, but given k_0, we will choose $(k_i)_{i\geqslant 1}$ to be as large as possible given (4.1); this will mean that $k_{i+1} = \lfloor k_i + \log_\beta k_i \rfloor$.

Write p_i for $\mathbb{P}(|X_{k_j}| \geqslant k_j\beta^{k_j} \text{ for every } j \leqslant i)$. Now $V_{k_{i+1}}$ consists of $b^{k_{i+1}-k_i}$ cosets of V_{k_i}, and $b^{k_{i+1}-k_i} \approx b^{\log_\beta k_i} = k_i^{\log_\beta b}$; note that $\log_\beta b > 1$.

The number of edges from X_{k_i} to any given coset xV_{k_i}, where $k_i < \ell(x) \leqslant k_{i+1}$, is given by a Poisson random variable with mean at least $\lambda|X_{k_i}|b^{k_i}\alpha^{-k_{i+1}}$. So if $|X_{k_i}| \geqslant k_i\beta^{k_i} \geqslant \beta^{k_{i+1}}$, then the probability that there is no such edge is at most

$$q_i := \exp\left(-\lambda \frac{b\beta}{\alpha}^{k_{i+1}} b^{k_i - k_{i+1}}\right).$$

Call a coset xV_{k_i} *good* if there is an edge from X_{k_i} to a component of size at least $k_i\beta^{k_i}$ in the graph restricted to xV_{k_i}. Thus each coset independently has probability at least $(1-q_i)p_i$ of being good, and $(1-q_i)p_i k_i^{\log_\beta b} > 2k_{i+1}$ provided $p_i > p$ and k_0 is sufficiently large. Consequently, by the multiplicative Chernoff bound the probability of at least k_{i+1} cosets being good is at least $1 - c^{k_{i+1}}$ for some absolute constant $c < 1$. This event is sufficient to imply that $|X_{k_{i+1}}| \geqslant k_{i+1}\beta^{k_{i+1}}$, so

$$p_{i+1} \geqslant p_i(1 - c^{k_{i+1}}) \quad \text{for each } i \geqslant 0,$$

and hence

$$p_i \geqslant p_0 \prod_{j=1}^{i} (1 - c^{k_j}) \geqslant p_0 \prod_{k=k_0}^{\infty} (1 - c^k).$$

We may choose k_0 sufficiently large that $\prod_{k=k_0}^{\infty}(1-c^k) > 1 - p$, and then if p_0 is sufficiently close to 1 this will imply $p_i > p$ for all i. But by choosing λ appropriately we may ensure p_0 is large enough. □

Thus we have proved that for generating measures of the form $\mu(y) = \alpha^{-\ell(y)}$, the value $\alpha = b^2$ is critical for the occurrence of percolation for large enough λ. Our next result 'zooms into' the critical case $\alpha = b^2$ by considering measures $\mu(y)$ that decay as $(b^2 - o(1))^{-\ell(y)}$. We give a class of such measures for which percolation does occur. Our previous proof of non-percolation in the critical case was based on the existence of arbitrarily large sets of constant 'boundary', *i.e.* finite

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

subsets S_k of G, namely those of the form $S_k := \{y \mid \ell(y) \leqslant k\}$, such that the total μ-measure of edges with exactly one endvertex in S_k is bounded above. We will show that in some cases there is no percolation even though the minimum size of the boundary of a set of size n tends to infinity with n.

Theorem 4.2. *Let $\mu(y) = b^{-2\ell(y)} f(\ell(y))$, where $f(\ell)$ is an increasing function. Then the percolation model $G_\mu(\lambda)$ satisfies*

(a) $\lambda_c < \infty$ *if* $f(\ell) = \Omega(a^{\sqrt{\ell}})$ *for some* $a > 0$,
(b) $\lambda_c = \infty$ *if* $f(\ell) = o(\log \ell)$ *and* $f(\ell)$ *is ultimately concave.*

Remark. In particular, if

$$f(\ell) = \frac{\log \ell}{\log \log \ell}$$

then percolation does not occur, even though large sets of constant boundary do not exist.

Proof. For (a), we follow a similar argument to Theorem 4.1. Define a sequence $(k_i)_{i \geqslant 0}$ where k_0 is to be chosen later, and for each $i \geqslant 0$ we choose k_{i+1} as large as possible so that $k_{i+1} - k_i \leqslant (\log_b a)\sqrt{k_{i+1}}/4$ (we will choose k_0 sufficiently large that $k_{i+1} > k_i$ for each i). We aim to show that $\mathbb{P}(|X_{k_i}| \geqslant b^{k_i} a^{-\sqrt{k_i}/4} \text{ for every } i) > p$ for some fixed $p > 0$. Provided $X_{k_i} \geqslant b^{k_i} a^{-\sqrt{k_i}/4}$, we again consider the cosets xV_{k_i} contained in $V_{k_{i+1}}$. Call xV_{k_i} *good* if there is an edge from X_{k_i} to a component of size at least $b^{k_i} a^{-\sqrt{k_i}/4}$ in the graph restricted to xV_{k_i}. Provided there are at least $a^{\sqrt{k_i}/4}$ good cosets, we must have $|X_{k_{i+1}}| \geqslant b^{k_i} \geqslant b^{k_{i+1}} a^{-\sqrt{k_{i+1}}/4}$.

Again, provided $X_{k_i} \geqslant b^{k_i} a^{-\sqrt{k_i}/4}$, the number of edges from X_{k_i} to any given coset of V_{k_i} within $V_{k_{i+1}}$ is given by a Poisson random variable with mean at least

$$\lambda b^{k_i} a^{-\sqrt{k_i}/4} b^{k_i} a^{\sqrt{k_{i+1}}} b^{-2k_{i+1}} > \lambda a^{3asqrtk_{i+1}/4};$$

again we write $q_i = \exp(-\lambda a^{3\sqrt{k_{i+1}}/4})$ and $p_i = \mathbb{P}(|X_{k_j}| \geqslant b^{k_j} a^{-\sqrt{k_j}/4} \text{ for every } j < i)$. Then the expected number of good cosets is at least $(1 - q_i) p_i b^{k_{i+1} - k_i}$.

Since $\lceil (\sqrt{k_i} + \log_b a/10)^2 \rceil < k_i + a\sqrt{k_i}/4$ provided k_0 (and hence k_i) is sufficiently large, by a suitable choice of k_0 we may ensure that $\sqrt{k_{i+1}} > \sqrt{k_i} + \log_b a/10$ for all i. We may also choose k_0 and p large enough that $(1 - q_i) p a^{1/10} > 1 + \varepsilon$ for all i, for some $\varepsilon > 0$. Then by the Chernoff bound the probability of at least $a^{\sqrt{k_i}/4}$ good cosets exceeds $1 - \exp(-\delta b^{k_{i+1} - k_i})$ for some $\delta > 0$. As before, we can find some sufficiently large k_0 and some p' sufficiently close to 1 that $p_0 > p'$ will ensure $p_i > p$ for each i. Since $p_0 > p'$ for some sufficiently large λ, this completes the proof of (a).

For (b), we use a similar approach to Lemma 3.1. In order for the component of the identity to be infinite, there must be an edge crossing e_k for every k; write E_k for the event that e_k is crossed. The total rate at which edges crossing e_k appear is

$$(b-1) \sum_{h \geqslant k} b^{k-h} f(h) = b \sum_{j \geqslant 1} (1 - 1/b)(1/b)^{j-1} f(k + j - 1) = b\mathbb{E}(f(k + J - 1)),$$

where J is a geometric random variable with mean $b/(b-1)$. Since f is ultimately concave, this is at most $bf(k+1)$, and so $\mathbb{P}(E_k) \leqslant 1 - e^{-\lambda b f(k+1)}$, for every sufficiently large k.

Define a random sequence k_i where k_0 is a constant to be chosen later. For each i, we look sequentially, first for edges which cross e_{k_i} but not e_{k_i+1}, then for edges which cross e_{k_i}, e_{k_i+1} but not e_{k_i+2}, and so on. If we eventually find such an edge, let $e_{k_{i+1}}$ be the first edge it does not cross and move on to $i+1$; otherwise stop the process. The event $E_{k_{i+1}}$ only depends on the presence

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

or absence of edges which have not yet been revealed, so it is independent of the previous process. Also, the rate at which edges crossing $e_{k_i}, \dots, e_{k_i+h}$ but not e_{k_i+h+1} appear is $(b-1)b^{-h}f(k_i+h)$. Since f is ultimately concave and increasing, $f(k+1) = (1+o(1))f(k)$, so provided k_0 is sufficiently large, edges crossing $e_{k_i}, \dots, e_{k_i+h}$ but not e_{k_i+h+1} appear at least $3/2$ times as often as edges that cross $e_{k_i}, \dots, e_{k_i+h+1}$ but not e_{k_i+h+2} (any constant between 1 and b would do). Thus $\mathbb{P}(k_{i+1} = k_i + h + 1) \geqslant 3\mathbb{P}(k_{i+1} = k_i + h + 2)/2$, and, conditional on the existence of k_{i+1}, we can bound the distribution of $k_{i+1} - k_i$ by a geometric random variable with fixed mean m. Thus the probability that k_i exists and exceeds $k_0 + 2im$ is bounded by the probability that a negative binomial random variable exceeds twice its mean, which is at most c^i for some constant $c < 1$ by the Chernoff bound. Consequently,

$$\mathbb{P}\left(\bigcup_{j\leqslant i} E_{k_j} \wedge (k_i > k_0 + 2im)\right) \leqslant c^i$$

and

$$\mathbb{P}\left(\bigcup_{j\leqslant i} E_{k_j} \wedge (k_i \leqslant k_0 + 2im)\right) \leqslant \mathbb{P}\left(\bigcup_{j<i} E_{k_j}\right)(1 - \exp(-\lambda bf(k_0 + 2im + 1)),$$

so, writing

$$P_i = \mathbb{P}\left(\bigcup_{j\leqslant i} E_{k_j}\right),$$

we have

$$P_i \leqslant (1 - \exp(-\lambda bf(k_0 + 2im + 1)))P_{i-1} + c^i.$$

Since $f(k) = o(\log k)$, and b and λ are fixed, if i is sufficiently large, $\exp(-\lambda bf(k_0 + 2im + 1)) > (k_0 + 2im + 1)^{-1}$. Thus $\sum_{i\geqslant 1} \exp(-\lambda bf(k_0 + 2im + 1)) = \infty$, and hence Lemma 4.3 below gives the required result. □

Lemma 4.3. *Let $(x_i)_{i\geqslant 0}$ and $(y_i)_{i\geqslant 1}$ be sequences of positive real numbers with $y_i < 1$ for each i. Define $(z_i)_{i\geqslant 0}$ as follows: $z_0 = x_0$ and $z_i = x_i + (1-y_i)z_{i-1}$ for $i > 0$. Then, provided $\sum_i x_i < \infty$ and $\sum_i y_i = \infty$, we have $\lim_{i\to\infty} z_i = 0$.*

Proof. Fix $\varepsilon > 0$. For some N, $\sum_{i\geqslant N} x_i < \varepsilon/2$. For each $i < N$, choose n_i sufficiently large that $x_i \prod_{j=i+1}^{n_i} (1-y_j) < \varepsilon/2N$; this is possible since $\prod_{j>i}(1-y_j) = 0$. Now if $n \geqslant \max\{N, \max_i\{n_i\}\}$, then

$$\begin{aligned} z_n &= \sum_{i\leqslant n}\left(x_i \prod_{j=i+1}^{n} (1-y_j)\right) \\ &\leqslant \sum_{i<N}\left(x_i \prod_{j=i+1}^{n_j} (1-y_j)\right) + \sum_{i\geqslant N} x_i \\ &< \varepsilon. \end{aligned}$$

□

5. Expected size of $G(\lambda)$

The aim of this section is to prove Theorem 1.2. Here we will prove the corresponding bounds for $C_\infty(\lambda)$, the component of v_0 in the Poisson edge model $G_\infty(\lambda)$ defined in Section 2.2, which we proved in Section 3 coincides with $G(\lambda)$.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

5.1 Lower bound

The lower bound of Theorem 1.2 is given by the following result.

Proposition 5.1. *There exists $k > 0$ such that $\mathbb{E}(|C_\infty(\lambda)|) = \Omega(e^{k\lambda})$ as $\lambda \to \infty$.*

In order to prove Proposition 5.1 we will need the following result on coupled variables with identical distributions.

Lemma 5.2. *Suppose X and Y are identically distributed variables taking values in $\mathbb{Z}^+$, coupled such that $\mathbb{P}(X = bY) \geqslant 1 - p$. Then $\mathbb{E}(X) \geqslant f(p)$, where*

$$f(1/n) = \frac{b^n - 1}{n(b-1)} \quad \text{for } n \in \mathbb{Z}^+$$

and $f(p)$ is linear between such points, so

$$f(p) \geqslant p\frac{b^{1/p} - 1}{b-1}.$$

For ease of reading we proceed directly with the proof of Proposition 5.1, proving Lemma 5.2 afterwards.

Proof of Proposition 5.1. We will prove that

$$\mathbb{E}(|C_\infty(\lambda)|) \geqslant \frac{(e^{(b-1)\lambda/b} - 1)b\log b}{(b-1)^2\lambda},$$

for λ sufficiently large.

As in Section 3, applying operation (1.1) to $G_\infty(\lambda)$ (without discarding other components) gives an identically distributed random graph $G'_\infty(\lambda)$, and the components of the roots $C_\infty(\lambda)$ and $C'_\infty(\lambda)$ are likewise identically distributed.

For $v \in B_1$, let $F(v)$ be the event that the subgraph of $G'_\infty(\lambda)$ induced by the children of v is disconnected. For $b = 2$ this is just the event that there are no edges between the two children of v, so has probability $e^{-\lambda/b}$. If $b > 2$ we may bound $\mathbb{P}(F(v))$ by the probability of the event that some child of v has no edges to the other children of v. By Bonferroni's inequality this occurs with probability at least $b\,e^{-(b-1)\lambda/b} - \binom{b}{2} e^{-(2b-3)\lambda/b} = (b - o(1))b\,e^{-(b-1)\lambda/b}$ as $\lambda \to \infty$ (since $2b - 3 > b - 1$). Thus, for any $b \geqslant 2$ and λ sufficiently large we have $\mathbb{P}(F(v)) \geqslant (b-1)\,e^{-(b-1)\lambda/b}$.

We consider two cases. Suppose first that

$$\mathbb{P}(\exists v \in C_\infty(\lambda)\colon F(v) \text{ occurs}) \geqslant \frac{b\log b}{(b-1)\lambda}.$$

Then, for λ sufficiently large,

$$\begin{aligned}\mathbb{P}(\exists v \in C_\infty(\lambda)\colon F(v) \text{ occurs}) &\leqslant \mathbb{E}(\#\{v \in C_\infty(\lambda)\colon F(v) \text{ occurs}\}) \\ &= \mathbb{E}(|C_\infty(\lambda)|)\mathbb{P}(F(v)) \\ &= \mathbb{E}(|C_\infty(\lambda)|)(b-1)\,e^{-(b-1)\lambda/b},\end{aligned}$$

since the events $(F(v))_{v\in B_1}$ are independent of $C_\infty(\lambda)$. So we must have

$$\mathbb{E}(|C_\infty(\lambda)|) \geqslant \frac{e^{(b-1)\lambda/b}b\log b}{(b-1)^2\lambda}.$$

The second case is where

$$\mathbb{P}(\exists v \in C_\infty(\lambda)\colon F(v) \text{ occurs}) \leqslant \frac{b\log b}{(b-1)\lambda}.$$

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Then the complementary event, that $F(v)$ fails for all $v \in C_\infty(\lambda)$, occurs with probability greater than

$$1 - \frac{b \log b}{(b-1)\lambda},$$

and in this event we have $|C'_\infty(\lambda)| = b|C_\infty(\lambda)|$. But these two variables are identically distributed, and so, by Lemma 5.2,

$$\mathbb{E}(|C_\infty(\lambda)|) \geqslant \frac{b^{((b-1)\lambda)/(b\log b)} - 1}{b-1} \frac{b\log b}{(b-1)\lambda} = \frac{(\mathrm{e}^{(b-1)\lambda/b} - 1) b \log b}{(b-1)^2 \lambda},$$

as required. □

Remark. The same argument can be used to prove that if $C^-_\infty(\lambda)$ is the component of v_0 in the subgraph of $G_\infty(\lambda)$ obtained by deleting all edges between siblings of T_∞, then $\mathbb{E}(C^-_\infty(\lambda))$ is also at least exponential in λ.

We now give the deferred proof of Lemma 5.2.

Proof of Lemma 5.2. For each m not divisible by b, write Z_m for the set $\{m, bm, b^2m, \ldots\}$. Suppose that $\mathbb{P}(X \in Z_m) > 0$ and $\mathbb{P}(X = bY \mid X \in Z_m) = 1 - q_m$.

First we claim $\max_{x \in Z_m} \mathbb{P}(X = x \mid X \in Z_m) \leqslant q_m$. Suppose not, and $\mathbb{P}(X = x \mid X \in Z_m) = q_m + \varepsilon$. For some $y > 0$ we have $\mathbb{P}(X = b^y x \mid X \in Z_m) < \varepsilon$. Now

$$\begin{aligned}
\mathbb{P}((X \neq bY) \wedge (X \in Z_m)) &\geqslant \sum_{z=0}^{y-1} \mathbb{P}(X = b^z x \wedge Y \neq b^{z+1}x) \\
&\geqslant \sum_{z=0}^{y-1} (\mathbb{P}(X = b^z x) - \mathbb{P}(Y = b^{z+1}x)) \\
&= \mathbb{P}(X = x) - \mathbb{P}(X = b^y x) \\
&> q_m \mathbb{P}(X \in Z_m).
\end{aligned}$$

But $\mathbb{P}((X \neq bY) \wedge (X \in Z_m)) = q_m \mathbb{P}(X \in Z_m)$. This proves the claim. Now it follows that $\mathbb{P}(X > b^k m \mid X \in Z_m) \geqslant 1 - kq_m$, so the variable $X \mid X \in Z_m$ dominates the variable which takes the value $b^k m$ with probability q_m if $k \leqslant \lfloor q_m^{-1} \rfloor$, 0 if $k > \lfloor q_m^{-1} \rfloor + 1$, and $1 - q_m \lfloor q_m^{-1} \rfloor$ if $k = \lfloor q_m^{-1} \rfloor$. This variable has expectation $mf(q_m)$.

Finally, $p = \sum_m \mathbb{P}(X \in Z_m) q_m$, and

$$\mathbb{E}(X) = \sum_m \mathbb{P}(X \in Z_m) m f(q_m) \geqslant \sum_m \mathbb{P}(X \in Z_m) f(q_m).$$

This is at least $f(p)$ by Jensen's inequality. □

5.2 Upper bound

In order to bound the expected size of $C_\infty(\lambda)$, we first show that this is equivalent to bounding the limit as $n \to \infty$ of the expected size of $C_n(\lambda)$.

Lemma 5.3. *For any $m > n$ (including the case $m = \infty$), we have*

$$\mathbb{E}(|C_n(\lambda)|) < \mathbb{E}(|C_m(\lambda)|) < (1 + o(1))\mathbb{E}(|C_n(\lambda)|)$$

as $n \to \infty$.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Remark. In particular, taking $m = \infty$ and n sufficiently large gives $\mathbb{E}(|C_\infty(\lambda)|) < \infty$.

Proof. The first inequality follows from the fact that $G_n(\lambda)$ can be obtained as an induced subgraph of $G_m(\lambda)$. For the second inequality, we can obtain $G_m(\lambda)$ by first running $G_n(\lambda)$ on each copy of T_n, then adding edges between copies. We know from Lemma 3.1 that the probability that the component (in $G_n(\lambda)$) of a randomly chosen vertex in a copy of T_n having an edge in the second stage is $o(1)$. If it does, the expected number of additional edges it has in the second stage is at most the expected number of edges from the whole of that copy of T_n, which is about λ. So the expected number of edges received by the component of a randomly chosen vertex in the second stage is also $o(1)$. We now consider a branching process: start from the component of v_0 in its copy of T_n; the next generation is the set of components we reach with one long edge, and so on. We can bound this from above by assuming that all edges go to new, different components. This is subcritical, since the expected number of offspring is $o(1)$, and hence the total expected number of components merged is $1 + o(1)$. Each component has expected size at most $\mathbb{E}(|C_n(\lambda)|)$, giving the required bound. □

The upper bound of Theorem 1.2 follows from the next result.

Proposition 5.4. $\log \mathbb{E}(|C_\infty(\lambda)|) = O(\mathrm{e}^\lambda \log \lambda)$.

Proof. By Lemma 5.3, it is sufficient to show that this bound applies to $\mathbb{E}(|C_h(\lambda)|)$ for every $h < \infty$. Write $d_h(\lambda) = \mathbb{E}(|C_h(\lambda)|)/|L_h|$. Then $d_h(\lambda) \to 0$ as $h \to \infty$, since by Lemma 3.1 there exists k such that the probability that the component escapes L_k is at most ε, and hence for h sufficiently large we have $d_h(\lambda) < \varepsilon + b^{k-h} < 2\varepsilon$.

In particular, let $k = \lceil b \log ((2b+4)\lambda)\, \mathrm{e}^\lambda \rceil$ and $h = k + \log_b ((2b+4)\lambda)$. Lemma 3.1 gives

$$\begin{aligned}\mathbb{P}(C_h \not\subseteq L_k) &< (1 - \mathrm{e}^{-\lambda}/b)^{b \log ((2b+4)\lambda)\, \mathrm{e}^\lambda} \\ &< (\mathrm{e}^{-1})^{\log ((2b+4)\lambda)} \\ &= \lambda^{-1}/(2b+4).\end{aligned}$$

Therefore $d_h(\lambda) < \lambda^{-1}/(2b+4) + b^{k-h} = \lambda^{-1}/(b+2)$.

Suppose $\lambda d_h(\lambda) < (b^j + b^2/(b+1))^{-1}$ for some $j \geqslant 0$. We can bound $d_{h+1}(\lambda)$ by first generating two independent copies of $G_h(\mathrm{Po}(\lambda))$, and then adding edges between the two, with each pair getting $\mathrm{Po}(b^{1-2h}\lambda)$ edges (call these the 'long' edges). Now we use a similar branching process argument to Lemma 5.3: the component of v_0 is $C_{h+1}(\lambda) = \bigcup_{i \geqslant 1} C_i$, where C_1 is the component of the root in the first half, C_2 is the union of all components in the second half having a long edge to C_1, C_3 is the union of components in the first half which are not in C_1 but are joined by a long edge to C_2, and so on. The expected number of long edges between C_i and C_{i+1}, and therefore also the expected number of components in C_{i+1}, is at most $\lambda b^{1-h}|C_i|$. The expected total size of C_{i+1} is therefore at most $b\lambda d_h(\lambda)|C_i|$, since every time we reach a new vertex, the expected size of the component containing that vertex is at most $b^h d_h(\lambda)$, as all edges are independent. So

$$\mathbb{E}(\lambda |C_i| b^{-h}) \leqslant \frac{1}{b}(b\lambda d_h(\lambda))^i,$$

and

$$\begin{aligned}\lambda d_{h+1}(\lambda) &= \frac{1}{b}\sum_{i \geqslant 1} \mathbb{E}(\lambda b^{-h}|C_i|) \\ &\leqslant \frac{1}{b^2}\sum_{i \geqslant 1} (b\lambda d_h(\lambda))^i\end{aligned}$$

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

$$= \frac{\lambda d_h(\lambda)}{b(1 - b\lambda d_h(\lambda))}$$
$$< \frac{1}{b^j + b^2/(b+1)} \cdot \frac{b^j + b^2/(b-1)}{b(b^j + b/(b-1))}$$
$$= \frac{1}{b^{j+1} + b^2/(b-1)}.$$

Consequently $\lambda d_{h+i}(\lambda) < (b^{j+i} + b^2/(b-1))^{-1}$, so

$$\mathbb{E}(C_\infty(\lambda)) \leqslant \lim_{i\to\infty} b^{h+i} d_{h+i}(\lambda)$$
$$< \lim \lambda^{-1} b^{h+i}/(b^{j+i} + b^2/(b-1))$$
$$= \lambda^{-1} b^{h-j}.$$

In particular we have $j = 0$ for

$$h = \lceil b \log((2b+4)\lambda)\, e^\lambda \rceil + \log_b((2b+4)\lambda),$$

and hence

$$\log \mathbb{E}(|C_\infty(\lambda)|) \leqslant \log(\lambda^{-1} b^h) = O(e^\lambda \log \lambda). \qquad \square$$

This completes the proof of Theorem 1.2.

6. Calculations on the Poisson particle model

In this section we provide some calculations concerning the behaviour of the random walks used in the definition of the Poisson particle model, which we will use in Section 7, where we study connectedness thresholds for the finite versions of the model. These calculations will also show that $G^\bullet_\infty(\lambda)$ may be coupled between $G_\infty(c\lambda)$ and $G_\infty(C\lambda)$ for some suitable $C > c > 0$, and hence the results of Section 5 also apply to the Poisson particle model.

Recall that edges of the Poisson particle model are given by random walks on T_h, for some fixed $h \in \mathbb{N} \cup \{\infty\}$, starting and ending in $L_h \subseteq L_\infty$. For $x, y \in L_\infty$, write $x \xrightarrow{w} y$ for the event that a random walk starting from x finishes at y. When discussing the probability of such events we use $\mathbb{P}_h$ to indicate that the random walk in question occurs in T_h. Note that the probability that a particle, starting from a vertex of height 1, reaches height $k \leqslant h$ before height 0 is $(b-1)/(b^k - 1)$. This is because, provided the walk has not yet reached height h, $b^{h(u_i)}$ is a martingale, where u_i is the position of the particle after i steps, and so if we run the process until the first time τ with $h(u_\tau) \in \{0, k\}$, we have

$$b^k \mathbb{P}(h(v_\tau) = k) + \mathbb{P}(h(u_\tau) = 0) = b.$$

Lemma 6.1. *Let x and y be two vertices with lowest common ancestor having height k (if $x = y$ we have $k = 1$). Then*

$$\mathbb{P}_\infty(x \xrightarrow{w} y) = \sum_{h \geqslant k} \frac{(b-1)^2}{(b^{h+1} - 1)(b^h - 1)}.$$

Proof. We condition on the maximum height H reached. As before, $\mathbb{P}(H \geqslant h) = (b-1)/(b^h - 1)$, so

$$\mathbb{P}(H = h) = \frac{b-1}{b^h - 1} - \frac{b-1}{b^{h+1} - 1} = \frac{(b-1)^2 b^h}{(b^h - 1)(b^{h+1} - 1)}.$$

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Now, given that $H = h$, any of the b^h vertices having a common ancestor with x at height h are equally probable endpoints of the walk. So

$$\begin{aligned}\mathbb{P}(x \xrightarrow{w} y) &= \sum_{h \geqslant k} \mathbb{P}(H = h)\mathbb{P}(x \xrightarrow{w} y \mid H = h) \\ &= \sum_{h \geqslant k} \frac{(b-1)^2}{(b^{h+1}-1)(b^h-1)}.\end{aligned}$$

□

Write ζ_h for $\mathbb{P}_\infty(x \xrightarrow{w} y)$, where x and y have lowest common ancestor at height h. We can use Lemma 6.1 to get good bounds on ζ_h.

Lemma 6.2. *For every $1 \leqslant i < h$, we have $\zeta_{i+1} < \zeta_i/b^2$.*

Proof. Note that

$$\begin{aligned}\zeta_i &= \sum_{j \geqslant i} \frac{(b-1)^2}{(b^j-1)(b^{j+1}-1)} \\ &< \sum_{j \geqslant i} \frac{(b-1)^2}{(b^j-b^{j-i})(b^{j+1}-b^{j-i})} \\ &= \sum_{j \geqslant i} \frac{(b-1)^2}{b^{2j-2i}(b^i-1)(b^{i+1}-1)} \\ &= \frac{(b-1)^2}{(b^i-1)(b^{i+1}-1)} \sum_{j \geqslant 0} b^{-2j} \\ &= \frac{(b-1)^2 b^2}{(b^i-1)(b^{i+1}-1)(b^2-1)}.\end{aligned}$$

Consequently

$$\zeta_i - \zeta_{i+1} = \frac{(b-1)^2}{(b^i-1)(b^{i+1}-1)} > \frac{b^2-1}{b^2}\zeta_i,$$

so $\zeta_{i+1} < \zeta_i/b^2$. □

Remark. If leaves $x_1, \ldots, x_b$ are siblings, then $\zeta_1 = \mathbb{P}(x_1 \xrightarrow{w} x_j)$ for each i, so we must have $\zeta_1 < 1/b$, whence $\zeta_i < b^{1-2i}$ for each i.

Our next lemma provides upper and lower bounds on ζ_h, showing that it decays roughly like b^{1-2h}.

Lemma 6.3. *For each $h \geqslant 1$, we have*

$$\frac{b-1}{b+1}(b^{1-2h} + b^{1-3h}) < \zeta_h < \left(1 + \frac{1}{b^h-1}\right)\left(1 + \frac{1}{b^{h+1}-1}\right)\frac{b-1}{b+1}b^{1-2h}.$$

Proof. By Lemma 6.1, we have

$$\zeta_h = \sum_{j \geqslant h}\left(1 + \frac{1}{b^j-1}\right)\left(1 + \frac{1}{b^{j+1}-1}\right)\frac{(b-1)^2}{b^{2j+1}}.$$

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

The first two factors are close to 1, and we bound them by their maximum values to get

$$\begin{aligned}\zeta_h &< \sum_{j\geqslant h}\left(1+\frac{1}{b^h-1}\right)\left(1+\frac{1}{b^{h+1}-1}\right)\frac{(b-1)^2}{b^{2j+1}}\\ &= \left(1+\frac{1}{b^h-1}\right)\left(1+\frac{1}{b^{h+1}-1}\right)\frac{b-1}{b+1}b^{1-2h}.\end{aligned}$$

Moreover, noting that

$$\frac{1}{m-1}=\frac{1}{m}+\frac{1}{m(m-1)},$$

we have

$$\begin{aligned}\zeta_h &= \sum_{j\geqslant h}\frac{(b-1)^2}{(b^j-1)(b^{j+1}-1)}\\ &= (b-1)^2\sum_{j\geqslant h}\left(\frac{1}{b^j}+\frac{1}{b^j(b^j-1)}\right)\left(\frac{1}{b^{j+1}}+\frac{1}{b^{j+1}(b^{j+1}-1)}\right).\end{aligned}$$

We may bound

$$\frac{1}{b^j(b^j-1)}>\frac{1}{b^{2j}}$$

and expand (neglecting the lowest-order term) to get

$$\begin{aligned}\zeta_h &> (b-1)^2\sum_{j\geqslant h}(b^{-2j-1}+b^{-3j-1}+b^{-3j-2})\\ &= \frac{b-1}{b+1}b^{1-2h}+b^{h-1}\frac{b-1}{b^2+b+1}(b^{2-3h}+b^{1-3h})\\ &> \frac{b-1}{b+1}(b^{1-2h}+b^{1-3h}).\end{aligned}$$

□

The independence between the particles implies that the number of particles going from x to y has a Poisson distribution with rate $(\lambda/2)\mathbb{P}(x\xrightarrow{w}y)$, and so the total number of edges between x and y is a Poisson random variable with mean $\lambda\mathbb{P}(x\xrightarrow{w}y)$. Furthermore, the total number of edges vw and the total number of edges xy are independent Poisson random variables whenever $\{v,w\}\neq\{x,y\}$.

Remark. Since both the lower and upper bounds for ζ_h are $\Theta(b^{1-2h})$, this proves our claim that by choosing the parameters λ appropriately, the Poisson edge model can dominate the Poisson particle model and *vice versa*.

In $G^\bullet_\infty(\lambda)$, the number of edges between x and y has distribution $\mathrm{Po}(\zeta_k\lambda)$, where k is the height of the lowest common ancestor of x,y, and consequently the number of edges from x to other vertices has distribution $\mathrm{Po}(\lambda\sum_{h>0}(b-1)b^{h-1}\zeta_h)$. Write Ξ_0 for $\sum_{h>0}(b-1)b^{h-1}\zeta_h$. Similarly, if we fix a vertex $v\in T_\infty$ with height k, write Ξ_k for the value such that the number of edges in $G^\bullet_\infty(\lambda)$ between descendants of v and the rest of the graph has distribution $\mathrm{Po}(\Xi_k\lambda)$. Each descendant of v has $\mathrm{Po}(\lambda\sum_{h>k}(b-1)b^{h-1}\zeta_h)$ edges to the rest of the graph, and there are b^k descendants, so $\Xi_k=(b-1)b^k\sum_{h>k}b^{h-1}\zeta_h$.

We will need some bounds on the value of Ξ_k.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Lemma 6.4. *Ξ_k is strictly decreasing with $\lim_k \Xi_k = (b-1)/(b+1)$. Moreover,*

$$\frac{b-1}{b+1}\left(1+\frac{b^{-k}}{b+1}\right) < \Xi_k < \frac{b-1}{b+1}\left(1+\frac{1}{b^{k+1}-1}\right)\left(1+\frac{1}{b^{k+2}-1}\right).$$

Proof. By Lemma 6.2 we have $b^{i+1}\zeta_{i+1} < b^i\zeta_i/b$ for every i, and hence

$$\sum_{h>k+1} b^h\zeta_h < \sum_{h>k} b^h\zeta_h/b,$$

that is,

$$\Xi_{k+1}/((b-1)b^{k+1}) < \Xi_k/((b-1)b^{k+1}).$$

So Ξ_k is strictly decreasing. Also, by Lemma 6.3 we have

$$\begin{aligned}\Xi_k &= (b-1)b^k \sum_{h>k} b^{h-1}\zeta_h \\ &< (b-1)b^k \sum_{h>k}\left(1+\frac{1}{b^h-1}\right)\left(1+\frac{1}{b^{h+1}-1}\right)\frac{b-1}{b+1}b^{-h}.\end{aligned}$$

We can easily calculate

$$\sum_{h>k}\frac{b-1}{b+1}b^{-h}.$$

The other factors are close to 1, so we bound them by their maximum values, which occur at $h=k+1$.

$$\begin{aligned}\Xi_k &< (b-1)b^k \sum_{h>k}\left(1+\frac{1}{b^{k+1}-1}\right)\left(1+\frac{1}{b^{k+2}-1}\right)\frac{b-1}{b+1}b^{-h} \\ &= \frac{b-1}{b+1}\left(1+\frac{1}{b^{k+1}-1}\right)\left(1+\frac{1}{b^{k+2}-1}\right).\end{aligned}$$

Finally, using the lower bound from Lemma 6.3, we have

$$\begin{aligned}\Xi_k &> (b-1)b^k\frac{b-1}{b+1}\sum_{h>k}(b^{-h}+b^{-2h}) \\ &= \frac{b-1}{b+1}\left(1+\frac{b^{-k}}{b+1}\right).\end{aligned}$$

Since these upper and lower bounds tend to $(b-1)/(b+1)$, so does Ξ_k. □

Remark. The upper bound gives

$$\Xi_k < \frac{b-1}{b+1}\frac{b}{b-1}\frac{b^2}{b^2-1} = \frac{b^3}{b^3+b^2-b-1} < 1 \quad \text{for all } k \geqslant 0.$$

Also, since

$$\frac{b(b-1)}{(b+1)^2}$$

is increasing and $b \geqslant 2$, the lower bound gives

$$\Xi_k > \frac{b-1}{b+1} + \frac{2}{9}b^{-k-1}.$$

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Finally, we prove bounds on the corresponding parameters for the finite model. Write $\xi_n^{(k)}$ for the value such that the number of edges in $G_n^\bullet(\lambda)$ from the descendants of a vertex at height k to other vertices is distributed $\mathrm{Po}(\lambda\xi_n^{(k)})$, that is, $\xi_n^{(k)}$ is b^k times the probability that a random walk starting from a descendant of v ends at a non-descendant.

Lemma 6.5.

$$\Xi_k - b^{2k-2n} < \xi_n^{(k)} < \Xi_k.$$

Proof. We prove that

$$\xi_n^{(k)} = \Xi_k - (b-1)\sum_{h>n} b^{h-n+2k-1}\zeta_h,$$

which clearly implies the upper bound. The lower bound follows since $\zeta_h < b^{1-2h}$, and so

$$\sum_{h>n} b^{h-n+2k-1}\zeta_h < b^{2k-n}\sum_{h>n} b^{-h} = \frac{b^{2k-2n}}{b-1}.$$

We couple random walks on T_n and T_∞ as follows. Start by running the two walks identically. At any point when the walk on T_∞ leaves T_n, pause the walk on T_n at the apex until the walk on T_∞ next reaches a vertex at height n. The descendants of this vertex form a copy of T_n; as long as the second walk stays within that copy, duplicate all its steps in the walk on T_n. Now the first random walk ends in $D(v)$ if and only if the second walk ends at a vertex in the translation of $D(v)$ to the subtree formed by the descendants of some vertex at height n. There are $(b-1)b^{h-n-1+k}$ such vertices whose lowest common ancestor with T_n is at height h, for every $h > n$, and hence for any $x \in D(v)$

$$\mathbb{P}_n(x \xrightarrow{\mathrm{w}} D(v)) = \mathbb{P}_\infty(x \xrightarrow{\mathrm{w}} D(v)) + \sum_{h>n}(b-1)b^{h-n-1+k}\zeta_h.$$

Since $\xi_n^{(k)} = b^k(1 - \mathbb{P}_n(x \xrightarrow{\mathrm{w}} D(v)))$ and $\Xi_k = b^k(1 - \mathbb{P}_\infty(x \xrightarrow{\mathrm{w}} D(v)))$, the result follows. □

In particular, $\xi_n^{(k)} \to \Xi_k$. Also, provided $2n \geqslant 3k+5$, we may combine the bounds of Lemmas 6.5 and 6.4 to get

$$\begin{aligned}\xi_n^{(k)} &> \Xi_k - b^{2k-2n}\\ &\geqslant \Xi_k - b^{-k-5}\\ &> \frac{b-1}{b+1} + \frac{2}{9}b^{-k-1} - b^{-k-5}\\ &> \frac{b-1}{b+1} + \frac{1}{4}b^{-k-1}.\end{aligned}$$

7. Phase transitions for our finite models

In this section we study the threshold λ at which our Poisson particle model $G_n^\bullet(\lambda)$ and Poisson edge model $G_n(\lambda)$ become connected for finite n. We will prove that there is a phase transition for the connectedness of $G_n^\bullet(\lambda)$ and $G_n(\lambda)$ occurring at $\lambda = \sigma_{\mathrm{crit}} n$, where σ_{crit} is a constant (different for the two models) depending on b only. This phase transition occurs in a window of width logarithmic in n (Theorems 7.6 and 7.7).

It turns out that for the Poisson edge model the asymptotic threshold for connectedness coincides with that for having no isolated vertices, just as in the Erdős–Rényi model [6], while for the

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Poisson particle model this is not the case (Theorem 7.5). We therefore answer Problem 8.2 of [7] in the negative.

We start with two lemmas that will be needed for the analysis of each of our models. We then establish the connectedness threshold for the Poisson particle model in Section 7.1, and the Poisson edge model is treated in Section 7.2 in a similar fashion.

The following lemma bounds the probability of isolated vertices in a wide range of random graph settings where different edges appear with different probability.

Lemma 7.1. *Let G be a random graph on vertex set $[n]$, where each edge ij is independently present with probability p_{ij} and absent with probability $q_{ij} = 1 - p_{ij}$. Write I_i for the event that i is isolated, and N for the number of isolated vertices. If $\mathbb{P}(I_i) = q$ for every i, then $\mathbb{P}(N = 0) < 2/(2 + nq)$.*

Proof. Note that $\mathbb{P}(I_i) = \prod_{j\neq i} q_{ij}$, so this product equals q for each i. Also,

$$\begin{aligned}\mathbb{E}(N^2) &= \sum_{i,j} \mathbb{P}(I_i \wedge I_j)\\ &= \sum_i \left(q + \sum_{j\neq i} \left(q \prod_{k\neq i,j} q_{jk} \right) \right)\\ &= \sum_i \left(q + \sum_{j\neq i} q^2 q_{ij}^{-1} \right).\end{aligned}$$

Now $q_{ij}^{-1} \geqslant 1$ for each $j \neq i$, and $\prod_{j\neq i} q_{ij}^{-1} = q^{-1}$, so $\sum_{j\neq i} q_{ij}^{-1}$ is maximized when one of the terms is q^{-1} and the others are all 1. To see this, note that if $q_{ij}^{-1}, q_{ik}^{-1} > 1$ for distinct j, k then we increase the sum by replacing $q_{ij}^{-1} + q_{ik}^{-1}$ with $1 + q_{ij}^{-1} q_{ik}^{-1}$; by a sequence of operations of this form we continue increasing the sum until only one term exceeds 1. So

$$\begin{aligned}\mathbb{E}(N^2) &\leqslant n(q + q^2(q^{-1} + n - 2))\\ &= n(2q - 2q^2 + nq^2)\\ &< 2nq + \mathbb{E}(N)^2,\end{aligned}$$

and thus $\mathrm{Var}\,(N) < 2nq$. Consequently, by Cantelli's inequality,

$$\mathbb{P}(N = 0) \leqslant \mathrm{Var}\,(N)/(\,\mathrm{Var}\,(N) + \mathbb{E}(N)^2) < 2/(2 + nq). \qquad \square$$

Remark. In particular, if $nq \to \infty$ as $n \to \infty$, $\mathbb{P}(N = 0) \to 0$. The converse is also true. Suppose $nq \leqslant c$ infinitely often. The event that an individual vertex meets some edges is an increasing event, and so, by Harris's inequality [9], all such events are positively correlated. So $\mathbb{P}(N = 0) \geqslant (1 - q)^n \geqslant (1 - c/n)^n$ infinitely often, and this bound approaches $\mathrm{e}^{-c} > 0$. Trivially, if $nq \to 0$ then $\mathbb{P}(N = 0) \to 1$; again the converse is true since a non-trivial lower bound on nq gives a non-trivial upper bound on $2/(2 + nq)$.

Remark. Suppose that each vertex has an additional probability r of being active, and we are interested in the number of active isolated vertices. Now the expected number is nqr and a similar analysis gives a bound of $(1 + r)/(1 + r + nqr) \leqslant 2/(2 + nqr)$.

We next prove a sufficient condition for connectedness which will be the key ingredient in establishing the supercritical region for each model.

First we introduce some notation. Let T be an arbitrary finite b-ary tree, and let v, w be two vertices of T. We say v is a *k-uncle* of w if a sibling of v is an ancestor of w at distance k or less. We

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

say v and w are *k-cousins* if there exists w' which is an ancestor of w at distance k or less such that w' is a k-uncle of v (note that this definition is symmetric, since some sibling v' of w' is a k-ancestor of v and k-uncle of w). If $(x_1, \dots, x_b)$ and $(y_1, \dots, y_b)$ are two disjoint b-tuples of siblings, then we say $(x_1, \dots, x_b)$ and $(y_1, \dots, y_b)$ are k-cousins if x_1 and y_1 are k-cousins (which will also imply that all other pairs are k-cousins).

Now let G be a graph whose vertex set is $L(T)$, the set of leaves of T. We say that two vertices $x, y \in V(T)$ are *linked* by G if there exist vertices $x', y' \in L_n$ such that x is an ancestor of x', y is an ancestor of y', and there is an edge (in G) from x' to y'. If $X = \{x_1, \dots, x_b\}$ is a set of siblings in T, we say X is *strongly linked* by G if the graph H_X on vertex set X with edges between linked pairs is connected, and *weakly linked* if H_X has two components and there exist i, j such that x_i and x_j are in different components and some z which is a k-uncle of x_i, x_j such that the pairs x_i, z and x_j, z are both linked.

Lemma 7.2. *Suppose that G has the following properties, for some fixed k.*

(i) *Every set of siblings in T is either strongly linked or weakly linked by G.*
(ii) *For any two sets of siblings which are k-cousins, at least one of them is strongly linked by G.*
(iii) *Any set of siblings within the top k layers of T are strongly linked by G.*

Then G is connected.

Proof. Define the *depth* of a vertex $v \in V(T)$ to be the distance from the apex of T, and the *height* $h(v)$ to be the difference between the depth of v and the maximum depth (note that this coincides with our earlier definition for $T = T_n$). We show by induction on j that for every vertex $v \in V(T)$ with $h(v) = j$, all the leaves of T which are descendants of v are in the same component of G. This is trivial for $j = 0$, and whenever v is a leaf. Suppose it is true for $0, \dots, j-1$, let v be a non-leaf vertex at height j, and let $x_1, \dots, x_b$ be its neighbours at height $j-1$. Every descendant of v is a descendant of some x_i, and for each i all descendants of x_i are in one component. If $x_1, \dots, x_b$ are strongly linked then there is an edge between a descendant of x_i and x_j for $x_i x_j \in E(H_X)$, and these connect all the components so all descendants of v are in the same component. If not, the descendants of v are in at most two components, and there is some z which is linked to both components and is a k-uncle of $x_1, \dots, x_b$. It is sufficient to prove that all descendants of z are in one component. If $h(z) < j$ then this is true by the induction hypothesis. If $h(z) \geqslant j$ we use a subsidiary induction to show that all descendants of z' are in the same component whenever z' is a descendant of z at height i; this is true for $i = j-1$, and may be extended to each successive i by noting that if $z'_1, \dots, z'_b$ are siblings which are descendants of z at height at least $j-1$ then $(x_1, \dots, x_b)$ and $(z'_1, \dots, z'_b)$ are k-cousins, so $(z'_1, \dots, z'_b)$ are strongly linked by (ii). □

7.1 Connectedness threshold for the Poisson particle model

We will show that, unlike the Erdős–Rényi random graph model, isolated vertices are not the main obstacle to connectedness in $G^{\bullet}_n(\lambda)$.

The bounds of Section 6 will allow us to give, for a fixed height k, a threshold dividing regimes where there is a component of the form $D(v)$ for some vertex v at height k with high probability from regimes where there is no such component with high probability. In fact, of all sets of vertices with lowest common ancestor v, $D(v)$ is the most likely to be separated from the rest of the graph, so we can show that in the latter regime there is actually no component with a common ancestor of height at most k with high probability. There may be such a component even if other descendants of the common ancestor are not disconnected from the rest of the graph, so this is a genuinely stronger statement.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Lemma 7.3. *Let $x \in L_k$ and $\varnothing \neq A \subset L_k$. If $k \leqslant m < n \leqslant \infty$ then $\mathbb{P}_m(x \xrightarrow{w} A) > \mathbb{P}_n(x \xrightarrow{w} A)$.*

Proof. We couple the random walks on T_m and T_n as in the proof of Lemma 6.5: duplicate every step made in the bottom m layers of T_n, but pause the walk in T_m while the walk in T_n is above that level. The walk in T_m reaches A if and only if the walk in T_n reaches the translation of A to the m-subtree it finishes in. □

Lemma 7.4. *Let A be any non-empty subset of L_k. Then the probability that there are no edges between A and the rest of $G_n^{\bullet}(\lambda)$ is maximized when $A = L_k$, for any $n \geqslant k$ or for $n = \infty$.*

Proof. We use induction on k; it is clearly true for $k = 0$ since there is no choice of A. Assume it is true for $k-1$. Note that

$$\mathbb{P}(e(A, V \setminus A) = 0) = \exp\left(-\lambda \sum_{x \in A} \mathbb{P}(x \xrightarrow{w} V \setminus A)\right),$$

since $e(A, V \setminus A)$ has a Poisson distribution with the appropriate mean. Write $P(A)$ for $\sum_{x \in A} \mathbb{P}(x \xrightarrow{w} V \setminus A)$. Now

$$P(A) - P(L_k) = \sum_{x \in L_k \setminus A} \left(\left(\sum_{y \in A} \mathbb{P}(y \xrightarrow{w} x) \right) - \mathbb{P}(x \xrightarrow{w} V \setminus L_k) \right).$$

By Lemma 7.3, $\mathbb{P}(y \xrightarrow{w} x)$ is minimized and $\mathbb{P}(x \xrightarrow{w} V \setminus L_k)$ is maximized in the case $n = \infty$, so it is sufficient to prove $P(A) - P(L_k) > 0$ in this case:

$$\begin{aligned}
P(A) - P(L_k) &\geqslant \sum_{x \in L_k \setminus A} \left(|A| \zeta_k - \sum_{y \notin L_k} \mathbb{P}(x \xrightarrow{w} y) \right) \\
&= \sum_{x \in L_k \setminus A} \left(|A| \zeta_k - \sum_{h > k} (b-1) b^{h-1} \zeta_h \right) \\
&> \sum_{x \in L_k \setminus A} \left(|A| \zeta_k - \sum_{h > k} (b-1) b^{h-1} b^{2k-2h} \zeta_k \right) \\
&= (b^k - |A|)(|A| - b^{k-1}) \zeta_k,
\end{aligned}$$

so the result follows for all A with $|A| \geqslant b^{k-1}$. If $|A| < b^{k-1}$, we claim that $\min_{|B|=|A|} P(B)$ is achieved when B consists of the first $|A|$ elements of L_k. Then, since $B \subset L_{k-1}$, we have $P(A) \geqslant P(B) \geqslant P(L_{k-1})$ by induction, and $P(L_{k-1}) > P(L_k)$ since $|L_{k-1}| = b^{k-1}$.

To prove the claim, note that

$$\begin{aligned}
P(B) &= |B| - \sum_{x \in B} \mathbb{P}(x \xrightarrow{w} B) \\
&= |B| - \sum_{x,y \in B} \mathbb{P}(x \xrightarrow{w} y) \\
&= |B| - \sum_{x,y \in B} \zeta_{h(x,y)},
\end{aligned}$$

where $h(x, y)$ is the height of the lowest common ancestor of x and y. Since ζ_h is decreasing, it is sufficient to prove that an initial segment maximizes $|\{(x, y) \in B^2 \colon h(x, y) \leqslant h\}|$ for each h. Separate L_k into chunks of length b^h, and write $b_1, \ldots, b_r$ for the number of vertices of b in each chunk. Then $|\{(x, y) \in B^2 \colon h(x, y) \leqslant h\}| = \sum_i b_i^2$. If $b^h > b_i \geqslant b_j > 0$ then this can be increased by

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

replacing b_i, b_j with $b_i + 1, b_j - 1$. Consequently $P(B)$ is minimal only if at most one chunk is neither full nor empty for every h. There is a unique (up to reordering of the chunks) sequence $b_1, \dots, b_r$ which achieves this, so any B which achieves this for every h has the same value of $P(B)$, which is minimal; in particular, the initial segment is one such B. □

Write D_k for the event that there is some $v \in T_n$ at height k such that the descendants of v are disconnected from the rest of $G_n^\bullet(\lambda)$. Write C_k for the event that $G_n^\bullet(\lambda)$ has some component with a common ancestor at height k (not necessarily the lowest common ancestor, so $C_k \subset C_{k+1}$).

Theorem 7.5. *There exist values $\sigma_0 < \sigma_1 < \cdots$ with*

$$\lim_{k\to\infty} \sigma_k = \sigma_{\text{crit}} = \frac{b+1}{b-1} \log b$$

such that

(a) $\lambda = \sigma_k n$ *is a sharp threshold for both C_k and D_k, and*
(b) $\lambda = \sigma_{\text{crit}} n$ *is a sharp threshold for connectedness of $G_n^\bullet(\lambda)$.*

Further, if $\lambda = \sigma_k n$ then $\mathbb{P}(D_k)$ and $\mathbb{P}(C_k)$ are bounded away from 0 *and* 1.

Remark. In particular, the sharp thresholds for the existence of isolated vertices ($\lambda = \sigma_0 n$) and for connectedness ($\lambda = \sigma_{\text{crit}} n$) do not coincide.

Proof. For (a), since $D_k \subseteq C_k$, it is sufficient to show that, for $\lambda = \sigma n$, D_k holds with high probability if $\sigma < \sigma_k$ and C_k fails with high probability if $\sigma > \sigma_k$.

Write $G_n^{(j)}(\lambda)$ to be the graph whose vertices are the vertices at height j in T_n, with the number of edges between v and w in $G_n^{(j)}(\lambda)$ being equal to the number of edges between descendants of v and descendants of w in $G_n^\bullet(\lambda)$. Then $G_n^{(0)}(\lambda) = G_n^\bullet(\lambda)$, and v is isolated in $G_n^{(h(v))}(\lambda)$ if and only if its descendants are disconnected from the rest of $G_n^\bullet(\lambda)$. Note that adjacencies in $G_n^{(j)}(\lambda)$ are independent events with varying probabilities.

Suppose $\sigma \Xi_k > \log b$. For n sufficiently large, also $\sigma \xi_n^{(k)} > \log b$. The probability that a given vertex v of height $h \leqslant k$ will be isolated in $G_n^{(h)}(\sigma n)$ is $\exp(-\xi_n^{(h)} \sigma n) \leqslant \exp(-\xi_n^{(k)} \sigma n) = o(b^{-n})$. By Lemma 7.4, each possible subset of $D(v)$ is a component of $G_n^\bullet(\sigma n)$ with probability $o(2^{-n})$. There are fewer than b^{n+1} such vertices v, and at most $2^{b^k} = O(1)$ subsets of $D(v)$ in each case, so $\mathbb{P}(C_k) = o(1)$.

Conversely, suppose $\sigma \Xi_k < \log b$. The probability that a given vertex v at height k will be isolated in $G_n^{(k)}(\sigma n)$ is $\exp(-\xi_n^{(k)} \sigma n)$. But $\sigma \xi_n^{(k)} < \log b$, so this probability is $\omega(b^{-n})$. There are b^{n-k} vertices in $G_n^{(k)}(\sigma n)$, and hence the expected number of isolated vertices is $\omega(1)$. By Lemma 7.1, $\mathbb{P}(D_k) = 1 - o(1)$. Thus (a) holds for $\sigma_k = \Xi_k^{-1} \log b$.

Now suppose $\sigma = \sigma_k = \Xi_k^{-1} \log b$. Write N_{D_k} for the number of vertices at height k which are isolated in $G_n^{(k)}(\sigma n)$. Then

$$\begin{aligned}
\mathbb{E}(N_{D_k}) &= b^{n-k} \exp(-\xi_n^{(k)} \sigma n) \\
&> b^{n-k} \exp(-(\Xi_k - b^{2k-2n}) \sigma n) \\
&= b^{-k} \exp(b^{2k-2n} \sigma n) \\
&= b^{-k}(1 + o(1)).
\end{aligned}$$

Consequently, by the remarks following Lemma 7.1, $\mathbb{P}(D_k)$ is bounded away from 0.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

For each non-empty set $S \subseteq L_k$, write N_S for the number of clones of S which are disconnected from the rest of the graph. By Lemma 7.4, for each such S we have $\mathbb{E}(N_S) \geqslant \mathbb{E}(N_{D_k}) = b^{-k}(1 + o(1))$. Also, we can consider this as counting isolated, active vertices in some graph. Draw a graph whose vertices are the vertices at height k in T_n, with an edge between two vertices if there is an edge in $G_n^\bullet(\sigma n)$ between the clones of S descended from them, and a vertex being active if there is no edge from the clone of S descended from it to any vertex which is not part of one of the clones of S. Thus, by the remarks following Lemma 7.1, since $\mathbb{E}(N_S)$ is bounded, we have $\mathbb{P}(N_S = 0) > \varepsilon > 0$ for some ε. Now the events $(N_S = 0)_{S \subseteq L_k}$ are all decreasing events, so positively correlated by Harris's inequality. Thus $\mathbb{P}(N_S = 0 \text{ for every } S) > \varepsilon^{2^{b^k}}$, and hence $\mathbb{P}(C_k)$ is bounded away from 1. Since $\mathbb{P}(C_k) > \mathbb{P}(D_k)$, both bounds apply to both events, as required.

Since

$$\Xi_k \to \frac{b-1}{b+1}$$

from above, we have

$$\sigma_{\text{crit}} = \lim_{k \to \infty} \sigma_k = \frac{b+1}{b-1} \log b.$$

If $\sigma < \sigma_{\text{crit}}$, then for some k we have $\sigma < \sigma_k$, so $G_n^\bullet(\sigma n)$ is disconnected with high probability. It remains to show that $G_n^\bullet(\sigma n)$ is connected with high probability for any $\sigma > \sigma_{\text{crit}}$.

Fix $\sigma > \sigma_{\text{crit}}$ and let $G = G_n^\bullet(\sigma n)$ and $T = T_n$. It suffices to show that there is some fixed k for which G satisfies (i), (ii) and (iii) of Lemma 7.2 with high probability. Let $\varepsilon > 0$ be such that $\sigma = (1 + \varepsilon)\sigma_{\text{crit}}$.

First we show that, with high probability, every group of siblings which is not strongly linked has one sibling that is not linked to any of the others, but all other pairs linked. It is sufficient to show that with high probability there is no group of siblings containing b unlinked pairs.

Fix siblings $x_1, \ldots, x_b$ at height $h - 1$; for any pair x_i', x_j' of descendants of x_i, x_j where $i \neq j$, the number of edges between x_i' and x_j' is $\text{Po}(\zeta_h \sigma n)$. Consequently the total number of such edges, over all possible pairs (x_i', x_j'), is $\text{Po}(b^{2h-2}\zeta_h \sigma n)$. Recalling our bounds on ζ_h from Lemma 6.2,

$$b^{2h-1}\zeta_h > \frac{b-1}{b+1}.$$

Consequently, the probability that x_i and x_j are not linked is

$$\exp(-b^{2h-2}\zeta_h \sigma n) < \exp\left(-\frac{b-1}{b(b+1)}\sigma n\right).$$

Therefore, writing $N_{x_1,\ldots,x_b}$ for the number of pairs $i \neq j$ such that x_i and x_j are not linked,

$$\begin{aligned}\mathbb{P}(N_{x_1,\ldots,x_b} \geqslant b) &\leqslant \binom{\binom{b}{2}}{b} \exp\left(-\frac{b-1}{b+1}\sigma n\right) \\ &= O(b^{-(1+\varepsilon)n}),\end{aligned}$$

so the probability that this holds true for some set of siblings is $O(b^{-\varepsilon n}) = o(1)$.

Similarly, the probability that any given set of siblings has one sibling that is not linked to any of the others, but all other pairs linked, is at most

$$b \exp\left(-\frac{(b-1)^2}{b(b+1)}\sigma n\right).$$

This probability is $o(1)$, and there are only a constant number of groups of siblings at height $n - k$ or above, so certainly (iii) is satisfied with high probability.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Suppose $x_1, \dots, x_b$ are siblings below this point, and write

$$\{z_1^{(1)}, \dots, z_{b-1}^{(1)}\}, \dots, \{z_1^{(k)}, \dots, z_{b-1}^{(k)}\}$$

for the groups of k-uncles of $x_1, \dots, x_b$ (in increasing order of height). The number of edges that link $z_i^{(j)}$ to x_m is $\mathrm{Po}(b^{2h+j-2}\zeta_{h+j-1}\sigma n)$, and hence

$$\begin{aligned}\mathbb{P}(x_m, z_i^{(j)} \text{ linked}) &= 1 - \exp\left(-b^{2h+j-2}\zeta_{h+j-1}\sigma n\right)\\ &> 1 - \exp\left(-\frac{b-1}{b^{j+1}(b+1)}\sigma n\right).\end{aligned}$$

Consequently, since edges between different pairs of vertices are independent,

$$\begin{aligned}\mathbb{P}(\exists m\colon x_m, z_i^{(j)} \text{ not linked}) &\leqslant 1 - \left(1 - \exp\left(-\frac{b-1}{b^{j+1}(b+1)}\sigma n\right)\right)^b\\ &< b\exp\left(-\frac{b-1}{b^{j+1}(b+1)}\sigma n\right).\end{aligned}$$

The event $X_{x_1,\dots,x_b}$ that $x_1, \dots, x_b$ are neither strongly linked nor weakly linked therefore satisfies

$$\begin{aligned}\mathbb{P}(X_{x_1,\dots,x_b}) &< b\exp\left(-\frac{(b-1)^2}{b(b+1)}\sigma n\right)\prod_{j=1}^{k}\prod_{i=1}^{b-1} b\exp\left(-\frac{(b-1)}{b^{j+1}(b+1)}\sigma n\right)\\ &= b\exp\left(-\frac{(b-1)^2}{b(b+1)}\sigma n\right)\prod_{j=1}^{k} b^{b-1}\exp\left(-\frac{(b-1)^2}{b^{j+1}(b+1)}\sigma n\right)\\ &= b^{kb-k+1}\exp\left(-\sum_{j=0}^{k}\frac{(b-1)^2}{b^{j+1}(b+1)}\sigma n\right)\\ &= b^{kb-k+1}\exp\left(-(1-b^{-k-1})\frac{(b-1)}{b+1}\sigma n\right)\\ &= b^{kb-k+1}\exp(-(1-b^{-k-1})(1+\varepsilon)n\log b).\end{aligned}$$

For some value of k that depends only on ε we have $\mathbb{P}(X_{x_1,\dots,x_b}) = O(b^{(1+\varepsilon/2)n})$. Thus the expected number of such events that occur is $O(b^{\varepsilon n/2})$, that is, the probability that (i) fails to be satisfied is $o(1)$.

Likewise, if the groups of siblings $x_1, \dots, x_b$ and $y_1, \dots, y_b$ are k-cousins, the probability that both groups have a single vertex not linked to any of the others is at most

$$b^2\exp\left(-\frac{(b-1)^2}{b(b+1)}\sigma n\right)$$

(since these are independent events), and for any given $x_1, \dots, x_b$ there are fewer than $k(b-1)b^k$ possible choices for $y_1, \dots, y_b$, since any k-cousins of $x_1, \dots, x_b$ are in a copy of T_k with apex a k-uncle of $x_1, \dots, x_b$. Consequently the probability that (ii) fails is at most

$$k(b-1)b^{n+k+2}\exp\left(-\frac{(b-1)^2}{b(b+1)}\sigma n\right).$$

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Since

$$\frac{(b-1)^2}{b(b+1)}\sigma \geqslant (1+\varepsilon)\log b,$$

this probability is $o(1)$, as required.

Thus the conditions of Lemma 7.2 hold with high probability for this choice of k, so (b) holds. □

We can improve the results of Theorem 7.5 to give a logarithmic window in which connectedness occurs.

Theorem 7.6. *For any fixed $\alpha > b+1$, $G_n^\bullet(\sigma_{\text{crit}}n + \alpha \log n)$ is connected with high probability and for any fixed*

$$\beta > \frac{b+1}{b-1},$$

$G_n^\bullet(\sigma_{\text{crit}}n - \beta \log n)$ is disconnected with high probability.

Proof. We follow the proof of Theorem 7.5, but instead of taking k constant, set $k = k_n = \log_b n$ (note that this logarithm is to base b but the log n in the statement of the theorem is natural). This is the right multiple to take: if $k = \gamma \log_b n$ then $\gamma < 1$ does not work and $\gamma > 1$ forces α, β to be higher.

For $G_n^\bullet(\sigma_{\text{crit}}n + \alpha \log n)$, we need to show that with high probability (i), (ii) and (iii) still apply. As before, for any group of siblings $x_1, \ldots, x_b$,

$$\begin{aligned}\mathbb{P}(N_{x_1,\ldots,x_b} \geqslant b) &\leqslant \binom{\binom{b}{2}}{b} \exp\left(-\frac{b-1}{b+1}(\sigma_{\text{crit}}n + \alpha \log n)\right)\\ &= O(b^{-n}n^{-\alpha(b-1)/(b+1)}),\end{aligned}$$

and thus with high probability no group of siblings fails to be strongly connected except by having one vertex not linked to any of the others. Also,

$$\begin{aligned}\mathbb{P}(X_{x_1,\ldots,x_b}) &< b^{kb-k+1} \exp\left(-(1-b^{-k-1})\frac{b-1}{b+1}(\sigma_{\text{crit}}n + \alpha \log n)\right)\\ &= bn^{b-1}b^{-n}n^{-\alpha(b-1)/(b+1)}b^{1/b}n^{o(1)}\\ &= b^{b+1-n}n^{(b-1)(1+o(1)-\alpha/(b+1))}.\end{aligned}$$

As $\alpha > b+1$, this is $o(b^{-n})$, so with high probability (i) holds for every group $x_1, \ldots, x_b$.

As before, the probability of a particular group of siblings not being strongly linked is at most

$$b\exp\left(-\frac{(b-1)^2}{b(b+1)}(\sigma_{\text{crit}}n + \alpha \log n)\right) < b^{-(b-1)n/b}.$$

Since there are fewer than $b^k = n$ pairs of siblings in the top k layers, the probability that some such pair are not linked is at most $nb^{-(b-1)n/b} = o(1)$, so (iii) holds with high probability.

Since the events of different groups of siblings being strongly linked are independent, for any groups of siblings $x_1, \ldots, x_b$ and $y_1, \ldots, y_b$ which are k-cousins, the probability that both groups have a vertex not linked to any of the others is at most

$$b^2\exp\left(-\frac{(b-1)^2}{b(b+1)}(\sigma_{\text{crit}}n + \alpha \log n)\right) \leqslant b^{-n}n^{-\alpha(b-1)^2/(b(b+1))}.$$

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Since $\alpha > b+1$ and $b \geqslant 2$, we have $\alpha(b-1)^2/(b(b+1)) - 1 = \delta > 0$. There are at most b^n choices for $x_1, \dots, x_b$, and for any such choice at most $k(b-1)b^k$ possible choices of $y_1, \dots, y_b$, so the probability that (ii) fails is $O(n^{-\delta} \log n) = o(1)$, as required.

For $G_n^\bullet(\sigma_{\text{crit}} n - \beta \log n)$, we show that with high probability some vertex of height k is isolated in $G_n^{(k)}(\sigma_{\text{crit}} n - \beta \log n)$. Write I_v for the event that a given vertex v is isolated. Noting that

$$\xi_n^{(k)} < \Xi_k < \frac{b-1}{b+1}\left(1 + \frac{b^{-k}}{b+1}\right),$$

we have

$$\begin{aligned}
\mathbb{P}(I_v) &= \exp(-\xi_n^{(k)}(\sigma_{\text{crit}} n - \beta \log n)) \\
&> \exp\left(-\frac{b-1}{b+1}\left(1 + \frac{n^{-1}}{b+1}\right)(\sigma_{\text{crit}} n - \beta \log n)\right) \\
&= \exp\left(-n \log b + \frac{\beta(b-1)}{b+1} \log n - \frac{\log b}{b+1} + O(n^{-1} \log n)\right) \\
&= \Theta(b^{-n} n^{\beta(b-1)/(b+1)}).
\end{aligned}$$

Since this is $\omega(b^{k-n})$, by Lemma 7.1 there is at least one such v with high probability. □

7.2 *Connectedness threshold for the Poisson edge model*

In this section we consider the connectedness threshold for the Poisson edge model restricted to T_n. This gives a random graph $G_n(\lambda)$. Now write $\hat{\xi}_n^{(k)}$ for the value such that the number of edges from the descendants of a vertex at height k to other vertices has distribution $\text{Po}(\lambda \hat{\xi}_n^{(k)})$. Now

$$\begin{aligned}
\hat{\xi}_n^{(k)} &= \sum_{h=k+1}^{n} b^k (b-1) b^{h-1} b^{1-2h} \\
&= (1 - b^{k-n}).
\end{aligned}$$

Consequently $\hat{\Xi}_k = \lim_{n\to\infty} \hat{\xi}_n^{(k)} = 1$ for each k, so all the thresholds seen in the previous section coincide. Adapting Theorem 7.6 to this setting gives the following result, saying in particular that $\lambda = n \log b$ is a tight threshold for connectivity.

Theorem 7.7. *Set $\lambda = n \log b + f(n)$. Then $G_n(\lambda)$ has isolated vertices with high probability if $f(n) \to -\infty$, and with probability bounded away from 0 and 1 if $f(n) = O(1)$, whereas if $f(n) \geqslant \alpha \log n$ then $G_n(\lambda)$ is connected with high probability.*

Proof. Write X for the expected number of isolated vertices. Then

$$\begin{aligned}
\mathbb{E}(X) &= b^n \exp(-(1 - b^{-n})\lambda) \\
&= b^{nb^{-n}} \exp(-(1 - b^{-n}) f(n)) \\
&= (1 + o(1)) \exp(-(1 - b^{-n}) f(n)),
\end{aligned}$$

so $(1 + o(1))\, \mathrm{e}^{-f(n)} < \mathbb{E}(X) < (1 + o(1))\, \mathrm{e}^{-(1-b^{-1})f(n)}$. Therefore $\mathbb{E}(X) \to 0$ if $f(n) \to -\infty$, and $\mathbb{E}(X) = \Theta(1)$ if $f(n) = O(1)$. In either case, by Lemma 7.1 we have the required result.

It remains to show that $G_n(n \log b + \alpha \log n)$ is connected with high probability. We follow the same approach as in proving Theorem 7.6: set $k = \log_b n$ and show that conditions (i), (ii) and (iii) of Lemma 7.2 are satisfied. In the Poisson edge model, the probability that siblings x_i, x_j at height

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

$h-1$ are not linked is $\mathrm{e}^{-\lambda/b}$, and hence for any set of siblings $x_1,\dots,x_b$ we have $\mathbb{P}(N_{x_1,\dots,x_b}\geqslant b)=O(\mathrm{e}^{-\lambda})=O(b^n n^\alpha)=o(b^n)$, so with high probability this does not occur for any set of siblings. The probability that $x_1,\dots,x_b$ has some vertex not linked to any of the others is at most $b\,\mathrm{e}^{-\lambda(b-1)/b}=o(n^{-1})$, so with high probability none of the $b^k=n$ sets in the top k layers fail to be strongly linked, and (iii) holds.

If $x_1,\dots,x_b$ are siblings below this point, by a similar calculation to that in Theorem 7.6 we get

$$\begin{aligned}\mathbb{P}(X_{x_1,\dots,x_b}) &< b^{kb-k+1}\exp\big(-(1-b^{-k-1})(n\log b+\alpha\log n)\big)\\ &= bn^{b-1}b^{-n}n^{-\alpha}b^{1/b}n^{o(1)}.\end{aligned}$$

Since $\alpha>b-1$, this probability is $o(b^{-n})$. Consequently, with high probability this does not occur for any set of siblings, so (i) holds.

Since the events of different groups of siblings being strongly linked are independent, for any groups of siblings $x_1,\dots,x_b$ and $y_1,\dots,y_b$ which are k-cousins, the probability that both groups have a vertex not linked to any of the others is at most $b^2\,\mathrm{e}^{-2\lambda(b-1)/b}=b^2(b^{-2n}n^{-2\alpha})^{(b-1)/b}$. There are at most b^n choices for $x_1,\dots,x_b$, and for any such choice at most $k(b-1)b^k=(b-1)n\log_b n$ possible choices of $y_1,\dots,y_b$. Since $(b-1)/b\geqslant 1/2$ and $\alpha>1$, we have that the expected number of pairs $(x_1,\dots,x_b),y_1,\dots,y_b$ which are k-cousins and neither of which are strongly linked is $o(1)$. Consequently (ii) also holds with high probability. □

8. Open problems

We have proved (Theorem 1.2) an exponential lower bound and a doubly exponential upper bound on the expected size $\chi(\lambda)$ of the random graph $G(\lambda)$ of Proposition 1.1. Simulations suggest that $\chi(\lambda)$ might be of order $\lambda^{c\lambda}$. We would be very interested in any progress.

Problem 8.1. Determine the growth of $\chi(\lambda)$, or provide better bounds.

We conclude with some rather general problems for percolation on groups. A well-known conjecture of Benjamini and Schramm [3] states that every Cayley graph of a group which is not virtually $\mathbb{Z}$ satisfies $p_c<1$. A proof by Duminil-Copin, Goswami, Raoufi, Severo and Yadin [5] appeared while these lines were being written. In a similar spirit, we ask the following.

Problem 8.2. Does every countable group admit a generating measure μ with $\lambda_c(\mu)<\infty$?

Here μ and λ_c are as introduced in Section 1.1. Note that we do not have to make an exception for groups that are virtually $\mathbb{Z}$ this time, because, as shown by known results on long range percolation (see Section 1.1), $\mathbb{Z}$ does admit such a μ. Thus Problem 8.2 is only open for groups having no infinite finitely generated subgroup.

Another conjecture of [3] states that almost transitive graphs with isoperimetric dimension greater than 1 satisfy $p_c<1$, where the *isoperimetric dimension* of a graph G is defined as

$$\dim(G):=\sup\left\{d>0\;\middle|\;\inf_{S\Subset V(G)}\frac{|\partial S|}{|S|^{1-1/d}}>0\right\}. \tag{8.1}$$

Here, $S\Subset V(G)$ means that S is a finite set of vertices of G, and $|\partial S|$ denotes the number of edges with exactly one endvertex in S. In our set-up, we can define the *isoperimetric dimension* $\dim(\mu)$ of a generating measure μ similarly, by letting $|\partial S|$ denote the total measure of the edges with exactly one endvertex in S, *i.e.* by letting $|\partial S|:=\sum_{g\in S,h\notin S}\mu(g^{-1}h)$, and otherwise leaving (8.1) unchanged. We can then ask the following.

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

Problem 8.3. Does $\dim(\mu) > 1$ imply $\lambda_c(\mu) < \infty$?

In fact, the example of long range percolation suggests that a much weaker isoperimetric condition might be sufficient for percolation (see [3, Conjecture 3]).

Problem 8.4. Let μ be generating measure on a group G such that $|\partial S| > c\log|S|$ holds for some constant $c > 0$ and every $S \Subset V(G)$. Must $\lambda_c(\mu) < \infty$ hold?

We can define the *percolation threshold* of a group G by

$$\lambda_c(G) := \inf_\mu \lambda_c(\mu),$$

where the infimum ranges over all generating measures μ on G. It is not hard to see that $\lambda_c(G) \geqslant 1$ by comparing with the Poisson branching process. Some nice features of $\lambda_c(G)$ are that it is a group invariant, monotone with respect to the subgroup and the quotient group relations. Yet, it is unclear whether this is a trivial concept.

Problem 8.5. Is there a group G with $\lambda_c(G) \neq 1$?

This problem is the opposite extreme of Problem 8.2, which asks whether $\lambda_c(G) < \infty$ for every G.

A generating measure μ also naturally defines a random walk on Γ, by letting the transition probability from an element g to an element h be $\mu(g^{-1}h)$. Identifying properties of this random walk that are determined by Γ and are independent from the choice of μ is an interesting and widely studied topic [12, 19]. It would be interesting to compare the behaviour of the random walk to that of our percolation model for the same μ. In fact, the same μ can be used to define and compare with other models of statistical mechanics, *e.g.* first passage percolation, and one can consider corresponding group invariants analogous to $\lambda_c(G)$. We hope to pursue these ideas in future research.

Acknowledgements

We thank Omer Angel and Gourab Ray for the important observations mentioned above.

References

[1] Aizenman, M. and Newman, C. M. (1984) Tree graph inequalities and critical behavior in percolation models. *J. Statist. Phys.* **36** 107–143.
[2] Aizenman, M. and Newman, C. M. (1986) Discontinuity of the percolation density in one-dimensional $1/|x-y|^2$ percolation models. *Comm. Math. Phys.* **107** 611–647.
[3] Benjamini, I. and Schramm, O. (1996) Percolation beyond $\mathbb{Z}^d$, many questions and a few answers. *Electron. Comm. Probab.* **1** 71–82.
[4] Benjamini, I. and Schramm, O. (2001) Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6** 23.
[5] Duminil-Copin, H., Goswami, S., Raoufi, A., Severo, F. and Yadin, A. (2018) Existence of phase transition for percolation using the Gaussian free field. arXiv:1806.07733
[6] Erdős, P. and Rényi, A. (1959) On random graphs I. *Publ. Math. Debrecen* **6** 290–297.
[7] Georgakopoulos, A. (2016) Group-walk random graphs. In *Groups, Graphs, and Random Walks* (T. Ceccherini-Silberstein, M. Salvatori and E. Sava-Huss, eds), Vol. 436 of the London Mathematical Society Lecture Note Series, Cambridge University Press, pp. 190–204.
[8] Georgakopoulos, A. and Panagiotis, C. (2018) Analyticity results in Bernoulli percolation. arXiv:1811.07404

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press

[9] Harris, T. E. (1960) A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.* **56** 13–20.
[10] van der Hofstad, R. and Nachmias, A. (2014) Unlacing hypercube percolation: A survey. *Metrika* **77** 23–50.
[11] van der Hofstad, R. and Nachmias, A. (2017) Hypercube percolation. *J. Eur. Math. Soc. (JEMS)* **19** 725–814.
[12] Kaimanovich, V. A. and Vershik, A. M. (1983) Random walks on discrete groups: Boundary and entropy. *Ann. Probab.* **11** 457–490.
[13] Kesten, H. (1981) Analyticity properties and power law estimates of functions in percolation theory. *J. Statist. Phys.* **25** 717–756.
[14] Lovász, L. and Szegedy, B. (2012) Random graphons and a weak Positivstellensatz for graphs. *J. Graph Theory* **70** 214–225.
[15] Newman, C. M. and Schulman, L. S. (1986) One dimensional $1/|j-i|^s$ percolation models: The existence of a transition for $s \leqslant 2$. *Comm. Math. Phys.* **104** 547–571.
[16] Penrose, M. (2003) *Random Geometric Graphs*, Vol. 5 of Oxford Studies in Probability, Oxford University Press.
[17] Schulman, L. S. (1983) Long range percolation in one dimension. *J. Phys. A* **16** L639–L641.
[18] Steif, J. E. (2009) A Survey of Dynamical Percolation. In *Fractal Geometry and Stochastics IV* (C. Bandt, M. Zähle and P. Mörters, eds), Vol. 61 of Progress in Probability, Birkhäuser, pp. 145–174.
[19] Woess, W. (2002) *Random Walks on Infinite Graphs and Groups*, Vol. 138 of Cambridge Tracts in Mathematics, Cambridge University Press.

Cite this article: Georgakopoulos A and Haslegrave J (2020). Percolation on an infinitely generated group. *Combinatorics, Probability and Computing* **29**, 587–615. https://doi.org/10.1017/S096354832000005X

https://doi.org/10.1017/S096354832000005X Published online by Cambridge University Press