COVER TIME FOR BRANCHING RANDOM WALKS ON REGULAR TREES

MATTHEW I. ROBERTS ⓘ,*University of Bath

Abstract

Let T be the regular tree in which every vertex has exactly $d \geq 3$ neighbours. Run a branching random walk on T, in which at each time step every particle gives birth to a random number of children with mean d and finite variance, and each of these children moves independently to a uniformly chosen neighbour of its parent. We show that, starting with one particle at some vertex 0 and conditionally on survival of the process, the time it takes for every vertex within distance r of 0 to be hit by a particle of the branching random walk is $r + (2/\log(3/2)) \log \log r + o(\log \log r)$.

Keywords: Branching random walk; tree; cover time; rightmost particle

2020 Mathematics Subject Classification: Primary 60J80
Secondary 60D05

1. Introduction and main result

Consider a *branching random walk* (BRW) on a graph G, beginning with one particle at some vertex, where each particle branches into a random number of offspring (independently and according to some fixed distribution), each of which jumps to a uniformly chosen neighbour. The behaviour of BRW when $G = \mathbb{Z}$ is a well-studied subject starting with Hammersley [12], Kingman [14], and several papers by Biggins; see for example [4], [5], and [6]. We also highlight an early paper of Bramson [8], which contrasts with more recent results of Aidekon [1] and Bramson, Ding, and Zeitouni [9].

In this article we consider instead the case when the underlying graph G is the regular tree in which every vertex has exactly $d \geq 3$ neighbours (of course, $G = \mathbb{Z}$ can be viewed as the case $d = 2$). We suppose that the expected number of offspring of each particle in the branching random walk is also d; this is *critical* in the geometric sense that the expected number of offspring moving to each neighbouring site has mean 1. We start with one particle at the root (an arbitrary vertex) of the tree, and ask for the cover time of a ball of radius r. That is, how long does it take before every site within distance r of the root has been visited by a particle of the BRW?

To state our result precisely, let T be the infinite d-ary tree in which every vertex has d neighbours, and fix a vertex which we label 0 and refer to as the root or origin. Suppose that μ is a probability measure on $\mathbb{Z}_+$ such that $\sum_{j \geq 0} j\mu(j) = d$ and $\sum_{j \geq 0} j^2\mu(j) < \infty$. Consider a branching random walk on T, starting with one particle at the root, in which at every time step:

256

(a) each particle at any site $x \in T$ dies and gives birth to a random number of children independently and with distribution μ;

(b) each of these offspring independently jumps to a neighbour of x, uniformly at random.

For each vertex $x \in T$, let $H(x)$ be the first time at which there is a particle at x. For $r \geq 0$, let $B(r) = \{x \in T : d(0, x) \leq r\}$ and $\partial B(r) = \{x \in T : d(0, x) = \lfloor r \rfloor\}$. We are interested in the cover time of $B(r)$, defined to be

$$T_{\mathrm{cov}}(r) = \max_{x \in B(r)} H(x),$$

when r is large. Of course, if $\mu(0) > 0$, there is a positive probability that the process will die out in finite time; however, since μ has mean larger than one and finite variance, there is strictly positive probability that the process does not die out in finite time [13]. In this case we say that the process *survives*.

Theorem 1.1. *For any $d \geq 3$, given that the process survives,*

$$\lim_{r \to \infty} \frac{T_{\mathrm{cov}}(r) - r}{\log \log r} = \frac{2}{\log(3/2)} \quad \textit{almost surely.}$$

This result is initially surprising for two reasons. The first is that the cover time is so close (within a constant times $\log \log r$) to its trivial lower bound of r. However, upon reading Bramson's article [8], one can see the reason for the $\log \log r$ term, and might guess convergence of the quantity in our theorem to $2/\log 2$. Indeed, in the case $d = 2$, this is the correct answer. The appearance of $2/\log(3/2)$ instead comes from the fact that there are exponentially many vertices in $\partial B(r)$, some of which are hit unusually late; although it is interesting then that the answer does not depend on the value of $d \geq 3$. We give a short heuristic in Section 2.

The article is set out as follows. In Section 2 we give some background to Theorem 1.1 as well as a heuristic walkthrough of the proof; we also state some related open problems. In Section 3 we introduce the key ingredient for our proof of Theorem 1.1, which is a variant of our BRW in which we freeze particles that do not move in a particular direction. We then prove the lower bound for Theorem 1.1 in Section 4 and the upper bound in Section 5.

2. Notation, background, heuristics and open problems

We write $f(n) \asymp g(n)$ to mean that there exist constants $0 < c \leq C < \infty$ such that $c \leq f(n)/g(n) \leq C$ for all large n. We write $\mathbb{P}_{n,x}$ for the probability measure under which we start with n particles at the vertex x. More generally, for a collection $\Gamma = (x_1, \ldots, x_n)$ of vertices, we write $\mathbb{P}_\Gamma$ for the probability measure under which we start with a particle at each of the vertices $x_1, \ldots, x_n$. For example, $\mathbb{P}_{(x,x,x)} = \mathbb{P}_{3,x}$ and $\mathbb{P} = \mathbb{P}_{(0)} = \mathbb{P}_{1,0}$.

2.1. Background

The problem of how fast a branching random walk spreads first appeared in the mid-1970s, with papers by Hammersley [12], Kingman [14], and Biggins [4–6] giving – amongst other results – the first-order behaviour of the particle at maximal (or minimal) distance from the origin after n steps. In 1978, Bramson [8] described a branching random walk on $\mathbb{Z}_+$, beginning with one particle at 0, in which at each time step each particle branched into an average of $m > 1$ new particles, each of which stayed at its previous location with probability $1/m$ and moved one step to the right with probability $1 - 1/m$. Letting M_n be the position of the minimal

particle after n steps, he showed that

$$M_n - \left\lceil \frac{\log \log n - \log (V + \mathrm{o}(1))}{\log 2} \right\rceil \to 0 \qquad (2.1)$$

almost surely, where V is some non-trivial random variable. One of the purposes of looking at this model was that it showed significantly different behaviour from Bramson's concurrent work on branching Brownian motion (BBM) [7], demonstrating that giving a result on the detailed behaviour of M_n was much harder in general for BRW than for BBM. In fact, results in the spirit of [7] were not given for BRW in $\mathbb{R}$ until relatively recently, by Aidekon [1] and then Bramson, Ding, and Zeitouni [9]. Bramson's result (2.1) is very closely related to the cover time problem in the case $d = 2$, and we will use elements of Bramson's proof in this article.

For branching random walks on other graphs, particularly trees, much of the existing literature is concerned with recurrence and transience and related questions: see for example [11], [16], [18], and [20]. The 'maximal particle' question mentioned above for BRW on $\mathbb{R}$ has no direct analogue on trees. One could ask for the maximal distance from the origin over all particles after n steps; it is easy to see that with our choice of parameters (d-ary tree and offspring distribution mean d), conditionally on survival, this is $n - \mathrm{O}(1)$ almost surely. Studying the cover time, or equivalently the largest ball that has been covered in n steps, is an equally natural alternative, and with our choice of parameters it is a much more delicate and interesting question.

For our proof we will need the following well-known result on critical Galton–Watson processes, which is originally due to Kolmogorov [15] under a third moment assumption. See for example [17, Theorem 12.7] for a modern proof.

Lemma 2.1. (Kolmogorov.) *Suppose that $(Z_n, n \geq 0)$ is a Galton–Watson process started from $Z_0 = 1$ satisfying $\mathbb{E}[Z_1] = 1$ and $\sigma^2 := \mathbb{E}[Z_1^2] - 1 < \infty$. Then*

$$n\mathbb{P}(Z_n > 0) \to 2/\sigma^2.$$

Key to our argument will be a result on the total progeny of a Galton–Watson process up to generation n due to Pakes [19].

Lemma 2.2. (Pakes.) *Suppose that $(Z_n, n \geq 0)$ is a Galton–Watson process started from $Z_0 = 1$ satisfying $\mathbb{E}[Z_1] = 1$ and $\sigma^2 := \mathbb{E}[Z_1^2] - 1 < \infty$. Let $S_n = \sum_{i=0}^{n} Z_i$. Then, for any $\gamma \in (0, \infty)$,*

$$\mathbb{P}(S_n \geq \gamma n^2 \mid Z_n > 0) \to 1 - F(\gamma),$$

where F satisfies

$$\int_0^\infty \mathrm{e}^{-\theta v} \mathrm{d}F(v) = \sqrt{2\sigma^2 \theta} \operatorname{cosech}\left(\sqrt{2\sigma^2 \theta}\right) \quad \text{for all } \theta \in [0, \infty).$$

We will not need the precise form of F, only that $F(\gamma)$ is strictly smaller than 1 for each finite γ. As a result, combining Lemmas 2.1 and 2.2 (using the same notation), and noting that

$$\mathbb{P}\big(S_n \geq \gamma n^2\big) \geq \mathbb{P}\big(\{S_n \geq \gamma n^2\} \cap \{Z_n > 0\}\big) = \mathbb{P}\big(S_n \geq \gamma n^2 \big| Z_n > 0\big)\mathbb{P}\big(Z_n > 0\big),$$

we obtain that for each $\gamma \in [0, \infty)$ there exists a constant $q(\gamma) > 0$ depending only on γ and σ^2 such that

$$\liminf_{n \to \infty} n\mathbb{P}\big(S_n \geq \gamma n^2\big) \geq q(\gamma). \qquad (2.2)$$

We will also use the following simple and well-known Chernoff bound. Suppose that X is a finite sum of independent Bernoulli random variables. Then

$$P\left(X \leq \frac{E[X]}{2}\right) \leq \exp\left(-\frac{E[X]}{8}\right). \tag{2.3}$$

2.2. Heuristics

We now give a heuristic for Theorem 1.1. We hope that it will provide useful intuition for our proof.

Fix a vertex y in $\partial B(r)$ for large r. In order to hit y by time close to r, some particles must make long 'runs' of consecutive steps towards y without taking any steps away from y. (We associate each particle with all its ancestors, so that although technically a particle only lives for one unit of time, when we talk about it making a run of length ℓ towards y, we mean that it and its last $\ell - 1$ ancestors all stepped towards y.)

Start from one particle at 0, and let Z_i be the number of particles at time i that have taken i steps towards y. Then $(Z_i, i \geq 0)$ forms a critical Galton–Watson process. Although this process will eventually die out (likely before any particle hits y), its total progeny has infinite mean. This gives rise to a potentially large number of particles that have taken exactly *one* step away from y. Suppose this number is A. Each of these A particles starts another critical Galton–Watson tree of particles that have taken $i - 1$ steps towards y at time i. It is known that if we run A independent critical Galton–Watson processes, or equivalently one critical Galton–Watson process starting with A initial particles, then with probability of order 1 it will survive for of order A generations, with a total progeny of order A^2. This gives rise to of order A^2 particles that have taken exactly *two* steps away from y. Repeating this argument k times suggests that we should expect (very roughly) $A^{2^{k-1}}$ particles that have taken exactly k steps away from y, giving rise to another critical Galton–Watson process starting with (very roughly) $A^{2^{k-1}}$ initial particles, which survives for (very roughly) $A^{2^{k-1}}$ generations.

As soon as one of these processes – say the kth – survives for $d(0, y) + k$ generations, then y must have been hit by a particle. At the moment it hits y, this particle will have taken k steps in the wrong direction, and therefore $d(0, y) + 2k$ steps in total. In other words, if $A^{2^{k-1}} > d(0, y) + k$, then $H(y) \leq d(0, y) + 2k$. The converse is not quite true, but the fact that $A^{2^{k-1}}$ grows so quickly means that it is *almost* true, in that the correction is of smaller order. This implies that $H(y) \approx d(0, y) + (2/\log 2) \log \log d(0, y)$, which can be made rigorous and holds with high probability for each y.

We might thus expect the cover time of the ball of radius r to be roughly $r + (2/\log 2) \log \log r$. Indeed, this is essentially the explanation for the $(1/\log 2) \log \log n$ term in (2.1) (the extra factor of 2 accounts for the fact that our particles cannot stay still but must either move towards or away from y at each step), and gives the correct answer to the cover time problem when $d = 2$. The fact that we instead see $r + (2/\log(3/2)) \log \log r$ when $d \geq 3$ boils down to the fact that while *most* vertices $y \in \partial B(r)$ are hit by time $r + (2/\log 2) \log \log r$, there are many vertices in $\partial B(r)$, and some are not hit until later.

To see how this happens, again suppose that we have A particles that have taken exactly one step away from y. The probability that the resulting critical Galton–Watson process starting with A initial particles survives for fewer than $A^{1/2+\varepsilon}$ generations (for some small $\varepsilon > 0$) is roughly $(1 - c/A^{1/2+\varepsilon})^A \approx \exp(-cA^{1/2-\varepsilon})$. In doing so the particles cover distance roughly $A^{1/2+\varepsilon}$, and therefore the number of possible vertices y at this distance from the origin that could see such behaviour is of order $(d-1)^{A^{1/2+\varepsilon}}$. Since, for A large and $d > 3$, we have

$(d-1)^{A^{1/2+\varepsilon}} \gg \exp(cA^{1/2-\varepsilon})$, we might expect that some vertices *do* see such behaviour. The total number of particles seen if this occurs is of order $A \cdot A^{1/2+\varepsilon} = A^{3/2+\varepsilon}$.

Continuing recursively, we might expect that some vertices y see only $A^{(3/2+\varepsilon)^k}$ particles that have taken k steps away from y. Following the same argument as above, we deduce that these vertices have hitting times satisfying

$$H(y) \approx d(0, y) + \frac{2}{\log(3/2 + \varepsilon)} \log \log d(0, y),$$

and since $\varepsilon > 0$ was arbitrarily small, this agrees with our desired result.

This argument gives us essentially a first moment estimate on the number of vertices whose hitting times are of the order stated in Theorem 1.1. Unfortunately a naive second moment bound does not work, since the processes seen from two different vertices in the tree are highly dependent, especially if the two vertices are near each other. To get around this we use the tree structure of the graph strongly. We fix $r' < r$ and show that many vertices in $\partial B(r')$ behave 'normally', in that they are not hit too early and the number of particles moving towards them is not too large. For each of these 'normal' vertices x we fix a vertex $z(x) \in \partial B(r)$ that is in the subtree rooted at x, by which we mean that any path from 0 to z must pass through x. Using the argument above, we estimate the probability that $z(x)$ is hit later than usual and has a relatively small number of particles moving towards it (given that x is normal), and use the tree structure to get independence of these events for different vertices x. In fact, rather than just carrying out this procedure for the desired choice of r, we carry out a multi-scale argument using the scales dictated by the argument above: very roughly, $r_k \approx A^{(1/2+\varepsilon)(3/2+\varepsilon)^k}$ for each k.

2.3. Criticality of μ

We mentioned in Section 1 that our model is critical in a certain geometric sense. To explain more thoroughly, we consider a branching random walk on T where the offspring distribution μ has mean m not necessarily equal to d. It is well known that the Galton–Watson tree underlying our BRW becomes extinct almost surely when $\mu \le 1$, and survives with positive probability when $\mu > 1$. When $\mu > 1$ we may condition on survival and ask whether the whole of T is eventually covered by particles of the BRW. A standard calculation shows that the spectral radius of the d-ary tree is $\rho = (2\sqrt{d-1})/d$, and a (much more general) result of Gantert and Müller [11, Theorem 3.2] implies that when $m \le 1/\rho$ there is local extinction and some vertices in T are never hit by particles of the BRW, whereas when $m > 1/\rho$ the system is recurrent on survival. In the latter case the cover time $T_{\mathrm{cov}}(r)$ is finite for all r. Since we consider simple random walk in discrete time, $T_{\mathrm{cov}}(r)$ is at least r, and when $m > d$ it is easy to see – for example via a simple second moment calculation – that $T_{\mathrm{cov}}(r) = r + O(1)$ almost surely given survival. We expect that for $m \in (1/\rho, d)$, conditionally on survival, T_{cov} satisfies

$$\lim_{r \to \infty} \frac{T_{\mathrm{cov}}(r) - c_1 r}{\log r} = c_2 \quad \text{almost surely}$$

for some constants c_1 and c_2; in fact, $c_1 = c_1(m, d) > 1$ should satisfy

$$c_1 \log m + \left(\frac{c_1 - 1}{2}\right) \log\left(\frac{(d-1)(c_1+1)}{c_1 - 1}\right) + c_1 \log\left(\frac{2c_1}{d(c_1+1)}\right) = 0. \tag{2.4}$$

Determining the value of c_2 may be possible using techniques related to those in this paper, together with tail estimates for the maximal particle in a random walk on a one-dimensional

lattice available from [9, Proposition 3.1]; see also [9, Section 5] for discussion of lattice branching random walks.

We therefore have the following categorisation.

- $m \leq 1$: global extinction.

- $m \in (1, d/(2\sqrt{d-1})]$: global survival with positive probability, but local extinction and $T_{\text{cov}}(r) = \infty$ for all large r, almost surely.

- $m \in (d/(2\sqrt{d-1}), d)$: conjecture that on survival,

$$\lim_{r \to \infty} \frac{T_{\text{cov}}(r) - c_1 r}{\log r} = c_2 \quad \text{almost surely}$$

for $c_1 = c_1(m, d) > 1$ satisfying (2.4) and some as yet unknown c_2.

- $m = d$: Theorem 1.1 states that for $d \geq 3$, on survival,

$$\lim_{r \to \infty} \frac{T_{\text{cov}}(r) - r}{\log \log r} = \frac{2}{\log(3/2)} \quad \text{almost surely,}$$

and it follows from (2.1) that when $d = 2$, on survival

$$\lim_{r \to \infty} \frac{T_{\text{cov}}(r) - r}{\log \log r} = \frac{2}{\log 2} \quad \text{almost surely.}$$

- $m > d$: $T_{\text{cov}}(r) = r + \mathrm{O}(1)$.

2.4. Open questions

One open question is whether our main result, Theorem 1.1, can be strengthened further, perhaps along the lines of the result (2.1) of Bramson [8]. It would also be interesting to give results when the offspring distribution has mean $m \in (d/(2\sqrt{d-1}), d)$, as discussed above.

Our model considers only the simple random walk in discrete time, and one could consider other random walks, either in discrete or continuous time. One important observation in continuous time is that the random walk is no longer restricted to any finite set by a fixed time t, and therefore we will not have $T_{\text{cov}}(r) = r - \mathrm{O}(1)$ when the offspring mean is larger than d; instead we expect that $\lim_{r \to \infty} T_{\text{cov}}(r)/r > 1$, with corrections of order $\log r$, as in the (geometrically) subcritical case. The critical case when the offspring mean equals d should also show a correction of order $\log r$ in this setting, unlike the behaviour identified in Theorem 1.1. This is more 'typical' behaviour, in that only branching random walks with carefully chosen parameters will see a correction of order $\log \log r$. One example that *can* be expected to demonstrate a delay of order $\log \log r$ is the (discrete-time) lazy branching random walk with offspring distribution mean $2d$, where particles remain at their parent's location with probability $1/2$, or jump to each neighbour with probability $1/(2d)$. In this case we should have (for $d \geq 3$)

$$\lim_{r \to \infty} \frac{T_{\text{cov}}(r) - r}{\log \log r} = \frac{1}{\log(3/2)} \quad \text{almost surely,} \tag{2.5}$$

since the 'mis-steps' that break up long runs are able to remain in place, rather than stepping in the wrong direction. Alternatively one may consider a system where particles give birth to

on average d children who each jump independently to a uniformly chosen neighbour, but the original particle remains in place and gives birth to another family at each subsequent time step. Again, since particles are able to remain in place rather than either jumping towards or away from their target, we expect (2.5) to hold for $d \geq 3$. A proof of (2.5) in either of these cases would require a relatively minor alteration to the strategy used in this paper.

Another option is to extend our results to other trees. For example, what is the cover time when the underlying graph G is itself a non-trivial Galton–Watson tree? This is a much more delicate question than the one in this paper. The speed of simple random walk on (non-trivial) Galton–Watson trees with mean offspring distribution $d - 1$ is slower than on regular d-ary trees, and our proof techniques no longer apply. Work is underway to at least partially address this question.

Further, one may ask for the cover time of tree-indexed random walk on trees, where the time tree has branching number d, and the space tree has branching number d'. See [2] for the study of tree-indexed random walks and the book [17] for background. Variants of BRW and tree-indexed random walks were used in [3] and [21] to study the embedding of trees into graphs.

Finally, this paper relies heavily on the fact that the underlying graph G is a tree, so that there is only one path between any two vertices. It seems natural to expect qualitatively similar behaviour in more general settings, for example on planar hyperbolic lattices.

3. Freezing particles

3.1. Freezing particles after one step in the wrong direction

Fix a vertex x in our d-ary tree T. Consider our usual BRW (with offspring mean d) on T, but freeze any particle (i.e. prevent it from moving or branching) as soon as it either (a) takes a step away from x, or (b) reaches x, whichever happens first. In this picture, let Y_x be the number of particles that hit x, let F_x be the number of particles that are frozen as they step away from x, and let S_x be the total number of particles ever seen (until the time that all particles have become frozen). If we start with any finite collection of particles, then S_x is finite since at each step any non-frozen particle must step either towards or away from x and particles are frozen as soon as they step away from x or reach x.

For $x \in T$ and $r \geq 0$, let $T(x, r)$ be the set of vertices at distance $\lfloor r \rfloor$ from x in the subtree rooted at x; that is, those vertices at distance $\lfloor r \rfloor$ from x and $\lfloor r \rfloor + d(0, x)$ from 0. We aim to provide upper and lower bounds on the number of particles in the freezing process outlined above. First we give an upper bound in the form of the following simple expectation calculation.

Lemma 3.1. *Fix $n \in \mathbb{N}$, $R \geq 1$, and $x \in T$. Suppose that Γ consists of vertices that are at distance at most R from x. Then*

$$\mathbb{E}_\Gamma[Y_x] = |\Gamma| \quad and \quad \mathbb{E}_\Gamma[S_x] \leq (d + 1)R|\Gamma|.$$

Proof. Fix a vertex $v \in \Gamma$ and label the vertices in the path from v to x as $v = v_0, v_1, \ldots, v_k = x$. Let Z_j be the number of particles that reach v_j after j steps, for each $j \leq k$. For $j < k$, each particle at v_j independently has a random number of children with distribution μ, which has mean d and finite variance, each of which moves to v_{j+1} with probability $1/d$. Thus the sequence $(Z_j, j = 0, \ldots, k)$ forms a critical Galton–Watson process with finite variance, stopped at generation k. As a result,

$$\mathbb{E}_{1,v}[Y_x] = \mathbb{E}_{1,v}[Z_k] = 1.$$

Now, the total number of particles seen is exactly those that contribute to Z_j for $0 \leq j \leq k-1$, together with their frozen children. Thus

$$\mathbb{E}_{1,v}[S_x] \leq (d+1)\mathbb{E}_{1,v}\left[\sum_{j=0}^{k-1} Z_j\right] = (d+1)k,$$

and since $v \in \Gamma$ we have $k = d(v, x) \leq R$ so $\mathbb{E}_{1,v}[S_x] \leq (d+1)R$. To complete the proof of the lemma we simply sum over $v \in \Gamma$. $\qquad\square$

For a lower bound it is easier to bound the number of unfrozen particles than the number of frozen particles. However, we will eventually need to bound the number of frozen particles, so we will need the following lemma, which checks that if the number of unfrozen particles is large then the number of frozen particles should be large.

Lemma 3.2. *There exists a constant* $v \in (0, 1/2)$ *such that, for any* $R, M \in \mathbb{N}$ *and* $v, x \in T$ *with* $d(v, x) \leq R$,

$$\mathbb{P}_{1,v}(F_x + Y_x \leq vS_x - RM) \leq Re^{-vM}.$$

Proof. As in the proof of Lemma 3.1, label the vertices in the path from v to x as $v = v_0, v_1, \ldots, v_k = x$. For each $j \leq k-1$, let Z_j be the number of (non-frozen) particles that reach v_j after j steps, and let W_j be the number of these particles that have at least one child that is frozen as it steps away from x.

Note that for each non-frozen particle at v_j, the event that it has *no* children that step away from x is independent of other non-frozen particles at v_j, and has probability

$$\mu(0) + \mu(1)(1/d) + \mu(2)(1/d)^2 + \cdots = \mathbb{E}[1/d^L] < 1,$$

where L is a random variable with distribution μ. Therefore, for each $j \leq k-1$,

$$\mathbb{E}_{1,v}[W_j \mid Z_j] = Z_j(1 - \mathbb{E}[1/d^L]).$$

Given the value of Z_j, W_j is the sum of Z_j independent Bernoulli random variables. Thus we can apply the Chernoff bound (2.3), giving

$$\mathbb{P}_{1,v}(W_j \leq \mathbb{E}_{1,v}[W_j \mid Z_j]/2 \mid Z_j) \leq e^{-\mathbb{E}_{1,v}[W_j|Z_j]/8}.$$

Setting $\kappa = 1 - \mathbb{E}[1/d^L] \in (0, 1)$ and combining the two equations above, we have

$$\mathbb{P}_{1,v}(W_j \leq \kappa Z_j/2 \mid Z_j) \leq e^{-\kappa Z_j/8},$$

and so

$$\mathbb{P}_{1,v}(Z_j \geq M \text{ and } W_j \leq \kappa Z_j/2) \leq e^{-\kappa M/8}.$$

By a union bound, since $k = d(v, x) \leq R$,

$$\mathbb{P}_{1,v}(\exists j \leq k-1: Z_j \geq M \text{ and } W_j \leq \kappa Z_j/2) \leq ke^{-\kappa M/8} \leq Re^{-\kappa M/8}. \tag{3.1}$$

Since $S_x = \sum_{j=0}^{k-1} Z_j + F_x + Y_x$, we have

$$\mathbb{P}_{1,v}(F_x + Y_x \leq v S_x - RM) = \mathbb{P}_{1,v}\left(F_x + Y_x \leq v \sum_{j=0}^{k-1} Z_j + v F_x + v Y_x - RM\right)$$

$$= \mathbb{P}_{1,v}\left((1-v)(F_x + Y_x) \leq v \sum_{j=0}^{k-1} Z_j - RM\right)$$

$$\leq \mathbb{P}_{1,v}\left((1-v)F_x \leq v \sum_{j=0}^{k-1} Z_j - (1-v)kM\right),$$

where the last line follows trivially since $Y_x \geq 0$ and $RM \geq kM \geq (1-v)kM$. Now note that $F_x \geq \sum_{j=0}^{k-1} W_j$, so following on from the above,

$$\mathbb{P}_{1,v}(F_x + Y_x \leq v S_x - RM) \leq \mathbb{P}_{1,v}\left((1-v)\sum_{j=0}^{k-1} W_j \leq v \sum_{j=0}^{k-1} Z_j - (1-v)kM\right)$$

$$= \mathbb{P}_{1,v}\left(\sum_{j=0}^{k-1}(W_j + M) \leq \frac{v}{1-v}\sum_{j=0}^{k-1} Z_j\right)$$

$$\leq \mathbb{P}_{1,v}\left(\exists j \leq k-1 : W_j + M \leq \frac{v}{1-v} Z_j\right).$$

Choosing $v \in (0, 1/2)$ such that $v/(1-v) \leq \kappa/2$ and $v \leq \kappa/8$, the result follows from (3.1). □

We can now give our lower bound on the number of frozen particles.

Lemma 3.3. *Suppose that $A, n \in \mathbb{N}$, $x \in T$, and $y \in T(x, A)$. Suppose also that Γ consists of at least n vertices, none of which is in the subtree rooted at x except possibly at x itself, and all of which at distance at most 2A from y. Then, provided that A is sufficiently large, we have*

$$\mathbb{P}_\Gamma(F_y + Y_y \leq \delta A n) \leq e^{-\delta n/A}$$

for some constant $\delta \in (0, 1]$ depending only on d and the variance of μ.

Proof of Lemma 3.3. Without loss of generality we may assume that $|\Gamma| = n$. Label the initial n particles from 1 to n and say that particle i starts from vertex v_i. Run the y-freezing process and set X_i equal to 1 if particle i has at least A^2 descendants in total, and $X_i = 0$ otherwise. Fix $v > 0$ as in Lemma 3.2 and let X'_i equal 1 if particle i has at least $vA^2/2$ frozen descendants (i.e. descendants that contribute to either F_y or Y_y), and $X'_i = 0$ otherwise.

By the argument in the proof of Lemma 3.1, the number of non-frozen descendants of particle i after $0, 1, 2, \ldots, d(v_i, y) - 1$ steps forms a critical Galton–Watson process with finite variance, stopped at generation $d(v_i, y) - 1$. Thus the total number of non-frozen descendants of particle i is distributed as the total progeny up to generation $d(v_i, y) - 1 \geq A - 1$ of a Galton–Watson process with mean offspring number 1 and finite variance. By (2.2), the probability that a critical Galton–Watson process with finite variance has total progeny up to generation $A - 1$ of at least A^2 is at least c/A, for some constant $c > 0$ depending on the variance. Thus, for each i,

$$\mathbb{P}_\Gamma(X_i = 1) \geq c/A.$$

Also, by Lemma 3.2, choosing $M = \nu A/4$ and $R = 2A$,

$$\mathbb{P}_\Gamma(X_i = 1 \text{ but } X'_i = 0) \leq 2Ae^{-\nu^2 A/4},$$

which for A sufficiently large is at most $c/(2A)$. Therefore, for A sufficiently large,

$$\mathbb{P}_\Gamma(X'_i = 1) \geq \mathbb{P}_\Gamma(X_i = 1) - \mathbb{P}_\Gamma(X = 1 \text{ but } X' = 0) \geq \frac{c}{A} - \frac{c}{2A} = \frac{c}{2A}.$$

Letting $X = \sum_{i=1}^n X'_i$, we have

$$\mathbb{E}_\Gamma[X] \geq \frac{cn}{2A}.$$

Since X is a sum of independent Bernoulli random variables we can apply the Chernoff bound (2.3), obtaining

$$\mathbb{P}_\Gamma\left(X \leq \frac{cn}{4A}\right) \leq \exp\left(-\frac{cn}{16A}\right).$$

But if $X > cn/(4A)$, then $F_y + Y_y$ must be at least

$$\frac{cn}{4A} \cdot \frac{\nu A^2}{2} = \frac{c\nu nA}{8}.$$

Since $\nu < 1/2$, choosing $\delta = \min\{c\nu/8, 1\}$ gives the result. $\qquad\square$

3.2. Ulam–Harris notation and the FKG inequality

For our next lemma we will need to introduce the Fortuin–Kasteleyn–Ginibre (FKG) inequality in the context of a certain class of marked forests. To make our discussion precise, we need to introduce some more notation, which will also be useful in the next section. We introduce the Ulam–Harris set

$$\Omega = \bigcup_{j \in \mathbb{N}} \mathbb{N}^j.$$

The particles (we talk about *particles* to avoid confusion with the *vertices* in our spatial tree T) in our forest will be elements of Ω, with (3,2,7), for example, representing the 7th child of the 2nd child of the 3rd root. For $u, v \in \Omega$, we write $u \circ v$ for the concatenation of u and v, so for example $(3, 2, 7) \circ (1, 5) = (3, 2, 7, 1, 5)$. We write $u \leq v$ and say that u is an ancestor of v if there exists $w \in \Omega$ such that $u \circ w = v$, and say that u is in generation j if it has j ancestors, or equivalently if $u \in \mathbb{N}^{j+1}$.

A *forest* is a subset $f \subset \Omega$ such that the following hold.

- There exists $k \in \mathbb{N}$ such that $(1), (2), \ldots, (k) \in f$, but $(\ell) \notin f$ for $\ell > k$. That is, there are $k \geq 1$ roots.

- For any $u, v \in \Omega$, if $u \circ v \in f$ then $u \in f$. That is, if f contains a particle then it contains all ancestors of that particle.

- For any $u \in f$, there exists $A_u \in \{0, 1, 2, \ldots\}$ such that $u \circ (1), u \circ (2), \ldots, u \circ (A_u) \in f$ but $u \circ (\ell) \notin f$ for $\ell > A_u$. That is, every particle in f has a finite number of children.

Say that an *up–down k-forest of height at most r* is a forest with k roots and no vertices of generation greater than r, in which every particle except the root in each tree is marked either 'up' or 'down', and internal particles (i.e. those with at least one child) can only be marked

'up'. In the freezing process outlined above, if we start with k particles at $y \in T$, we can think of r as $d(y, x)$, the number of roots in the forest being k, with particles that step towards x being marked as 'up', and particles that step away from x being marked as 'down'.

Indeed, we define a probability measure $P_{k,r}$ on the set of up–down k-forests of height at most r, under which each root independently has j children in generation 1 with probability $\mu(j)$, and each child is independently marked 'up' with probability $1/d$ and 'down' otherwise. Then each of the generation 1 particles marked 'up' independently has j children in generation 2 with probability $\mu(j)$, each of which is independently marked 'up' with probability $1/d$ and 'down' otherwise. (Particles marked 'down' do not have any children.) This continues until we reach generation r, when all particles have 0 children.

For two up–down k-forests of height at most r, say F_1 and F_2, we define $F_1 \wedge F_2$ to be the up–down k-forest of height at most r given by taking the intersection of the two forests (as subsets of Ω), and marking each non-root vertex 'up' only if it is marked 'up' in both F_1 and F_2. Also define $F_1 \vee F_2$ to be the up–down k-forest of height at most r given by taking the union of the two forests (as subsets of Ω) and marking a vertex 'up' if it is marked 'up' in either F_1 or F_2. Moreover, introduce the partial order $\preceq$ by setting $F_1 \preceq F_2$ if $F_1 \wedge F_2 = F_1$.

Say that an event A is decreasing if $\mathbb{1}_A(F_1) \geq \mathbb{1}_A(F_2)$ whenever $F_1 \preceq F_2$. The FKG inequality [10] says that if A and B are both decreasing events, then

$$P_{k,r}(A \cap B) \geq P_{k,r}(A)P_{k,r}(B). \tag{3.2}$$

It is easy to check that $P_{k,r}$ satisfies the FKG lattice condition with equality, since $P_{k,r}(F)$ can be written as a product over the particles in F of the probability that that particle has the required number of children and the required mark. Indeed, the collection of offspring numbers and marks of F_1 and F_2 together equals the collection of offspring numbers and marks of $F_1 \vee F_2$ and $F_1 \wedge F_2$ together, so $P_{k,r}(F_1)P_{k,r}(F_2) = P_{k,r}(F_1 \vee F_2)P_{k,r}(F_1 \wedge F_2)$.

We deduce the following lemma, which will be needed in the next section.

Lemma 3.4. *Suppose that $y \in T$ and $x \in T(y, r)$. Then, for any $k, N \in \mathbb{N}$,*

$$\mathbb{P}_{k,y}(Y_x = 0 \text{ and } S_x \leq N) \geq \mathbb{P}_{k,y}(Y_x = 0)\mathbb{P}_{k,y}(S_x \leq N).$$

Proof. As mentioned above, the freezing process started with k particles at y corresponds to an up–down k-forest of height at most $d(y, x) = r$, where particles that step towards x are marked 'up' and particles that step away from x are marked 'down'. For F an up–down k-forest of height at most r, let $\tilde{Y}(F)$ be the number of 'up' particles in generation r and $\tilde{S}(F)$ the total number of particles in F.

Then Y_x, the number of particles frozen at x, has the same distribution as $\tilde{Y}(F)$ under $P_{k,r}$, and S_x, the total number of particles seen until all particles are frozen, has the same distribution as $\tilde{S}(F)$ under $P_{k,r}$. Thus

$$\mathbb{P}_{k,y}(Y_x = 0 \text{ and } S_x \leq N) = P_{k,r}(\tilde{Y}(F) = 0 \text{ and } \tilde{S}(F) \leq N),$$

and since the events $\{\tilde{Y}(F) = 0\}$ and $\{\tilde{S}(F) \leq N\}$ are decreasing, by (3.2) we have

$$P_{k,r}(\tilde{Y}(F) = 0 \text{ and } \tilde{S}(F) \leq N) \geq P_{k,r}(\tilde{Y}(F) = 0)P_{k,r}(\tilde{S}(F) \leq N)$$

$$= \mathbb{P}_{k,y}(Y_x = 0)\mathbb{P}_{k,y}(S_x \leq N),$$

which completes the proof. $\square$

3.3. Freezing particles after k steps in the wrong direction

Fix a vertex $x \in T$ and $k \in \mathbb{N}$, and consider our original branching random walk. We would like to generalise the discussion in the previous section, allowing ourselves to freeze any particle (i.e. prevent it from moving or branching) as soon as it either (a) takes its kth step away from x, or (b) hits x, whichever happens first; we call this the (x, k)-freezing process. However, we would like to do this in a consistent way for all vertices x, and for all k, so that the analogues of Y_x, F_x, and S_x are coupled in a natural way.

To do this we will use the Ulam–Harris labels $\Omega = \bigcup_{j \in \mathbb{N}} \mathbb{N}^j$ and the definition of a forest given in the previous section, recalling that each particle is an element of Ω, and $(3,2,7)$, for example, represents the 7th child of the 2nd child of the 3rd initial particle. Recall also that A_u is the number of children of particle u. We can then interpret our original (i.e. unfrozen) branching random walk as a labelled forest, where every particle has a label $X(u) \in T$ corresponding to its position (recall that T is our d-ary tree).

Let τ be the (random) forest generated by our branching random walk. For each particle $u \in \tau$, let $\overleftarrow{u}$ be the parent of u. To each particle $u \in \tau$ and vertex $x \in T$ we associate the following additional labels: $J_x(u)$ counts the number of steps that u has taken away from x, and $H_x(u)$ counts the number of times that u has hit x. We note that these are functions of $X(w)$ for the ancestors w of u; they do not include any new information. Then define

$$Y_x^{(k)} = \#\{u \in \tau : X(u) = x, \ J_x(u) \leq k - 1\},$$

$$F_x^{(k)} = \#\{u \in \tau : J_x(u) = k, \ J_x(\overleftarrow{u}) = k - 1, \ H_x(u) = 0\},$$

and

$$S_x^{(k)} = \#\{u \in \tau : J_x(u) \leq k - 1\} + F_x^{(k)}.$$

In the (x, k)-freezing process described above, $Y_x^{(k)}$ corresponds to the number of particles that are frozen at x, $F_x^{(k)}$ the number of particles that are frozen as they take their kth step away from x, and $S_x^{(k)}$ the total number of particles ever seen (until all particles are frozen). In particular, $Y_x^{(1)}$, $F_x^{(1)}$, and $S_x^{(1)}$ have the same distributions as the quantities Y_x, F_x, and S_x described in the previous section.

We also let $\mathcal{F}_x^k$ be the σ-algebra generated by the (x, k)-freezing process. More precisely,

$$\mathcal{F}_x^k = \sigma\left(\{A_u : J_x(u) \leq k - 1, \ H_x(u) = 0\} \cup \{X(u) : J_x(\overleftarrow{u}) \leq k - 1, \ H_x(\overleftarrow{u}) = 0\}\right).$$

In words, $\mathcal{F}_x^k$ knows everything about particles that have taken strictly fewer than k steps away from x and have not hit x, and it knows the position of particles that have just taken their kth step away from x, or just hit x for the first time (after taking fewer than k steps away from x).

We will use freezing processes in both the lower and upper bounds for Theorem 1.1. In the remainder of this section we aim to prove the following proposition, which uses Lemma 3.1 and will be used in the lower bound for Theorem 1.1.

Proposition 3.1. *Suppose that $k, n \in \mathbb{N}$, $k \leq n^{1/2}$, and $x \in T(0, n^{1/2})$. Then, on the event*

$$\left\{F_x^{(k-1)} \leq n \text{ and } Y_x^{(k-1)} = 0\right\},$$

we have

$$\mathbb{P}\left(\exists z \in T(x, An^{1/2}) : Y_z^{(k)} = 0 \text{ and } F_z^{(k)} \leq An^{3/2} \mid \mathcal{F}_x^{k-1}\right) \geq 1 - c/A^{1/2}$$

for some constant c (depending only on d and the variance of μ) and all large A and n.

We now aim to prove this result. Take $a \geq 5d$. We will use a two-stage argument, first showing that there are, with high probability, many vertices y in $T(x, an^{1/2})$ that satisfy $Y_y^{(k-1)} = 0$, $Y_y^{(k)} < an$, and $F_y^{(k)} < 3a^2n^{3/2}$. We call such vertices 'good'. Then, in the second stage, we will show that with high probability there is a vertex z in $T(y, an^{1/2})$ with our desired properties for at least one of the good vertices y.

To make this argument rigorous, define

$$M_n(x) = \left\{ y \in T(x, an^{1/2}) : Y_y^{(k-1)} = 0, \ Y_y^{(k)} < an \text{ and } F_y^{(k)} < 3a^2n^{3/2} \right\}$$

and

$$\bar{M}_n(x) = \left\{ y \in T(x, an^{1/2}) : Y_y^{(k-1)} > 0 \text{ or } Y_y^{(k)} \geq an \text{ or } F_y^{(k)} \geq 3a^2n^{3/2} \right\}.$$

That is, $M_n(x)$ is the set of good vertices and $\bar{M}_n(x)$ is its complement in $T(x, an^{1/2})$.

Our first aim is to show that $M_n(x)$ is large with high probability, on the event that $F_x^{(k-1)} \leq n$ and $Y_x^{(k-1)} = 0$.

Lemma 3.5. *Suppose that $k, n \in \mathbb{N}$, $k \leq n^{1/2}$, $a \geq 4d$, and $x \in T(0, n^{1/2})$. Then, on the event*

$$\left\{ F_x^{(k-1)} \leq n \text{ and } Y_x^{(k-1)} = 0 \right\},$$

we have

$$\mathbb{P}\left(|M_n(x)| \geq \frac{1}{2} \left| T(x, an^{1/2}) \right| \ \middle| \ \mathcal{F}_x^{k-1} \right) \geq 1 - \frac{4d}{a}.$$

Proof. Fix $y \in T(x, an^{1/2})$. Note that if we start from one particle at 0, then in order to be (y, k)-frozen, a particle must first be $(x, k-1)$ frozen. If furthermore $Y_x^{(k-1)} = 0$, then all $(x, k-1)$-frozen particles may take exactly one more step away from y before they are (y, k)-frozen. Thus, if $Y_x^{(k-1)} = 0$, running the $(y,1)$-freezing process from the starting configuration consisting of the $(x, k-1)$-frozen particles is equivalent to running the (y, k)-freezing process. Thus, letting $\Gamma = (v_1, \ldots, v_m)$ consist of the locations of the $(x, k-1)$-frozen particles, we have

$$\mathbb{P}\left(Y_y^{(k)} \geq an \text{ or } F_y^{(k)} \geq 3a^2n^{3/2} \mid \mathcal{F}_x^{k-1} \right) \mathbb{1}_{\{Y_x^{(k-1)}=0\}} = \mathbb{P}_\Gamma\left(Y_y^{(1)} \geq an \text{ or } F_y^{(1)} \geq 3a^2n^{3/2} \right). \quad (3.3)$$

Since the $(x, k-1)$-frozen particles have taken at most $k-1$ steps away from x (in fact, if $Y_x^{(k-1)} = 0$, then they have taken exactly $k-1$ steps away from x), we know that $d(v_i, y) \leq k + d(0, x) + d(x, y) \leq (1 + 1 + a)n^{1/2} \leq 3an^{1/2}$. By Lemma 3.1, we have

$$\mathbb{E}_\Gamma\left[Y_y^{(1)} \right] = \mathbb{E}_\Gamma[Y_y] = m \quad \text{and} \quad \mathbb{E}_\Gamma\left[F_y^{(1)} \right] \leq \mathbb{E}_\Gamma[S_y] \leq (d+1) \cdot 3an^{1/2} \cdot m,$$

and therefore, by Markov's inequality,

$$\mathbb{P}_\Gamma\left(Y_y^{(1)} \geq an \text{ or } F_y^{(1)} \geq 3a^2n^{3/2} \right) \leq \frac{\mathbb{E}_\Gamma\left[Y_y^{(1)} \right]}{an} + \frac{\mathbb{E}_\Gamma\left[F_y^{(1)} \right]}{3a^2n^{3/2}} \leq \frac{m}{an} + \frac{3a(d+1)mn^{1/2}}{3a^2n^{3/2}}.$$

Note that on the event

$$\left\{ F_x^{(k-1)} \leq n \text{ and } Y_x^{(k-1)} = 0 \right\}$$

we have $|\Gamma| = m \leq n$, so combining the above with (3.3), we have

$$\mathbb{P}\left(Y_y^{(k)} \geq an \text{ or } F_y^{(k)} \geq 3a^2n^{3/2} \middle| \mathcal{F}_x^{k-1} \right) \mathbb{1}_{\left\{ F_x^{(k-1)} \leq n \text{ and } Y_x^{(k-1)}=0 \right\}} \leq (2+d)/a \leq 2d/a.$$

Then, since $Y_x^{(k-1)} = 0$ implies that $Y_y^{(k-1)} = 0$, we have

$$\mathbb{E}\big[|\bar{M}_n(x)| \big| \mathcal{F}_x^{k-1}\big] \mathbb{1}_{\big\{F_x^{(k-1)} \leq n \text{ and } Y_x^{(k-1)} = 0\big\}}$$

$$= \big|T\big(x, an^{1/2}\big)\big| \cdot \mathbb{P}\big(Y_y^{(k)} \geq an \text{ or } F_y^{(k)} \geq 3a^2 n^{3/2}\big| \mathcal{F}_x^{k-1}\big) \mathbb{1}_{\big\{F_x^{(k-1)} \leq n \text{ and } Y_x^{(k-1)} = 0\big\}}$$

$$\leq \frac{2d}{a} |T(x, an^{1/2})|.$$

Thus, applying Markov's inequality again,

$$\mathbb{P}\left(|\bar{M}_n(x)| \geq \frac{1}{2} \, \big| \, T\big(x, an^{1/2}\big)\big| \, \Big| \, \mathcal{F}_x^{k-1}\right) \mathbb{1}_{\big\{F_x^{(k-1)} \leq n \text{ and } Y_x^{(k-1)} = 0\big\}} \leq \frac{4d}{a},$$

and therefore, since $|M_n(x)| + |\bar{M}_n(x)| = \big|T\big(x, an^{1/2}\big)\big|$, we have

$$\mathbb{P}\left(|M_n(x)| \geq \frac{1}{2} \, \big| \, T\big(x, an^{1/2}\big)\big| \, \Big| \, \mathcal{F}_x^{k-1}\right) \mathbb{1}_{\big\{F_x^{(k-1)} \leq n \text{ and } Y_x^{(k-1)} = 0\big\}} \geq 1 - \frac{4d}{a},$$

as required. $\qquad\qquad\square$

Next we aim to bound from below the probability that, if y is a good vertex, then there is a vertex z in $T(y, an^{1/2})$ such that $Y_z = 0$.

Lemma 3.6. *Suppose that* $k, n \in \mathbb{N}$, $k \leq n^{1/2}$, $a \geq 4d$, *and* $x \in T(0, n^{1/2})$. *There exists* $c > 0$ *such that if* $y \in M_n(x)$ *and* n *is large, then for any* $z \in T(y, an^{1/2})$ *we have*

$$\mathbb{P}\big(Y_{z(y)}^{(k)} = 0 \text{ and } F_{z(y)}^{(k)} - F_y^{(k)} \leq 2(d+1)a^2 n^{3/2}\big| \mathcal{F}_y^k\big) \geq \frac{1}{2} e^{-2cn^{1/2}}.$$

Proof. Fix $z \in T(y, an^{1/2})$. Note that if we start with one particle at 0, then in order for a particle to be (z, k)-frozen, it must also be (y, k)-frozen. In fact, to contribute to either $Y_z^{(k)}$ or $F_z^{(k)} - F_y^{(k)}$, a particle must be (y, k)-frozen at y specifically. Also, if $y \in M_n(x)$, then all particles that are (y, k)-frozen at y have taken exactly $k-1$ steps away from y, since $Y_y^{(k-1)} = 0$. Thus, if $y \in M_n(x)$, then

$$\mathbb{P}\big(Y_z^{(k)} = 0 \text{ and } F_z^{(k)} - F_y^{(k)} \leq 2(d+1)a^2 n^{3/2}\big| \mathcal{F}_y^k\big)$$

$$\geq \mathbb{P}_{\lfloor an \rfloor, y}(Y_z = 0 \text{ and } F_z \leq 2(d+1)a^2 n^{3/2}).$$

Of course $F_z \leq S_z$, so for $y \in M_n(x)$ we have

$$\mathbb{P}\big(Y_z^{(k)} = 0 \text{ and } F_z^{(k)} - F_y^{(k)} \leq 2(d+1)a^2 n^{3/2}\big| \mathcal{F}_y^k\big)$$

$$\geq \mathbb{P}_{\lfloor an \rfloor, y}(Y_z = 0 \text{ and } S_z \leq 2(d+1)a^2 n^{3/2}).$$

Applying Lemma 3.4, we deduce that for $y \in M_n(x)$,

$$\mathbb{P}\big(Y_z^{(k)} = 0 \text{ and } F_z^{(k)} - F_y^{(k)} \leq 2(d+1)a^2 n^{3/2}\big| \mathcal{F}_y^k\big)$$

$$\geq \mathbb{P}_{\lfloor an \rfloor, y}(Y_z = 0)\mathbb{P}_{\lfloor an \rfloor, y}(S_z \leq 2(d+1)a^2 n^{3/2}). \qquad (3.4)$$

Note that, starting with any number $j \in \mathbb{N}$ of particles at y,

$$\mathbb{P}_{j,y}(Y_z = 0) = \mathbb{P}_{1,y}(Y_z = 0)^j.$$

Recall from the proof of Lemma 3.3 that under $\mathbb{P}_{1,y}$, the event that Y_z equals zero is the event that a critical Galton–Watson tree (with finite variance) survives for fewer than $\lfloor an^{1/2} \rfloor$ generations; this has probability at least $1 - c/(an^{1/2})$ for some finite constant c by Lemma 2.1. Thus, for large n,

$$\mathbb{P}_{j,y}(Y_z = 0) \geq \left(1 - \frac{c}{an^{1/2}}\right)^j \geq \exp\left(-\frac{2cj}{an^{1/2}}\right),$$

where for the second inequality we used the fact that $1 - u \geq e^{-2u}$ for $0 \leq u \leq 1/2$. Also

$$\mathbb{P}_{j,y}(S_z \leq 2(d+1)a^2 n^{3/2}) = 1 - \mathbb{P}_{j,y}(S_z > 2(d+1)a^2 n^{3/2}) \geq 1 - \frac{\mathbb{E}_{j,y}[S_z]}{2(d+1)a^2 n^{3/2}},$$

so applying Lemma 3.1,

$$\mathbb{P}_{j,y}(S_z \leq 2(d+1)a^2 n^{3/2}) \geq 1 - \frac{(d+1) \cdot an^{1/2} \cdot j}{2(d+1)a^2 n^{3/2}} = 1 - \frac{j}{2an}.$$

Substituting these bounds back into (3.4), we get that for $y \in M_n(x)$,

$$\mathbb{P}(Y_z^{(k)} = 0 \text{ and } F_z^{(k)} - F_y^{(k)} \leq 2(d+1)a^2 n^{3/2} \mid \mathcal{F}_y^k) \geq \exp\left(-\frac{2can}{an^{1/2}}\right) \cdot \left(1 - \frac{an}{2an}\right)$$

$$= \frac{1}{2}e^{-2cn^{1/2}},$$

which completes the proof. $\square$

We now state a simple lemma that will be used in the proof of Proposition 3.1. Recall that $\overleftarrow{u}$ denotes the parent of particle u.

Lemma 3.7. *Suppose that $x, y \in T$ and x is not in the subtree rooted at y. If $X(u) = y$ and $H_y(\overleftarrow{u}) = 0$, then $J_x(u) \geq J_y(u) + 1$.*

Proof. Let $u_0 \leq u_1 \leq \cdots \leq u_k = u$ be the ancestors of u. Take any $j \in \{1, \ldots, k-1\}$ such that $J_y(u_j) = J_y(u_{j-1}) + 1$, i.e. such that u_j jumps away from y. Since T is a tree and $X(u) = y$, every step along the path of u that moves away from y must later be retraced. That is, there must be $\ell > j$ such that $X(u_\ell) = X(u_{j-1})$ and $X(u_{\ell-1}) = X(u_j)$, and of course this step is towards y, so $J_y(u_\ell) = J_y(u_{\ell-1})$. But then one of the two steps identified, either from $X(u_{j-1})$ to $X(u_j)$ or from $X(u_j)$ to $X(u_{j-1})$, must be away from x. In other words, either $J_x(u_j) = J_x(u_{j-1}) + 1$ or $J_x(u_\ell) = J_x(u_{\ell-1}) + 1$. We deduce that, over the jth and ℓth steps together, the path of u makes one step towards y and one step away from y, and one step towards x and another away from x. That is,

$$(J_y(u_j) - J_y(u_{j-1})) + (J_y(u_\ell) - J_y(u_{\ell-1})) = (J_x(u_j) - J_x(u_{j-1})) + (J_x(u_\ell) - J_x(u_{\ell-1})).$$

This is true for any step where $J_y(u_j)$ increases, and $J_x(u_j)$ is non-decreasing in j, so for other steps

$$J_y(u_j) - J_y(u_{j-1}) = 0 \leq J_x(u_j) - J_x(u_{j-1}).$$

As a result,

$$J_y(u_{k-1}) = \sum_{j=1}^{k-1} (J_y(u_j) - J_y(u_{j-1})) \le \sum_{j=1}^{k-1} (J_x(u_j) - J_x(u_j)) = J_x(u_{k-1}).$$

On the other hand, at the final step u steps towards y and away from x, so

$$J_y(u) = J_y(u_{k-1}) \le J_x(u_{k-1}) = J_x(u) - 1.$$

This completes the proof. □

Now we prove Proposition 3.1 by putting the estimates from Lemmas 3.5 and 3.6 together, and using Lemma 3.7 to ensure that there is enough independence between different freezing processes.

Proof of Proposition 3.1. Take $a \ge 5d$. Let

$$\mathcal{G}_x^{k,r} = \sigma\left(\bigcup_{y \in T(x,r)} \mathcal{F}_y^k\right),$$

the σ-algebra generated by the (y, k)-freezing processes for all $y \in T(x, r)$. For each $y \in T(x, an^{1/2})$, arbitrarily choose a vertex $z(y) \in T(y, an^{1/2})$. Say that $y \in T(x, an^{1/2})$ is 'special' if $Y_{z(y)}^{(k)} = 0$ and $F_{z(y)}^{(k)} - F_y^{(k)} \le 2(d+1)a^2 n^{3/2}$.

If $y \in M_n(x)$, then since $Y_y^{(k-1)} = 0$, any particle u that is (y, k)-frozen at y satisfies $J_y(u) = k - 1$ and (obviously) $X(u) = y$. Thus, by Lemma 3.7, for any other $y' \in T(x, \lfloor an^{1/2} \rfloor)$, we must have $J_{y'}(u) \ge k$. This means that no information about descendants of u is included in $\mathcal{G}_x^{k, \lfloor an^{1/2} \rfloor}$. Since only particles that are (y, k)-frozen at y can contribute to either $Y_{z(y)}^{(k)}$ or $F_{z(y)}^{(k)} - F_y^{(k)}$, we deduce that

(a) the events $\{\{y \text{ is special}\}: y \in M_n(x)\}$ are independent given $\mathcal{G}_x^{k, \lfloor an^{1/2} \rfloor}$, and

(b) $\mathbb{P}(y \text{ is special} \mid \mathcal{G}_x^{k, \lfloor an^{1/2} \rfloor}) = \mathbb{P}(y \text{ is special} | \mathcal{F}_y^k)$ for all $y \in M_n(x)$.

We note next that if y is both good and special, then $Y_{z(y)}^{(k)} = 0$ and $F_{z(y)}^{(k)} \le (2d+5)a^2 n^{3/2}$. Thus, using also that $\mathcal{F}_x^{k-1} \subset \mathcal{G}_x^{k,r}$ for any $r \ge 0$,

$$\mathbb{P}\left(\exists z \in T(x, 2an^{1/2}): Y_z^{(k)} = 0 \text{ and } F_z^{(k)} \le (2d+5)a^2 n^{3/2} \Big| \mathcal{F}_x^{k-1}\right)$$

$$\ge \mathbb{P}\left(|M_n(x)| \ge \frac{1}{2} |T(x, an^{1/2})|, \ \exists y \in M_n(x): y \text{ is special} \Big| \mathcal{F}_x^{k-1}\right)$$

$$= \mathbb{E}\left[\mathbb{1}_{\left\{|M_n(x)| \ge \frac{1}{2}|T(x, an^{1/2})|\right\}} \mathbb{P}(\exists y \in M_n(x): y \text{ is special} \mid \mathcal{G}_x^{k, \lfloor an^{1/2} \rfloor}) \Big| \mathcal{F}_x^{k-1}\right].$$

By (a) the above is at least

$$\mathbb{E}\left[\mathbb{1}_{\left\{|M_n(x)| \ge \frac{1}{2}|T(x, an^{1/2})|\right\}} \left(1 - \prod_{y \in M_n(x)} \mathbb{P}(y \text{ is not special} \mid \mathcal{G}_x^{k, \lfloor an^{1/2} \rfloor})\right) \Big| \mathcal{F}_x^{k-1}\right].$$

By (b) and Lemma 3.6, for any $y \in M_n(x)$ we have

$$\mathbb{P}\Big(y \text{ is not special}\,\Big|\,\mathcal{G}_x^{k,\lfloor an^{1/2}\rfloor}\Big) = \mathbb{P}\big(y \text{ is not special} \mid \mathcal{F}_y^k\big)$$

$$\leq 1 - \mathrm{e}^{-2cn^{1/2}}/2 \leq \exp\big(-\mathrm{e}^{-2cn^{1/2}}/2\big).$$

Putting all this together we have shown that

$$\mathbb{P}\big(\exists z \in T(x, 2an^{1/2})\colon Y_z^{(k)} = 0 \text{ and } F_z^{(k)} \leq (2d+5)a^2 n^{3/2}\big|\mathcal{F}_x^{k-1}\big)$$

$$\geq \mathbb{E}\left[\mathbb{1}_{\left\{|M_n(x)|\geq\frac{1}{2}|T(x,an^{1/2})|\right\}}\left(1 - \prod_{y\in M_n(x)} \exp\left(-\mathrm{e}^{-2cn^{1/2}}/2\right)\right)\Bigg|\,\mathcal{F}_x^{k-1}\right]$$

$$\geq \mathbb{P}\left(|M_n(x)| \geq \frac{1}{2}|T(x, an^{1/2})|\Big|\mathcal{F}_x^{k-1}\right)\big(1 - \exp\big(-|T(x, an^{1/2})|\mathrm{e}^{-2cn^{1/2}}/4\big)\big)$$

By Lemma 3.5, on the event

$$\big\{F_x^{(k-1)} \leq n \text{ and } Y_x^{(k-1)} = 0\big\},$$

this is at least

$$\left(1 - \frac{4d}{a}\right)\big(1 - \exp(-\mathrm{e}^{-2cn^{1/2}}|T(x, an^{1/2})|/4)\big).$$

Since $|T(y, an^{1/2})| \geq (d-1)^{an^{1/2}}$, we have $\exp\big(-\mathrm{e}^{-2cn^{1/2}}|T(y, an^{1/2})|/4\big) \leq 1/a$ provided that a and n are large, so the above is at least $1 - 5d/a$. And of course if $Y_z^{(k)} = 0$ and $F_z^{(k)} \leq (2d+5)a^2 n^{3/2}$, then for any $j \geq 0$, any vertex $v \in T(z, j) \subset T(x, 2an^{1/2} + j)$ also has $Y_v^{(k)} = 0$ and $F_v^{(k)} \leq (2d+5)a^2 n^{3/2}$. Thus

$$\mathbb{P}\big(\exists v \in T\big(x, (2d+5)a^2 n^{1/2}\big)\colon Y_v^{(k)} = 0 \text{ and } F_v^{(k)} \leq (2d+5)a^2 n^{3/2} \mid \mathcal{F}_x^{k-1}\big) \geq 1 - 5d/a.$$

Writing $A = (2d+5)a^2$ completes the proof. $\qquad\square$

4. Proof of the lower bound in Theorem 1.1

We aim to prove that, for any $\eta > 0$,

$$\mathbb{P}\left(\liminf_{n\to\infty} \frac{T_{\mathrm{cov}}(n) - n}{\log\log n} < \frac{2}{\log(3/2)} - \eta\right) = 0.$$

Take $\delta \in (0, 1/4)$ small and M large, both to be fixed later. For $k \geq 1$, let

$$n_k = M^{(3/2+\delta)^k}, \quad p_k = M^{-\delta(3/2+\delta)^{k-1}/2}, \quad R_k = \sum_{j=0}^{k-1} M^{(1/2+\delta)(3/2+\delta)^j}.$$

Say that $x \in \partial B(R_k)$ is *slow* if $Y_x^{(k)} = 0$ and $F_x^{(k)} \leq n_k$. Define $\mathcal{A}_k$ to be the event that there is at least one slow vertex in $\partial B(R_k)$, that is,

$$\mathcal{A}_k = \big\{\exists z \in \partial B(R_k)\colon Y_z^{(k)} = 0 \text{ and } F_z^{(k)} \leq n_k\big\}.$$

Let X_k be a uniformly chosen slow vertex in $\partial B(R_k)$, or $X_k = 0$ if there are no such vertices.

Note that when M is large, $R_{k-1} \leq n_{k-1}^{1/2}$ and $k \leq n_k^{1/2}$. Thus, setting $A = n_{k-1}^{\delta}$ so that $n_k = An_{k-1}^{3/2}$, $R_k - R_{k-1} = An_{k-1}^{1/2}$, and $p_k = A^{-1/2}$, Proposition 3.1 tells us that, for any slow $x \in \partial B(R_{k-1})$, we have

$$\mathbb{P}\big(\exists z \in T(x, R_k - R_{k-1}): Y_z^{(k)} = 0 \text{ and } F_z^{(k)} \leq n_k \mid \mathcal{F}_x^{k-1}\big) \geq 1 - cp_k.$$

In particular,

$$\mathbb{P}(\mathcal{A}_k \cap \mathcal{A}_{k-1}) = \mathbb{E}\Big[\mathbb{P}\big(\mathcal{A}_k \big| \mathcal{F}_{X_{k-1}}^{k-1}\big)\mathbb{1}_{\mathcal{A}_{k-1}}\Big] \geq (1 - cp_k)\mathbb{P}(\mathcal{A}_{k-1})$$

and therefore

$$\mathbb{P}(\mathcal{A}_k^c \cap \mathcal{A}_{k-1}) \leq cp_k\mathbb{P}(\mathcal{A}_{k-1}) \leq cp_k.$$

Since

$$\mathbb{P}\left(\bigcup_{j=1}^k \mathcal{A}_j^c\right) \leq \mathbb{P}\left(\bigcup_{j=1}^{k-1} \mathcal{A}_j^c\right) + \mathbb{P}(\mathcal{A}_k^c, \mathcal{A}_{k-1}) \leq \mathbb{P}\left(\bigcup_{j=1}^{k-1} \mathcal{A}_j^c\right) + cp_k,$$

by induction we have

$$\mathbb{P}\left(\bigcup_{j=1}^k \mathcal{A}_j^c\right) \leq c\sum_{j=2}^k p_j + \mathbb{P}(\mathcal{A}_1^c) \leq c\sum_{j=2}^\infty p_j + \mathbb{P}(\mathcal{A}_1^c).$$

Note that we can make $\mathbb{P}(\mathcal{A}_1^c)$ arbitrarily small by choosing M sufficiently large, since for large enough r *any* vertex $z \in \partial B(r)$ satisfies $Y_z = 0$ and $F_z \leq r$ with probability at least $1 - \varepsilon$.

Choose $\varepsilon > 0$ and $\delta > 0$ arbitrarily small, and M large enough that $c\sum_{j=2}^\infty p_j + \mathbb{P}(\mathcal{A}_1^c) < \varepsilon$. Then $\mathbb{P}(\bigcup_k \mathcal{A}_k^c) < \varepsilon$. On the event $\mathcal{A}_k$, there is a vertex $z \in \partial B(R_k)$ such that no particles hit z without first taking at least k steps away from z. In this case the first hitting time of z (and all its descendants in the tree) must be at least $R_k + 2k$. We deduce that

$$\mathbb{P}(T_{\text{cov}}(r) \geq r + 2k \text{ for all } r \geq R_k \text{ and } k \geq 1) > 1 - \varepsilon.$$

All that remains now is to invert R_k. Note that if

$$k \leq \frac{\log \log n}{\log (3/2 + 3\delta)}$$

then $\exp((3/2 + 3\delta)^k) \leq n$, so if n is large then $M^{(3/2+2\delta)^k} \leq n$ and indeed $R_k \leq n$. Therefore

$$\mathbb{P}\left(T_{\text{cov}}(n) \geq n + \frac{2\log \log n}{\log (3/2 + 3\delta)} \text{ for all large } n\right) > 1 - \varepsilon,$$

and since δ and ε were arbitrary, this completes the proof of the lower bound in Theorem 1.1.

5. Proof of the upper bound in Theorem 1.1

We want to show that for any $\eta > 0$

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{T_{\text{cov}}(n) - n}{\log \log n} > \frac{2}{\log (3/2)} + \eta\right) = 0.$$

Suppose that $x \in T$ and $y \in T(x, A)$, for some large $A \in \mathbb{N}$. Lemma 3.3 tells us that if Γ consists of at least n particles, none of which is in the subtree rooted at x except possibly at x itself, and none of which have distance greater than $2A$ from y, then

$$\mathbb{P}_\Gamma(F_y + Y_y \leq \delta An) \leq e^{-\delta n/A},$$

where $\delta \in (0, 1]$ is some fixed constant.

Now fix $N \in \mathbb{N}$. If we start with N particles all at 0, then (for any $k \geq 2$) none of the $(x, k-1)$-frozen particles are within the subtree rooted at x except possibly at x itself, and all have distance at most $d(0, y) + k$ from y. Thus, if $d(0, y) + k \leq 2A$, then recalling that $\mathcal{F}_x^{k-1}$ is the σ-algebra generated by the $(x, k - 1)$-freezing process, on the event

$$\{F_x^{(k-1)} + Y_x^{(k-1)} \geq n\},$$

we have

$$\mathbb{P}_{N,0}\big(F_y^{(k)} + Y_y^{(k)} \leq \delta An \mid \mathcal{F}_x^{k-1}\big) \leq e^{-\delta n/A}.$$

For each $k \in \mathbb{N}$, set

$$N_k = \frac{e^{(3/2)^k}}{\delta^2 a^2} \quad \text{and} \quad R_k = \sum_{j=1}^{k-1} aN_j^{1/2}$$

for some small $a > 0$ to be chosen later, and for $z \in T$ define the event

$$\mathcal{B}_z^k = \Big\{F_z^{(k)} + Y_z^{(k)} \geq N_k\Big\}.$$

Note that provided a is sufficiently small, we have $R_k + k \leq 2aN_{k-1}^{1/2}$ for all $k \in \mathbb{N}$. Thus, by the argument above, if $k \geq 2$, $x \in \partial B(R_{k-1})$ and $y \in T(x, aN_{k-1}^{1/2})$, then on the event $\mathcal{B}_x^{k-1}$ we have

$$\mathbb{P}_{N,0}\big(F_y^{(k)} + Y_y^{(k)} \leq \delta aN_{k-1}^{3/2} \mid \mathcal{F}_x^{k-1}\big) \leq \exp\big(-\delta N_{k-1}/aN_{k-1}^{1/2}\big).$$

Since $\delta aN_{k-1}^{3/2} = N_k$, we deduce that

$$\mathbb{P}_{N,0}\big((\mathcal{B}_y^k)^c \cap \mathcal{B}_x^{k-1}\big) \leq \exp\big(-\delta N_{k-1}^{1/2}/a\big).$$

Let $\mathcal{B}_k = \bigcap_{x \in \partial B(R_k)} \mathcal{B}_x^k$. There are $d(d-1)^{R_k-1}$ vertices in $\partial B(R_k)$, so a union bound gives

$$\mathbb{P}_{N,0}(\mathcal{B}_k^c \cap \mathcal{B}_{k-1}) \leq d(d-1)^{R_k-1} \exp\big(-\delta N_{k-1}^{1/2}/a\big).$$

Since

$$\mathbb{P}_{N,0}\left(\bigcup_{j=1}^k \mathcal{B}_j^c\right) \leq \mathbb{P}_{N,0}\left(\bigcup_{j=1}^{k-1} \mathcal{B}_j^c\right) + \mathbb{P}_{N,0}(\mathcal{B}_k^c, \mathcal{B}_{k-1}),$$

by induction we have

$$\mathbb{P}_{N,0}\left(\bigcup_{j=1}^k \mathcal{B}_j^c\right) \leq \sum_{j=2}^k d(d-1)^{R_j-1} \exp\big(-\delta N_{j-1}^{1/2}/a\big) + \mathbb{P}_{N,0}(\mathcal{B}_1^c). \tag{5.1}$$

Now fix $k \geq 1$ and suppose that $\mathcal{B}_k$ occurs, so for each $z \in \partial B(R_k)$, there are at least N_k particles that are (z, k)-frozen. All such particles have taken at most k steps away from z when

they are frozen. Each of these particles has distance at most $R_k + k$ from z, and therefore the probability that it has a descendant that hits z without taking any more steps away from z is bounded from below by the probability that a critical Galton–Watson process with finite variance survives for $R_k + k \le 2R_k$ generations. This is at least c/R_k for some constant c by Lemma 2.1. Thus the probability that *none* of the (z, k)-frozen particles has a descendant that hits z without taking any more steps away from z is at most $(1 - c/R_k)^{N_k} \le \exp(-cN_k/R_k)$. Therefore, if $H(z)$ is the first hitting time of z, we have

$$\mathbb{P}_{N,0}(H(z) > R_k + 2k \mid \mathcal{B}_k) \le \exp(-cN_k/R_k).$$

Now, if a particle hits z without taking more than k steps away from z, then for every x on the path from 0 to z, x is hit by time $d(0, x) + 2k$. Thus

$$
\begin{aligned}
&\mathbb{P}_{N,0}(\exists r \le R_k \colon T_{\mathrm{cov}}(r) > r + 2k \mid \mathcal{B}_k) \\
&= \mathbb{P}_{N,0}(\exists r \le R_k, \ x \in \partial B(r) \colon H(x) > r + 2k \mid \mathcal{B}_k) \\
&\le \mathbb{P}_{N,0}(\exists z \in \partial B(R_k) \colon H(z) > R_k + 2k \mid \mathcal{B}_k) \\
&\le d(d-1)^{R_k-1} \exp(-cN_k/R_k).
\end{aligned}
$$

(5.2)

Then

$$
\begin{aligned}
&\mathbb{P}_{N,0}\Big(\exists k \colon \max_{r \le R_k} (T_{\mathrm{cov}}(r) - r) > 2k\Big) \\
&\le \mathbb{P}_{N,0}\left(\bigcup_{j=1}^{\infty} \mathcal{B}_j^c\right) + \mathbb{P}_{N,0}\left(\{\exists k, \ r \le R_k \colon T_{\mathrm{cov}}(r) > r + 2k\} \cap \bigcap_{j=1}^{\infty} \mathcal{B}_j\right) \\
&\le \mathbb{P}_{N,0}\left(\bigcup_{j=1}^{\infty} \mathcal{B}_j^c\right) + \sum_{k=1}^{\infty} \mathbb{P}_{N,0}(\{\exists r \le R_k \colon T_{\mathrm{cov}}(r) > r + 2k\} \cap \mathcal{B}_k) \\
&\le \mathbb{P}_{N,0}\left(\bigcup_{j=1}^{\infty} \mathcal{B}_j^c\right) + \sum_{k=1}^{\infty} \mathbb{P}_{N,0}(\exists r \le R_k \colon T_{\mathrm{cov}}(r) > r + 2k \mid \mathcal{B}_k).
\end{aligned}
$$

By (5.1) and (5.2), this is at most

$$\sum_{j=2}^{\infty} d(d-1)^{R_j-1} e^{-\delta N_{j-1}^{1/2}/a} + \mathbb{P}_{N,0}(\mathcal{B}_1^c) + \sum_{k=1}^{\infty} d(d-1)^{R_k-1} \exp(-cN_k/R_k).$$

Recalling that for $k \ge 1$

$$N_k = \frac{e^{(3/2)^k}}{\delta^2 a^2} \quad \text{and} \quad R_k = \sum_{j=1}^{k-1} aN_j^{1/2},$$

we note that for any $\varepsilon > 0$, by choosing a sufficiently small, we can ensure that

$$\sum_{j=2}^{\infty} d(d-1)^{R_j-1} e^{-\delta N_{j-1}^{1/2}/a} + \sum_{k=1}^{\infty} d(d-1)^{R_k-1} \exp(-cN_k/R_k) < \varepsilon$$

and thus

$$\mathbb{P}_{N,0}\Big(\exists k\colon \max_{r\leq R_k}(T_{\mathrm{cov}}(r)-r)>2k\Big)\leq \mathbb{P}_{N,0}(\mathcal{B}_1^c)+\varepsilon.$$

Since $R_1=0$ and $N_1=\mathrm{e}^{3/2}/(\delta^2 a^2)$, we have

$$\mathcal{B}_1=\big\{F_0^{(1)}+Y_0^{(1)}\geq \mathrm{e}^{3/2}/(\delta^2 a^2)\big\}.$$

However, under $\mathbb{P}_{N,0}$, we have $F_0^{(1)}=0$ and $Y_0^{(1)}=N$, so for $N\geq \mathrm{e}^{3/2}/(\delta^2 a^2)$ we have $\mathbb{P}_{N,0}(\mathcal{B}_1)=1$ and therefore

$$\mathbb{P}_{N,0}\Big(\exists k\colon \max_{r\leq R_k}(T_{\mathrm{cov}}(r)-r)>2k\Big)<\varepsilon.$$

In our original model we started with 1 particle (rather than N particles) at the origin, but by waiting until the first time at which the number of particles at 0 is at least $\mathrm{e}^{3/2}/(\delta^2 a^2)$, which is almost surely finite given that the process survives, we may choose t large enough such that

$$\mathbb{P}_{1,0}\Big(\exists k\colon \max_{r\leq R_k}(T_{\mathrm{cov}}(r)-r)>2k+t\mid \mathrm{survival}\Big)<2\varepsilon.$$

Since $\varepsilon>0$ was arbitrary, applying this with

$$k=\frac{(1+\eta)\log\log n}{\log(3/2)}$$

for arbitrarily small $\eta>0$ completes the proof of the upper bound in Theorem 1.1.

Acknowledgements

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] AIDEKON, E. (2013). Convergence in law of the minimum of a branching random walk. *Ann. Prob.* **41**, 1362–1426.
[2] BENJAMINI, I. AND PERES, Y. (1994). Markov chains indexed by trees. *Ann. Prob.* **22**, 219–243.
[3] BENJAMINI, I. AND SCHRAMM, O. (1997). Every graph with a positive Cheeger constant contains a tree with a positive Cheeger constant. *Geom. Funct. Anal.* **7**, 403–419.
[4] BIGGINS, J. D. (1976). The first- and last-birth problems for a multitype age-dependent branching process. *Adv. Appl. Prob.* **8**, 446–459.
[5] BIGGINS, J. D. (1977). Chernoff's theorem in the branching random walk. *J. Appl. Prob.* **14**, 630–636.
[6] BIGGINS, J. D. (1978). The asymptotic shape of the branching random walk. *Adv. Appl. Prob.* **10**, 62–84.

[7] BRAMSON, M. D. (1978). Maximal displacement of branching Brownian motion. *Commun. Pure Appl. Math.* **31**, 531–581.

[8] BRAMSON, M. D. (1978). Minimal displacement of branching random walk. *Z. Wahrscheinlichkeitsth.* **45**, 89–108.

[9] BRAMSON, M. D., DING, J. AND ZEITOUNI, O. (2016). Convergence in law of the maximum of nonlattice branching random walk. *Ann. Inst. H. Poincaré Prob. Statist.* **52**, 1897–1924.

[10] FORTUIN, C., KASTELEYN, P. AND GINIBRE, J. (1971). Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.* **22**, 89–103.

[11] GANTERT, N. AND MÜLLER, S. (2006). The critical branching Markov chain is transient. *Markov Process. Relat. Fields* **12**, 805–814.

[12] HAMMERSLEY, J. M. (1974). Postulates for subadditive processes. *Ann. Prob.* **2**, 652–680.

[13] KESTEN, H. AND STIGUM, B. P. (1966). A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37**, 1211–1223.

[14] KINGMAN, J. (1975). The first birth problem for an age-dependent branching process. *Ann. Prob.* **3**, 790–801.

[15] KOLMOGOROV, A. N. (1938). On the solution of a biological problem. *Proc. Tomsk University* **2**, 7–12.

[16] LIGGETT, T. (1996). Branching random walks and contact processes on homogeneous trees. *Prob. Theory Relat. Fields* **106**, 495–519.

[17] LYONS, R. AND PERES, Y. (2016). *Probability on Trees and Networks* (Cambridge Series in Statistical and Probabilistic Mathematics **42**). Cambridge University Press, New York.

[18] MADRAS, N. AND SCHINAZI, R. (1992). Branching random walks on trees. *Stoch. Process. Appl.* **42**, 255–267.

[19] PAKES, A. G. (1971). Some limit theorems for the total progeny of a branching process. *Adv. Appl. Prob.* **3**, 176–192.

[20] PEMANTLE, R. AND STACEY, A. M. (2001). The branching random walk and contact process on Galton–Watson and nonhomogeneous trees. *Ann. Prob.* **29**, 1563–1590.

[21] SUDAKOV, B. AND VONDRÁK, J. (2007). Nearly optimal embeddings of trees. Available at arXiv:0707.2079.