

THE BRANCHING RANDOM WALK AND CONTACT PROCESS ON GALTON–WATSON AND NONHOMOGENEOUS TREES

BY ROBIN PEMANTLE¹ AND ALAN M. STACEY²

Ohio State University and Cambridge University

We show that the branching random walk on a Galton–Watson tree may have one or two phase transitions, depending on the relative sizes of the mean degree and the maximum degree. We show that there are some Galton–Watson trees on which the branching random walk has one phase transition while the contact process has two; this contradicts a conjecture of Madras and Schinazi. We show that the contact process has only one phase transition on some trees of uniformly exponential growth and bounded degree, contradicting a conjecture of Pemantle.

1. Introduction. There has been considerable interest, in recent years, in the behavior of certain stochastic processes on much more general graphs than the usual \mathbf{Z}^d . Work has been done on percolation processes, the Ising model and related models and others, as well as the processes we consider in this paper (see [13] for a recent survey of various processes and [9] for an up-to-date account of the contact process). For the contact process and branching random walk, most of this work has involved studying the processes on homogeneous trees, where they can behave in a quite different way from their behavior on \mathbf{Z}^d (see, e.g., [7, 10, 11, 16]). It is thought that the basic results for the processes on homogeneous trees should extend to some more general class of graphs, although a recent body of work on the contact process on more general graphs (see [17, 18, 20]) as well as the results of this article, does show that any such class must be more restrictive than was previously thought.

Our investigation of the processes on more general graphs than homogeneous trees begins with a partial characterization of the behavior of the branching random walk on a Galton–Watson tree in terms of the underlying offspring distribution; this enables us to demonstrate that the behavior of the processes there is not as straightforward as has been conjectured. Later we find that on another reasonably well-behaved class—spherically symmetric nonamenable trees of bounded degree—the contact process behaves unexpectedly.

Received October 1999; revised December 2000.

¹Supported in part by NSF Grant DMS-93-00191, by a Sloan Foundation Fellowship and by a Presidential Faculty Fellowship.

²Supported in part by NSF Grant DMS-97-04197 and Engineering and Physical Sciences Research Council Grant GR/L15425; carried out partly while at the University of California, Los Angeles.

AMS 2000 subject classification. 60K35.

Key words and phrases. Tree, branching random walk, contact process, phase transition, spectral radius.

We begin by defining the contact process and branching random walk on a reasonably general graph. Throughout this paper, any graph $G = (V, E)$, consisting of a set of *vertices* or *sites* V and a set of (undirected) *edges* $E \subseteq V^{(2)}$, will be assumed to be connected and of bounded degree, where, as usual, the *degree* $d(v)$ of a vertex v is the number of neighbors it has. Note also that a graph is said to be *locally finite* if $d(v) < \infty$ for every vertex v . We are always interested in infinite graphs, although in Section 4 we consider finite graphs in order to prove results about infinite ones.

The *contact process* on G is a continuous-time Markov process with state space $\{0, 1\}^V$, evolving according to the following rules: for each $v \in V$,

$$\begin{aligned} \eta_t(v) &\rightarrow 0 && \text{at rate } 1; \\ \eta_t(v) &\rightarrow 1 && \text{at rate } \lambda \cdot \#\{u \in V: u \sim v \text{ and } \eta_t(u) = 1\}; \end{aligned}$$

where λ is a fixed parameter and $u \sim v$ denotes that u is a neighbor of v . The fact that these transition rates define a unique process, given a starting state η_0 , is well known; for further details see [8]. One thinks of a site v with $\eta_t(v) = 1$ as being infected (or occupied by a particle) at time t ; one with $\eta_t(v) = 0$ is healthy (or unoccupied).

The branching random walk on G is defined similarly but now more than one infection (or particle) is permitted at a site. The state space is a subset of $\mathbf{N}^V = \{0, 1, 2, \dots\}^V$ (see remarks following) and the transition rates are

$$\begin{aligned} \eta_t(v) &\rightarrow \eta_t(v) - 1 && \text{at rate } \eta_t(v); \\ \eta_t(v) &\rightarrow \eta_t(v) + 1 && \text{at rate } \lambda \sum_{u: u \sim v} \eta_t(u); \end{aligned}$$

so if transitions outside $\{0, 1\}^V$ were forbidden, the branching random walk would reduce to the contact process. The fact that these rates determine a *unique* process does not seem to follow easily from standard results, so it appears to us that a little more work must be done to specify the branching random walk rigorously. There are various equivalent approaches to this, which we now discuss briefly. One can construct a process with these rates using the graphical representation (see the start of Section 3) and simply define the branching random walk to be this process. Another approach would be to approximate by the well-defined process on the compact space $\{0, 1, 2, \dots, n\}^V$ (again see [8]) and then let $n \rightarrow \infty$ (carefully) using monotonicity. A third approach takes advantage of an important independence property of the branching random walk (which generally makes it easier to analyze than the contact process): if ξ_t, η_t are independent branching random walks, then $\xi_t + \eta_t$ is a branching random walk with starting state $\xi_0 + \eta_0$. To make use of this property, first note that, starting from a single infection [i.e., $\sum_{v \in V} \eta_0(v) = 1$], one can define the process in the usual way as a continuous-time Markov process on the *countable* state space $\{\eta: \sum_{v \in V} \eta(v) < \infty\}$; it is not hard to see that it is nonexplosive. One then defines the process for a more general starting state by the superposition of independent processes.

Whatever approach is adopted to the definition of the branching random walk, it is worth pointing out that some restriction on the state space is necessary to prevent explosion of the process at some single site. With a little work one can show that a restriction which suffices is that for some fixed vertex v_0 (equivalently all v_0) and all $c > 0$,

$$\exp(-c \operatorname{dist}(v, v_0))\eta(v) \rightarrow 0 \quad \text{as } \operatorname{dist}(v, v_0) \rightarrow \infty.$$

We note also at this point that, as usual, \mathcal{F}_t denotes the σ -field generated by the process in question up to time t .

The most basic questions about our processes concern survival. One says that the branching random walk *survives globally* if

$$\mathbb{P}_{\delta_o}(\eta_t \equiv 0 \text{ for sufficiently large } t) < 1$$

and that the branching random walk *survives locally* if

$$\mathbb{P}_{\delta_o}(\eta_t(0) = 0 \text{ for sufficiently large } t) < 1,$$

where \mathbb{P}_{δ_o} denotes the measure for the branching random walk starting from δ_o , a single infection at a distinguished vertex O , the *root* or the *origin*. One makes similar definitions for survival of the contact process. The next proposition states some well-known basic facts. [The penultimate statement follows easily from consideration of the branching random walk on a one-vertex graph and on the connected graph with two vertices. The remaining facts can be found in ([9], I.1), once one observes that certain results for the contact process extend, essentially unchanged, to the branching random walk.]

PROPOSITION 1.1. *For the branching random walk on any connected graph of bounded degree, G , there exists $\lambda_l = \lambda_l(G) \in (0, \infty]$ such that local survival occurs for $\lambda > \lambda_l$ and not for $\lambda < \lambda_l$. There is also a $\lambda_g(G) \in (0, \lambda_l]$ such that global survival occurs for $\lambda > \lambda_g$ and not for $\lambda < \lambda_g$.*

For the contact process on any connected graph of bounded degree, G , there exists $\lambda_2 = \lambda_2(G) \in (0, \infty]$ such that local survival occurs for $\lambda > \lambda_2$ and not for $\lambda < \lambda_2$. There is also a $\lambda_1(G) \in (0, \lambda_2]$ such that global survival occurs for $\lambda > \lambda_1$ and not for $\lambda < \lambda_1$.

Moreover,

$$\begin{aligned} \lambda_l(G) < \infty &\Leftrightarrow \lambda_g(G) < \infty &\Leftrightarrow G \text{ has at least two vertices;} \\ \lambda_2(G) < \infty &\Leftrightarrow \lambda_1(G) < \infty &\Leftrightarrow G \text{ is infinite.} \end{aligned}$$

As mentioned at the start, the contact process has recently been studied extensively on the homogeneous tree: we use T_d to denote the tree in which every vertex has $d + 1$ neighbors; we shall use T'_d to denote the rooted tree in which every vertex has d children, so all vertices except the root have degree $d + 1$. The principal reason for the study of the process on trees has been that on T_d ($d \geq 3$) one has $\lambda_1 < \lambda_2$, so there is a phase of *weak survival* in which the process survives globally but not locally. This was shown by Pemantle [16] for $d \geq 4$; Liggett [10] proved this for $d = 3$ and a simpler proof was then

found by Stacey [22] which covered $d \geq 3$ and a class of nonhomogeneous trees. Other works (e.g., [7, 11, 19]) have characterized the various phases and studied their properties. The weak survival phase does not occur in the more traditional setting of \mathbf{Z}^d ; the proof [1] of this fact is far from trivial. These matters are discussed in detail in [9]. The branching random walk has also been studied on trees (e.g., [6, 11, 12, 15]), perhaps mainly because it is usually easier to analyze than the contact process but expected to exhibit similar behavior in many respects. A number of results, including the fact that $\lambda_g < \lambda_l$ for T_2 (see [15]), were proved for the branching random walk before they were proved for the contact process.

In the first half of this paper we shall mainly be concerned with processes on Galton–Watson trees. Given a probability generating function $f(z) = \sum_{n=0}^{\infty} p_n z^n$ we consider the standard branching process with offspring distribution given by f . To each realization of the branching process there naturally corresponds a rooted tree; a tree arising in this way is known as a *Galton–Watson tree*; f induces a measure, which we denote \mathbf{GW} or \mathbf{GW}_f , on the space of locally finite rooted trees. For simplicity we shall assume that $p_0 = 0$ (no extinction) and that the distribution is bounded, that is $d = \sup\{n: p_n > 0\} < \infty$. However, it would not be difficult to relax these restrictions: we can replace the first assumption by the assumption that the mean $m = \sum n p_n$ is greater than 1 and condition on survival of the process; also, it will be clear from our results what happens in the case $d = \infty$. In what follows, we shall make use of the size-biased mean m' which we now define. Given f with $0 < m < \infty$, we define the *size-biased distribution* to be that with generating function $g(z) = \frac{1}{m} \sum_{n=0}^{\infty} n p_n z^n$. The mean of this distribution is the *size-biased mean*, denoted m' . Note that $m' = \frac{1}{m} \sum n^2 p_n = g'(1) = \frac{1}{m} (zf(z))'(1) = 1 + f''(1)/f'(1)$.

Although Galton–Watson trees are inhomogeneous, they are stochastically regular, and in consequence processes on them are often better behaved and easier to study than on more general trees. They were considered in the first paper on the contact process on trees [16]; various people have attempted (without success) to show that $\lambda_1 < \lambda_2$ for Galton–Watson trees; and other processes on Galton–Watson trees have also been studied (e.g., [14]). The following result and similar results for other processes are easy and well known. (A slick proof, for the contact process, is given in [16].)

PROPOSITION 1.2. *For any bounded offspring distribution f , with $f(0) = 0$, there exist $0 < \lambda_l(f), \lambda_g(f), \lambda_1(f), \lambda_2(f) < \infty$ such that for \mathbf{GW} -almost all T , $\lambda_l(T) = \lambda_l(f)$, $\lambda_g(T) = \lambda_g(f)$, $\lambda_1(T) = \lambda_1(f)$ and $\lambda_2(T) = \lambda_2(f)$.*

When one is considering random processes on random graphs, one needs to consider a number of different measures. In order to avoid confusion, we now establish notation for these measures. We shall use \mathbb{P} to denote the measure associated with a process on a single graph, T , and use \mathbb{E} to denote the associated expectation. Since the graph in question will usually be clear, any suffix (e.g., \mathbb{P}_{δ_0}) indicates the starting distribution. (If T is itself a random variable

then \mathbb{P} and \mathbb{E} become random measures, one for each possible T ; for example, in the proof of Proposition 2.5 we show that, **GW**-almost surely, \mathbb{E} has a particular property.) When we are considering a Galton–Watson measure on trees we shall use the notation \mathbf{P}_{GW} and \mathbf{E}_{GW} . When we consider the measure corresponding to a Galton–Watson distribution on trees together with a process on the trees, we shall use the notation \mathbf{P}_{\otimes} and \mathbf{E}_{\otimes} . Therefore \mathbf{P}_{\otimes} is a measure on the space of all pairs (T, η) , with T a tree of bounded degree and η a possible realization of the process on T ; the marginal distribution of the first coordinate under \mathbf{P}_{\otimes} is \mathbf{P}_{GW} ; and the conditional distribution of the second coordinate given the first is (\otimes -almost surely) some \mathbb{P} .

Positive results about the processes for nonhomogeneous trees are rather few. It was conjectured by Pemantle [16] that the critical values for the contact process are distinct when nonzero, under a minimum growth hypothesis. To state this precisely, we need to make some further definitions. Given $G = (V, E)$ and $K \subset V$, define $\partial_E K = \{\{x, y\} \in E: x \in K, y \notin K\}$. Define the *edge-isoperimetric constant* to be

$$\iota_E(G) = \inf \left\{ \frac{|\partial_E K|}{|K|} : K \subset V \text{ is finite} \right\}.$$

Then G is said to be *amenable* if $\iota_E(G) = 0$. For many processes on well-behaved graphs, the behavior of the process depends on whether or not the graph is amenable; see [13]. To understand amenability better for a tree T , define a *shrub* to be a finite subgraph which is a component of $T \setminus \{e\}$ for some edge e . Then T is non-amenable if it contains neither arbitrarily large shrubs, nor arbitrarily long paths consisting entirely of vertices of degree (in T) two. Pemantle [16] used yet another, though equivalent, definition, called strongly exponential growth. Conjecture 2 of [16] was that for a non-amenable (locally finite) tree, one either has $\lambda_1 = \lambda_2 = 0$ (which can happen if the tree has unbounded degree) or one has $\lambda_1 < \lambda_2$. It was conjectured by Madras and Schinazi [15] that the branching random walk on a graph has distinct critical values exactly when the contact process on that graph does. (Strictly speaking, the conjecture was stated as a one-way implication, but it was equally thought to hold the other way [21].) We find below (Theorems 2.3 and 2.4) that both of these conjectures are false, and, in fact, false on fairly simple graphs. Indeed, there is a bounded offspring distribution such that the resulting Galton–Watson tree almost surely satisfies $\lambda_g = \lambda_l$ and $\lambda_1 < \lambda_2$. Our investigation gives conditions under which the branching random walk has distinct critical values in the case of a Galton–Watson tree. We show that it depends on whether the mean m is less or greater than approximately $2\sqrt{d}$, where d , as above, is the maximal number of offspring.

Conjectures 1 and 3 of [16] also generalize the behavior of the contact process on homogeneous trees to more general trees and were shown to be false in [18] and [20].

2. Statements of main results. Our first result gives a general criterion for the branching random walk on a Galton–Watson tree to have two phase transitions and a criterion under which it has only one.

THEOREM 2.1. *Let λ_g and λ_l be the critical values for global and local survival (respectively) for the branching random walk on a Galton–Watson tree whose offspring distribution has mean $m > 1$ and maximum d .*

(i) *If $m + 1 \geq 2\sqrt{d}$ then $\lambda_g < \lambda_l$.*

On the other hand,

(ii) *if $m(1 + m'/d) < 2\sqrt{d}$ where m' is the size-biased mean, then $\lambda_g = \lambda_l$.*

Since $0 < m'/d \leq 1$, this says that the boundary between one and two phase transitions occurs somewhere between $m = \sqrt{d}$ and $m = 2\sqrt{d}$. The bound in part (i) appears simpler and one might wonder whether it is sharp and whether the appearance of m' in part (ii) is spurious and nonsharp. A partial answer to this is given by Theorem 2.2, showing that part (ii) is in some sense also sharp.

THEOREM 2.2. *For any $\varepsilon > 0$ there is a $\delta > 0$ such that for any offspring distribution with maximum degree d , $\lambda_g < \lambda_l$ as long as $m \geq (2 - \delta)\sqrt{d}$ and $m' \geq \varepsilon d$.*

The proof of the following result, which shows that the branching random walk and the contact process can exhibit quite different behavior on a tree of bounded degree, relies heavily on Theorem 2.1.

THEOREM 2.3. *There is a bounded offspring distribution such that $\lambda_g = \lambda_l$ but $\lambda_1 < \lambda_2$. In particular there is a tree T of bounded degree on which $\lambda_g = \lambda_l$ but such that the contact process has two phase transitions.*

Our last main result is Theorem 2.4. Its proof makes use of detailed knowledge of the behavior of the contact process on homogeneous trees. A rooted tree is said to be *spherically symmetric* if any two vertices at the same distance from the root have the same number of children.

THEOREM 2.4. *Let λ_1 and λ_2 be the critical values for global and local survival (respectively) of the contact process. Then there is a nonamenable spherically symmetric (rooted) tree of bounded degree for which $0 < \lambda_1 = \lambda_2 < \infty$.*

It turns out to be rather easy to obtain the exact value of λ_l for a Galton–Watson tree. The same does not seem to be true for λ_g , but we do obtain, in Proposition 2.6, an upper bound which is sufficient to establish 2.1(i).

PROPOSITION 2.5. *The critical probability for local survival on a Galton–Watson tree is given by $\lambda_l = 1/(2\sqrt{d})$.*

PROPOSITION 2.6. *The critical probability for global survival on a Galton–Watson tree, λ_g , satisfies $\lambda_g \leq 1/(m + 1)$. If, furthermore, the offspring distribution is nontrivial (i.e., not concentrated entirely on a single value) then this inequality is strict.*

A supermartingale argument, given in the following section, is necessary to establish Theorem 2.1(ii). The proofs of the two preceding propositions and Theorem 2.2, being about general Galton–Watson trees, are also given in the next section. Theorems 2.3 and 2.4 are of a different flavor since they give specific trees with particular properties; they are proved in Section 4.

3. Proofs for branching random walks on Galton–Watson trees. We assume familiarity with the well-known graphical representation for the contact process on a fixed graph $G = (V, E)$. Briefly, a Poisson process of rate 1 is associated with each vertex of the graph, corresponding to deaths. Two Poisson processes, each of rate λ , are associated to each edge; these processes correspond to births, one for each direction. This was introduced by Harris [5]; for details see [4] or [8].

Given the representation for the contact process, it is not too hard to see that something similar can be done for the branching random walk, but the situation becomes rather more complicated since multiple infections at a given vertex are permitted, each independently dying and giving rise to births at neighbouring vertices. One needs a multiplicity of Poisson processes at each vertex and edge. Since this does not seem to appear in the literature, and since we shall make heavy use of the graphical representation for the branching random walk, we briefly give the details of one way the representation could be set up.

For each vertex $v \in V$, let $N(v)$ denote the set of neighbors of v ; let us choose some injection $\phi_v: N(v) \times \mathbf{Z} \times \mathbf{N} \rightarrow \mathbf{N}$. For each $m \in \mathbf{Z}$ let $\delta_{v,m}^0, \delta_{v,m}^1, \delta_{v,m}^2, \dots$ be the arrival times of a Poisson process with intensity 1. For each ordered pair (v, w) of neighbors and $m \in \mathbf{Z}$, let $a_{(v,w),m}^0, a_{(v,w),m}^1, \dots$ be the arrival times of a Poisson process of rate λ . All these Poisson processes are taken to be independent.

We shall think of each infection that occurs at a site as having some index m , with different infections at the same site having different indices; in fact initial infections (those present at time zero) will have negative indices and new infections will have nonnegative ones. The $\delta_{v,m}$ process gives the times at which the infection with index m at site v can die; the $a_{(v,w),m}$ process corresponds to possible new infections at site w arising from infection m at site v . To be precise, from a fixed starting state η_0 , an *infection trail* to (v, t) (for $v \in V$ and $t \geq 0$) is a choice of $v_0 \in V$ and m_0 with $-\eta_0(v_0) \leq m_0 < 0$,

together with a sequence

$$(3.1) \quad 0 \leq a_{(v_0, v_1), m_0}^{k_0} < a_{(v_1, v_2), m_1}^{k_1} < \dots < a_{(v_{l-1}, v), m_{l-1}}^{k_{l-1}} \leq t,$$

such that for each $i < l - 1$, $\phi_{v_{i+1}}(v_i, m_i, k_i) = m_{i+1}$; $\delta_{v_0, m_0}^0 > a_{(v_0, v_1), m_0}$; for $1 \leq i < l$ there is no j with $a_{(v_{i-1}, v_i), m_{i-1}}^{k_{i-1}} \leq \delta_{v_i, m_i}^j \leq a_{(v_i, v_{i+1}), m_i}^{k_i}$; and there is no j with $a_{(v_{l-1}, v), m_{l-1}}^{k_{l-1}} \leq \delta_{v, m_l}^j \leq t$. [If the sequence (3.1) is null, these conditions become that $v_0 = v$ and that $\delta_{v, m_0}^0 > t$.] The *length* of this infection trail is l . Although the definition of an infection trail appears rather complex, the idea is essentially the same as for the contact process, where an infection trail is a path that travels along arrows and up vertical lines, without going through any deaths; the added complication for the branching random walk comes from the fact that each infection at a site has its own arrows and deltas, and arrows must match up appropriately for an infection trail to be valid.

Having defined an infection trail, we set $\eta_t(v)$ to be the cardinality of the set of infection trails from η_0 to (v, t) . We then have that η_t is a construction of the branching random walk on G , with parameter λ and starting state η_0 .

We define a *walk* of length n , from a vertex v to a vertex w , to be a sequence of (not necessarily distinct) vertices v_i and edges e_i , $v = v_0 e_1 v_1 e_2 v_2 \dots e_n v_n = w$, such that each e_i joins v_{i-1} to v_i . To each infection trail there naturally corresponds a walk, and it is easy to see that the expected number of infection trails from $\eta_{\delta_{v_0}}$ to w (ending at any time) corresponding to any given walk of length n from v_0 to w is exactly λ^n . This calculation also applies to the counting of infection trails for the contact process, but it is not so useful there since different infection trails in the contact process do not, in general, give rise to different infections.

Counting of particular walks will be important in the next two lemmas. We let $M(v, n)$ be the number of walks of length n starting and ending at v . Note that $M = \lim_{n \rightarrow \infty} M(v, 2n)^{1/2n} = \sup_n M(v, 2n)^{1/2n}$ exists for any graph and any vertex v by an easy supermultiplicativity argument, and does not depend on the choice of v (recall that all graphs are assumed to be connected and of bounded degree).

LEMMA 3.1. *Let $G = (V, E)$ be a graph with $M = \lim_{n \rightarrow \infty} M(v, 2n)^{1/2n} = \sup_n M(v, 2n)^{1/2n}$ as above. Then the critical probability for local survival of the branching random walk is given by*

$$\lambda_l = 1/M.$$

PROOF. Fix some vertex O and consider the branching random walk on G with parameter λ , starting from a single infection at O . To begin, we consider the case $\lambda > 1/M$.

Let X_n be the number of infection trails of length n from $(O, 0)$ back to O (ending at any time). It is easy to see that the process $(X_{2kn})_{k=0}^\infty$ dominates a branching process with offspring distribution equal to the distribution of X_{2n} . However, for n sufficiently large, $M(O, 2n) > (M - \varepsilon)^{2n}$, where ε is chosen so

that $(M - \varepsilon)\lambda > 1$. Then $\mathbb{E}(X_{2n}) = \lambda^{2n} M(O, 2n) > 1$, so this branching process is supercritical. Hence there is a positive probability that there exist arbitrarily long infection trails from $(O, 0)$ back to the root, and this is well known to imply local survival. [The proof of this implication is easy once one observes, via a straightforward calculation, that the expected number of infection trails, $a_1 \cdots a_n$, with $a_n \leq T$, starting from a single infection and corresponding to a fixed walk of length n , is equal to

$$\frac{\lambda^n}{(n - 1)!} \int_0^T e^{-s} s^{n-1} ds,$$

which is less than $\lambda^n T^n/n!$, and in particular is superexponentially small in n for fixed T .]

The proof that local survival does not occur for $\lambda < 1/M$ is even easier. The expected number of infection trails from O to O of any length is

$$\sum_{n=1}^{\infty} \lambda^n M(O, n) < \infty.$$

Hence there are almost surely only finitely many infection trail back to O which implies local nonsurvival. \square

We now apply this lemma to the finite rooted tree of height h in which every vertex, other than those in generation h , has d children; we denote this finite tree by T_d^h .

LEMMA 3.2. *The critical value for local survival of the branching random walk on T_d^h satisfies*

$$\lim_{h \rightarrow \infty} \lambda_l(T_d^h) = \frac{1}{2\sqrt{d}}.$$

REMARK. This matter, and the rate of convergence in particular, is examined in more detail in [12].

PROOF. Clearly $\lambda_l(T_d^h)$ is decreasing in h and is bounded below by $\lambda_l(T_d) = 1/2\sqrt{d}$ (see [15], or the estimate on M for T_d given below). Therefore, by Lemma 3.1, we must show that given any $\varepsilon > 0$, we can find some $h > 0$ and $n > 0$ such that on T_d^h for some vertex v , $M(v, 2n)^{1/2n} > 2\sqrt{d} - \varepsilon$.

On the infinite tree T_d , a path of length $2n$ from O to O is specified by choosing n of the $2n$ steps to be down, and for each down step choosing one of the d possibilities: there are $\binom{2n}{n} d^n$ such paths; note that $\binom{2n}{n} d^n \sim c(2\sqrt{d})^{2n}/\sqrt{n}$ for some $c > 0$. Therefore for some $n > 0$ we have at least $(2\sqrt{d} - \varepsilon)^{2n}$ paths from O to O on T_d . Note that none of these paths can travel more than distance n from O . Therefore, if we consider a vertex v at generation n in T_d^{2n} we have that $M(v, 2n)^{1/2n} \geq 2\sqrt{d} - \varepsilon$, precisely as required. \square

PROOF OF PROPOSITION 2.5. We certainly have $\lambda_l \geq 1/2\sqrt{d}$ by comparison with T_d for which $\lambda_l = 1/2\sqrt{d}$ (see comment at the start of the proof of Lemma 3.2).

Now consider a Galton–Watson tree T generated by offspring distribution f whose maximum is d . Given any $k > 0$ there is, **GW**-almost surely, a vertex v such that each descendant of v up to k generations has exactly d children; that is, T contains a copy of T_d^k . Therefore, by Lemma 3.2 one has $\lambda_l(T) \leq 1/2\sqrt{d}$ almost surely, as required. \square

In the proof of Proposition 2.6 we will also need to count paths on trees, although, as we shall see, the ideas are rather different from Proposition 2.5. We shall need a lemma which counts certain walks on \mathbf{Z} ; as in the proof of Lemma 3.2, we will count walks on trees by first considering possible sequences of “ups” and “downs.” As usual, a *walk of length n on \mathbf{Z}* will be a sequence $w_0 w_1 \cdots w_n$ of integers, with $|w_i - w_{i+1}| = 1$ (for $0 \leq i < n$). In accordance with the way these walks are used to count walks on trees, we will refer to a step with $w_{i+1} > w_i$ by a *D*; if $w_{i+1} < w_i$ we refer to it by a *U*. So the walk 01012 can be described by *DUDD* (although this description forgets the starting vertex). The following walk-counting lemma is partly motivated by the crude estimate that, if the mean number of children is m , then in a typical walk on a Galton–Watson tree there should be m down steps for every up step.

LEMMA 3.3. *Let $m > 1$ be given. For $k \in \mathbf{N}$ and $\varepsilon > 0$ let $H(k, \varepsilon)$ be the number of walks, $w_0 \cdots w_l$ on \mathbf{Z} of length $l(k) = \lfloor km \rfloor + k$ with the following three properties:*

- (i) *The walk starts at 0 and ends at $n = \lfloor km \rfloor - k$ (i.e., $w_0 = 0$ and $w_l = n$), so there are $\lfloor km \rfloor$ *D*s and k *U*s.*
- (ii) *$0 \leq w_i \leq n$ for all $0 \leq i \leq l$.*
- (iii) *The walk contains at least εk *UD* pairs (i.e., values of i with $w_{i-1} = w_i + 1 = w_{i+1}$).*

Then, for some $\varepsilon > 0$, there exist $C > 0$ and γ so that for all sufficiently large k

$$(3.2) \quad H(k, \varepsilon) \geq (m+1)^k \left(\frac{m+1}{m} \right)^{km} \frac{C}{k^\gamma}.$$

PROOF. We shall construct a random walk of length l , starting at 0 and estimate the probability that it has properties (i)–(iii). The first and last $\lceil c \log k \rceil$ steps of the walk, for some c to be specified later, will always be *D*. The remaining $r = l - 2\lceil c \log k \rceil$ steps of the walk are independent and random; they are “*D*” steps with probability $p = (r - k)/r$ and are “*U*” steps with probability $q = k/r$. Note that we assume that k is large enough for all these values to be positive, and that for large k one has $p \approx m/(m+1) > 1/2$.

Let A be the event that the resulting path does not go above the starting point, that is, $w_i \geq 0$ for all i . This is just the event that the random part of the

path never goes more than $\lceil c \log k \rceil$ steps above its starting point; by standard estimates about hitting probabilities of simple (biased) random walks, one has

$$(3.3) \quad \begin{aligned} \mathbb{P}(A) &\geq 1 - (q/p)^{\lceil c \log k \rceil} \\ &\geq 1 - \frac{1}{k^{c \log(p/q)}}, \end{aligned}$$

where $(p/q) \rightarrow m$ as $k \rightarrow \infty$.

Let B be the event that the walk does not ever go below its finishing point (i.e., $w_i \leq w_l$ for all $0 \leq i \leq l$). We likewise have

$$(3.4) \quad \mathbb{P}(B) \geq 1 - \frac{1}{k^{c \log(p/q)}}.$$

Now let C_ε be the event that the walk satisfies condition (iii). We find a crude lower bound for $\mathbb{P}(C_\varepsilon)$ by considering $\lfloor r/2 \rfloor$ disjoint pairs of adjacent steps in the random part of the walk. Each such pair has probability pq to be “UD,” and these $\lfloor r/2 \rfloor$ events are independent. If C_ε does not happen, then there must be some subset of the pairs, of size exactly $\lceil \varepsilon k \rceil - 1$ with the property that all pairs outside this subset (and perhaps some others) are not UD’s. Therefore, we can bound the probability of the complement of C_ε by

$$\mathbb{P}(C_\varepsilon^c) \leq \binom{\lfloor r/2 \rfloor}{\lceil \varepsilon k \rceil - 1} (1 - pq)^{\lfloor r/2 \rfloor - \lceil \varepsilon k \rceil + 1}.$$

Using the standard estimate of $\binom{a}{b} \leq \left(\frac{ea}{b}\right)^b$ (see, e.g., [2]), this becomes

$$\mathbb{P}(C_\varepsilon^c) \leq \left(\frac{er}{2(\varepsilon k - 1)}\right)^{\varepsilon k} (1 - pq)^{r/2 - \varepsilon k + 1},$$

which becomes, for $\varepsilon < 1$ and large (dependent on ε) k ,

$$\begin{aligned} &\leq \left(\frac{er}{2(\varepsilon k - 1)}\right)^{\varepsilon k} (1 - pq)^{km/2} \\ &\leq \left[\left(\frac{e(m+1)}{2\varepsilon}\right)^\varepsilon (1 - pq)^{m/2}\right]^k. \end{aligned}$$

Now as $\varepsilon \rightarrow 0$, $\left(\frac{e(m+1)}{2\varepsilon}\right)^\varepsilon \rightarrow 1$, so for ε sufficiently small (and large k) one has

$$(3.5) \quad \mathbb{P}(C_\varepsilon^c) \leq \beta^k,$$

for $\beta = \beta(\varepsilon) < 1$.

Our final estimate is of the probability that $w_l = n$, that is, that there are precisely k U ’s and $r - k$ D ’s in the random part of the path. Denote this event

by D . One has

$$\begin{aligned}
 \mathbb{P}(D) &= \binom{r}{k} q^k p^{r-k} \\
 (3.6) \quad &\geq \frac{1}{2\sqrt{r}} \left(\frac{r}{r-k}\right)^{r-k} \left(\frac{r}{k}\right)^k q^k p^{r-k} \\
 &= \frac{1}{2\sqrt{r}},
 \end{aligned}$$

where the inequality just uses a standard lower bound on binomial coefficients (see [2]).

Combining (3.3), (3.4), (3.5) and (3.6) we obtain

$$(3.7) \quad \mathbb{P}(A \cap B \cap C \cap D) \geq \frac{1}{2\sqrt{r}} - \frac{2}{k^{c \log(p/q)}} - \beta^k.$$

We wish to now choose c so that $\sqrt{r} = o(k^{c \log(p/q)})$ as $k \rightarrow \infty$; since $r \sim k(m+1)$ and $p/q \rightarrow m$ we just take $c = 1/\log m$. Then for sufficiently large k we have

$$(3.8) \quad \mathbb{P}(A \cap B \cap C \cap D) \geq \frac{1}{3\sqrt{r}}.$$

The event $A \cap B \cap C \cap D$ is a set of paths, each of which satisfies conditions (i)–(iii) in the statement of the lemma. Each such path has probability precisely $p^{r-k} q^k$. Therefore the total number of such paths, using (3.8), is at least (again, assuming k sufficiently large throughout)

$$\begin{aligned}
 &\frac{1}{3\sqrt{r}} / p^{r-k} q^k \\
 &= \frac{1}{3\sqrt{r}} \left(\frac{l - 2\lceil c \log k \rceil}{l - k - 2\lceil c \log k \rceil}\right)^{r-k} \left(\frac{l - 2\lceil c \log k \rceil}{k}\right)^k \\
 &\geq \frac{1}{3\sqrt{k(m+1)}} \left(\frac{m+1}{m}\right)^{r-k} \left(\frac{k(m+1) - 2c \log k - 2}{k}\right)^k \\
 &\geq \frac{1}{3\sqrt{k(m+1)}} \left(\frac{m+1}{m}\right)^{km} \left(\frac{m+1}{m}\right)^{-3c \log k} \left(m+1 - \frac{(2c \log k + 2)}{k}\right)^k \\
 &\geq \frac{1}{3\sqrt{k(m+1)}} \left(\frac{m+1}{m}\right)^{km} k^{-3c \log((m+1)/m)} (m+1)^k e^{-3c \log k/(m+1)} \\
 &\geq (m+1)^k \left(\frac{m+1}{m}\right)^{km} \frac{C}{k^\gamma},
 \end{aligned}$$

[with $\gamma = \frac{1}{2} + 3c \log(\frac{m+1}{m}) + \frac{3c}{m+1}$], as required. \square

We now turn to the proof of the upper bound on the critical rate for global survival. Note that although we include the case of a degenerate offspring distribution for completeness, that case is already well known (see [15]).

PROOF OF PROPOSITION 2.6. For a nontrivial offspring distribution we must show that global survival occurs for some $\lambda < 1/(m + 1)$ and hence $\lambda_g < 1/(m + 1)$. By Proposition 2.5 we may assume that $1/(m + 1) \leq 2\sqrt{d}$. For the case of a degenerate distribution we need only show global survival for all $\lambda > 1/(m + 1)$ and we may assume $1/(m + 1) < 2\sqrt{d}$ (and, indeed, this assumption will hold since $m = d > 1$!).

We consider the branching random walk on a Galton–Watson tree and let $X_{l,n}$ be the number of sites at generation n which can be reached from $(O, 0)$ by an infection trail of length precisely l which does not go below generation n at any point (or above generation 0!). Unlike in the proof of Lemma 3.1 where X_n was defined for a fixed G (which was subsequently found embedded in a random graph), the random variable $X_{l,n}$ is defined on the space of pairs (T, η) described in Section 1; that is, there is just one random variable $X_{l,n}$ for a given offspring distribution f .

Since each site can be regarded as the root of a new Galton–Watson tree, with different sites giving rise to independent trees, we see that $(X_{kl, kn})_{k=0}^\infty$ dominates a branching process with offspring distribution equal to the distribution of $X_{l,n}$. It is crucial that $X_{l,n}$ counts sites, not infections, since two infections at the same site do not give rise to independent offspring; the (random) trees below the two infections will not be independent. The fact that we have to count sites, rather than infections, is a significant complication. If we can show that for some $l, n > 0$, $\mathbf{E}_\otimes X_{l,n} > 1$, then with positive probability there are infinite infection trails from the root at time 0, which implies global survival.

It is convenient to construct the measure **GW** by starting with a rooted tree T'_d in which every vertex has d children and then, for each vertex v , removing a random subset of the edges leading away from the root so that the size of the remaining subset has the offspring distribution and so that the subsets are exchangeable and are independent of each other. Let E' denote the random set of edges not removed. The component T containing O in the subgraph with edge set E' then has law **GW**.

Consider one of the d^n sites, v , at generation n in T'_d . We have that

$$(3.9) \quad \mathbf{E}_\otimes X_{l,n} = d^n \mathbf{P}_\otimes(X_{l,v} > 0),$$

where $X_{l,v}$ is equal to the number of infection trails of length l , starting from $(O, 0)$ and ending at v lying entirely between generations 0 and n of T . Note that when v is not itself a vertex of T the random variable $X_{l,v}$ is zero.

Our first calculation is to estimate $\mathbf{E}_\otimes X_{l,v}$, for the same choice of l and n as in Lemma 3.3 (see remarks preceding this lemma for motivation), that is, $l = \lfloor km \rfloor + k$ and $n = \lfloor km \rfloor - k$, for some large k to be determined.

In order to choose a valid walk of length $\lfloor km \rfloor + k$ to generation $\lfloor km \rfloor - k$ we must first choose $\lfloor km \rfloor$ of the steps to be down, with the restriction that we stay between generations 0 and n . If we impose an additional condition (which we will use later), namely condition (iii) of Lemma 3.3, the number of ways to choose these down steps is precisely the $H(k, \varepsilon)$ of Lemma 3.3.

Let us fix $\varepsilon > 0$ so that $H(k, \varepsilon) = H(k)$ satisfies (3.2). Now for each of these $H(k)$ choices, there are $d^{\lfloor km \rfloor}$ valid walks in T'_d . Let us fix one of the $H(k)$ choices and pick one of these $d^{\lfloor km \rfloor}$ walks uniformly at random by choosing, for each down step, one of the d possibilities with equal probability. What is the probability that this random walk consists entirely of edges lying in the random tree T ? At any up step, the walk must remain in T if it has remained in T before that point. The probability that the i th down step of the random walk lies in T (or *survives* the edge deletion process) conditional on survival so far is equal to the expectation of the conditional survival probability given the first $i - 1$ down steps on the event of survival so far. A sufficient statistic for this is j , the number of children of the present vertex visited up to that point. The conditional probability, μ_j , is the conditional expectation of the proportion of children of the present vertex which lie in T , given that the first j samples do. Since μ_j is increasing in j with $\mu_0 = m/d$, the conditional survival probability at each down step, given survival so far, is at least m/d . Hence the probability our random walk lies in T is at least

$$(m/d)^{\lfloor km \rfloor}.$$

However, for a nondegenerate offspring distribution we can improve on this. Whenever a U is followed immediately by a D , j is guaranteed to be at least 1. For each of the $H(k)$ walks on \mathbf{Z} , UD occurs at least εk times [by condition (iii) of Lemma 3.3], so the probability our random walks lies in T is at least

$$(3.10) \quad (m/d)^{\lfloor km \rfloor} \alpha^k,$$

where $\alpha = (\mu_1/\mu_0)^\varepsilon$ and, for a nondegenerate distribution, $\alpha > 1$. Therefore the expected number of walks from O to generation n , lying entirely within generations 0 to n of T , is at least $H(k)d^{\lfloor km \rfloor}(m/d)^{\lfloor km \rfloor}\alpha^k$, that is, $H(k)m^{\lfloor km \rfloor}\alpha^k$. Since the $d^{\lfloor km \rfloor - k}$ vertices at level $n = \lfloor km \rfloor - k$ are equivalent, we see that the expected number of valid walks to the particular vertex v is at least $H(k)m^{\lfloor km \rfloor}\alpha^k/d^{\lfloor km \rfloor - k}$. Hence

$$(3.11) \quad \mathbf{E}_\otimes X_{l,v} \geq \frac{\lambda^l H(k)m^{\lfloor km \rfloor}\alpha^k}{d^{\lfloor km \rfloor - k}} \geq \frac{(\lambda(m+1)\alpha^{1/(m+1)} + o(1))^{k(m+1)}}{d^{k(m-1)}} \quad \text{as } k \rightarrow \infty.$$

It follows that, for any $\lambda > 1/(\alpha^{1/(m+1)}(m+1))$, $d^n \mathbf{E}_\otimes X_{l,v}$ tends to infinity exponentially fast as $k \rightarrow \infty$; furthermore, for a nontrivial distribution, $\alpha > 1$. Of course, we wish to show that this holds with $\mathbf{E}_\otimes X_{l,v}$ replaced by $\mathbf{P}_\otimes(X_{l,v} > 0)$.

In order to estimate $\mathbf{P}_\otimes(X_{l,v} > 0)$ in terms of $\mathbf{E}_\otimes X_{l,v}$ we shall obtain bounds on $\mathbf{E}_\otimes X_{l,v}^2$. The quantity $X_{l,v}^2$ counts ordered pairs of infection trails from $(O, 0)$ to v . Given two such infection trails, t_1 and t_2 , of length l , there will be some i ($0 \leq i \leq l$) such that the first i birth arrows of t_1 are identical to the first i birth arrows of t_2 and subsequently all birth arrows are distinct; in this case we say that their *last common infection* is *after i steps*.

Letting $Y_{l,v}^i$ denote the number of pairs (t_1, t_2) of valid infection trails to v whose last common infection is after i steps, we certainly have

$$(3.12) \quad \mathbf{E}_\otimes X_{l,v}^2 = \sum_{i=0}^n \mathbf{E}_\otimes (Y_{l,v}^i).$$

We can count $\mathbf{E}_\otimes Y_{l,v}^i$ by

$$(3.13) \quad \begin{aligned} \mathbf{E}_\otimes Y_{l,v}^i &= \sum_{e_1 \cdots e_i} \lambda^i \mathbf{P}_{\mathbf{GW}}(e_1, \dots, e_i \in T) \\ &\times \sum_{\substack{e_{i+1} \cdots e_l \\ e'_{i+1} \cdots e'_l}} \mathbf{P}_{\mathbf{GW}}(e_{i+1}, \dots, e_l, e'_{i+1}, \dots, e'_l \in T \mid e_1, \dots, e_i \in T) \\ &\times \gamma_{e_{i+1}e'_{i+1}} (\lambda^{l-i-1})^2; \end{aligned}$$

the first sum is over all possible choices for the first i edges, e_1, \dots, e_i , of a valid walk from O to v ; the second sum is over pairs of choices for the remaining $l - i$ edges. The quantity $\gamma_{e_{i+1}e'_{i+1}}$ counts the expected number of pairs of distinct birth arrows (a, a') , with a along e_{i+1} in the direction of the walk and a' likewise along e'_{i+1} , before a single infection (at the common initial vertex of these two edges) dies. An easy Poisson process calculation gives

$$\gamma_{e_{i+1}e'_{i+1}} = \begin{cases} \lambda^2, & \text{if } e_{i+1} = e'_{i+1}, \\ 2\lambda^2, & \text{if } e_{i+1} \neq e'_{i+1}. \end{cases}$$

By using $2\lambda^2$ as a bound on γ , and more significantly assuming that $e'_{i+1} \cdots e'_l$ lie in T if all the other edges do, we can replace (3.13) with an upper bound on $\mathbf{E}_\otimes Y_{l,v}^i$,

$$\begin{aligned} \mathbf{E}_\otimes Y_{l,v}^i &\leq \sum_{e_1 \cdots e_i} \lambda^i \mathbf{P}_{\mathbf{GW}}(e_1, \dots, e_i \in T) \\ &\times \sum_{e_{i+1} \cdots e_l} \mathbf{P}_{\mathbf{GW}}(e_{i+1}, \dots, e_l \in T \mid e_1, \dots, e_i \in T) 2\lambda^2 (\lambda^{l-i-1})^2 N(e_i); \end{aligned}$$

where $N(e_i)$ is the number of possible choices for e'_{i+1}, \dots, e'_l . From any given vertex (in particular from the end vertex of e_i) the number of walks of length $l - i$ which end at v is at most $2^{l-i} d^{(l-i)/2}$; one can see this from the fact that there are at most two choices for whether each step goes away from, or toward, v , then at most d choices to be made for each of the $\lceil (l - i)/2 \rceil$ steps away from v . Therefore,

$$(3.14) \quad \begin{aligned} \mathbf{E}_\otimes Y_{l,v}^i &\leq \sum_{e_1 \cdots e_l} 2\lambda^{2l-i} \mathbf{P}_{\mathbf{GW}}(e_1, \dots, e_l \in T) (2\sqrt{d})^{l-i} \\ &= \sum_{e_1 \cdots e_l} 2\mathbf{P}_{\mathbf{GW}}(e_1, \dots, e_l \in T) \lambda^l (2\sqrt{d})^{l-i} \lambda^{l-i} \\ &= 2(2\sqrt{d}\lambda)^{l-i} \mathbf{E}_\otimes X_{l,v}. \end{aligned}$$

Combining (3.12) with (3.14) gives the bound on $\mathbf{E}_\otimes X_{l,v}^2$ of

$$\begin{aligned}
 \mathbf{E}_\otimes X_{l,v}^2 &\leq \mathbf{E}_\otimes X_{l,v} \sum_{i=0}^l 2(2\sqrt{d}\lambda)^{l-i} \\
 &\leq \frac{2}{1 - 2\sqrt{d}\lambda} \mathbf{E}_\otimes X_{l,v},
 \end{aligned}
 \tag{3.15}$$

the last inequality holding provided $\lambda < 1/2\sqrt{d}$. We shall apply (3.15) in the case $\lambda \in I$, where $I = (1/(\alpha^{1/(m+1)}(m+1)), 1/2\sqrt{d})$; see the remark following (3.11) for an explanation of the lower limit of I . Note that I is nonempty [by our assumption at the start of this proof that $1/(m+1) \leq 1/2\sqrt{d}$, that for the degenerate case $1/(m+1) < 1/2\sqrt{d}$ and that for the nondegenerate case $\alpha > 1$]. Note also that it is sufficient to prove global survival for $\lambda \in I$.

Letting $c = \frac{2}{1-2\sqrt{d}\lambda}$ (for $\lambda \in I$), (3.15) implies that

$$\mathbf{P}_\otimes(X_{l,v} > 0) \geq \frac{\mathbf{E}_\otimes X_{l,v}}{c}.
 \tag{3.16}$$

Combining this with (3.9) and (3.11) we obtain

$$\begin{aligned}
 \mathbf{E}_\otimes X_{l,n} &= d^n \mathbf{P}_\otimes(X_{l,v} > 0) \\
 &\geq d^n \frac{(\lambda(m+1)\alpha^{1/(m+1)} + o(1))^{k(m+1)}}{cd^{k(m-1)}} \quad \text{as } k \rightarrow \infty.
 \end{aligned}
 \tag{3.17}$$

Since $\lambda \in I$ (and recalling that $n = \lfloor km \rfloor - k$) the r.h.s. of (3.17) tends to infinity as $k \rightarrow \infty$, so for some choice of k we do indeed obtain $\mathbf{E}_\otimes X_{l,n} > 1$; this implies, by the remarks in the third paragraph of this proof, that one has global survival, as required. \square

Before we are in a position to prove Theorem 2.1 we need four further lemmas.

LEMMA 3.4. *Let η_t be a branching random walk with parameter λ and bounded starting state g , on a tree T of bounded degree. Let W be a nonnegative function on the vertices of T and suppose that $\sum_{v \in V(T)} W(v) < \infty$. Then for all t , $\mathbb{E}_g \sum W(v)\eta_t(v) < \infty$. If*

$$\lambda \sum_{w \sim v} W(w) \leq W(v)
 \tag{3.18}$$

for all v , where \sim denotes the neighbor relation, then $X_t := \sum_{v \in V(T)} W(v)\eta_t(v)$ is a supermartingale. If furthermore,

$$\lambda \sum_{w \sim v} W(w) \leq (1 - \delta)W(v)
 \tag{3.19}$$

for some $\delta > 0$, then $\mathbb{E}X_t \rightarrow 0$ exponentially fast and $X_t \rightarrow 0$ almost surely.

PROOF. Suppose that (3.18) holds. For any g satisfying $\mathbb{E}_g \sum W(v)\eta_t(v) < \infty$ (for all t),

$$\begin{aligned}
 \frac{d}{dt} \mathbb{E}_g X_t \Big|_{t=0} &= \sum_v W(v) \left[\left(\sum_{w:w\sim v} \lambda g(w) \right) - g(v) \right] \\
 &= \sum_v g(v) \left[\left(\sum_{w:w\sim v} \lambda W(w) \right) - W(v) \right] \\
 &\leq 0.
 \end{aligned}
 \tag{3.20}$$

The inequality follows from (3.18). To prove the first equality, which exchanges the order of an infinite sum and a derivative, first approximate the branching random walk by a process confined to the state space $\{0, 1, \dots, n\}^V$, as described in the Introduction; for this process the Hille–Yosida theorem (see [8]) gives the derivative of X_t for general t ; this derivative converges uniformly (as $n \rightarrow \infty$) on compact intervals, to the required value.

The above inequality shows that X_t is a supermartingale: one uses the fact that for any t one has $\sum_v W(v)\eta_t(v) < \infty$ almost surely so (3.20) holds with g replaced by η_s .

Similarly if (3.19) holds, one has $\frac{d}{dt} \mathbb{E}_g X_t \leq -\delta \mathbb{E}_g X_t$ for any t so one has $\mathbb{E} X_t \rightarrow \infty$ exponentially fast. Furthermore one has (much as above) that $(1 - \delta)^{-t} X_t$ is a supermartingale, so an application of the martingale convergence theorem implies that $X_t \rightarrow 0$ almost surely. \square

LEMMA 3.5. *Let η_t be a branching random walk on a tree T of bounded degree. Then*

$$\mathbb{E}_{\delta_0} \sum_{v \in V(T)} \eta_t(v) = \mathbb{E}_1 \eta_t(0),$$

where $\mathbf{1}$ is the constant function 1.

PROOF. For each path from w to v of length k , the expected number of offspring located at v at time t from a single particle at w at time 0 is $(\lambda t)^k e^{-t}/k!$. Since this is symmetric in w and v , we see that $\mathbb{E}_{\delta_w} \eta_t(v) = \mathbb{E}_{\delta_v} \eta_t(w)$. Thus by superposition (see discussion following the definition of the branching random walk in Section 1),

$$\mathbb{E}_1 \eta_t(0) = \sum_{v \in V(T)} \mathbb{E}_{\delta_v} \eta_t(0) = \sum_{v \in V(T)} \mathbb{E}_{\delta_0} \eta_t(v). \quad \square$$

Note that this generalizes to give a duality relation which holds for any (deterministic, bounded) starting states, η_0, ξ_0 ,

$$\mathbb{E}_{\eta_0} \left(\sum_v \eta_t(v) \xi_0(v) \right) = \mathbb{E}_{\xi_0} \left(\sum_v \xi_t(v) \eta_0(v) \right).$$

The next lemma defines the particular weighting function W that we shall use. Although W may appear at first to be somewhat opaque, it can be regarded as a refinement of the simpler weighting function $W_O(v) = (1/\sqrt{d})^{\text{dist}(O,v)}$.

W_O is sufficient to prove a weaker version of Theorem 2.1(ii), valid when $m < \sqrt{d}$. This would still provide a class of Galton–Watson trees for which $\lambda_g = \lambda_l$, but it would be insufficient for our proof of Theorem 2.3.

LEMMA 3.6. *On a rooted tree T , let $w \geq v$ denote that w is a descendant of (or equal to) v . Define*

$$L(v) := \sum_{w \geq v} d^{-\text{dist}(v, w)-1/2} 2^{-\text{dist}(v, w)-1}.$$

Define

$$W(v) := \prod_{0 < w \leq v} L(w)$$

with $W(0) = 1$. Then, provided no vertex of T has more than d children (in particular, almost surely- \mathbf{GW}_f , if f has maximum d), W satisfies (3.18) with $\lambda = 1/(2\sqrt{d})$.

PROOF. Define $S(v) = \sum_{v \rightarrow w} L(w)$, where the sum is over children of v . From the definitions of L and S , we see that for any $v \in V(T)$,

$$\begin{aligned} L(v) &= \frac{1}{2\sqrt{d}} + \sum_{v \rightarrow w} \sum_{y \geq w} d^{-\text{dist}(y, w)-3/2} 2^{-\text{dist}(y, w)-2} \\ &= \frac{1}{2\sqrt{d}} + \frac{S(v)}{2d}. \end{aligned}$$

Observe that $L(v) \leq \sum_{n=0}^{\infty} d^n d^{-n-1/2} 2^{-n-1} = d^{-1/2}$ and hence that $S(v) \leq \sqrt{d}$. The function $\phi(x) := (2\sqrt{d} - x)^{-1}$ is convex on $[0, \sqrt{d}]$ and hence bounded above by its chord,

$$\phi(0) + \frac{x}{\sqrt{d}}(\phi(\sqrt{d}) - \phi(0)) = \frac{1}{2\sqrt{d}} + \frac{x}{2d}.$$

Hence

$$(3.21) \quad \frac{1}{2\sqrt{d} - S(v)} \leq \frac{1}{2\sqrt{d}} + \frac{S(v)}{2d} = L(v).$$

Let y be any vertex of T with a child v and let w_1, \dots, w_r be the children of v . From (3.21), we then see that

$$\frac{W(y)}{W(v)} = \frac{1}{L(v)} \leq 2\sqrt{d} - S(v) = 2\sqrt{d} - \sum_{j=1}^r \frac{W(w_j)}{W(v)}.$$

Multiplying through by $W(v)/(2\sqrt{d})$ then establishes (3.18). \square

Lemma 3.6 is sufficient to show that for any tree of maximum degree at most d , $\lambda_l \geq 1/(2\sqrt{d})$. However, in order to apply Lemma 3.4 to show that $\lambda_g \geq 1/(2\sqrt{d})$ we need to know that $\sum_v W(v) < \infty$; the fact that this holds

for Galton–Watson trees, under certain conditions, is the content of the next lemma.

LEMMA 3.7. *Let m' be the size-biased mean of the offspring distribution; let W be as in the statement of Lemma 3.6. When*

$$m\left(1 + \frac{m'}{d}\right) < 2\sqrt{d}$$

*then $\sum_{v \in V(T)} W(v) < \infty$ almost surely-**GW**. In particular, the conclusion holds eventually for any family of offspring distributions with $m \leq (2 - \varepsilon)\sqrt{d}$ and $m' = o(d)$ as $d \rightarrow \infty$.*

PROOF. As described in the proof of Proposition 2.6, it is convenient to start with T'_d and delete edges at random, leaving an edge set E' in such a way that the component, T , containing the origin has law **GW**.

To show that $\sum_v W(v)$ is almost surely finite, it suffices to show it has finite expectation, which in turn will follow from an exponential bound

$$\mathbf{E}_{\mathbf{GW}} \sum_{v: \text{dist}(0,v)=n} W(v) \leq e^{-cn}.$$

From the construction of **GW** on the d -ary tree, we see that

$$(3.22) \quad \mathbf{E}_{\mathbf{GW}} \sum_{v: \text{dist}(0,v)=n} W(v) = d^n \mathbf{E}_{\mathbf{GW}} [W(v_n) \mathbf{1}(v_n \in T)],$$

where v_n is the leftmost vertex in generation n . Note that the value of W on vertices not in T need not be defined for this to make sense.

To bound this expectation above, we define a function $L(n)$ that will serve as an upper bound for $L(v_n)$. Let $D(n)$ be the number of children of v_n , that is, the number of edges adjacent to v_n in E' that lead away from the root. Let $D^{(2)}(n)$ be the number of grandchildren of v_n that are not children of v_{n+1} . Define

$$L(n) = \frac{1}{2\sqrt{d}} + \frac{1}{4d^{3/2}}(D(n) + 1) + \frac{1}{4d^{5/2}}D^{(2)}(n).$$

To see that $L(n)$ is an upper bound for $L(v_n)$, note that $L(n) = L(v_n)$ when E' contains every edge below v_{n+1} and every edge below every grandchild of v .

The value of $L(n)$ depends on the presence or absence of edges leading out of v_n and out of the children of v_n other than v_{n+1} , where edges are directed away from the root. The presence or absence of different edges are independent unless the edges lead out of the same vertex. If $n \neq m$, the variables $L(n)$ and $L(m)$ are measurable with respect to edges leading out of disjoint sets of vertices, hence the variables $\{L(n)\}$ are independent. Similar reasoning shows they are conditionally independent given $\{v_N \in T\}$ for any $N > n$. Recall that

$$\mathbf{E}_{\mathbf{GW}} W(v_n) \mathbf{1}(v_N \in T) \leq \mathbf{E}_{\mathbf{GW}} \mathbf{1}(v_N \in T) \prod_{j=1}^n L(j) = \frac{m^n}{d^n} \prod_{j=1}^n \mathbf{E}_{\mathbf{GW}}(L(j) | v_N \in T).$$

Write

$$\begin{aligned} \mathbf{E}_{\mathbf{GW}}(L(n) | v_N \in T) &= \frac{1}{2\sqrt{d}} + \frac{1}{4d^{3/2}} \mathbf{E}_{\mathbf{GW}}(D(n) + 1 | v_N \in T) \\ &\quad + \frac{1}{4d^{5/2}} \mathbf{E}_{\mathbf{GW}}(D^{(2)}(n) | v_N \in T). \end{aligned}$$

It is not hard to see that conditioning on $v_N \in T$ yields a random subtree in which vertices $0, v_1, \dots, v_{N-1}$ have numbers of offspring that are independent picks from the *size-biasing* of the offspring distribution and all other vertices have offspring picked from the (not size-biased) offspring distribution. Letting m' denote the mean of the size-biased distribution, this gives

$$\mathbf{E}_{\mathbf{GW}}(L(n) | v_N \in T) = \frac{1}{2\sqrt{d}} + \frac{1 + m'}{4d^{3/2}} + \frac{m(m' - 1)}{4d^{5/2}},$$

and hence

$$\mathbf{E}_{\mathbf{GW}} W(v_n) \mathbf{1}(v_N \in T) \leq \frac{m^n}{d^n} \prod_{j=1}^n \frac{1}{2\sqrt{d}} \left(1 + \frac{1 + m' + (m' - 1)m/d}{2d} \right).$$

From (3.22) we then have

$$\left[\mathbf{E}_{\mathbf{GW}} \sum_{v: \text{dist}(0, v)=n} W(v) \right]^{1/n} \leq \frac{m}{2\sqrt{d}} \left(1 + \frac{m'}{d} \right).$$

Thus the hypothesis

$$\left(1 + \frac{m'}{d} \right) m < 2\sqrt{d}$$

implies the exponential decrease of $\mathbf{E}_{\mathbf{GW}} \sum_{v: \text{dist}(0, v)=n} W(v)$, which implies the conclusion of the lemma. \square

Having proved Proposition 2.5 to Lemma 3.7, the proof of Theorem 2.1 is now quite short.

PROOF OF THEOREM 2.1. Part (i) follows directly from Proposition 2.5 and Proposition 2.6, so we need only see how part (ii) follows from the remaining lemmas. Assume that $m(1 + m'/d) < 2\sqrt{d}$ and pick any $\lambda < \lambda_l = 1/(2\sqrt{d})$. By Lemma 3.6, the function W defined therein satisfies (3.19) for some $\delta > 0$. By Lemma 3.7, $\sum_{v \in V(T)} W(v) < \infty$ (\mathbf{GW} -almost surely), so by Lemma 3.4, for any bounded starting state g , X_t is a supermartingale converging exponentially fast to zero. This implies that $\mathbb{E}_1 \eta_t(0)$ converges exponentially fast to zero, which by Lemma 3.5 implies that $\mathbb{E}_{\delta_0} \sum_{v \in V(T)} \eta_t(v)$ converges exponentially fast to zero. This implies, by Borel–Cantelli, that the process started from δ_0 dies out almost surely, proving the theorem. \square

We now turn to the final proof in this section, which proceeds by a careful path-counting argument.

PROOF OF THEOREM 2.2. Consider a Galton–Watson tree, T , arising from an offspring distribution with maximum d , mean m and size-biased mean m' , where $m \geq (2-\delta)\sqrt{d}$, $m' \geq \varepsilon d$, with $\varepsilon > 0$ arbitrary and δ to be described later. As in the proof of Proposition 2.6 we obtain T by starting with T'_d , removing a random subset of edges and taking the component containing the root O . Also much as in that proof, we wish to show that for some $\lambda < \lambda_l = 1/(2\sqrt{d})$ we have $\mathbf{E}_\otimes X_{l,n} > 1$; by the same second moment argument as before it suffices to show that $d^n \mathbf{E}_\otimes X_{l,v}$ can be made arbitrarily large, for some choice of $l(n)$, where v is an arbitrary vertex at generation n of T'_d . We have

$$(3.23) \quad d^n \mathbf{E}_\otimes X_{l,v} = d^n \lambda^l \mathbf{P}_{\mathbf{GW}}(v \in T) \mathbf{E}_{\mathbf{GW}}(W_{l,v} | v \in T)$$

$$(3.24) \quad = m^n \lambda^l \mathbf{E}_{\mathbf{GW}}(W_{l,v} | v \in T),$$

where $W_{l,v}$ is the number of walks of length l , from O to v , within T , which do not go below generation n .

It will be convenient to obtain a lower bound on $d^n \mathbf{E}_{\mathbf{GW}}(W_{l,v} | v \in T) = m^n \mathbf{E}_{\mathbf{GW}}(W_{l,v} | v \in T)$, for some suitable l . Let $O = v_0, v_1, \dots, v_{n-1}$ be the ancestors of v , and let $A(v)$ be the (random) set consisting of those ancestors which have at least $\varepsilon d/3$ children. As in the proof of Lemma 3.7, when one conditions on $v \in T$ each of the n ancestors of v has offspring with size-biased distribution, and hence mean number of offspring $m' \geq \varepsilon d$. Thus with probability (conditional on $v \in T$) at least $\varepsilon/3$, $|A(v)| \geq \varepsilon n/3$. Now we obtain a lower bound on $\mathbf{E}_{\mathbf{GW}} W_{n(1+2\delta),v}$ by considering walks from O to v which make δn side-trips of length 2 from distinct ancestors $v_j \in A(v)$. (For simplicity we assume δ is rational and that n is chosen so that δn is an integer.) If $|A(v)| \geq \varepsilon n/3$, the number of such walks is at least $\binom{c_1 n}{\delta n} (c_1 d)^{\delta n}$ where $c_1 = \varepsilon/3$. So we see that

$$(3.25) \quad d^n \mathbf{E}_{\mathbf{GW}} W_{n(1+2\delta),v} \geq m^n c_1 \binom{c_1 n}{\delta n} (c_1 d)^{\delta n}.$$

Now by standard estimates (see, e.g., [2]) one has

$$(3.26) \quad \binom{c_1 n}{\delta n} \geq \left(\frac{c_1}{\delta}\right)^{\delta n} \left(\frac{c_1}{c_1 - \delta}\right)^{(c_1 - \delta)n} \frac{1}{2\sqrt{c_1 n}}.$$

Combining (3.25) and (3.26) one has

$$\begin{aligned} \frac{\log(d^n \mathbf{E}_{\mathbf{GW}} W_{n(1+2\delta),v})}{n} &\geq \log m + c_2/n + \delta \log(c_1 d) + \delta \log\left(\frac{c_1}{\delta}\right) \\ &\quad + (c_1 - \delta) \log\left(\frac{c_1}{c_1 - \delta}\right) + (c_3 - \log n)/2n, \end{aligned}$$

provided $n\delta$ is an integer and $\delta < c_1$. Hence if $m \geq (2 - \delta)\sqrt{d}$ one has

$$(3.27) \quad \limsup_{n \rightarrow \infty} \frac{\log(d^n \mathbf{E}_{\mathbf{GW}} W_{n(1+2\delta),v})}{n(1+2\delta)} \geq \phi(\delta),$$

where

$$(3.28) \quad \begin{aligned} \phi(\delta) = & (\log((2 - \delta)\sqrt{d}) + 2\delta \log(\sqrt{c_1 d}) + \delta \log(c_1/\delta) \\ & + (c_1 - \delta) \log(c_1/(c_1 - \delta))) / (1 + 2\delta). \end{aligned}$$

It is easily seen that $\phi(\delta) = \log(2\sqrt{d}) + \psi(\delta)$ where $\psi(\delta)$ is independent of d and $\psi(0) = 0$; $\psi'(0) = +\infty$ since the derivative of $\delta \log(c_1/\delta)$ is infinite while all other contributions are finite. Now (3.27) implies that

$$\limsup_{n \rightarrow \infty} \frac{\log(d^n \mathbf{E}_{\otimes} X_{n(1+2\delta), v})}{n(1+2\delta)} \geq \phi(\delta) + \log \lambda,$$

and hence for sufficiently small $\delta > 0$ (independent of d), this limit is strictly positive for some $\lambda < 1/(2\sqrt{d})$, showing that for such a λ , $d^n \mathbf{E}_{\otimes} X_{n(1+2\delta), v} \rightarrow \infty$ as required. \square

4. Construction of counterexamples. We now turn to consider the matter of proving Theorem 2.3. If we wish to directly construct a tree of exponential growth on which the branching random walk has only one phase transition, we have considerable leeway to do so. We might, for example, consider rooted trees where every vertex has either m or d children, where $d > m \geq 2$ and $m(1 + m/d) < 2\sqrt{d}$. If, for every k , the tree contains a vertex all of whose descendants for k generations have d children, then the tree will have $\lambda_l = 1/2\sqrt{d}$. If, furthermore, vertices with d children are sufficiently rare (in some sense) in the tree as a whole then also $\lambda_g = 1/2\sqrt{d}$, much as in Theorem 2.1(ii). Now one might hope to show, roughly as follows, that the contact process on such a tree can have two phase transitions. On any finite graph the contact process (unlike the branching random walk) dies out for any parameter value. It seems reasonable to think, therefore, that making finitely many changes (roughly speaking) to a graph will not alter the contact processes critical values. Similarly, if in the graph construction described above, vertices with d children are sufficiently rare, then the contact processes critical values should equal (or at least be arbitrarily close to) those for T_m . For T_m ($m \geq 2$) it is known [10, 16, 22] that λ_1 and λ_2 are distinct.

It does not seem to be an easy matter, however, to make the argument of the last paragraph rigorous. Indeed, it is an open problem to show that the addition of a single edge to a graph does not change the contact processes critical values (although this can be shown in a number of special cases). Instead, therefore, of pursuing the argument just outlined, we prove Theorem 2.3 in a more direct fashion, making use of Theorem 2.1 and known bounds on λ_1 and λ_2 .

PROOF OF THEOREM 2.3. Consider a Galton–Watson tree T with offspring distribution satisfying $p_c, p_d > 0$, $p_c + p_d = 1$, for some $d > c \geq 2$. If $c(1 + c/d) < 2\sqrt{d}$ then, for p_d sufficiently small, Theorem 2.1 implies that $\lambda_g = \lambda_l$ for the branching random walk. On the other hand, for the contact process a

trivial coupling argument gives that on any such tree,

$$(4.1) \quad \lambda_1(T) \leq \lambda_1(T'_c) \quad \text{and} \quad \lambda_2(T) \geq \lambda_2(T'_d).$$

Hence if c, d , satisfy $\lambda_1(T'_c) < \lambda_2(T'_d)$ then the contact process on T will certainly have two phase transitions. Therefore we will establish the theorem if we can find c, d with $c(1 + c/d) < 2\sqrt{d}$ and $\lambda_1(T'_c) < \lambda_2(T'_d)$.

In [16] various bounds are given on λ_1 and λ_2 for homogeneous trees, T_d . It is not hard to see that the existence of two phase transitions for such trees ($d \geq 2$) implies that the critical values are exactly the same if we consider the corresponding rooted tree, T'_d , in which one vertex has a lower degree than all the others.

The upper bound on λ_1 that we shall use is the simple [16]

$$(4.2) \quad \lambda_1(T'_c) < \frac{1}{c-1}.$$

The crudest lower bound on $\lambda_2(T_d)$ is $1/2\sqrt{d}$, obtained by comparison with the branching random walk, and this is not sufficient for our purposes. We use instead [16]

$$(4.3) \quad \lambda_2(T_d) \geq \frac{4 + 2/(\sqrt{d} - 1) - \sqrt{8 + 16/(\sqrt{d} - 1) + 4/(\sqrt{d} - 1)^2}}{2(\sqrt{d} - 1)}.$$

Denoting the r.h.s. of (4.3) by $F(d)$ we see that

$$(4.4) \quad F(d) \sim \frac{4 - \sqrt{8}}{2\sqrt{d}} \quad \text{as } d \rightarrow \infty.$$

Since $4 - \sqrt{8} > 1$, this certainly implies that we can find c, d , with both

$$(4.5) \quad \frac{1}{c-1} < F(d) \quad \text{and} \quad \frac{1}{c(1 + c/d)} > \frac{1}{2\sqrt{d}};$$

the first inequality ensures [by (4.1), (4.2) and (4.3)] that $\lambda_1(T) < \lambda_2(T)$; the second inequality ensures that, for p_d small, $\lambda_g = \lambda_l$ (**GW**-almost surely). To be specific, we may take $c = 50$ and $d = 729$. It is easily checked that (4.5) holds. \square

To construct a tree T satisfying Theorem 2.4, we will again consider rooted trees in which every vertex has either c children or d children; to be specific, we take $c = 2$ and $d = 100$. Before stating precisely what tree we shall use and proving the required properties, we explain the key ideas.

Consider T_d^h , the finite rooted tree of height h in which every vertex—other than those in generation h —has d children. The contact process on any finite graph eventually dies out, but the length of time for which it survives on T_d^h is closely related to its behavior on the corresponding infinite tree, T_d . Roughly speaking, if $\lambda < \lambda_2(T_d)$ then the process should survive for a time which is at most linear in the height h . On the other hand, if $\lambda > \lambda_2(T_d)$ then the

process on T_d^h has a significant probability of survival for a time which is doubly exponential in h .

Now suppose that T —all of whose vertices have either 2 children or 100 children—contains copies of T_{100}^h for arbitrarily large h and that these subgraphs are *fairly* sparse. Then, for $\lambda > \lambda_2(T_{100})$, the process will survive for an extremely long time on the copies of T_{100}^h for h large; this gives the process many chances to spread from one copy of T_{100}^h to another and thereby for local survival to occur. On the other hand, if $\lambda < \lambda_2(T_{100})$, then the process dies out on each T_d^h in a linear time; since most vertices have 2 children it then behaves much as it would on T_2 . However, $\lambda_2(T_{100}) < \lambda_1(T_2)$ so for $\lambda < \lambda_2(T_{100})$ the process does not even survive globally.

We now turn to the matter of making these ideas precise. We begin by quoting a theorem from [23] about the survival of the process on finite trees.

THEOREM 4.1. *Let $d \geq 2, a < 1, \lambda > \lambda_2(T_d)$. Then there exists $p > 0, c > 0, \alpha > 1$ so that, for the contact process on T_d^h starting from a single infection at the root, with probability at least p the process survives for time at least*

$$c\alpha^{(da)^h}. \quad \square$$

PROOF OF THEOREM 2.4. Let T be the rooted tree in which every vertex has either 2 children or 100 children in the following way. Every vertex in the same generation has the same number of children. If a vertex is in generation $2^h n + k$ for some $h \geq 10, n \geq 1, 0 \leq k < h$, then it has 100 children; otherwise it has 2 children. Note that, roughly speaking, for each $h \geq 10$, this means that every 2^h generations there are h successive generations with 100 children per vertex.

First, we show that if $\lambda < \lambda_2(T_{100})$ then the contact process does not survive globally on T . An upper bound for the number of generations up to level 2^h in which vertices have 100 children each is

$$\begin{aligned} & (h-1) + (h-2)2 + (h-3)2^2 + (h-4)2^3 + \dots + 10 \cdot 2^{h-11} \\ (4.6) \quad & \leq 10 \cdot 2^{h-11} (1 + 2 \cdot 2^{-1} + 3 \cdot 2^{-2} + 4 \cdot 2^{-3} + \dots + (h-10) \cdot 2^{-(h-11)}) \\ & \leq 10 \cdot 2^{h-9}. \end{aligned}$$

Now, for the contact process on T_{100} , with parameter $\lambda < \lambda_2(T_{100})$, starting with a single infection at the root, the probability that a fixed vertex at generation l is *ever* infected is at most $(1/10)^l$ [7]. Therefore by a trivial coupling, the probability that a particular vertex of T at level 2^h is ever infected is at most $(1/10)^{2^h}$. However, by (4.6) the number of vertices of T at level 2^h is at most

$$\begin{aligned} 2^{2^h} \left(\frac{100}{2}\right)^{10 \cdot 2^{h-9}} &= \left(50^{10/2^9} \cdot 2\right)^{2^h} \\ &\leq (2.16)^{2^h}. \end{aligned}$$

Therefore the probability that any of the vertices at level 2^h is ever infected is at most $(2.16/10)^{2^h}$ which tends to 0 as $h \rightarrow \infty$; hence, with probability one, the infection does not reach arbitrarily deep levels, and this implies global nonsurvival.

Now suppose that $\lambda > \lambda_2(T_{100})$. We must show that local survival occurs for the contact process on T . Take constants $p > 0, c > 0$ and $\alpha > 1$ (which depend on λ) as in the statement of Theorem 4.1. Now choose further constants as follows. Choose N such that $(1 - p^2/8)^N < 1/2$. Let $\delta = e^{-1}(1 - e^{-\lambda})$; note that if a vertex v is infected at time t and w is a neighboring vertex, then the probability (conditional on \mathcal{F}_t) that w is infected at time $t + 1$ is at least δ ; we will also need the (more trivial) fact that δ is a lower bound on the conditional probability that v is still infected at time $t + 1$. Now choose $H \geq 10$ so that $2^{2^H} \geq N$ and

$$(4.7) \quad \left(1 - \delta^{N(2^H+H)}\right)^{\lfloor c\alpha^{99^H}/(2^H+H) \rfloor} \leq \frac{p}{2}.$$

Note that $\exp(-\delta^{N(2^H+H)}\lfloor c\alpha^{99^H}/(2^H+H) \rfloor)$ is an upper bound for the left-hand side of this inequality, and (since $99 > 2$) this $\rightarrow 0$ as $H \rightarrow \infty$.

Let us define another rooted tree T' as follows. If a vertex of T' is in generation $2^H n + k$ for some $n \geq 0, 0 \leq k < H$, then it has 100 children; all other vertices have 2 children. Note that T' implicitly depends on the choice of λ (via H) and also that T contains a subgraph isomorphic to T' .

We shall show the following for the contact process on T' by induction on j . For any $j \geq 0$, starting from a single infection at the root, with probability $\rho = p^2/8$ the origin is infected at some time after $2j(2^H + H)$. This will establish that on T' , and hence on T , the process survives locally.

The result trivially holds for $j = 0$. Now let us fix some $j > 0$ and suppose the result holds for smaller values.

Let E_l be the event that at time $l(2^H + H)$ some vertex at generation H or above is infected, and let $E = \bigcap_{l=0}^L E_l$ where $L = \lfloor c\alpha^{99^H}/(2^H + H) \rfloor$. By (4.1) $\mathbb{P}(E) \geq p$. Let v_1, \dots, v_N be N distinct vertices at generation 2^H , chosen arbitrarily in advance; note that N such exist by choice of H . Let v be any vertex within distance H of the root. Note that the distance from v to any given v_i is at most $2^H + H$. So for any v_i , the event G_i that there is an infection trail from (v, t) to $(v_i, t + (2^H + H))$ has probability at least δ^{2^H+H} . The events $\{G_i\}_{i=1}^N$ are monotone increasing and hence positively correlated (see [8]). Therefore if v is infected at time t then, with probability (conditional on \mathcal{F}_t) at least $\delta^{N(2^H+H)}$, v_1, \dots, v_N are all infected [via infection trails from (v, t)] at time $t + (2^H + H)$.

For simplicity, let $\tilde{\mathcal{F}}_l$ denote $\mathcal{F}_{l(2^H+H)}$, the σ -field generated by the process up to time $l(2^H + H)$; let A_l be the event that v_1, \dots, v_N are all infected at time $l(2^H + H)$. Then A_l is $\tilde{\mathcal{F}}_l$ -measurable and $\mathbb{P}(A_{l+1} | \tilde{\mathcal{F}}_l) \geq \delta^{N(2^H+H)}$ on E , for

$0 \leq l \leq L$. This implies (with a little care) that

$$(4.8) \quad \mathbb{P}\left(\bigcap_{l=0}^L A_{l+1}^c\right) \leq (1 - \mathbb{P}(E)) + \left(1 - \delta^{N(2^H+H)}\right)^{L+1} \\ \leq 1 - p/2;$$

that is,

$$\mathbb{P}\left(\bigcup_{l=1}^{L+1} A_l\right) \geq p/2.$$

Now let l_0 be the least $l \geq 1$ such that A_l occurs (if there is such). Each v_i is at the top of a copy of T' so, suppressing infections outside each of these N subtrees, we see by the induction hypothesis that, conditional on $\tilde{\mathcal{F}}_{l_0}$, each v_i has probability at least ρ of infection at some time at least $l_0(2^H + H) + 2(j-1)(2^H + H)$, with each v_i behaving independently in this respect. So the (nonconditional) probability that some v_i is infected at time beyond $(2j-1)(2^H + H)$ is at least

$$(4.9) \quad \mathbb{P}\left(\bigcup_{l=1}^{L+1} A_l\right)(1 - (1 - \rho)^N) \geq p/4,$$

the inequality following from (4.8) and the choice of N .

So, if we let

$$t_0 = \inf\{t \geq (2j-1)(2^H + H): \text{some } v_i \text{ is infected at time } t\},$$

then $\mathbb{P}(t_0 < \infty) \geq p/4$. On this event let i_0 be the least i with v_i infected at time t_0 . Conditional on $t_0 < \infty$, with probability p the copy of T_{100}^H which has v_{i_0} as its root remains infected until time $t_0 + c\alpha^{99^H}$. Letting this event play the same role as E above, one can see that if this happens one has essentially L independent opportunities to reinfect the origin, each with probability δ^{2^H+H} of success. Exactly as in the proof of (4.8), one obtains that, conditional on the event $t_0 < \infty$, the probability that O is infected at time beyond $t_0 + (2^H + H)$ is at least $p/2$. Hence the (non-conditional) probability that O is infected at time beyond $2j(2^H + H)$ is at least $(p/4)(p/2) = \rho$; as required. \square

We remark that the overall scheme of the last part of this proof was inspired by the proof of Proposition 1 in [19].

5. Conclusion and open problems. Although we have found that the contact process and branching random walk are not as well behaved as expected on general graphs, it is natural to try to prove positive results for a more restricted class of graphs. One reasonable class to consider is that of the quasi-transitive graphs. A graph is said to be *transitive* if the automorphism group acts transitively on the set of vertices (i.e., has only one orbit); it is said to be *quasi-transitive* if the action of the automorphism group has only finitely many orbits. These concepts often arise in the context of other

processes. For example it is now well known (and follows from the proofs of [3]) that, for percolation on a quasi-transitive amenable graph, there can be at most one infinite cluster.

The example exhibited to prove Theorem 2.4 is not quasi-transitive, although it is as well behaved as possible without being so: it is spherically symmetric, non-amenable, and of bounded degree, indeed with every vertex having either 2 children or 100.

Since processes on quasi-transitive graphs are often better behaved than on more general graphs, we do not claim to be the first to make the following conjecture. It is a more restricted version of the conjectures disproved in this paper.

CONJECTURE 5.1. *For a quasi-transitive connected infinite locally finite (and hence bounded degree) graph G , the following are equivalent:*

- (i) G is nonamenable.
- (ii) $\lambda_1(G) < \lambda_2(G)$.
- (iii) $\lambda_g(G) < \lambda_l(G)$.

For the contact process the best result of this kind in one direction is in [22], which gives a class of trees (and treelike graphs), *isotropic block trees*, for which $\lambda_1 < \lambda_2$; the proof is easily adapted to the branching random walk (with some simplifications in the preliminary lemmas). However this class does not even include all quasi-transitive nonamenable trees. In the other direction even less is known. Results of Bezuidenhout and Grimmett [1] imply that $\lambda_1 = \lambda_2$ for \mathbf{Z}^d and could no doubt be adapted to certain other lattices. Their proofs are rather hard, but show a good deal more. It would be very interesting to find a simpler proof of $\lambda_1 = \lambda_2$ that could be adapted to a general class of graphs.

Of course Galton–Watson trees, although widely studied, are far from being quasi-transitive. The following question, which originally motivated the work of this paper, remains open.

PROBLEM 5.2. *Is there a bounded offspring distribution, with mean $m > 1$, for which $\lambda_1 = \lambda_2$?*

We make a conjecture about one further problem, of a rather different flavor, to which we referred prior to the proof of Theorem 2.4.

CONJECTURE 5.3. *Suppose that $G = (V, E)$ is a connected locally finite graph, and let $G' = (V, E \cup \{e\})$ be a graph with the same vertex set and one additional edge. Then $\lambda_1(G) = \lambda_1(G')$ and $\lambda_2(G) = \lambda_2(G')$.*

Note that this conjecture is blatantly false with the contact process's critical values replaced by those for the branching random walk, but then the branching random walk, unlike the contact process, can survive on a finite graph.

REFERENCES

- [1] BEZUIDENHOUT, C. and GRIMMETT, G. (1990). The critical contact process dies out. *Ann. Probab.* **18** 1462–1482.
- [2] BOLLOBÁS, B. (1979). *Graph Theory*. Springer, New York.
- [3] BURTON, R. M. and KEANE, M. (1989). Density and uniqueness in percolation. *Comm. Math. Phys.* **121** 501–505.
- [4] DURRETT, R. (1988). *Lectures Notes on Particle Systems and Percolation*. Wadsworth, Belmont, CA.
- [5] HARRIS, T. E. (1978). Additive set-valued Markov processes and graphical methods. *Ann. Probab.* **6** 355–378.
- [6] HUETER, I. and LALLEY, S. (2000). Anisotropic branching random walks on homogeneous trees. *Probab. Theory Related Fields* **116** 57–88.
- [7] LALLEY, S. and SELLKE, T. (1998). Limit set of a weakly supercritical contact process on a homogeneous tree. *Ann. Probab.* **26** 644–657.
- [8] LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, New York.
- [9] LIGGETT, T. M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer, New York.
- [10] LIGGETT, T. M. (1996). Multiple transition points for the contact process on the binary tree. *Ann. Probab.* **24** 1675–1710.
- [11] LIGGETT, T. M. (1996). Branching random walks and contact processes on homogeneous trees. *Probab. Theory Related Fields* **106** 495–519.
- [12] LIGGETT, T. M. (1999). Branching random walks on finite trees. In *Perplexing Problems in Probability: Papers in Honor of Harry Kesten* (M. Bramson and R. Durrett, eds.) 315–330. Birkhäuser, Boston.
- [13] LYONS, R. (2000). Phase transitions on nonamenable graphs. *J. Math. Phys.* **41** 1099–1126.
- [14] LYONS, R., PEMANTLE, R. and PERES, Y. (1996). Biased random walks on Galton–Watson trees. *Probab. Theory Related Fields* **106** 249–264.
- [15] MADRAS, N. and SCHINAZI, R. (1992). Branching random walks on trees. *Stochastic Process. Appl.* **42** 255–267.
- [16] PEMANTLE, R. (1992). The contact process on trees. *Ann. Probab.* **20** 2089–2116.
- [17] SALZANO, M. (1999). Infinitely many contact process transitions on a tree. *J. Statist. Phys.* **97** 817–826.
- [18] SALZANO, M. and SCHONMANN, R. H. (1997). The second lowest extremal invariant measure of the contact process. *Ann. Probab.* **25** 1846–1871.
- [19] SALZANO, M. and SCHONMANN, R. H. (1998). A new proof that for the contact process on homogeneous trees local survival implies complete convergence. *Ann. Probab.* **26** 1251–1258.
- [20] SALZANO, M. and SCHONMANN, R. H. (1999). The second lowest extremal invariant measure of the contact process II. *Ann. Probab.* **27** 845–875.
- [21] SCHINAZI, R. Private communication.
- [22] STACEY, A. M. (1996). The existence of an intermediate phase for the contact process on trees. *Ann. Probab.* **24** 1711–1726.
- [23] STACEY, A. M. (2000). The contact process on finite trees. *Probab. Theory Related Fields*. To appear.

DEPARTMENT OF MATHEMATICS
OHIO STATE UNIVERSITY
231 W. 18TH AVENUE
COLUMBUS, OHIO 43210-4065

STATISTICAL LABORATORY
DEPARTMENT OF PURE MATHEMATICS
AND MATHEMATICAL STATISTICS
WILBERFORCE ROAD
CAMBRIDGE CB3 0WB
UNITED KINGDOM