

## THE WILLIAMS–BJERKNES MODEL ON REGULAR TREES

BY OREN LOUIDOR<sup>1</sup>, RAN TESSLER AND  
 ALEXANDER VANDENBERG-RODES

*Technion, Hebrew University of Jerusalem and University of California, Irvine*

We consider the Williams–Bjerknes model, also known as the biased voter model on the  $d$ -regular tree  $\mathbb{T}^d$ , where  $d \geq 3$ . Starting from an initial configuration of “healthy” and “infected” vertices, infected vertices infect their neighbors at Poisson rate  $\lambda \geq 1$ , while healthy vertices heal their neighbors at Poisson rate 1. All vertices act independently. It is well known that starting from a configuration with a positive but finite number of infected vertices, infected vertices will continue to exist at all time with positive probability if and only if  $\lambda > 1$ . We show that there exists a threshold  $\lambda_c \in (1, \infty)$  such that if  $\lambda > \lambda_c$  then in the above setting with positive probability, all vertices will become eventually infected forever, while if  $\lambda < \lambda_c$ , all vertices will become eventually healthy with probability 1. In particular, this yields a complete convergence theorem for the model and its dual, a certain branching coalescing random walk on  $\mathbb{T}^d$ —above  $\lambda_c$ . We also treat the case of initial configurations chosen according to a distribution which is invariant or ergodic with respect to the group of automorphisms of  $\mathbb{T}^d$ .

**1. Introduction and results.** We study the *Williams–Bjerknes model* (henceforth WB process), also known as the *biased voter model*, on the  $d$ -regular tree  $\mathbb{T} = \mathbb{T}^d$  for  $d \geq 3$ . This is a continuous time Markov process whose state space is  $\mathcal{X} := \{-, +\}^{\mathbb{T}}$ , that is, the set of all configurations (assignments) of  $\pm$  to the vertices of the tree. “+ vertices” will be thought of as *infected*, while “− vertices” as *healthy*. Starting from some initial configuration  $\xi_0 \in \mathcal{X}$ , infected vertices infect each of their neighbors at Poisson rate  $\lambda$ , where  $\lambda \geq 1$  is the *infection rate* parameter, while healthy vertices heal each of their neighbors at Poisson rate 1. All vertices act independently. We shall denote by  $\xi_t^{\xi_0, \mathbb{T}, \lambda}$  the state of this process at time  $t$  and often omit some or all of the superscripts when they are clear or irrelevant.

Formally,  $(\xi_t^{\xi_0, \mathbb{T}, \lambda} : t \geq 0)$  is a Markov spin-system whose generator is the closure in  $\mathcal{C}(\mathcal{X})$  of the operator [defined on a suitable sub-space of  $\mathcal{C}(\mathcal{X})$ ].

$$\begin{aligned}
 \mathcal{L}f(\xi) &= \sum_x (1_{\{\xi(x)=+\}} + \lambda 1_{\{\xi(x)=-\}}) \{y \sim x : \xi(y) \neq \xi(x)\} \\
 (1.1) \quad &\quad \times [f(\xi^x) - f(\xi)],
 \end{aligned}$$

---

Received February 2013; revised August 2013.

<sup>1</sup>Supported by a Simons Postdoctoral Fellowship.

*MSC2010 subject classifications.* Primary 60K35; secondary 82C41.

*Key words and phrases.* Williams–Bjerknes, biased voter, local survival, fixation.

where  $\xi^x$  is equal to  $\xi$  except at the vertex  $x$  where it has the opposite sign, and  $x \sim y$  means that  $x$  and  $y$  are neighboring vertices in  $\mathbb{T}$ . We shall identify a configuration  $\xi$  with the subset of vertices which are infected under it, that is, the set  $\{x \in \mathbb{T} : \xi(x) = +\}$ .

This process was introduced in 1972 by Williams and Bjerknes [22] as a model for tumor growth and independently by Schwartz [19] in 1977 as an example of a particle system with an increasing dual. It is closely related to both the *voter model* (the case of  $\lambda = 1$ ) and the *contact process* (healing rates are fixed and do not depend on the number of healthy neighbors). As such it exhibits behavior which is similar to both models (this will be further discussed below). For standard texts on all these models, see [11, 12].

The main question in this model, both from a mathematical and a biological point of view, is that of survival. Namely, starting from a finite nonempty initial configuration  $\xi_0$ , that is,  $0 < |\xi_0| < \infty$  (where  $|\xi_0|$  is the cardinality of  $\xi_0$ ), whether infected sites will continue to exist at all times or become extinct. As was noticed by Williams and Bjerknes, observed at the times of transition:  $0 = \tau_0, \tau_1, \tau_2, \dots$ , the process  $(|\xi_{\tau_k}| : k = 0, 1, \dots)$  is just a nearest-neighbor random walk on  $\mathbb{Z}_+$  with an absorbing state at 0 and drift

$$(1.2) \quad -1 \frac{1}{\lambda + 1} + 1 \frac{\lambda}{\lambda + 1} = \frac{\lambda - 1}{\lambda + 1}.$$

Therefore *global* survival, that is,

$$(1.3) \quad \Omega_g^{\xi_0} := \{\sup\{t \geq 0 : \xi_t^{\xi_0} \neq \emptyset\} = \infty\},$$

has probability

$$(1.4) \quad \mathbb{P}(\Omega_g^{\xi_0}) = 1 - \lambda^{-|\xi_0|},$$

which for finite, nonempty  $\xi_0$ , is positive if and only if  $\lambda > 1$  (the reason for the term “global” will become apparent shortly). In other words, the threshold for the possibility of global survival is  $\lambda_g = 1$  regardless of the underlying graph, as long as it is infinite, connected and has a bounded degree (this can be relaxed, but some restrictions are needed to ensure that the process is well defined).

In the lattice case, based on numerical simulations Williams and Bjerknes predicted that once the infection survives, the set of infected sites will “roughly” look like an ever growing “blob” around the initially infected vertex. This was proved by Bramson and Griffeath in 1980 [2, 3] who gave a shape theorem with a linear rate for the subset of infected sites—for  $\mathbb{Z}^d$  in all  $d \geq 1$  and any  $\lambda > 1$ . (This is similar to the shape theorem for the Richardson Growth Model, which was proved by Richardson [17] and Kesten [4].) Thus, in particular on  $\mathbb{Z}^d$  for all  $\lambda > 1$ , global survival implies *complete* survival, namely

$$(1.5) \quad \Omega_c^{\xi_0} := \{\sup\{t \geq 0 : \xi_t^{\xi_0} \not\ni x\} < \infty\},$$

where  $x$  is any vertex of  $\mathbb{Z}^d$ . Notice that except for an event of zero probability, on  $\Omega_c^{\xi_0}$  every vertex will eventually become permanently infected, and hence the choice of  $x$  is immaterial in the above definition.

On the  $d$ -regular tree, the situation is more intricate and so far was less understood. Madras, Schinazi (and Durrett) [14] showed that for  $d \geq 3$ , survival can be global but not complete. More precisely, for  $\xi_0 \in \mathcal{X}$  and any  $x \in \mathbb{T}$ , define *local survival* as the event

$$(1.6) \quad \Omega_l^{\xi_0} := \{\sup\{t \geq 0 : \xi_t^{\xi_0} \ni x\} = \infty\}.$$

Then for all  $d \geq 3$ , there exists  $\lambda'$  strictly higher than  $\lambda_g = 1$ , such that for any finite nonempty  $\xi_0$ , if  $\lambda \in (1, \lambda')$ , then  $\mathbb{P}(\Omega_l^{\xi_0}) = 0$ . In other words, if  $\lambda \in (1, \lambda')$  the infection can survive, but it must eventually “drift to infinity.” Letting

$$(1.7) \quad \lambda_l(\mathbb{T}) := \inf\{\lambda > 0 : \mathbb{P}(\Omega_l^{\xi_0}) > 0\}$$

denote the threshold value for local survival, where  $\xi_0$  is any finite nonempty configuration, whose precise value is immaterial, they are able to show that

$$(1.8) \quad \lambda_l(\mathbb{T}^d) \geq \frac{d}{2\sqrt{d-1}}.$$

However, it was not clear whether  $\lambda_l(\mathbb{T}) < \infty$  nor what exactly happens above this threshold. More precisely, if we define

$$(1.9) \quad \lambda_c(\mathbb{T}) := \inf\{\lambda > 0 : \mathbb{P}(\Omega_c^{\xi_0}) > 0\},$$

then it is not clear whether  $\lambda_c(\mathbb{T}) < \infty$  and whether its value coincides with that of  $\lambda_l(\mathbb{T})$  or is strictly larger than it. Had  $\lambda_l(\mathbb{T}) < \lambda_c(\mathbb{T}) < \infty$  been the case, there would have been three phases for the model: global but not local nor complete survival, local but not complete survival and finally, complete survival.

The notion of local survival and the existence of an intermediate phase where survival is global but not local was first observed by Pemantle in the context of the contact process on trees [16]. By finding upper and lower bounds on the infection thresholds for global (resp., local) survival, he was able to conclude that there is an intermediate regime for  $\mathbb{T}^d$  when  $d \geq 4$ . Liggett [10] and then [20] showed that this is also true for  $d = 3$ .

By adapting the martingale methods of Pemantle, one can fairly easily obtain bounds on the threshold values  $\lambda_l(\mathbb{T})$  and  $\lambda_c(\mathbb{T})$ .

PROPOSITION 1.1. *Let  $d \geq 3$ .*

- (1)  $\frac{d}{2\sqrt{d-1}} \leq \lambda_l(\mathbb{T}^d) \leq \min(2d, \frac{4d}{(\sqrt{d-1}-4)\vee 0})$ .
- (2)  $\lambda_c(\mathbb{T}^d) \leq (d-1) \vee \lambda_l(\mathbb{T}^d)$ .

This shows that both local and complete survival occur for large enough values of  $\lambda$ , but does not settle the question of whether there is a second intermediate phase of local but not complete survival. The lower bound in part (1) is the same as the one obtained in [14]. However, the martingale approach used here seems more robust, as it does not rely on the tree isotropy, which is exploited in [14]. Therefore, it could be used to handle other tree-like graphs which are less regular (e.g., a realization of a super-critical Galton–Watson process). It should be noted that the argument leading to part (2) of the proposition can be applied to  $\mathbb{Z}^d$  as well. In this case one gets  $\lambda_c(\mathbb{Z}^d) = \lambda_g = 1$  for all  $d \geq 1$ , thereby providing a very short proof for (1.5), albeit without a shape theorem.

It requires much more work to show:

**THEOREM 1.2.** *For all  $d \geq 3$  we have  $\lambda_c(\mathbb{T}^d) = \lambda_l(\mathbb{T}^d)$ .*

The proof of this theorem constitutes the main part of this paper. The theorem implies that the only possibility for local but not complete survival is when  $\lambda = \lambda_l(\mathbb{T}) = \lambda_c(\mathbb{T})$ . We conjecture that this is not the case and that in fact at this  $\lambda$  survival can only be global. This is the case in the contact process [9, 23]. As an immediate corollary of Theorem 1.2 we get the following characterization of all possible weak limits of  $\xi_t$ . In what follows, we naturally endow the space  $\mathcal{X}$  with the product topology and product  $\sigma$ -algebra.

**COROLLARY 1.3.** *Let  $d \geq 3$  and  $\lambda > \lambda_l(\mathbb{T}^d)$ . For all  $\xi_0 \in \mathcal{X}$  as  $t \rightarrow \infty$ ,*

$$(1.10) \quad \mathbb{P}(\xi_t^{\xi_0} \in \cdot) \Rightarrow \mathbb{P}(\Omega_g^{\xi_0})\delta_{\mathbb{T}} + (1 - \mathbb{P}(\Omega_g^{\xi_0}))\delta_{\emptyset}.$$

*In particular  $\delta_{\emptyset}$  and  $\delta_{\mathbb{T}}$  are the only extremal invariant measures for the model above  $\lambda_l(\mathbb{T})$ .*

Equation (1.10) is an analog of the complete convergence theorem for the contact process on  $\mathbb{T}$ , conjectured by Pemantle and first proved by Zhang [23] and reproved in a simpler way by Salzano and Schonmann [18]. Here  $\delta_{\mathbb{T}}$  plays the role of the upper invariant measure of the contact process (i.e., the limiting measure of the process when started from the all + configuration). When  $\lambda \in (1, \lambda_l(\mathbb{T}))$  it is not clear whether aside from  $\delta_{\emptyset}$  and  $\delta_{\mathbb{T}}$  there are other extremal invariant measures. We conjecture that this is the case, as it is for the contact process [5] below the threshold for local survival and for the nonbiased voter model, that is, the case when  $\lambda = 1$ , [11].

Another consequence of Theorem 1.2 is the process “mostly” fixates. More formally, for  $\xi_0 \in \mathcal{X}$  and  $x \in \mathbb{T}$  define the *fixation* event as

$$(1.11) \quad \Omega_f^{\xi_0} := \{\sup\{t \geq 0 : \xi_{t-}^{\xi_0}(x) \neq \xi_t^{\xi_0}(x)\} < \infty\}.$$

Then we have the following:

COROLLARY 1.4. Fix  $d \geq 3$ .

- (1) If  $\lambda > \lambda_l(\mathbb{T}^d)$ , then  $\mathbb{P}(\Omega_f^{\xi_0}) = 1$ , for any  $\xi_0 \in \mathcal{X}$ .
- (2) If  $1 \leq \lambda < \lambda_l(\mathbb{T}^d)$ , then  $\mathbb{P}(\Omega_f^{\xi_0}) = 1$ , for any finite  $\xi_0 \in \mathcal{X}$ .

Duality plays an important role in the analysis of particle systems; see, for example [6, 7]. At the same time, the dual processes are often of interest by themselves. When  $\lambda \geq 1$ , a dual for the WB process, which was exploited time and again in the past, is a certain (continuous time) branching coalescing random walk (henceforth the BCRW process), which we now describe. Like  $\xi$ , this process takes value in the space  $\mathcal{X}$  of all  $\pm$  configurations on  $\mathbb{T}$ . However, this time we interpret a “+ vertex” as occupied by a particle, while a “− vertex” as vacant. Starting from an initial configuration  $\hat{\xi}_0$ , particles independently move to each of their neighbors at rate 1 and give birth (branch) to a new particle at each of their neighbors at rate  $\lambda - 1$ . If a vertex to which a particle moved or branched was already occupied by a particle, the two particles coalesce. We shall denote this process by  $(\hat{\xi}_t^{\hat{\xi}_0, \mathbb{T}, \lambda} : t \geq 0)$ . Formally, its generator is the closure of

$$\begin{aligned}
 & \hat{\mathcal{L}}f(\hat{\xi}) \\
 (1.12) \quad & = \sum_{x \sim y} (1_{\{\hat{\xi}(x)=+, \hat{\xi}(y)=-\}} ([f(\hat{\xi}^{xy}) - f(\hat{\xi})] + (\lambda - 1)[f(\hat{\xi}^y) - f(\hat{\xi})]) \\
 & \quad + 1_{\{\hat{\xi}(x)=+, \hat{\xi}(y)=+\}} ([f(\hat{\xi}^x) - f(\hat{\xi})] + [f(\hat{\xi}^y) - f(\hat{\xi})])).
 \end{aligned}$$

As before  $\hat{\xi}^x$  is  $\hat{\xi}$  with the sign at  $x$  flipped, while  $\hat{\xi}^{xy}$  is  $\hat{\xi}$  with the sign flipped both at  $x$  and at  $y$ .

There are two known duality relations between  $\xi$  and  $\hat{\xi}_t$ . The first one which is more standard, can be read immediately from the graphical representation of the model. The second was discovered by Sudbury and Lloyd [21] and involves  $p$ -thinning of configurations, whereby each + vertex becomes a − vertex with probability  $1 - p$  and kept + with probability  $p$ , independently of other vertices. We shall make use of both of these relations in the proofs, but they are not needed in order to state the results concerning  $\hat{\xi}$ , and therefore we shall defer their precise formulation to Section 2.5.

For  $p \in [0, 1]$ , let  $\nu_p$  denote the Bernoulli( $p$ )-product measure on  $\mathcal{X}$ . Using either of the two duality relations, the previous results on  $\hat{\xi}$  immediately give

THEOREM 1.5. Let  $d \geq 3$ . If  $\lambda \in [1, \lambda_l(\mathbb{T}^d))$ , then for any finite  $\hat{\xi}_0 \in \mathcal{X}$  as  $t \rightarrow \infty$ ,

$$(1.13) \quad \mathbb{P}(\hat{\xi}_t^{\hat{\xi}_0} \in \cdot) \Rightarrow \delta_{\emptyset}.$$

If  $\lambda > \lambda_l(\mathbb{T}^d)$ , then for any  $\hat{\xi}_0 \in \mathcal{X}$  as  $t \rightarrow \infty$ ,

$$(1.14) \quad \mathbb{P}(\hat{\xi}_t^{\hat{\xi}_0} \in \cdot) \Rightarrow 1_{\{\hat{\xi}_0 \neq \emptyset\}} \nu_{1-1/\lambda} + 1_{\{\hat{\xi}_0 = \emptyset\}} \delta_{\emptyset}.$$

*In particular the only extremal invariant measures for BCRW above  $\lambda_l(\mathbb{T})$  are  $\delta_\emptyset$  and  $\nu_{1-1/\lambda}$ .*

The cases of  $\lambda = \lambda_l(\mathbb{T}^d)$ —any initial configuration  $\hat{\xi}_0$ —and  $\lambda \in (1, \lambda_l(\mathbb{T}^d))$  with infinite  $\hat{\xi}_0$ —remain open, as they do for  $\xi$ .

As mentioned, when  $\lambda \in (1, \lambda_l(\mathbb{T}^d))$  and the initial configuration is chosen according to a distribution which puts mass on infinite configurations, then it is an open problem to characterize the set of possible weak limits for both  $\xi$  and  $\hat{\xi}$ . Nevertheless, if the initial configuration is invariant or even ergodic, with respect to the group of automorphisms of  $\mathbb{T}^d$ , then such a characterization is possible.

More precisely, denote by  $\mathcal{I}$  the set of probability measures on  $\mathcal{X}$  which are invariant under all automorphisms of  $\mathbb{T}^d$ . The subset of  $\mathcal{I}$  of all measures which are in addition ergodic will be denoted by  $\mathcal{E}$ . For a configuration  $\xi$ , a (*connected*) *component* is a maximal subset of vertices  $U$  of  $\mathbb{T}$ , for which the induced subgraph is connected and such that all vertices in  $U$  have the same sign under  $\xi$ . We shall call a component infected if its vertices are infected under  $\xi$ . Then we have the following:

**THEOREM 1.6.** *Let  $d \geq 3$  and  $\lambda > 1$ .*

(1) *If  $\mathbb{P}(\xi_0 \in \cdot) \in \mathcal{I}$ , then*

$$(1.15) \quad \mathbb{P}(\Omega_f^{\xi_0}) = 1.$$

*In particular, any automorphism-invariant stationary distribution for  $\xi$  is a convex combination of  $\delta_\emptyset$  and  $\delta_{\mathbb{T}}$ .*

(2) *If  $\mathbb{P}(\xi_0 \in \cdot) \in \mathcal{E} \setminus \{\delta_\emptyset\}$ , then*

$$(1.16) \quad \mathbb{P}(\Omega_c^{\xi_0}) = 1.$$

*In this case, infinite infected components are formed in finite time  $\mathbb{P}$ -almost surely. In particular, the only automorphism-ergodic stationary distributions for  $\xi$  are  $\delta_\emptyset$  and  $\delta_{\mathbb{T}}$ .*

It should be noted that the proof of Theorem 1.6 applies to a much larger class of vertex transitive graphs.

1.1. *Outline of the paper.* The remainder of the paper is organized as follows. In Section 2 we recall some known facts about the WB and BCRW processes as well as introduce most of the notation which will be used later in the proofs. In Section 3 we prove Theorems 1.2 and 1.5 as well as Corollaries 1.3 and 1.4. Section 4 includes the proof of Proposition 1.1, and finally Section 5 is devoted to the proof of Theorem 1.6.

**2. Preliminaries and notation.** In this section we set up some additional notation which will often be used in the sequel as well as collect some well-known facts about the process and its dual. Any future use of these facts will be accompanied by a proper reference to this section. Consequently, the reader who is familiar with the model can skim through this section quickly or skip it altogether, without much risk of getting lost later on.

2.1. *Graphs.* We will mostly be concerned with the  $d$ -regular tree  $\mathbb{T}$ , although occasionally we shall use other graphs  $\mathbb{G} = (V, E)$ . We shall identify sub-graphs with their corresponding edge-set and vertex-set. For example, for a set of vertices  $U$  we may write  $U \subseteq \mathbb{T}$  to indicate that  $U$  is a subset of the vertex set of  $\mathbb{T}$  and in this case treat  $U$  also as the sub-graph of  $\mathbb{T}$  induced by the vertices in  $U$ . We shall distinguish one vertex of  $\mathbb{T}$  to be called the origin and denoted  $0$ .

Although in the definition of  $\xi$ . the underlying graph need not be directed, it will be convenient to think of the edges of  $\mathbb{T}$  as oriented such that each vertex will have exactly one predecessor, its *parent* and  $d - 1$  successors—its *children* (formally, we fix an end of  $\mathbb{T}$  and define the parent of  $x$  as the first vertex after  $x$  on the ray from  $x$  which belongs to this end). For a vertex  $x \in \mathbb{T}$ , we let  $\mathbb{T}_x$  denote the subtree, in the above orientation, rooted at  $x$ .

The graph distance will be denoted by  $\rho$ . For  $x \in \mathbb{G}$  and  $r > 0$  we denote by  $B_x(r)$  the closed ball of radius  $r$  around  $x$  in this metric, namely  $B_x(r) := \{y \in \mathbb{G} : \rho(x, y) \leq r\}$  and set  $S_x(r) := B_x(r) \setminus B_x(r - 1)$ . Given a subset of vertices  $U \subseteq \mathbb{G}$ , we denote by  $\partial_{\mathbb{G}}U$  the set of edges in  $\mathbb{G}$  with exactly one endpoint in  $U$ .

2.2. *WB and BCRW on general underlying graphs and boundary conditions.* The definition of  $\xi$ . in (1.1) and  $\hat{\xi}$ . in (1.12) can, of course, be extended to any underlying graph  $\mathbb{G} = (V, E)$  with a bounded degree (as mentioned, this can be relaxed). In this case the state space is  $\mathcal{X}^{\mathbb{G}} := \{+, -\}^{\mathbb{G}}$  and the Williams–Bjerknes process for such graph, initial configuration  $\xi_0 \in \mathcal{X}^{\mathbb{G}}$  and infection parameter  $\lambda \geq 1$  will be denoted by  $(\xi_t^{\xi_0, \mathbb{G}, \lambda} : t \geq 0)$ . Similarly, the corresponding branching coalescing random walk will be denoted by  $(\hat{\xi}_t^{\xi_0, \mathbb{G}, \lambda} : t \geq 0)$ . As mentioned in the [Introduction](#), we shall often omit some or all of the superscripts.

The *inclusion time* of a subset of vertices  $U \subseteq \mathbb{G}$  will be used often. For  $\xi$ . it is defined as

$$(2.1) \quad \tau_U := \inf\{t \geq 0 : \xi_t \supseteq U\}.$$

Similarly  $\hat{\tau}_U$  will denote the inclusion time of  $U$  by  $\hat{\xi}$ .

We shall often treat several instances of  $\xi$ . and  $\hat{\xi}$ . corresponding to different  $(\xi_0, \mathbb{G}, \lambda)$  at the same time. In this case, it will be useful to decorate all events and random variables pertaining to a certain instance with the same superscripts and accents used to denote the process itself. For example, we may write  $\Omega_g^{\xi_0, \mathbb{G}, \lambda}$  for the event of global survival for  $\xi$ . <sup>$\xi_0, \mathbb{G}, \lambda$</sup>  or  $\hat{\tau}_x^{\xi_0, \mathbb{G}}$  for the inclusion time of  $x$  by  $\hat{\xi}$ . <sup>$\xi_0, \mathbb{G}$</sup> .

If  $\mathbb{G}$  is a sub-graph of a larger graph  $\mathbb{G}'$ , we may often want to put *boundary conditions* on the vertices of  $\mathbb{G}' \setminus \mathbb{G}$ . Given  $\zeta \in \mathcal{X}^{\mathbb{G}' \setminus \mathbb{G}} = \{-, +\}^{\mathbb{G}' \setminus \mathbb{G}}$ , the process  $\xi_\zeta$  with boundary conditions  $\zeta$  evolves as before, only that the sign of vertices in  $\mathbb{G}' \setminus \mathbb{G}$  remain fixed according to  $\zeta$ . Thus vertices in  $\mathbb{G}' \setminus \mathbb{G}$  cannot be infected nor healed, but they continue to infect or heal their neighboring vertices in  $\mathbb{G}$  at the usual rates.

It will be convenient to suppose that sub-graphs can possibly “come” with boundary conditions and we shall write  $\mathbb{G}^\zeta$  to mean that  $\mathbb{G}$  “comes” with boundary conditions  $\zeta$  on  $\mathbb{G}' \setminus \mathbb{G}$ . Writing just  $\mathbb{G}$  means that there are no boundary conditions associated with  $\mathbb{G}$ . In practice, we shall only use either the  $+$  *boundary conditions*, by which we mean that  $\zeta = \delta_{\mathbb{G}' \setminus \mathbb{G}}$  or the  $-$  *boundary conditions*, by which we mean that  $\zeta = \delta_\emptyset$ . In these cases we shall write either  $\mathbb{G}^+$  or  $\mathbb{G}^-$ . Furthermore, if  $\mathbb{G}'$  is not specified it will be assumed to be  $\mathbb{T}$ . For example,  $\xi_{\xi_0, \mathbb{T}_0^+, \lambda}$  is the WB process on  $\mathbb{T}_0$  with  $+$  boundary conditions on  $\mathbb{T} \setminus \mathbb{T}_0$ .

Boundary conditions will also be used for  $\hat{\xi}$ , although here we need to clarify what they mean exactly. Given  $\zeta \in \mathcal{X}^{\mathbb{G}' \setminus \mathbb{G}}$  as before,  $\hat{\xi}_\zeta$  on  $\mathbb{G}$  with boundary conditions  $\zeta$  evolves as  $\hat{\xi}$  does, only that particles which reach a  $-$  vertex in  $\mathbb{G}' \setminus \mathbb{G}$  disappear, while particles which reach a  $+$  vertex in  $\mathbb{G}' \setminus \mathbb{G}$  stay there forever. No particles are initially placed in any of the vertices of  $\mathbb{G}' \setminus \mathbb{G}$ . We shall see in Section 2.5 why this definition is useful.

*2.3. The graphical representation.* The use of a graphical representation for describing the evolution of particle systems, originally due to Harris [8], is now a standard tool in their analysis. A more detailed account of this construction can be found in [6]. Let a graph  $\mathbb{G} = (V, E)$  and an infection parameter  $\lambda \geq 1$  be given. Consider the set  $D_{\mathbb{G}} = V \times \mathbb{R}_+$  which we think of as embedded in the plane as a disjoint collection of vertical rays, one for each vertex in  $V$ , starting at some point on the  $x$ -axis and going upwards. An element  $(v, t)$  of  $D_{\mathbb{G}}$  where  $v \in V$  and  $t \geq 0$  is therefore identified with the point on the ray corresponding to  $v$  at height  $t$  above the  $x$ -axis. We think of the second coordinate  $t$  as time.

With each ordered pair of neighboring vertices  $u \sim v$  in  $\mathbb{G}$ , we associate two Poisson point processes on  $\mathbb{R}_+$ :  $\mathcal{N}_{u,v}^\bullet, \mathcal{N}_{u,v}^\circ$ . The former has intensity measure  $(\lambda - 1)dt$  and the latter  $1 dt$ . Now fix a realization of all these processes. For each point  $t$  in  $\mathcal{N}_{u,v}^\bullet$  we add to  $D_{\mathbb{G}}$  a horizontal segment between  $(u, t)$  and  $(v, t)$ , which we think of as oriented from  $(u, t)$  to  $(v, t)$ . Similarly, for each point  $s$  in  $\mathcal{N}_{u,v}^\circ$  we add a horizontal segment between  $(u, s)$  and  $(v, s)$ , which we think of as oriented from  $(u, s)$  to  $(v, s)$ , but just below  $(v, s)$  we make a hole in the ray corresponding to  $v$ . In both cases the segments are considered to only intersect the rays emanating from  $u$  and  $v$ . The set  $D_{\mathbb{G}}$  along with all oriented segments and holes will be denoted  $D_{\mathbb{G}, \lambda}$ .

Given a realization of  $D_{\mathbb{G}, \lambda}$ , a path from  $(u, s)$  to  $(v, t)$ , where  $0 \leq s \leq t$  and  $u, v \in V$ , is a self-avoiding curve from  $(u, s)$  to  $(v, t)$  which is also a subset of



$D_{\mathbb{G},\lambda}$  and adheres to the orientation of all rays and segments. In other words, it can only go upwards on a ray and in the direction of the segment on a segment and cannot pass through holes. If  $\gamma$  is such a path, we shall write  $\gamma : (u, s) \xrightarrow{D_{\mathbb{G},\lambda}} (v, t)$ .

The following easy-to-see relation explains the connection between  $\xi_t$  and this graphical representation. Recall that for a subset of vertices  $A \subseteq \mathbb{G}$  the WB process on  $\mathbb{G}$  with infection parameter  $\lambda$  and initial configuration  $\xi_0 = A$  is denoted by  $\xi^{A,\mathbb{G},\lambda}$ . Then, for any  $A, B \subseteq \mathbb{G}$ ,

$$(2.2) \quad \mathbb{P}(\xi_t^{A,\mathbb{G},\lambda} \cap B \neq \emptyset) = \mathbb{P}(\exists \gamma : (u, 0) \xrightarrow{D_{\mathbb{G},\lambda}} (v, t) \text{ such that } u \in A, v \in B).$$

In other words, if we set

$$(2.3) \quad \xi_t^{A,\mathbb{G},\lambda} := \{v \in \mathbb{G} : \exists \gamma : (u, 0) \xrightarrow{D_{\mathbb{G},\lambda}} (v, t) \text{ such that } u \in A\},$$

then  $(\xi_t^{A,\mathbb{G},\lambda} : t \geq 0)$  is the Williams–Bjerknes process for  $A, \mathbb{G}, \lambda$ .

If  $\mathbb{G} \subset \mathbb{G}'$  has boundary conditions  $\zeta \in \mathcal{X}^{\mathbb{G}' \setminus \mathbb{G}}$ , then with the ordered neighbors  $u \sim v$  with  $u \in \mathbb{G}' \setminus \mathbb{G}, v \in \mathbb{G}$  we also associate the point processes  $\mathcal{N}_{u,v}^\bullet, \mathcal{N}_{u,v}^\circ$ . Using the same construction as above we define the set  $D_{\mathbb{G}',\lambda}$ , and now

$$(2.4) \quad \xi_t^{A,\mathbb{G}',\lambda} := \{v \in \mathbb{G} : \exists \gamma : (u, 0) \xrightarrow{D_{\mathbb{G}',\lambda}} (v, t) \text{ such that } u \in A \cup \zeta\}$$

is the Williams–Bjerknes process with boundary conditions  $\zeta$ . The usefulness of this graphical representation will become apparent in the next subsections.

**2.4. Coupling.** The graphical representation gives rise to a natural coupling between instances of  $\xi^{A,\mathbb{G},\lambda}$  for different initial configurations  $A$ , underlying subgraphs  $\mathbb{G} \subseteq \mathbb{G}'$  and infection parameters  $\lambda \geq 1$ . This is because there is a natural way to couple  $D_{\mathbb{G},\lambda}$  for different  $\mathbb{G}$ 's and  $\lambda$ 's and in light of (2.3). From this, for example, one can immediately get the following *monotonicity* (or *attractiveness*) property. If  $1 \leq \lambda \leq \lambda'$  and  $A \subseteq A' \subseteq \mathbb{G}$ , then under the above coupling,

$$(2.5) \quad \xi_t^{A,\mathbb{G},\lambda} \leq \xi_t^{A',\mathbb{G},\lambda'} \quad \text{for all } t \geq 0,$$

where the comparison is by the standard partial ordering on  $\{-, +\}^{\mathbb{G}}$ . This also extends to the case of graphs with boundary conditions in an obvious way. The monotonicity property will be used so frequently in the proofs to follow that we shall often not explicitly state it.

**2.5. Duality.** If instead of (2.3) we set

$$(2.6) \quad \hat{\xi}_t^{B,\mathbb{G},\lambda} := \{u \in \mathbb{G} : \exists \gamma : (u, 0) \xrightarrow{D_{\mathbb{G},\lambda}} (v, t) \text{ such that } v \in B\},$$

then (2.2) can be rewritten as

$$(2.7) \quad \mathbb{P}(\hat{\xi}_t^{B,\mathbb{G},\lambda} \cap A \neq \emptyset) = \mathbb{P}(\exists \gamma : (u, 0) \xrightarrow{D_{\mathbb{G},\lambda}} (v, t) \text{ such that } u \in A, v \in B)$$

and therefore

$$(2.8) \quad \mathbb{P}(\xi_t^{A, \mathbb{G}, \lambda} \cap B \neq \emptyset) = \mathbb{P}(\hat{\xi}_t^{B, \mathbb{G}, \lambda} \cap A \neq \emptyset) \quad \text{for any } A, B \subseteq \mathbb{G}.$$

Since the distribution of  $\mathcal{N}_{u,v}^\bullet$  and  $\mathcal{N}_{u,v}^\circ$  is invariant under time reversal, reading  $D_{\mathbb{G}, \lambda}$  from time  $t$  down to time 0 and with the orientation of the segments reversed, we see that  $\hat{\xi}_t^{B, \mathbb{G}, \lambda}$  is distributed as the configuration at time  $t$  of the (continuous-time) branching coalescing random walk whose generator was described in (1.12), with underlying graph  $\mathbb{G}$ , initial configuration  $B$  and parameter  $\lambda$ . Thus (2.8) gives one duality relation between  $\xi$  and  $\hat{\xi}$ .

In the presence of boundary conditions  $\zeta$  on  $\mathbb{G}' \setminus \mathbb{G}$  where  $\mathbb{G}' \supseteq \mathbb{G}$ , we can set for  $B \subseteq \mathbb{G}$

$$(2.9) \quad \hat{\xi}_t^{B, \mathbb{G}^\zeta, \lambda} := \{u \in \mathbb{G} \cup \zeta : \exists \gamma : (u, 0) \xrightarrow{D_{\mathbb{G}', \lambda}} (v, t) \text{ such that } v \in B\},$$

which yields a process whose distribution is that of the BCRW in the presence of boundary conditions, as described in the end of Section 2.2. In this case relation (2.8) becomes

$$(2.10) \quad \mathbb{P}(\xi_t^{A, \mathbb{G}^\zeta, \lambda} \cap B \neq \emptyset) = \mathbb{P}(\hat{\xi}_t^{B, \mathbb{G}^\zeta, \lambda} \cap (A \cup \zeta) \neq \emptyset)$$

for any  $A, B \subseteq \mathbb{G}$ .

In particular for  $-$ , boundary conditions (2.8) is still valid (with  $\mathbb{G}^-$  replacing  $\mathbb{G}$ ), while for  $+$  boundary conditions, we can rewrite (2.10) as

$$(2.11) \quad \mathbb{P}(\xi_t^{A, \mathbb{G}^+, \lambda} \cap B \neq \emptyset) = \mathbb{P}(\hat{\xi}_t^{B, \mathbb{G}^+, \lambda} \cap A \neq \emptyset \text{ or } \exists s \leq t : \hat{\xi}_s^{B, \mathbb{G}^+, \lambda} \not\subseteq \mathbb{G}),$$

$A, B \subseteq \mathbb{G}$ .

To describe the second duality relation between  $\xi$  and its dual, we have to define the notion of *thinning*. Fix  $p \in [0, 1]$ . For a configuration  $\xi \in \mathcal{X}$  we define the  $p$ -thinning  $\xi^{(p)}$  of  $\xi$  as the random configuration obtained from  $\xi$  by independently flipping the sign of every  $+$  vertex with probability  $1 - p$  and retaining it with probability  $p$ .

The following remarkable relation is due to Sudbury and Lloyd [21], Theorem 13. For any  $\lambda \geq 1$ ,

$$(2.12) \quad \hat{\xi}_t^{(\xi_0^{(p)})} \stackrel{d}{=} (\hat{\xi}_t^{\xi_0})^{(p)} \quad \text{where } p = 1 - \lambda^{-1}.$$

Note that  $p = \mathbb{P}(\Omega_g^0)$  by (1.4).

2.6. *Additional notation.* As usual,  $C, C'$  will denote positive constants whose value may change from one use to another.

**3. Proof of Theorems 1.2, 1.5 and corollaries.** In this section we prove Theorem 1.2, 1.5 and Corollaries 1.3 and 1.4. The proof of Theorem 1.2 is essentially linear. It consists of a sequence of lemmas, one derived from the other with the theorem following from the last. Nevertheless, to put some hierarchical structure in the proof, we have split it into two main steps which are stated in the next subsection as key lemmas. They are of interest on their own. The proofs of these lemmas are deferred to Sections 3.3 and 3.4, so that we can first show how the theorem follows from them; this is done in the Section 3.2. In this subsection we also prove the two corollaries and Theorem 1.5. They are only a short step once the theorem is established.

3.1. *Key lemmas.* The first key step is an analog of Zhang’s lemma for the contact process on regular trees [23], Proposition 5. It is the main step in Zhang’s proof for the Complete Convergence theorem in this setting. The proof was later simplified by Schonmann and Salzano [18], Proposition 1, and our arguments are essentially an adaption of the latter to this model.

LEMMA 3.1 (An analog of Zhang’s lemma). *Fix  $d \geq 3$ . For  $\lambda > \lambda_l(\mathbb{T})$ ,*

$$(3.1) \quad \inf_{t \geq 0} \mathbb{P}(0 \in \hat{\xi}_t^{0, \mathbb{T}_0^-}) > 0.$$

Next, we need tail estimates on the distribution of the inclusion time of a neighboring vertex for  $\hat{\xi}_t$ . The lemma shows that the tail of this distribution decays faster than any polynomial. We believe that this is not optimal, but for the sake of showing complete survival, this is enough.

LEMMA 3.2 (Super-polynomial decay for inclusion times). *Fix  $d \geq 3$  and  $\lambda > \lambda_l(\mathbb{T})$ . For any  $x \sim y$  neighboring vertices of  $\mathbb{T}$ ,*

$$(3.2) \quad \lim_{t \rightarrow \infty} \frac{\log \mathbb{P}(\hat{\tau}_y^{x, \mathbb{T}} > t)}{\log t} = -\infty.$$

3.2. *Proof of the theorem and corollaries.*

PROOF OF THEOREM 1.2. Fix  $d \geq 3$  and  $\lambda > \lambda_l(\mathbb{T})$ . It is clearly enough to prove

$$(3.3) \quad \mathbb{P}(\xi_t^{0, \mathbb{T}} \ni 0 \text{ for all } t \geq 0) > 0.$$

Indeed, since  $\{\xi_t^{0, \mathbb{T}} \ni 0 \text{ for all } t \geq 0\} \subseteq (\Omega_c^{0, \mathbb{T}})$ , this shows that  $\mathbb{P}(\Omega_c^{0, \mathbb{T}}) > 0$  as required.

Equation (3.3) follows from the seemingly weaker statement

$$(3.4) \quad \mathbb{P}(\xi_t^{0, \mathbb{T}_0^+} \ni 0 \text{ for all } t \geq 0 \text{ large enough}) > 0.$$

To see this, notice that (3.4) implies that there exists  $\varepsilon > 0, s > 0$  such that

$$(3.5) \quad \mathbb{P}(\xi_t^{0, \mathbb{T}_0^+} \ni 0 \text{ for all } t \geq s) > \varepsilon.$$

Therefore, writing  $\tilde{\mathbb{T}}$  for  $\mathbb{T}$  rooted at 0, and recalling Section 2.3, we have

$$(3.6) \quad \begin{aligned} & \mathbb{P}(\xi_t^{0, \mathbb{T}} \ni 0 \text{ for all } t \geq 0) \\ & \geq \mathbb{P}\left(\bigcup_{x \sim 0} \{\xi_t^{x, \tilde{\mathbb{T}}_x^+} \ni x \text{ for all } t \geq s, \mathcal{N}_{x,0}^\circ([0, s]) = 0\}\right) \\ & \geq (\varepsilon e^{-s})^d > 0. \end{aligned}$$

To establish (3.4) we will show that for all  $s \geq 0$

$$(3.7) \quad \mathbb{P}(\xi_t^{0, \mathbb{T}_0^+} \not\ni 0 \text{ for some } t \in [s^2, (s+1)^2]) \leq C s^{-2}.$$

Since these probabilities are summable in  $s = 1, 2, \dots$ , the Borel–Cantelli lemma will imply that (3.4) holds (with probability 1).

To this end, fix  $s \in \mathbb{N}$ , and let  $t_0 < t_1 < \dots < t_{s^4}$  be a partition of  $[s^2, (s+1)^2]$  into  $s^4$  sub-intervals of equal length. That is,  $t_0 = s^2, t_{s^4} = (s+1)^2$  and  $t_{k+1} - t_k = (2s+1)/s^4 \leq 3s^{-3}$ . The left-hand side in (3.7) can be bounded above by

$$(3.8) \quad \begin{aligned} & \mathbb{P}(\exists k \in 0, \dots, s^4 : \xi_{t_k}^{0, \mathbb{T}_0^+} \not\ni 0) \\ & + \mathbb{P}\left(\exists k \in 0, \dots, s^4 - 1 : \sum_{x \sim 0} \mathcal{N}_{x,0}^\circ([t_k, t_{k+1})) + \mathcal{N}_{x,0}^\bullet([t_k, t_{k+1})) \geq 2\right), \end{aligned}$$

where the second term is a bound on the probability that a site is infected and then healed during any time interval  $[t_k, t_{k+1}]$ . Using the Union Bound (and the tail of the Poisson distribution), this second term is bounded above by

$$(3.9) \quad C s^4 (t_{k+1} - t_k)^2 \leq C' s^{-2}.$$

The first term can be bounded above by

$$(3.10) \quad s^4 \sup_{t \in [s^2, (s+1)^2]} \mathbb{P}(\xi_t^{0, \mathbb{T}_0^+} \not\ni 0),$$

and it remains to bound  $\mathbb{P}(\xi_t^{0, \mathbb{T}_0^+} \not\ni 0)$ .

Let  $y$  be the parent of  $\mathbb{T}_0$ . By the duality relation (2.11) and Lemma 3.2, for all  $t \geq 0$  large enough,

$$(3.11) \quad \mathbb{P}(\xi_t^{0, \mathbb{T}_0^+} \ni 0) = \mathbb{P}(\hat{\xi}_t^{0, \mathbb{T}_0^+} \ni 0 \text{ or } \hat{\tau}_y^{0, \mathbb{T}_0^+} \leq t) \geq \mathbb{P}(\hat{\tau}_y^{0, \mathbb{T}} \leq t) \geq 1 - C t^{-3}.$$

Therefore (3.10) is bounded above by  $C s^{-2}$ . Combining this with (3.9) we see that (3.7) holds as desired. This completes the proof of the theorem.  $\square$

In fact, the following lemma, which is required for the proofs of the corollaries, shows that above  $\lambda_c(\mathbb{T})$ , global and complete survival are equivalent up to an event with zero probability.

LEMMA 3.3. *Let  $d \geq 3$  and  $\lambda > \lambda_c(\mathbb{T})$ . Then for any finite  $\xi_0 \in \mathcal{X}$ ,*

$$(3.12) \quad \mathbb{P}(\Omega_c^{\xi_0}) = \mathbb{P}(\Omega_g^{\xi_0}).$$

PROOF. Recall that definition (1.5) of  $\Omega_c^{\xi_0}$  does not depend on the observed vertex  $x \in \mathbb{T}$  and we can therefore choose  $x = 0$ . For  $s \geq 0$  and  $u \geq 0$  define

$$(3.13) \quad \mathcal{A}^{\xi_0}(s, u) := \{\xi_t^{\xi_0} \ni 0 \text{ for all } t \in [s, s + u]\}.$$

Since  $\Omega_c^{\xi_0} = \bigcup_{s \geq 0} \bigcap_{u \geq 0} \mathcal{A}^{\xi_0}(s, u)$  we have

$$(3.14) \quad \lim_{s \rightarrow \infty} \lim_{u \rightarrow \infty} \mathbb{P}(\mathcal{A}^{\xi_0}(s, u)) = \mathbb{P}(\Omega_c^{\xi_0}).$$

Now, we claim that as  $r \uparrow \infty$ ,

$$(3.15) \quad \mathbb{P}(\Omega_c^{B_0(r)}) \uparrow \alpha = 1.$$

Indeed, by monotonicity the limit exists, and so we may write

$$(3.16) \quad \begin{aligned} \mathbb{P}(\Omega_c^{\xi_0}) &= \lim_{s \rightarrow \infty} \lim_{u \rightarrow \infty} \mathbb{P}(\Omega_c^{\xi_0} \cap \mathcal{A}^{\xi_0}(s, u)) \\ &= \lim_{s \rightarrow \infty} \lim_{u \rightarrow \infty} \mathbb{P}(\mathcal{A}^{\xi_0}(s, u)) \mathbb{E}[\mathbb{P}(\Omega_c^{\xi_0} | \xi_{s+u}^{\xi_0}) | \mathcal{A}^{\xi_0}(s, u)] \\ &\leq \mathbb{P}(\Omega_c^{\xi_0}) \alpha. \end{aligned}$$

The last inequality follows from monotonicity again, since  $\xi_{s+u}^{\xi_0}$  must be included in some  $B_0(r)$  for  $r$  large enough. Since  $\mathbb{P}(\Omega_c^{\xi_0}) > 0$  it follows that  $\alpha$  must be 1.

Now if  $\lambda > \lambda_l(\mathbb{T})$ , then for any vertex  $x \in \mathbb{T}$  and  $r \geq 0$ , there exists  $s_{x,r} < \infty$  such that

$$(3.17) \quad \mathbb{P}(\tau_{B_0(r)}^x < s_{x,r}) \geq \frac{1}{2} \mathbb{P}(\Omega_l^x) > \varepsilon$$

for some  $\varepsilon > 0$  independent of  $x$  or  $r$ . Since on  $\Omega_g^{\xi_0}$  infected vertices exist at all times, it follows from monotonicity and Markov's property that for all  $r \geq 0$ ,

$$(3.18) \quad \mathbb{P}(\Omega_g^{\xi_0} \cap \{\tau_{B_0(r)}^{\xi_0} = \infty\}) = 0.$$

Consequently we may write

$$(3.19) \quad \mathbb{P}(\Omega_c^{\xi_0} | \Omega_g^{\xi_0}) \geq \mathbb{P}(\Omega_c^{\xi_0} | \tau_{B_0(r)}^{\xi_0} < \infty) \mathbb{P}(\tau_{B_0(r)}^{\xi_0} < \infty | \Omega_g^{\xi_0}) \geq \mathbb{P}(\Omega_c^{B_0(r)}).$$

Taking  $r \rightarrow \infty$  and using (3.15) we get  $\mathbb{P}(\Omega_c^{\xi_0} | \Omega_g^{\xi_0}) = 1$ , as desired.  $\square$

PROOF OF COROLLARY 1.3. This is an immediate consequence of Theorem 1.2 and Lemma 3.3. Indeed once  $\lambda > \lambda_l(\mathbb{T})$  we have for all  $\xi_0 \in \mathcal{X}$ ,

$$(3.20) \quad \lim_{t \rightarrow \infty} \xi_t^{\xi_0} = 1_{\Omega_g^{\xi_0}} \mathbb{T} + (1 - 1_{\Omega_g^{\xi_0}}) \emptyset, \quad \mathbb{P}\text{-almost surely,}$$

where we recall that the topology in  $\mathcal{X}$ , viewed as the space of functions on  $\mathbb{T}$ , is that of pointwise convergence. This immediately gives (1.10) and shows that any invariant measure must be a convex combination of  $\delta_{\mathbb{T}}$  and  $\delta_{\emptyset}$ .  $\square$

PROOF OF COROLLARY 1.4. Part (1) follows immediately from (3.20). Part (2) holds because once  $\lambda < \lambda_l(\mathbb{T})$ , starting from a finite configuration, the infection either dies out, or survives globally but not locally. In both cases, every vertex will eventually become healthy and therefore fixate.  $\square$

PROOF OF THEOREM 1.5. Any of the duality relations can be used to prove this theorem. When  $\lambda > \lambda_l(\mathbb{T})$ , Corollary 1.3, relation (2.8) and (1.4) imply that for all nonempty  $\hat{\xi}_0$  and any  $A \subseteq \mathbb{T}$  as  $t \rightarrow \infty$ ,

$$(3.21) \quad \mathbb{P}(\hat{\xi}_t^{\hat{\xi}_0} \cap A \neq \emptyset) = \mathbb{P}(\xi_t^A \cap \hat{\xi}_0 \neq \emptyset) \rightarrow \mathbb{P}(\Omega_g^A) = 1 - \lambda^{-|A|}.$$

This shows (1.14). On the other hand, if  $\lambda \in [1, \lambda_l(\mathbb{T}))$ , then for all finite  $A \subseteq \mathbb{T}$  and finite  $\hat{\xi}_0 \in \mathcal{X}$ , the above becomes

$$(3.22) \quad \mathbb{P}(\hat{\xi}_t^{\hat{\xi}_0} \cap A \neq \emptyset) = \mathbb{P}(\xi_t^A \cap \hat{\xi}_0 \neq \emptyset) \rightarrow 0$$

as  $t \rightarrow \infty$ . This shows (1.13).  $\square$

3.3. *Proof of Lemma 3.1.* The proof will be carried out using a number of lemmas. For an infinite connected bounded-degree graph  $\mathbb{G} = (V, E)$ , possibly with associated boundary conditions, we shall write  $\lambda_g(\mathbb{G})$  and  $\lambda_l(\mathbb{G})$  for the threshold value of  $\lambda$  for the possibility of global and local survival for  $\xi$  when the underlying graph is  $\mathbb{G}$ . Formally,

$$(3.23) \quad \begin{aligned} \lambda_g(\mathbb{G}) &:= \inf\{\lambda > 0 : \mathbb{P}(\Omega_g^{\xi_0, \mathbb{G}}) > 0\}, \\ \lambda_l(\mathbb{G}) &:= \inf\{\lambda > 0 : \mathbb{P}(\Omega_l^{\xi_0, \mathbb{G}}) > 0\}, \end{aligned}$$

where  $\Omega_g^{\xi_0, \mathbb{G}}$  and  $\Omega_l^{\xi_0, \mathbb{G}}$  are defined as in (1.3) and (1.6) with  $\mathbb{G}$  being the underlying graph, and  $\xi_0$  is any finite nonempty initial configuration. Notice that as  $\mathbb{G}$  may have associated boundary conditions, it is no longer clear that  $\lambda_g(\mathbb{G}) = 1$ .

Recall that  $\mathbb{T}_0$  is the sub-tree of  $\mathbb{T}$  rooted at 0. Our first lemma shows that if a graph  $\mathbb{G}$  (with or without boundary conditions) contains a copy of this sub-tree which is accessible only through its root, then its threshold values are at least as small as those of  $\mathbb{T}$ . Note by the monotonicity statement (2.5) it is enough to show this for  $\mathbb{G} = \mathbb{T}_0^-$ .

LEMMA 3.4. *Fix  $d \geq 3$ . If  $\mathbb{G}$  is any infinite connected bounded-degree graph, possibly with associated boundary conditions, that contains a copy of  $\mathbb{T}_0$ , which is connected to the rest of the graph only through its root 0, then*

$$(3.24) \quad \lambda_g(\mathbb{G}) \leq \lambda_g(\mathbb{T}) = 1,$$

$$(3.25) \quad \lambda_l(\mathbb{G}) \leq \lambda_l(\mathbb{T}).$$

PROOF. By monotonicity it is enough to show this for  $\mathbb{G} = \mathbb{T}_0^-$ . Fix any  $\lambda$  such that

$$(3.26) \quad \mathbb{P}(\Omega_g^{0,\mathbb{T}} \setminus \Omega_l^{0,\mathbb{T}}) > 0.$$

This is always possible, since  $\lambda_g(\mathbb{T}) < \lambda_l(\mathbb{T})$  for all  $d \geq 3$ , as shown in part (1) of Proposition 1.1. Clearly, the distribution of  $\xi_t^{0,\mathbb{T}}$  on  $\mathbb{T}_x$  is the same for any neighbor  $x$  of 0. Also, at any time  $t$  we have  $|\xi_t^{0,\mathbb{T}}| < \infty$ . These two facts, along with (3.26), imply that we may find  $T > 0$  and a finite subset of vertices  $A \subseteq \mathbb{T}_x \setminus \{x\}$  such that the following event has positive probability:

$$(3.27) \quad \{\xi_T^{0,\mathbb{T}} \cap \mathbb{T}_x = A \text{ and } \xi_t^{0,\mathbb{T}} \cap \mathbb{T}_x \neq \emptyset, x \notin \xi_t^{0,\mathbb{T}} \text{ for all } t \geq T\}.$$

By the Markov property, it follows that

$$(3.28) \quad \mathbb{P}(\xi_t^{A,\mathbb{T}} \neq \emptyset \text{ and } x \notin \xi_t^{A,\mathbb{T}} \text{ for all } t \geq 0) > 0.$$

Observe that this probability does not change if we add  $-$  boundary conditions on  $\mathbb{T} \setminus \mathbb{T}_x$ . Since, in addition, any two finite configurations are obtainable from each other using a finite number of transitions, we arrive at

$$(3.29) \quad \mathbb{P}(\xi_t^{x,\mathbb{T}_x^-} \neq \emptyset \text{ for all } t \geq 0) > 0.$$

This shows (3.24) as  $\mathbb{T}_x^-$  is isomorphic to  $\mathbb{T}_0^-$ , and  $\lambda$  can be chosen arbitrarily close to  $\lambda_g(\mathbb{T})$ .

Next, suppose that  $\lambda' > \lambda_l(\mathbb{T})$ . Since  $\Omega_l^{0,\mathbb{T},\lambda'}$  occurs with positive probability, there must exist  $\delta > 0$  and  $T_x > 0$  for all  $x \in \mathbb{T}$  such that

$$(3.30) \quad \mathbb{P}(\tau_0^{x,\mathbb{T},\lambda'} < T_x) > \delta \quad \forall x \in \mathbb{T}.$$

This still holds, under  $-$  boundary conditions, that is,

$$(3.31) \quad \mathbb{P}(\tau_0^{x,\mathbb{T}_0^-, \lambda'} < T_x) > \delta \quad \forall x \in \mathbb{T}_0^-.$$

Since  $\lambda' > \lambda$ , it follows from (3.29) via monotonicity that

$$(3.32) \quad \{\xi_t^{0,\mathbb{T}_0^-, \lambda'} \neq \emptyset; t \geq 0\}$$

occurs with positive probability. But on this event, by (3.31) and monotonicity, at all times  $t \geq 0$  there will be a vertex  $x \in \xi_t^{0,\mathbb{T}_0^-, \lambda'}$ , from which there is at least

$\delta$  probability of reinfecting the origin within  $T_x$  time. It follows then from the Markov property that the probability of (3.32) and the origin being infected only finitely many times is 0. Consequently  $\mathbb{P}(\Omega_l^{0, \mathbb{T}_0^-, \lambda'}) > 0$  which implies  $\lambda_l(\mathbb{T}_0^-) < \lambda'$  and since  $\lambda'$  is arbitrarily close to  $\lambda_l(\mathbb{T})$ , inequality (3.25) follows.  $\square$

For  $x \in \mathbb{T}$  and  $y \in \mathbb{T}_x$  we let  $\mathbb{T}_{xy} := (\mathbb{T}_x \setminus \mathbb{T}_y) \cup \{y\}$ . Then we have

LEMMA 3.5. *Fix  $d \geq 3$ . Let  $x \in \mathbb{T}$  and  $y \in \mathbb{T}_x$ . Then:*

- (1)  $\lambda_g(\mathbb{T}_x^-) = \lambda_g(\mathbb{T}_{xy}^-) = \lambda_g(\mathbb{T}) = 1$ ;
- (2)  $\lambda_l(\mathbb{T}_x^-) = \lambda_l(\mathbb{T}_{xy}^-) = \lambda_l(\mathbb{T})$ ,

where for the boundary conditions, both  $\mathbb{T}_x$  and  $\mathbb{T}_{xy}$ , are treated as subgraphs of  $\mathbb{T}$ .

PROOF. Monotonicity implies that

$$(3.33) \quad \lambda_g(\mathbb{T}_{xy}^-) \geq \lambda_g(\mathbb{T}_x^-) \geq \lambda_g(\mathbb{T}) = 1 \quad \text{and} \quad \lambda_l(\mathbb{T}_{xy}^-) \geq \lambda_l(\mathbb{T}_x^-) \geq \lambda_l(\mathbb{T}).$$

On the other hand, both  $\mathbb{T}_x^-$  and  $\mathbb{T}_{xy}^-$  contain a copy of  $\mathbb{T}_0$  which is connected to the rest of the graph only through its root. Therefore the opposite inequalities are a consequence of Lemma 3.4.  $\square$

PROOF OF LEMMA 3.1. Fix  $\lambda > \lambda_l(\mathbb{T})$ . By the duality relation (2.8) showing (3.1) is equivalent to showing

$$(3.34) \quad \inf_{t \geq 0} \mathbb{P}(0 \in \xi_t^{0, \mathbb{T}_0^-}) > 0.$$

We first argue that there exists  $r > 0, s > 0$  and

$$(3.35) \quad p > 1/\sqrt{d-1}$$

such that for any vertex  $x$  of  $\mathbb{T}_0$  whose distance from 0 is  $r$ , we have

$$(3.36) \quad \mathbb{P}(\xi_s^{0, \mathbb{T}_{0x}^-} \ni x) > p^r.$$

Indeed, by Lemma 3.5 we know that  $\lambda > \lambda_l(\mathbb{T}_0^-) = \lambda_l(\mathbb{T})$ . Therefore, we may find  $p' > 1/\sqrt{d-1}$  and integer  $r' > 0$  large enough such that

$$(3.37) \quad (p')^{r'} < \frac{1}{2} \mathbb{P}(\Omega_l^{0, \mathbb{T}_0^-}).$$

Enumerating the vertices on some path going down from 0 as

$$(3.38) \quad 0 = x_0, x_1, x_2, \dots,$$

there exists  $s' > 0$  such that

$$(3.39) \quad \mathbb{P}(\tau_{x_{r'}}^{0, \mathbb{T}_0^-} \leq s') > (p')^{r'}.$$



Using monotonicity and the Markov property we may iterate the above to get for all  $k \geq 1$

$$(3.40) \quad \mathbb{P}(\tau_{x_{kr'}}^{0, \mathbb{T}_{0x_{kr'}}^-} \leq ks') = \mathbb{P}(\tau_{x_{kr'}}^{0, \mathbb{T}_0^-} \leq ks') > (p')^{kr'}$$

Now, write

$$(3.41) \quad \begin{aligned} & (p')^{kr'} e^{-d} \\ & < \mathbb{P}(\tau_{x_{kr'}}^{0, \mathbb{T}_{0x_{kr'}}^-} \leq ks', \xi_t^{0, \mathbb{T}_{0x_{kr'}}^-} \ni x_{kr'} \text{ for all } t \in [\tau_{x_{kr'}}^{0, \mathbb{T}_{0x_{kr'}}^-}, \tau_{x_{kr'}}^{0, \mathbb{T}_{0x_{kr'}}^-} + 1]) \\ & \leq \sum_{t=0,1,\dots}^{\lfloor ks'+1 \rfloor} \mathbb{P}(\xi_t^{0, \mathbb{T}_{0x_{kr'}}^-} \ni x_{kr'}) \leq (ks' + 2) \mathbb{P}(\xi_{s''}^{0, \mathbb{T}_{0x_{kr'}}^-} \ni x_{kr'}), \end{aligned}$$

where  $s''$  is the index of the maximal term in the last sum. Therefore (for  $k$  large enough),

$$(3.42) \quad \mathbb{P}(\xi_{s''}^{0, \mathbb{T}_{0x_{kr'}}^-} \ni x_{kr'}) > C(p')^{kr'} / (ks'),$$

which implies that (3.36) holds with  $r := kr'$ ,  $s := s''$ ,  $x = x_r$  and some  $p \in (1/\sqrt{d-1}, p')$ , once we choose  $k$  large enough. Finally, since the choice of path in (3.38) is arbitrary,  $x_r$  can be replaced with any vertex whose distance from 0 is  $r$ .

Next we introduce a modified version of  $\xi^{0, \mathbb{T}_0^-}$  which we denote by  $\xi'_t$ . The process  $\xi'_t$  is still Markovian and takes values in the space of all configurations on  $\mathbb{T}_0$ . It starts from a single infection at the origin and evolves exactly as  $\xi^{0, \mathbb{T}_0^-}$  does, only that at times  $t = ks$  where  $k = 1, \dots$ , we heal all vertices whose distance from the origin is greater than  $kr$ , and we heal and keep healed the ones whose distance from the origin is less than  $kr$ . Formally, we set

$$(3.43) \quad \begin{aligned} \xi'_{ks}(x) &= - && \text{for } x \text{ s.t. } \rho(0, x) > kr, \\ \xi'_t(x) &= - && \text{for } x \text{ s.t. } \rho(0, x) < kr \text{ and all } t \geq ks. \end{aligned}$$

By monotonicity  $\xi_t^{0, \mathbb{T}_0^-}$  stochastically dominates  $\xi'_t$  for all  $t$ . At the same time, it is easy to see that the process  $Z_k := |\xi'_{tk}|$  where  $k = 0, \dots$ , is a branching process with mean reproduction  $\mu := \mathbb{E}Z_1 = (d-1)^r p^r > (d-1)^{r/2} > 1$  and  $\mathbb{E}(Z_1)^2 < \infty$ . Therefore, there exists  $\varepsilon > 0$  such that for all  $k$ ,

$$(3.44) \quad \mathbb{P}(|\xi_{ks}^{0, \mathbb{T}_0^-} \cap S_0(kr)| > \frac{1}{2}((d-1)p)^{kr}) \geq \mathbb{P}(Z_k > \frac{1}{2}\mu^k) > \varepsilon,$$

where we recall  $S_x(r) := \{y \in \mathbb{T} : \rho(x, y) = \lfloor r \rfloor\}$  for  $x \in \mathbb{T}$  and  $r \geq 0$ .

Now for  $k$  whose precise value will be chosen later, set  $\tilde{s} := ks$ ,  $\tilde{r} := kr$  and  $\tilde{p} := p^{\tilde{r}}$ . We now prove by induction that for all positive and even  $n \in \mathbb{N}$

$$(3.45) \quad p_n := \mathbb{P}(\xi_{n\tilde{s}}^{0, \mathbb{T}_0^-} \ni 0) > \frac{1}{2}\varepsilon\tilde{p}.$$

For  $n = 2$ , iterating (3.36)  $k$  times along the path from  $x \in S_0(\tilde{r})$  to 0 and using (3.44),

$$(3.46) \quad p_2 \geq \mathbb{P}(|\xi_{\tilde{s}}^{0, \mathbb{T}_0^-} \cap S_0(\tilde{r})| \geq 1) \mathbb{P}(\xi_{2\tilde{s}}^{0, \mathbb{T}_0^-} \ni 0 | |\xi_{\tilde{s}}^{0, \mathbb{T}_0^-} \cap S_0(\tilde{r})| \geq 1) \geq \varepsilon \tilde{p}.$$

For  $n + 2 > 2$ , we can bound  $p_{n+2}$  below by

$$(3.47) \quad \begin{aligned} & \mathbb{P}(|\xi_{\tilde{s}}^{0, \mathbb{T}_0^-} \cap S_0(\tilde{r})| > \frac{1}{2}((d-1)p)^{\tilde{r}}) \\ & \times \mathbb{P}(|\xi_{(n+1)\tilde{s}}^{0, \mathbb{T}_0^-} \cap S_0(\tilde{r})| \geq 1 | |\xi_{\tilde{s}}^{0, \mathbb{T}_0^-} \cap S_0(\tilde{r})| > \frac{1}{2}((d-1)p)^{\tilde{r}}) \\ & \times \mathbb{P}(\xi_{(n+2)\tilde{s}}^{0, \mathbb{T}_0^-} \ni 0 | |\xi_{(n+1)\tilde{s}}^{0, \mathbb{T}_0^-} \cap S_0(\tilde{r})| \geq 1, |\xi_{\tilde{s}}^{0, \mathbb{T}_0^-} \cap S_0(\tilde{r})| > \frac{1}{2}((d-1)p)^{\tilde{r}}). \end{aligned}$$

We can bound below the first term by  $\varepsilon$  using (3.44) and the last term by  $\tilde{p}$  using the argument in (3.46). For the middle one, we use the fact that for any configuration  $\eta$  on  $\mathbb{T}_0$

$$(3.48) \quad \xi_t^{\eta, \mathbb{T}_0^-} \geq \bigvee_{x \in S_0(\tilde{r})} \xi_t^{\eta_x, \mathbb{T}_x^-} \quad \text{for all } t \geq 0,$$

where  $(\xi^{\eta_x, \mathbb{T}_x^- : x \in S_0(\tilde{r}))$  are independent,  $\eta_x$  is the restriction of  $\eta$  to  $\mathbb{T}_x$ . This is because of monotonicity (2.5), the tree-structure and the choice of boundary conditions. Using this and the induction hypothesis for  $p_n$ , we can bound below the second term in (3.47) by

$$(3.49) \quad 1 - (1 - p_n)^{(1/2)((d-1)p)^{\tilde{r}}} \geq 1 - \exp(-\frac{1}{4}\varepsilon((d-1)p^2)^{kr}).$$

In light of (3.35), by choosing  $k$  large enough we can guarantee that the right-hand-side above is at least  $\frac{1}{2}$  and conclude that (3.46) holds for  $p_{n+2}$  as well.

Once we have (3.45) for all even positive  $n$ , it is only a short step to complete the proof of the lemma. Indeed, for any  $t \geq 0$ , find a positive even  $n$  such that  $n\tilde{s} \leq t < (n+2)\tilde{s}$ , and write

$$(3.50) \quad \begin{aligned} \mathbb{P}(\xi_t^{0, \mathbb{T}_0^-} \ni 0) & \geq \mathbb{P}(\xi_{n\tilde{s}}^{0, \mathbb{T}_0^-} \ni 0) \mathbb{P}(\mathcal{N}_{x,0}^c([n\tilde{s}, t]) = 0 \text{ for all } x \sim 0) \\ & > \frac{1}{2}\varepsilon \tilde{p} e^{-d} > 0. \end{aligned}$$

This shows (3.34) and completes the proof.  $\square$

3.4. Proof of Lemma 3.2. The proof will consist of a sequence of lemmas.

LEMMA 3.6. Fix  $d \geq 3$  and  $\lambda \geq 1$ . For all  $\delta_1 > 0$  there exists  $b > 0$  such that for all  $t \geq 0$  large enough,

$$(3.51) \quad \mathbb{P}(\hat{\xi}_t^{0, \mathbb{T}} \subseteq B_0(bt)) \geq 1 - e^{-\delta_1 t}.$$

PROOF. The proof will follow by coupling of  $\hat{\xi}_t^{0,\mathbb{T}}$  with a (continuous time) branching random walk on  $\mathbb{R}$ , whose growth rate is well controlled. (Alternatively, one can use a comparison to last passage percolation on  $\mathbb{T}$ , or just prove this via elementary methods.) To this end, we first introduce the following variant of  $\hat{\xi}_t^{0,\mathbb{T}}$  which we denote by  $(\tilde{\xi}_t^{0,\mathbb{T}} : t \geq 0)$ . The process  $\tilde{\xi}_t^{0,\mathbb{T}}$  starts from a single particle at 0 and evolves as  $\hat{\xi}_t^{0,\mathbb{T}}$  does, except for two differences. First, there are no coalescences, that is, more than one particle can share a single vertex. Second, whenever a particle at vertex  $v$  moves to (rate 1) or produces a particle at (rate  $\lambda - 1$ ) its parent  $u \sim v$ , one of its children  $w \sim v$  are chosen (according to some fixed method) instead of  $u$ .

If  $\hat{R}_t^{0,\mathbb{T}}$ ,  $\tilde{R}_t^{0,\mathbb{T}}$ , denote the maximal graph-distance of a particle in  $\hat{\xi}_t^{0,\mathbb{T}}$ , respectively,  $\tilde{\xi}_t^{0,\mathbb{T}}$ , from the origin, then by a straightforward coupling,

$$(3.52) \quad \hat{R}_t^{0,\mathbb{T}} \leq_s \tilde{R}_t^{0,\mathbb{T}}.$$

At the same time, the process  $(N_t : t \geq 0)$  defined as

$$(3.53) \quad N_t := \sum_{x \in \tilde{\xi}_t^{0,\mathbb{T}}} \delta_{\rho(0,x)},$$

is a continuous time branching random walk on  $\mathbb{R}_+$  with  $N_0 = \delta_0$  and whose reproduction measure on  $\mathbb{R}$  is  $\delta_1$  at rate  $d$  and  $\delta_1 + \delta_0$  at rate  $(\lambda - 1)d$ . That is, a particle at displacement  $r \geq 0$  is replaced by a particle at displacement  $r + 1$  at rate  $d$  and by two particles: at  $r$  and  $r + 1$ , at rate  $(\lambda - 1)d$ .

Writing

$$(3.54) \quad \begin{aligned} \mathbb{P}(\hat{\xi}_t^{0,\mathbb{T}} \subseteq B_0(bt)) &= \mathbb{P}(\hat{R}_t^{0,\mathbb{T}} \leq bt) \geq \mathbb{P}(\tilde{R}_t^{0,\mathbb{T}} \leq bt) \\ &= 1 - \mathbb{P}(N_t(bt, \infty) \geq 1) \geq 1 - \mathbb{E}(N_t(bt, \infty)). \end{aligned}$$

Now Theorem 4 in [1] says (note that “nonlattice” there refers to the distribution of times between reproductions, not the support of the reproduction measures)

$$(3.55) \quad \frac{\log \mathbb{E}(N_t[bt, \infty))}{t} \rightarrow \alpha^*(b),$$

where  $\alpha^*(b)$  depends on the Laplace transform of the reproduction measures (an analog of the Legendre transform in Cramér’s theorem) and in the case of reproduction measures with finite support and exponential reproduction times, can be made arbitrarily small by choosing  $b$  large enough.  $\square$

LEMMA 3.7. *Let  $d \geq 3$  and  $\lambda > 1$ . There exist  $a > 0$ ,  $\delta_2 > 0$  such that for all  $t \geq 0$  large enough,*

$$(3.56) \quad \mathbb{P}(|\hat{\xi}_t^{0,\mathbb{T}}| \geq e^{at}) \geq 1 - e^{-\delta_2 t}.$$

PROOF. We shall omit the superscript  $\mathbb{T}$  as all processes in this proof run on  $\mathbb{T}$ . By the thinning relationship (2.12) with initial state  $\delta_0$ , we can write for  $p = \mathbb{P}(\Omega_g^0)$  and any  $a > 0$ ,

$$(3.57) \quad \mathbb{P}(|(\xi_t^0)^{(p)}| \geq e^{at}) = \mathbb{P}(|\hat{\xi}_t^{(\delta_0^{(p)})}| \geq e^{at}) = p\mathbb{P}(|\hat{\xi}_t^0| \geq e^{at}).$$

Since every infected vertex in  $\xi_t^0$  stays infected in  $(\xi_t^0)^{(p)}$  with probability  $p$  independently of other vertices, it follows from Cramér’s theorem applied to this sequence of Bernoulli( $p$ ) random variables that

$$(3.58) \quad \begin{aligned} \mathbb{P}(|(\xi_t^0)^{(p)}| \geq e^{at}) &\geq \mathbb{P}\left(|\xi_t^0| \geq \frac{2}{p}e^{at}\right)\mathbb{P}\left(|(\xi_t^0)^{(p)}| \geq e^{at} \mid |\xi_t^0| \geq \frac{2}{p}e^{at}\right) \\ &\geq \mathbb{P}(|\xi_t^0| \geq e^{a't})(1 - e^{-C'e^{Cat}}), \end{aligned}$$

where  $a' = 2a$  and  $t \geq 0$  is large enough. Therefore it is enough to show that for some  $\delta > 0$  and all large  $t$ ,

$$(3.59) \quad p^{-1}\mathbb{P}(|\xi_t^0| \geq e^{a't}) \geq \mathbb{P}(|\xi_t^0| \geq e^{a't} \mid \Omega_g^0) \geq 1 - e^{-\delta t}.$$

As discussed in the **Introduction** [see the discussion above (1.2)], the transitions of  $|\xi_t^0|$  are those of a nearest neighbor random walk with drift  $\frac{\lambda-1}{\lambda+1}$  and an absorbing state at 0. These transitions occur at rate

$$(3.60) \quad (\lambda + 1)|\partial_{\mathbb{T}}\xi_t^0| \geq (\lambda + 1)|\xi_t^0|,$$

where we recall that  $\partial_{\mathbb{T}}\xi_t^0$  denotes the set of edges of  $\mathbb{T}$  with exactly one vertex in  $\xi_t^0$ , and the last inequality holds since  $|\partial_{\mathbb{T}}A| \geq (d - 2)|A|$  for any finite  $A \subset \mathbb{T}$ .

It follows that we can couple  $(|\xi_t^0| : t \geq 0)$  with a continuous time birth-and-death process  $(Y_t : t \geq 0)$  on  $\mathbb{N}$  with birth rates  $p(y) = \lambda y$  and death rates  $q(y) = y$  such that both processes start from 1, make the same transitions and that times between successive transitions of  $|\xi_t^0|$  are less or equal than the corresponding ones of  $Y_t$ . Thus if we define

$$(3.61) \quad T := \inf\{s \geq 0 : |\xi_s^0| \geq 2e^{a's}\} \quad \text{and} \quad S := \inf\{s \geq 0 : Y_s \geq 2e^{a's}\},$$

it follows from this coupling that for all  $t \geq 0$ ,

$$(3.62) \quad \mathbb{P}(T \leq t \mid \Omega_g^0) \geq \mathbb{P}(S \leq t \mid Y_s > 0, s \geq 0).$$

For  $Y_t$ , either explicit calculation or, for example, [15] shows that there exist  $a' > 0$  and  $\delta' > 0$  such that for all large  $t \geq 0$ ,

$$(3.63) \quad \mathbb{P}(S \leq t \mid Y_s > 0, s \geq 0) \geq \mathbb{P}(Y_t \geq 2e^{a't} \mid Y_s > 0, s \geq 0) \geq 1 - e^{-\delta't}.$$

At the same time,

$$(3.64) \quad \begin{aligned} &\mathbb{P}(|\xi_t^0| > e^{a't} \mid \Omega_g^0, T \leq t) \\ &\geq \mathbb{P}(|\xi_{T+s}^0| - |\xi_T^0| > -e^{a't}, \forall s \geq 0 \mid \Omega_g^0, T \leq t) \\ &\geq \mathbb{P}(|\xi_{T+s}^0| - |\xi_T^0| > -e^{a't}, \forall s \geq 0 \mid T \leq t) = 1 - \lambda^{-e^{a't}}, \end{aligned}$$

where  $1 - \lambda^{-e^{a't}}$  is the standard gambler-ruin probability.

Combining (3.62), (3.63) and (3.64) we arrive at

$$(3.65) \quad \mathbb{P}(|\hat{\xi}_t^0| > e^{a't} | \Omega_g^0) \geq 1 - e^{-\delta't} - \lambda^{-e^{a't}} \geq 1 - e^{-\delta t}$$

for a suitable  $0 < \delta < \delta'$ , as required in (3.59). The result follows.  $\square$

As an immediate consequence of Lemmas 3.6 and 3.7, we get

LEMMA 3.8. *Fix  $d \geq 3$  and  $\lambda > 1$ . There exists  $\delta > 0$ ,  $a > 0$ ,  $b > 0$  such that for all  $t \geq 0$  large enough,*

$$(3.66) \quad \mathbb{P}(|\hat{\xi}_t^{0, \mathbb{T}} \cap B_0(bt)| \geq e^{at}) \geq 1 - e^{-\delta t}.$$

PROOF. Use Lemmas 3.6 and 3.7 and the union bound.  $\square$

The next lemma shows that when  $\lambda > \lambda_l(\mathbb{T})$ , starting from a single occupied vertex  $x$ , any neighboring vertex  $y \sim x$  will eventually become occupied, with high probability, even if we restrict the underlying graph to a finite subset of  $\mathbb{T}$ , but as long as this sub-graph is large enough.

LEMMA 3.9. *Fix  $d \geq 3$  and  $\lambda > \lambda_l(\mathbb{T})$ . Let  $x$  and  $y$  be neighboring vertices in  $\mathbb{T}$ . For all  $\beta > 0$ , there exists  $r > 0$ ,  $u > 0$  such that*

$$(3.67) \quad \mathbb{P}(\hat{\tau}_y^{x, B_x^-(r)} \leq u) \geq e^{-\beta},$$

where  $B_x(r)$  is the ball of radius  $r$  (in the graph-distance) around  $x$ , viewed as a sub-graph of  $\mathbb{T}$ .

PROOF. Fix  $\lambda > \lambda_l(\mathbb{T})$ . Without loss of generality, we can assume that  $x = 0$  and that  $y$  is the parent of 0. We first show that

$$(3.68) \quad \mathbb{P}(\hat{\tau}_y^{0, \mathbb{T}} < \infty) = 1.$$

This will be a consequence of the existence of  $\varepsilon > 0$  and  $T_z < \infty$  for all  $z \in \mathbb{T}$  such that

$$(3.69) \quad \mathbb{P}(\hat{\tau}_y^{z, \mathbb{T}} \leq T_z) \geq \varepsilon \quad \forall z \in \mathbb{T}.$$

Indeed, since  $\hat{\xi}^{0, \mathbb{T}}$  never dies, at any time  $t \geq 0$ , there will be at least one occupied vertex  $z$ , which (by monotonicity and the Markov property) will give rise to a particle at  $y$  within  $T_z$  time with probability at least  $\varepsilon$ . It follows that the probability of  $y$  never being occupied is 0, which is what we need for (3.68). To see (3.69), write

$$(3.70) \quad \begin{aligned} \mathbb{P}(\hat{\tau}_y^{z, \mathbb{T}} \leq T_z) &\geq \mathbb{P}(\hat{\xi}_{T_z}^{z, \mathbb{T}} \ni y) = \mathbb{P}(\xi_{T_z}^{y, \mathbb{T}} \ni z) \\ &= \mathbb{P}(\xi_{T_z}^{y, \mathbb{T}} \ni z, \tau_z^{y, \mathbb{T}} \leq T_z) \geq \mathbb{E}(\mathbb{P}(\xi_{T_z - \tau_z^{y, \mathbb{T}}}^{z, \mathbb{T}} \ni z) 1_{\{\tau_z^{y, \mathbb{T}} \leq T_z\}}). \end{aligned}$$

The first factor in the last expectation is bounded below by some  $\delta_1 > 0$  independently of  $z$  due to Lemma 3.1, duality and monotonicity. At the same time, as in (3.30) there exists  $\delta_2 > 0$  and a choice for  $\{T_z > 0, z \in \mathbb{T}\}$  such that

$$(3.71) \quad \mathbb{P}(\tau_z^{y, \mathbb{T}} \leq T_z) \geq \frac{1}{2} \mathbb{P}(\Omega_t^{y, \mathbb{T}}) \geq \delta_2 \quad \forall z \in \mathbb{T}.$$

Using this in (3.70) we obtain (3.69) with  $\varepsilon = \delta_1 \delta_2$  and this choice for  $(T_z)_z$ . Therefore also (3.68) holds.

Since at all times  $t$  we have  $|\hat{\xi}_t^{0, \mathbb{T}}| < \infty$ ,

$$(3.72) \quad \{\hat{\tau}_y^{0, \mathbb{T}} < \infty\} = \bigcup_{r=1}^{\infty} \bigcup_{u=1}^{\infty} \{\hat{\tau}_y^{0, \mathbb{T}} \leq u, \hat{\xi}_t \subseteq B_0(r); t \in [0, \hat{\tau}_y^{0, \mathbb{T}}]\},$$

and as the sequence of events on the right-hand side is monotone increasing in  $(r, u)$ , it follows that

$$(3.73) \quad \lim_{r \rightarrow \infty} \lim_{u \rightarrow \infty} \mathbb{P}(\hat{\tau}_y^{0, \mathbb{T}} \leq u, \hat{\xi}_t \subseteq B_0(r); t \in [0, \hat{\tau}_y^{0, \mathbb{T}}]) = 1.$$

As a consequence we get that for any  $\beta > 0$  there exist  $u > 0, r > 0$  large enough such that

$$(3.74) \quad \mathbb{P}(\hat{\tau}_y^{0, \mathbb{T}} \leq u, \hat{\xi}_t \subseteq B_0(r - 1); t \in [0, \hat{\tau}_y^{0, \mathbb{T}}]) \geq e^{-\beta},$$

but the above event is equivalent to that in (3.67) for  $x = 0$ .  $\square$

**PROOF OF LEMMA 3.2.** Without loss of generality, we can assume that  $x = 0$ , and that  $y$  is the parent of 0 in  $\mathbb{T}$ . Let  $\alpha > 0$  be arbitrarily large and  $\delta, a, b$  be given by Lemma 3.8. Setting  $a' := (\alpha a) \wedge 1, b' := \alpha b$ , Lemma 3.8 implies that for all  $s \geq 0$  large enough, the event

$$(3.75) \quad \hat{\mathcal{A}} := \{|\hat{\xi}_{\alpha s}^{0, \mathbb{T}} \cap B_0(b's)| \geq e^{a's}\}$$

satisfies

$$(3.76) \quad \mathbb{P}(\hat{\mathcal{A}}) \geq 1 - e^{-\alpha \delta s}.$$

For any vertex  $z \in \mathbb{T}$ , let  $\gamma_z$  denote the set of vertices on the unique path from  $z$  to  $y$ . For  $r > 0$  let  $\Gamma_z(r) := \bigcup_{w \in \gamma_z} B_w(r)$ . With

$$(3.77) \quad \beta := \frac{a'}{5b'},$$

where  $a'$  and  $b'$  are as above and with  $r, u$  given by Lemma 3.9, set

$$(3.78) \quad \hat{\mathcal{B}}^z := \{\hat{\tau}_y^{z, \Gamma_z^-(r)} \leq 2b'us\}.$$

By iterating Lemma 3.9, if  $z \in B_0(b's)$ , we have

$$(3.79) \quad \mathbb{P}(\hat{\mathcal{B}}^z) \geq \mathbb{P}(\hat{\tau}_y^{z, \Gamma_z^-(r)} \leq (b's + 1)u) \geq e^{-\beta(b's+1)} \geq e^{-2\beta b's}.$$

Now suppose that  $\hat{A}_t$  occurs, and pick  $z_0 \in \hat{\xi}_{\alpha s}^{0, \mathbb{T}} \cap B_0(b's)$ . Henceforth, we shall assume that there is some fixed order among all vertices of  $\mathbb{T}$  and that every time we arbitrarily pick a vertex from a subset of  $\mathbb{T}$ , we pick the minimal one with respect to this order. Notice that if the set from which we choose is random, the chosen vertex, for example,  $z_0$  above, is a random variable.

For what is coming, it will be useful to employ the following notation. For any  $t \geq 0$ , we denote  $\theta_t$  the “shifting forward of time by  $t$ ,” that is,  $\theta_t$  acts on the underlying sample space, such that

$$(3.80) \quad \begin{aligned} \mathcal{N}^\bullet(\cdot)_{u,v} \circ \theta_t &= \mathcal{N}^\bullet_{u,v}(\cdot + t), \\ \mathcal{N}^\circ(\cdot)_{u,v} \circ \theta_t &= \mathcal{N}^\circ_{u,v}(\cdot + t) \end{aligned}$$

for any neighboring vertices  $u \sim v$  of  $\mathbb{T}$ .

Then using (3.79), the Markov property and monotonicity we may write

$$(3.81) \quad \mathbb{P}(\hat{\mathcal{B}}^{z_0} \circ \theta_{\alpha s} | \hat{A}) \geq e^{-2\beta b's}.$$

On  $\hat{A} \cap (\hat{\mathcal{B}}^{z_0} \circ \theta_{\alpha s})$  we have  $\hat{\tau}_y^{0, \mathbb{T}} < \alpha s + 2b's u$ , and if this indeed happens, we stop. If not, we pick  $z_1 \in \hat{\xi}_{\alpha s}^{0, \mathbb{T}} \cap B_0(b's) \setminus \Gamma_{z_0}(r)$  (which must exist, as argued below). Notice that by “removing” (from consideration) all vertices in  $\Gamma_{z_0}(r)$ , conditioned on  $\hat{A}$  and the choice of  $z_1$ , the process  $(\hat{\xi}_t^{z_1, \mathbb{T}_{z_1}^-} \circ \theta_{\alpha s} : t \geq 0)$  is independent of  $\hat{\mathcal{B}}^{z_0} \circ \theta_{\alpha s}$ . Therefore, Lemma 3.1 guarantees that there exists a universal  $\varepsilon > 0$  such that

$$(3.82) \quad \mathbb{P}(\hat{\xi}_{\alpha s + 2b'us}^{z_1, \mathbb{T}_{z_1}^-} \ni z_1 | \hat{A}, (\hat{\mathcal{B}}^{z_0} \circ \theta_{\alpha s})^c) \geq \mathbb{P}(\hat{\xi}_{2b'us}^{z_1, \mathbb{T}_{z_1}^-} \ni z_1) \geq \varepsilon.$$

The event in the first term can be written as  $\hat{\mathcal{C}}^{z_1} \circ \theta_{\alpha s + 2b'us}$  once we set for  $z \in \mathbb{T}$ ,

$$(3.83) \quad \hat{\mathcal{C}}^z := \{\hat{\xi}_0^{z, \mathbb{T}_z^-} \ni z\}.$$

If this happens, then by monotonicity  $z_1 \in \hat{\xi}_{\alpha s + 2b'us}^{0, \mathbb{T}}$ , and we now check whether  $\hat{\mathcal{B}}^{z_1} \circ \theta_{\alpha s + 2b'us}$  occurs. Since we are only conditioning on events that depend on (measurable w.r.t.), the process up to time  $\alpha s + 2b'us$ , by the Markov property we still have by virtue of (3.79)

$$(3.84) \quad \mathbb{P}(\hat{\mathcal{B}}^{z_1} \circ \theta_{\alpha s + 2b'us} | \hat{A}, (\hat{\mathcal{B}}^{z_0} \circ \theta_{\alpha s})^c, \hat{\mathcal{C}}^{z_1} \circ \theta_{\alpha s + 2b'us}) \geq e^{-2\beta b's}.$$

On the intersection of all the events in (3.84), then we have  $\hat{\tau}_y^{0, \mathbb{T}} < \alpha s + 4b'us$ . If either the event in (3.82) or the event in (3.84) fail, we pick a new vertex  $z_2 \in \hat{\xi}_{\alpha s}^{0, \mathbb{T}} \cap B_0(b's) \setminus (\Gamma_{z_0}(r) \cup \Gamma_{z_1}(r))$ .

Proceeding in this fashion we obtain vertices  $z_0, z_1, \dots, z_n$ , for  $n$  to be defined later. Indeed, vertex  $z_k$  is chosen from

$$(3.85) \quad \hat{\xi}_{\alpha s}^{0, \mathbb{T}} \cap B_0(b's) \setminus \bigcup_{0 \leq l < k} \Gamma_{z_l}(r),$$

and at time  $\alpha s + 2kb'us$  we have [as in (3.82), (3.84)]

$$(3.86) \quad \mathbb{P}\left(\left(\hat{\mathcal{C}}^{z_k}, \hat{\mathcal{B}}^{z_k}\right) \circ \theta_{\alpha s + 2kb'us} \Big| \hat{\mathcal{A}}, \bigcap_{l < k} \left(\left(\hat{\mathcal{C}}^{z_l}, \hat{\mathcal{B}}^{z_l}\right)^c \circ \theta_{\alpha s + 2lb'us}\right)\right) \geq \varepsilon e^{-2\beta b's},$$

and if the above event occurs, then  $\hat{\tau}_y^{0, \mathbb{T}} < \alpha s + 2(k + 1)b'us$ , and we stop. To make sure that we do not “run out” of vertices in (3.85), observe that on  $\hat{\mathcal{A}}$  when we pick  $z_k$ ,

$$(3.87) \quad \left| \hat{\xi}_{\alpha s}^{0, \mathbb{T}} \cap B_0(b's) \setminus \bigcup_{0 \leq l < k} \Gamma_{z_l}(r) \right| \geq e^{a's} - kd^r(b's + 1).$$

Thus, choosing  $n := e^{a's/2}$  guarantees that the above is positive for all  $k \leq n$ .

Combining (3.76), (3.77) and (3.86), we obtain for all  $s$  large enough,

$$(3.88) \quad \begin{aligned} & \mathbb{P}(\hat{\tau}_y^{0, \mathbb{T}} < e^{a's/3}) \\ & \geq \mathbb{P}(\hat{\tau}_y^{0, \mathbb{T}} < \alpha s + 2(n + 1)b'us) \\ & \geq \mathbb{P}\left(\bigcup_{0 \leq l \leq n} \left(\left(\hat{\mathcal{C}}^{z_l}, \hat{\mathcal{B}}^{z_l}\right) \circ \theta_{\alpha s + 2lb'us}\right) \Big| \hat{\mathcal{A}}\right) \mathbb{P}(\hat{\mathcal{A}}) \\ & \geq 1 - (1 - \varepsilon e^{-2\beta b's})^{n+1} - e^{-\alpha\delta s} \geq 1 - e^{-\varepsilon e^{(a'/2 - 2\beta b')s}} - e^{-\alpha\delta s} \\ & \geq 1 - C e^{-\alpha\delta s} = 1 - C(e^{a's/3})^{-3\alpha\delta/a'}. \end{aligned}$$

Substituting  $t := e^{a's/3}$  and  $q := 3\alpha\delta/a'$  we rewrite the above as

$$(3.89) \quad \mathbb{P}(\hat{\tau}_y^{0, \mathbb{T}} > t) \leq C t^{-q}$$

for all  $t$  large enough. But since  $\alpha$  was arbitrary,  $\delta$  is fixed and independent of it and  $a' \leq 1$  we can make  $q$  as large as we want. This shows (3.2).  $\square$

**4. Proof of Proposition 1.1.** Beginning with part (1), the upper bound easily follows from a comparison with the contact process. Indeed, consider the contact process  $(\zeta_t^{\zeta_0} : t \geq 0)$  on  $\mathbb{T}^d$  starting from some  $\zeta_0 \in \mathcal{X}$ , where each site heals independently at rate  $d$  and infected sites infect their neighbors at rate  $\lambda$ . Employing a standard coupling (e.g., using the graphical representations of both models), we can have

$$(4.1) \quad \zeta_t^{\zeta_0}(x) \leq \xi_t^{\zeta_0}(x), \quad t \geq 0, x \in \mathbb{T}.$$

Then local survival of  $\zeta^{\zeta_0}$  implies local survival of  $\xi^{\zeta_0}$ . By the first inequality in part (vi) of Theorem 2.2 in [16], this occurs with positive probability for  $\zeta^{\zeta_0}$  as soon as  $\lambda > \min(2d, \frac{4d}{(\sqrt{d-1}-4)\vee 0})$ . Observe that in the notation there  $n = d - 1$  and that all threshold values should be multiplied by  $d$  as the healing rate there is  $d$  times smaller.



Before proving the lower bound in part (1), let us first address part (2) of the proposition, as its proof is similar and at the same time slightly simpler. For  $\alpha \geq 0$  consider the function  $f : \mathcal{X} \rightarrow \mathbb{R}_+$  given by

$$(4.2) \quad f(\xi) = \sum_{x \in \xi} \alpha^{\rho(0,x)}.$$

If  $\alpha < (d - 1)^{-1}$  this function is uniformly bounded. At the same time, we claim that if  $\lambda^{-1} < \alpha < \lambda$ , then  $M_t^{\xi_0} := f(\xi_t^{\xi_0})$  is a sub-martingale for all finite nonempty  $\xi_0 \in \mathcal{X}$  (this is still true even if  $\xi_0$  is infinite, but our argument below needs to be modified in this case). Indeed, let  $t \geq 0$  and consider the next transition of  $\xi_t^{\xi_0}$  after  $t$ . Such transition must be due to a pair of neighboring vertices  $u \sim v \in \mathbb{T}$  such that  $\xi_t^{\xi_0}(u) = +$  while  $\xi_t^{\xi_0}(v) = -$ , for which there is either an ‘‘arrival’’ of an infection event, whereby  $u$  infects  $v$ , or an ‘‘arrival’’ of a healing event, whereby  $v$  heals  $u$  and in any case this arrival is the first to occur after time  $t$ .

Conditioned on  $u, v$  being the vertices for which one of these arrivals occurred first, the probability that it is an infection event is  $\frac{\lambda}{\lambda+1}$ , while the probability that it is a healing event is  $\frac{1}{\lambda+1}$ . There are two cases to consider: either  $\rho(0, u) = \rho(0, v) + 1$  or  $\rho(0, u) = \rho(0, v) - 1$ . The expected change in  $M_t^{\xi_0}$  due to this transition is then, respectively,

$$(4.3) \quad \alpha^{\rho(0,v)} \left( \frac{\lambda}{1+\lambda} - \frac{\alpha}{1+\lambda} \right) \quad \text{and} \quad \alpha^{\rho(0,u)} \left( \frac{\lambda\alpha}{1+\lambda} - \frac{1}{1+\lambda} \right).$$

Both are nonnegative when  $\lambda^{-1} < \alpha < \lambda$ , and hence  $M_t^{\xi_0}$  is a sub-martingale.

If  $\lambda > d - 1$ , we may find  $\lambda^{-1} < \alpha < (d - 1)^{-1}$  for which  $(M_t^{\xi_0} : t \geq 0)$  is a bounded sub-martingale and therefore must converge to some finite value with probability 1. Consequently, for any vertex  $v$  there is some time  $s_v$  after which it is either always infected or never infected; otherwise there would have been an unbounded sequence of times at which  $M_t^{\xi_0}$  changed by  $\alpha^{\rho(0,v)}$ , a contradiction to the convergence. Furthermore, two neighboring vertices cannot have different limits, and hence almost-surely  $\xi_t^{\xi_0}$  converges either to  $\mathbb{T}$  or  $\emptyset$ . The latter cannot happen once  $\Omega_t^{\xi_0}$  occurs, but  $\Omega_t^{\xi_0}$  occurs with positive probability when  $\lambda > \lambda_l(\mathbb{T})$ . Thus, when  $\lambda > (d - 1) \vee \lambda_l(\mathbb{T})$ , then with positive probability  $\xi_0^{\xi_0}$  converges to  $\mathbb{T}$  and the process survives completely.

We now come back to the proof of the lower bound in part (1). Let  $u \in \mathbb{T}$ , and set  $w \in \mathbb{T}$  to be the first common ancestor of  $u$  and 0. Note that it might be 0 or  $u$ . Now define the ‘‘height’’ of  $u$  relative to 0 as  $h(u) := \rho(w, u) - \rho(w, 0)$ . Next, for  $\alpha > 0$  let  $f(u) := \alpha^{h(u)}$  and for a subset of vertices  $U \subseteq \mathbb{T}$  we set  $f(U) := \sum_{u \in U} f(u)$ . Notice that  $f(u)$  may be  $\infty$ . Finally, define the process  $M_t := f(\xi_t^0)$ .

We would like to claim that as soon as  $\alpha$  and  $\lambda$  satisfy

$$(4.4) \quad \lambda(d - 1)\alpha^2 - d\alpha + \lambda \leq 0,$$

then  $(M_t : t \geq 0)$  is a nonnegative super-martingale and therefore must have an almost-sure finite limit  $M_\infty$ . As before this will imply that  $\xi_t^0$  converges to either  $\emptyset$  or  $\mathbb{T}$  with probability 1. Since  $\mathbb{E}M_\infty \leq \mathbb{E}M_0 < \infty$ , it must be that  $M_\infty$  is finite almost-surely, which is only possible if the convergence of  $\xi_t^0$  is to  $\emptyset$ . Thus, the probability of local survival is 0 and hence  $\lambda$  is a lower bound for  $\lambda_l(\mathbb{T})$ . It is not difficult to see that (4.4) has a solution with  $\alpha > 0$  if and only if  $\lambda \leq \frac{d}{2\sqrt{d-1}}$ . This gives the lower bound in part (1).

To see that once (4.4) holds,  $M_t$  is a super-martingale, we argue as before. Given  $\xi_t^0$  for some  $t \geq 0$ , the expected change in  $M_t$  at the next transition of  $\xi_t^0$  is

$$(4.5) \quad \frac{1}{(\lambda + 1)|\partial_{\mathbb{T}}\xi_t^0|} \sum_{\substack{u \sim v \\ v \in \xi_t, u \notin \xi_t}} \lambda f(u) - f(v).$$

It is therefore enough to argue that for any finite  $U \subseteq \mathbb{T}$ ,

$$(4.6) \quad \sum_{\substack{u \sim v \\ v \in U, u \notin U}} \lambda f(u) - f(v) \leq 0,$$

and by linearity, we can also assume that  $U$  is connected.

The proof of (4.6) will follow by induction on the size of  $U$ . It is easy to check that (4.4) is necessary and sufficient for (4.6) to hold for  $U = \{x\}$ . Now suppose that it holds for all  $U$  with  $|U| < n$ , and let  $U$  be a set with  $n$  vertices. Choose some leaf  $w \in U$  and set  $U' := U \setminus \{w\}$ . From the induction hypothesis we know

$$(4.7) \quad \sum_{\substack{u \sim v \\ v \in U', u \notin U'}} \lambda f(u) - f(v) \leq 0 \quad \text{and} \quad \sum_{u \sim w} \lambda f(u) - f(w) \leq 0.$$

Adding the two inequalities and letting  $w'$  denote the parent of  $w$ , we obtain

$$(4.8) \quad (\lambda - 1)(f(w) + f(w')) + \sum_{\substack{u \sim v \\ v \in U, u \notin U}} \lambda f(u) - f(v) \leq 0,$$

but since  $\lambda \geq 1$ , this implies that (4.6) holds as desired.

**5. Proof of Theorem 1.6.** Starting with part (1), for  $u \sim v \in \mathbb{T}$  and  $t \geq 0$ , denote by  $e_t^+(u, v)$  the number of times  $v$  becomes infected because of  $u$ , up to time  $t$ , and let  $E_t^+(u, v) := \mathbb{E}e_t^+(u, v)$ . If  $\mathbb{P}(\xi_0 \in \cdot) \in \mathcal{I}$ , the latter does not depend on the choice of  $u, v$  and we shall therefore just write  $E_t^+$ . Similarly, define  $e_t^-(u, v)$  as the number of times  $v$  becomes healthy because of  $u$ , up to time  $t$  and let  $E_t^- = E_t^-(u, v) := \mathbb{E}e_t^-(u, v)$ . Since for all  $v \in \mathbb{T}$ ,

$$(5.1) \quad \sum_{u:u \sim v} e_t^+(u, v) - \sum_{u:u \sim v} e_t^-(u, v) \in \{-1, 0, +1\},$$

by taking expectations we see that  $-\frac{1}{d} \leq E_t^+ - E_t^- \leq \frac{1}{d}$ . On the other hand,  $E_t^+(u, v) = \lambda E_t^-(v, u)$ , as every time  $\xi_t^{\xi_0}(u) = +$  but  $\xi_t^{\xi_0}(v) = -$  and an infection or healing event occurs along the edge  $\{u, v\}$ , the probability of an infection of  $v$  by  $u$  is larger by a factor of  $\lambda$  than the probability of a healing of  $u$  by  $v$  (formally, this is just a simple martingale argument). Thus  $(1 - \lambda^{-1})E_t^+ \leq 1/d$ , and so  $E_t^+$  is bounded by a constant that does not depend on  $t$ .

Consequently,  $e_\infty^+(u, v) := \lim_{t \rightarrow \infty} e_t^+(u, v)$ , which exists by monotonicity, has a bounded expectation and therefore must itself be bounded almost-surely. The same holds for  $e_\infty^-(u, v)$ , and we see that the number of sign flips at any vertex  $v$  must be finite almost surely, as desired.

Turning to part (2). Let  $x, y$  be two neighboring vertices in  $\mathbb{T}$  and  $t \geq 0$ . Define

$$(5.2) \quad \rho_t := \mathbb{P}(\xi_t^{\xi_0}(x) = +), \quad \delta_t := \mathbb{P}(\xi_t^{\xi_0}(x) = +, \xi_t^{\xi_0}(y) = -).$$

Clearly, the above does not depend on the choice of  $x, y$  since the initial distribution is automorphism-invariant. We claim that for all  $t \geq 0$ ,

$$(5.3) \quad \frac{d\rho_t}{dt} = d(\lambda - 1)\delta_t.$$

To see this, fix some  $t, h \geq 0$ , and for  $u \sim v \in \mathbb{T}$ , let

$$(5.4) \quad \begin{aligned} f(u, v) &:= [e_{t+h}^+(u, v) - e_t^+(u, v)] - [e_{t+h}^-(u, v) - e_t^-(u, v)], \\ F(u, v) &:= \mathbb{E}f(u, v). \end{aligned}$$

Since  $\mathbb{T}$  is transitive and unimodular, by the mass-transport principle (see, e.g., [13]) for all  $w \in \mathbb{T}$ ,

$$(5.5) \quad \sum_{u:u \sim w} F(u, w) = \sum_{v:v \sim w} F(w, v).$$

The left-hand side above is equal to  $\mathbb{E}\frac{1}{2}(\xi_{t+h}(w) - \xi_t(w)) = \rho_{t+h} - \rho_t$ . The right-hand side is

$$(5.6) \quad \begin{aligned} &\sum_{v:v \sim w} \mathbb{E}[e_{t+h}^+(w, v) - e_t^+(w, v)] - \sum_{v:v \sim w} \mathbb{E}[e_{t+h}^-(w, v) - e_t^-(w, v)] \\ &= d\delta_t h(\lambda - 1) + O(h^2). \end{aligned}$$

Equating the two sides, dividing by  $h$  and taking  $h \rightarrow 0$ , we obtain (5.3).

Now from the previous part we know that  $\xi_\infty^{\xi_0} := \lim_{t \rightarrow \infty} \xi_t^{\xi_0}$  exists almost-surely, and its distribution is supported on  $\{\emptyset, \mathbb{T}\}$ . Since ergodicity is preserved in the (strong) limit, it follows that  $\xi_\infty^{\xi_0}$  is either  $\emptyset$  a.s. or  $\mathbb{T}$  a.s. If  $\mathbb{P}(\xi_0 \in \cdot) \neq \delta_\emptyset$ , then  $\rho_0 > 0$ , and since the right-hand side of (5.3) is nonnegative it must be that also  $\rho_\infty := \mathbb{P}(\xi_\infty^{\xi_0}(x) = +) > 0$ . This leaves only the option  $\xi_\infty^{\xi_0} = \mathbb{T}$  almost surely, which implies  $\mathbb{P}(\Omega_c^{\xi_0}) = 1$ .

It remains to show that infinite infected components must form in finite time. We first show that if at time  $t \geq 0$  all components are finite almost-surely, then

$$(5.7) \quad \delta_t \geq C\rho_t$$

for some  $C > 0$  (which is independent of  $t$ ). This follows again from the mass-transport principle. Denote by  $\partial_{\mathbb{T}}^- U$  the internal vertex boundary of  $U$ , that is, the set of vertices in  $U$  with neighbors that are not in  $U$ . Let every infected vertex  $v$  under  $\xi_t^{\xi_0}$ , send one unit of mass to the interior vertex boundary of its component, divided equally among these boundary vertices. The expected amount of mass a vertex sends is  $\rho_t$ . Thus, by the mass-transport principle, the expected amount of mass a vertex receives is  $\rho_t$  as well.

On the other hand, a vertex receives a positive mass if and only if it is infected and lies in  $\partial_{\mathbb{T}}^- U$  where  $U$  is the infected component to which it belongs. In this case it receives a total mass of  $|U|/|\partial_{\mathbb{T}}^- U|$ . Since  $\mathbb{T}$  is nonamenable, this ratio is bounded above by a constant  $C' < \infty$ . If we let  $\beta_t$  be the probability a given vertex is in the interior boundary of an infected component of  $\xi_t$ , then these considerations lead to the inequality  $\beta_t \geq C'^{-1}\rho_t$ . Since  $\delta_t \geq \beta_t/d$ , we have shown (5.7).

Now suppose that  $\mathbb{P}(\xi_0 \in \cdot)$  is as in part (2) of the theorem and that for all time  $t \geq 0$ , the probability of the existence of an infinite infected component in  $\xi_t^{\xi_0}$  is 0. Then (5.3) and (5.7) together imply that  $\rho_t \geq \rho_0 e^{C''t}$  for some  $C'' > 0$ . Since  $\rho_0 > 0$ , this would imply  $\rho_t > 1$  for some  $t > 0$ , which is a contradiction. Therefore, there is some  $t \geq 0$ , for which  $\xi_t^{\xi_0}$  has infinite components with positive probability. Since the latter event is automorphism-invariant and since ergodicity is carried over to  $\xi_t^{\xi_0}$  for every  $t > 0$ , the latter probability must be 1.

**Acknowledgments.** The authors would like to thank the anonymous referees for their thorough review of the manuscript, and Tom Liggett, Noam Berger and Elchanan Mossel for many useful discussions.

## REFERENCES

- [1] BIGGINS, J. D. (1995). The growth and spread of the general branching random walk. *Ann. Appl. Probab.* **5** 1008–1024. [MR1384364](#)
- [2] BRAMSON, M. and GRIFFEATH, D. (1980). On the Williams–Bjerknes tumour growth model. II. *Math. Proc. Cambridge Philos. Soc.* **88** 339–357. [MR0578279](#)
- [3] BRAMSON, M. and GRIFFEATH, D. (1981). On the Williams–Bjerknes tumour growth model. I. *Ann. Probab.* **9** 173–185. [MR0606980](#)
- [4] BURKHOLDER, D. L., DALEY, D., KESTEN, H., NEY, P., SPITZER, F., HAMMERSLEY, J. M. and KINGMAN, J. F. C. (1973). Discussion on “Subadditive ergodic theory” by J. F. C. Kingman. *Ann. Probab.* **1** 900–909.
- [5] DURRETT, R. and SCHINAZI, R. (1995). Intermediate phase for the contact process on a tree. *Ann. Probab.* **23** 668–673. [MR1334165](#)
- [6] GRIFFEATH, D. (1979). *Additive and Cancellative Interacting Particle Systems. Lecture Notes in Math.* **724**. Springer, Berlin. [MR0538077](#)

- [7] HARRIS, T. E. (1976). On a class of set-valued Markov processes. *Ann. Probab.* **4** 175–194. [MR0400468](#)
- [8] HARRIS, T. E. (1978). Additive set-valued Markov processes and graphical methods. *Ann. Probab.* **6** 355–378. [MR0488377](#)
- [9] LALLEY, S. P. and SELKE, T. (1998). Limit set of a weakly supercritical contact process on a homogeneous tree. *Ann. Probab.* **26** 644–657. [MR1626499](#)
- [10] LIGGETT, T. M. (1996). Multiple transition points for the contact process on the binary tree. *Ann. Probab.* **24** 1675–1710. [MR1415225](#)
- [11] LIGGETT, T. M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Grundlehren der Mathematischen Wissenschaften* **324**. Springer, Berlin. [MR1717346](#)
- [12] LIGGETT, T. M. (2004). *Interacting Particle Systems*. Springer, Berlin.
- [13] LYONS, R. and PERES, Y. (2014). *Probability on Trees and Networks*. Cambridge Univ. Press. To appear.
- [14] MADRAS, N. and SCHINAZI, R. (1992). Branching random walks on trees. *Stochastic Process. Appl.* **42** 255–267. [MR1176500](#)
- [15] MORGAN, B. J. T. (1979). Four approaches to solving the linear birth-and-death (and similar) processes. *Internat. J. Math. Ed. Sci. Tech.* **10** 51–64. [MR0529525](#)
- [16] PEMANTLE, R. (1992). The contact process on trees. *Ann. Probab.* **20** 2089–2116. [MR1188054](#)
- [17] RICHARDSON, D. (1973). Random growth in a tessellation. *Math. Proc. Cambridge Philos. Soc.* **74** 515–528. [MR0329079](#)
- [18] 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. [MR1640345](#)
- [19] SCHWARTZ, D. L. (1977). Applications of duality to a class of Markov processes. *Ann. Probab.* **5** 522–532. [MR0448631](#)
- [20] STACEY, A. M. (1996). The existence of an intermediate phase for the contact process on trees. *Ann. Probab.* **24** 1711–1726. [MR1415226](#)
- [21] SUDBURY, A. and LLOYD, P. (1997). Quantum operators in classical probability theory. IV. Quasi-duality and thinnings of interacting particle systems. *Ann. Probab.* **25** 96–114. [MR1428501](#)
- [22] WILLIAMS, T. and BJERKNES, R. (1971). *A Stochastic Model for the Spread of an Abnormal Clone Through the Basal Layer of the Epithelium*. Symp. Tobacco Research Council, London.
- [23] ZHANG, Y. (1996). The complete convergence theorem of the contact process on trees. *Ann. Probab.* **24** 1408–1443. [MR1411500](#)

O. LOUIDOR  
 INDUSTRIAL ENGINEERING AND MANAGEMENT  
 TECHNION—ISRAEL'S INSTITUTE OF TECHNOLOGY  
 HAIFA 32000  
 ISRAEL  
 E-MAIL: [oren.louidor@gmail.com](mailto:oren.louidor@gmail.com)

R. TESSLER  
 EINSTEIN INSTITUTE OF MATHEMATICS  
 HEBREW UNIVERSITY OF JERUSALEM  
 JERUSALEM, 91904  
 ISRAEL  
 E-MAIL: [ran.tessler@mail.huji.ac.il](mailto:ran.tessler@mail.huji.ac.il)

A. VANDENBERG-RODES  
 MATHEMATICS DEPARTMENT  
 UNIVERSITY OF CALIFORNIA, IRVINE  
 IRVINE, CALIFORNIA 92697-3875  
 USA  
 E-MAIL: [vandenbe@math.uci.edu](mailto:vandenbe@math.uci.edu)