First passage percolation in hostile environment is not monotone[*]

Elisabetta Candellero[†]   Alexandre Stauffer[‡]

Abstract

We study a natural growth process with competition, modeled by two first passage percolation processes, FPP_1 and FPP_λ, spreading on a graph. FPP_1 starts at the origin and spreads at rate 1, whereas FPP_λ starts from a random set of *inactive seeds* distributed as Bernoulli percolation of parameter $\mu \in (0, 1)$. A seed of FPP_λ gets activated when one of the two processes attempts to occupy its location, and from this moment onwards spreads at some fixed rate $\lambda > 0$. In previous works [17, 3, 7] it has been shown that when both μ or λ are small enough, then FPP_1 *survives* (i.e., it occupies an infinite set of vertices) with positive probability. It might seem intuitive that decreasing μ or λ is beneficial to FPP_1. However, we prove that, in general, this is indeed false by constructing a graph for which the probability that FPP_1 survives is not a monotone function of μ or λ, implying the existence of multiple phase transitions. This behavior contrasts with other natural growth processes such as the 2-type Richardson model.

1 Introduction

We consider a natural growth process known as *First passage percolation in a hostile environment* (FPPHE), which consists of two first passage percolation processes that spread at different rates, which we shall call FPP_1 and FPP_λ. Such processes are competing against each other to infect vertices of the underlying graph, according to

[†]Università Roma Tre, Dipartimento di Matematica e Fisica, Largo S. Murialdo 1, 00146, Rome, Italy.
   E-mail: elisabetta.candellero@uniroma3.it
[‡]University of Bath, Dept of Mathematical Sciences, BA2 7AY Bath, UK & Università Roma Tre, Dipartimento di Matematica e Fisica, Largo S. Murialdo 1, 00146, Rome, Italy. E-mail: a.stauffer@bath.ac.uk

the following dynamics. Consider a connected infinite graph of bounded degree, choose a reference point on it (call it *the origin*), and place a FPP_1 particle there. Then fix a parameter $\mu \in (0, 1)$, and at each vertex other than the origin independently place a FPP_λ particle with probability μ and nothing with probability $1 - \mu$. We shall call such initial FPP_λ particles *seeds*.

At time zero, all seeds of FPP_λ are *inactive*, and FPP_1 starts spreading from the origin as a first passage percolation process of rate 1. Whenever there is an attempt to occupy a site that is currently occupied by an inactive seed, then the seed gets activated and starts to spread as a first passage percolation of rate λ. Whenever a site gets occupied by one of the two processes, it will remain so forever. Thus the two processes compete with one another for space. We see FPP_1 and FPP_λ as infection processes, so we will say that a vertex is *infected* by FPP_1 or FPP_λ to mean that the vertex is occupied by the given process.

FPPHE is a natural model for the growth of two processes with competition and, as we will see below, it has many interesting aspects and phase transitions. Nonetheless, it was introduced by Sidoravicius and Stauffer [17] as an auxiliary model for analyzing a notoriously challenging model called Multiparticle Diffusion Limited Aggregation (MDLA). The use of FPPHE as a tool of analysis has been further explored in Dauvergne and Sly [4] to analyze a non-equilibrium process for the spread of an infection. The way in which FPPHE is used to analyze such processes with non-equilibrium dynamics is by making the growth of FPP_1 represent the propagation of the front as it spreads through "typically good regions", while the activation of a FPP_λ seed represents places where the front behaves "atypically bad". The influence of discovering such bad events is represented by the spread of FPP_λ from the activated seed. In this context, one is interested to understand whether FPP_1 grows indefinitely and leaves behind any cluster of activated seeds of FPP_λ.

There are several challenging questions about FPPHE. One example is the *survival* of FPP_1 (resp. FPP_λ), meaning that FPP_1 (resp. FPP_λ) infects an infinite connected region of the graph. Note that it is natural to require connectivity of the infected region in the definition of survival since FPP_λ already starts from an infinite set of seeds. Another fundamental question is that of *coexistence* of FPP_1 and FPP_λ, which corresponds to both processes surviving simultaneously.

The first paper about FPPHE [17] established that on $\mathbb{Z}^d$ if $\lambda < 1$, then for all μ small enough FPP_1 survives with positive probability and produces a *bulky* shape (cf. [17, Theorems 1.3]). Simple arguments show that, on $\mathbb{Z}^d$, FPP_1 dies out almost surely if $\lambda \geq 1$ or μ large enough. The justification of the first case can be found in Finn's PhD thesis [6, Chapter 4]. The latter case follows from an percolation argument. Very recently, [7] showed the opposite direction that FPP_1 survives with positive probability for any μ for which non-seeds percolate on $\mathbb{Z}^d$ provided λ is small enough. This seems analogous to the work of [17], but actually FPPHE behaves rather differently in this setting. For example, the result of [7] implies a coexistence regime on $\mathbb{Z}^d$ for $d \geq 3$, and even a regime where both types occupy a set of positive density. Coexistence is a challenging phenomenon to establish. Even in more classical models, such as the two-type Richardson model, the regime of coexistence has not been fully determined (cf. [11], and [9] for a more general analysis of the model, not restricted to exponential passage times). For FPPHE on $\mathbb{Z}^d$, it remains entirely open whether coexistence occurs when $d = 2$. Before that, coexistence for FPPHE was established on hyperbolic non-amenable graphs by [3], which also showed that FPPHE behaves quite different in such graphs when compared to $\mathbb{Z}^d$. In fact, [3] shows that FPP_λ survives almost surely, for any fixed pair of values for the parameters $0 < \mu \leq 1$ and $\lambda > 0$.

The core of this work will focus on the behavior of FPPHE on a type of graph that is

hyperbolic, non-amenable, quasi-transitive and is endowed with a "tree-like structure". However, our result does *not* apply to trees, because of the lack of loops in their structure. Although this construction naturally leads to a specific kind of graph, we are confident that the main ideas developed in this paper will be applicable to other contexts, at least to other growth models. A work that follows a similar flavor is [5], in which the authors investigate competing Richardson processes on graphs. They construct specific graphs where the set of values for λ that allow coexistence has very counter-intuitive properties. However, a crucial difference is that the Richardson model is a monotone model, in the sense that increasing the rate of one type can only increase the probability that said type survives. The focus of [5] is on the event of coexistence, while we are interested in studying monotonicity properties of the FPPHE model on its own.

It is natural to believe that by increasing the density μ or the rate λ one would favor FPP_λ. Such monotonicity property would be quite useful as it plays a crucial role in the analysis of several other related models, such as the contact process and the two-type Richardson model [14, 11, 10, 8, 12]. However, it is possible to show that there are choices for the locations of the seeds and for the passage times such that the standard coupling fails to be monotone. In particular, on graphs whose structure present a lot of loops of different lengths, such configurations can be engineered so that adding a seed is beneficial to FPP_1. The reason is that when the seed is activated, the process gets slowed down locally, and this can delay the activation of seeds that are further away in that direction. This could be beneficial for FPP_1 as it grows along nearby directions.

The lack of knowledge of such monotonicity gives rise to serious challenges; the papers [17, 3, 7] develop quite involved multi-scale arguments to analyze FPPHE without using monotonicity. The aim of the present article is to show that FPPHE is indeed *not* monotone in general, meaning that there are graphs for which the probability of survival of FPP_1 is not a monotone function of the parameters. The graph that we obtain is rather different from a d-dimensional lattice; roughly speaking, it looks like a homogeneous tree of large degree, where all edges are replaced by identical copies of a finite graph that we call *tile*. (A formal definition is given in Section 1.1).

We believe that on regular graphs such as $\mathbb{Z}^d$ ($d \geq 1$) the process should be monotone. However, the challenges in this type of analysis stem from several facts. For example, geodesics on $\mathbb{Z}^d$ can go very far apart from each other, making it more difficult to control the spread of FPPHE. Moreover, there are arbitrarily long loops, allowing FPP_1 (resp. FPP_λ) to "surround" arbitrarily large clusters of FPP_λ (resp. FPP_1) within a short time. At the moment a general result still eludes us, but we hope that this work will trigger a more general analysis of this question.

For any event E and considering λ fixed, let $\mathbb{P}_\mu(E)$ denote the probability of E, given that the initial density of seeds of FPP_λ equals μ.

Theorem 1.1. *There is a connected, infinite graph of bounded degree and values μ_1, μ_2 with $0 < \mu_1 < \mu_2 < 1$, such that, whenever the rate λ is small enough,*

$$\mathbb{P}_{\mu_1}(\text{FPP}_1 \text{ survives}) = 0 \quad and \quad \mathbb{P}_{\mu_2}(\text{FPP}_1 \text{ survives}) > 0.$$

Remark 1.2. In our proof of Theorem 1.1, we will construct a graph (which we denote $\mathbb{G}^\infty$) with the above properties. We note that $\mathbb{G}^\infty$ is even a *quasi transitive* graph. Recall that a graph is quasi transitive if its vertex set can be partitioned into a finite number of sets $V_1, V_2, \ldots, V_\ell$ such that for any k and any two vertices $u, v \in V_k$, there exists an automorphism of the graph that maps u onto v. Theorem 1.1 then implies at least three phase transitions as λ is fixed and μ varies. It is possible to get more involved constructions that lead to a larger number of phase transitions, and a similar non-monotonicity result can be obtained for μ fixed and λ varying (see Theorem 2.4 and Remark 2.5).

In our analysis of whether FPP_1 survives we will use the following classical fact regarding percolation on regular trees (cf. [15, Chapter 5] for more details).

Theorem 1.3. *[15, Theorem 5.15]. Let T denote a homogeneous tree of degree $b \geq 3$. Then,*

$$p_c(T) = \frac{1}{b-1},$$

where $p_c(T)$ denotes the percolation threshold for the appearance of an infinite cluster for Bernoulli percolation on T.

Note that in the case of trees the threshold for bond and for site percolation is the same.

1.1 Idea of the proof

The proof proceeds through various steps, which we outline here.

Construction of the graph The graph $\mathbb{G}^\infty$ should be thought of as a sort of infinite tree where each vertex has degree $\Phi + 1$, where Φ is a large integer. We replace the edges of the tree by identical copies of a finite graph that we call "*tile*". The structure of a tile plays a key role in determining which process gets (with reasonable probability) to advance more quickly. Below we give an overall view on the construction of the tiles, leaving the details to Section 2.1.

The origin of the tile is a vertex from where two edges depart, we shall say that they begin the *upper part* and the *lower part* of the tile (refer to Figure 2). Following the lower part, we find that the endpoint of the first edge is connected with a finite binary tree of some (large) height H. The last generation of this tree is then identified into one vertex, which represents the start of a path of length R (for some large integer R). Following the upper part, we see that the endpoint of the edge departing from the origin is connected with a finite D-ary tree (for some large integer $D > 2$) of some height L. Again, the last generation of the D-ary tree is identified into one vertex, to which one edge is attached. Finally, the two loose ends that we are left with (the end of the path in the lower part and the endpoint of the edge in the upper part) are identified into one vertex.

Analysis of FPPHE for $\mu = \mu_1$ and $\mu = \mu_2$ The next step consists in investigating the typical behavior of FPPHE on a tile, hence we split the study of the process into two parts.

In Section 3.1 we set the initial density of seeds to be some properly chosen value $\mu = \mu_1 \in (0, 1/2)$, and we show that the first process to infect the opposite endpoint of the tile is likely to be FPP_λ. The idea is that for small values of μ, the seeds do not block the advance of FPP_1 through the two trees (the D-ary tree in the upper part and the binary tree in the lower part). Therefore it is likely that FPP_1 is the first process to reach the end of both trees. By letting H much smaller than L, we obtain that FPP_1 reaches the end of the binary tree first. At this point though, in the lower part it is likely that FPP_1 will encounter a seed along the path of length R. Hence we will have that the endpoint of the tile can be infected either by FPP_1 (which is making its way along the upper part of the tile) or by FPP_λ (started by a seed activated in the lower part). The construction of the tile, in particular by properly choosing the height of the D-ary tree and the binary tree, makes it more likely that FPP_λ (through the lower part) arrives first.

We note that implementing the idea above requires a delicate analysis of FPPHE on a finite tree. The main challenge is that even if FPP_1 spreads quickly along a tree it may

so happen that, near the last levels of the tree, seeds are located on the fastest route of FPP_1. Therefore, it is not enough to look at the geodesics of first passage percolation on trees: we also need to control paths that are "near optimal" for FPP_1. We do this via a comparison with branching random walks, for which strong concentration results are available. This is the most technical part of the paper. It is established in Proposition 2.1, whose proof is given in Section 5.

The second part is carried out in Section 3.2, where we set the initial density to some properly chosen value $\mu = \mu_2 \in (1/2, 1 - 1/D)$. This will imply that while in the upper part FPP_1 continues to be the process that reaches the last generation of the D-ary tree, in the lower part it will be likely that FPP_1 will be blocked by the seeds in the binary tree (as a straightforward consequence of Theorem 1.3). This delays the progress of the process through the lower part, making it likely that the first process to reach the opposite endpoint of the tile is FPP_1 (through the upper part).

Conclusion of the proof   To conclude, recall that Φ was a large integer appearing in the construction of $\mathbb{G}^\infty$ as the "degree" of each vertex. Then we show that the event

$$\mathcal{E} := \{\text{the first process to infect the endpoint of the tile is } \mathrm{FPP}_1\}$$

is such that $\mathbb{P}_{\mu_1}(\mathcal{E}) < \frac{1}{\Phi}$ whereas $\mathbb{P}_{\mu_2}(\mathcal{E}) > \frac{1}{\Phi}$. Then, by viewing each tile where $\mathcal{E}$ holds as an open edge of a suitable Bernoulli percolation process on a Φ-ary tree and using Theorem 1.3, we conclude the proof.

Organization of the paper   The paper is organized as follows. Section 2 deals with the construction of $\mathbb{G}^\infty$ with vertex set $V(\mathbb{G}^\infty)$ and edge set $E(\mathbb{G}^\infty)$ and states a result (Proposition 2.3) which is the main step of the proof, giving the typical behavior of FPP_1 on a tile. Using Proposition 2.3, we conclude Section 2 with the proof of Theorem 1.1. In Section 3 we prove auxiliary lemmas, and in Section 4 we prove Proposition 2.3 using these lemmas and Proposition 2.1. Section 5 is devoted to the proof of Proposition 2.1.

1.2 Passage times

A path π of length n is a sequence of n consecutive edges $\{e_{\{0,1\}}, e_{\{1,2\}}, \ldots, e_{\{n-1,n\}}\}$. Fix a value $\gamma > 0$. For every path $\pi = \{e_{\{0,1\}}, \ldots, e_{\{n-1,n\}}\}$ define its *passage time* by

$$T^{(\gamma)}(\pi) := \sum_{j=0}^{n-1} t_{\{j,j+1\}}^{(\gamma)},$$

where $t_{\{j,j+1\}}^{(\gamma)}$ is a random variable associated to edge $e_{\{j,j+1\}}$ and distributed according to $\mathrm{Exp}\,(\gamma)$. For any given γ the variables $t_{\{j,j+1\}}^{(\gamma)}$ are independent and are often called *passage times* of the corresponding edges. In particular, when we write FPP_1 we mean first passage percolation with $\gamma = 1$ whereas when we write FPP_λ we mean first passage percolation with $\gamma = \lambda$.

Definition 1.4 (First Passage Times). *At this point we can define the* passage time *between two vertices x and y of the graph as*

$$T^{(\gamma)}(x \to y) := \min_{\pi:\, x \to y} T^{(\gamma)}(\pi),$$

where the minimum is taken over all paths π that connect x to y. Throughout the paper we will also need a "mixed" first passage time. Given a path π from a vertex x to a vertex y, if there exists a vertex of π that hosts a seed, let w be the first such vertex and let $k \leq n$ denote the distance between x and w along π; note that for fixed x, y and π, the

value k is a random variable as it depends on whether vertices of π host a seed. Then the passage time of π is defined as

$$T^{(1,\lambda)}(\pi) := \sum_{j=0}^{k-1} t^{(1)}_{\{j,j+1\}} + \sum_{j=k}^{n-1} t^{(\lambda)}_{\{j,j+1\}};$$

otherwise, if no vertex of π hosts a seed simply define $T^{(1,\lambda)}(\pi) := T^{(1)}(\pi)$. Furthermore, for any two vertices $x, y \in V(G)$ that are cut-points[1] in $\mathbb{G}^\infty$, we also define the "mixed" passage time between two vertices x and y as

$$T^{(1,\lambda)}(x \to y) := \min_{\pi : x \to y} T^{(1,\lambda)}(\pi),$$

where the minimum is taken over all paths π that connect x to y. Since x and y are cut-points in $\mathbb{G}^\infty$, there are only finitely many paths between x and y.

Remark 1.5. We can couple the passage times for different values of γ as follows. For each edge e, let $t^{(1)}_e$ be independent exponential random variables of rate 1. Then, for any $\gamma > 0$, $\gamma \neq 1$ and any edge e, we set $t^{(\gamma)}_e := t^{(1)}_e / \gamma$.

We will need a result that is based on the known correspondence between the Gamma distribution and a Poisson process. For a complete proof the reader is referred to [3, Lemma 2.5]. In the following statement we denote by $\mathbb{P}$ the probability measure defined on the considered graph where seeds are not present. More specifically, we consider a FPP process with (generic) rate $\gamma > 0$ started at a unique source. We write the statement for a first passage percolation process of rate $\gamma > 0$, as we shall be interested in applying this result to the values $\gamma = 1$ for FPP$_1$ and $\gamma = \lambda$ for FPP$_\lambda$. For a more clear statement, in the following lemma we set $\gamma = 1$, but the statement hold for arbitrary $\gamma > 0$. In fact, it suffices to observe that for any path π, we have that $T^{(1)}(\pi)$ has the same distribution as $\gamma T^{(\gamma)}(\pi)$.

Lemma 1.6. *[3, Lemma 2.5 + Remark 2.6] Consider an infinite graph with bounded degree, and let o denote a reference vertex. For any constant $\mathbf{c_1} > 0$, there is a constant $\mathbf{c}_{\mathrm{out}} = \mathbf{c}_{\mathrm{out}}(\mathbf{c_1}) > 1$ so that for all $k \geq 1$*

$$\mathbb{P}\left[\exists \text{ a path } \pi_k \text{ of length } k \text{ started at } o, \text{ such that } T^{(1)}(\pi_k) \leq \frac{k}{\mathbf{c}_{\mathrm{out}}}\right] \leq e^{-\mathbf{c_1} k}. \quad (1.1)$$

For any constant $0 < \mathbf{c_0} < 1$ there is a constant $\mathbf{c}_{\mathrm{in}} = \mathbf{c}_{\mathrm{in}}(\mathbf{c_0}) < 1$ so that for all $k \geq 1$

$$\mathbb{P}\left[\exists \text{ a path } \pi_k \text{ of length } k \text{ started at } o, \text{ such that } T^{(1)}(\pi_k) \geq \frac{k}{\mathbf{c}_{\mathrm{in}}}\right] \leq e^{-\mathbf{c_0} k}. \quad (1.2)$$

2 Construction of the graph

The construction of the sought graph is done by steps, we start by describing the "tiles" that we will use. The procedure about how to "attach" them together in a suitable manner is explained in Section 2.3.

2.1 The tile G: construction, notation and some facts

Roughly speaking, a "tile" G with vertex set $V(G)$ and edge set $E(G)$ is a connected graph consisting of different types of trees. For any integer $d \geq 1$ let $\mathbb{T}_d$ be the infinite tree where all its nodes (including the root) have exactly d children. For any $h \in \mathbb{N}$ we let $\mathbb{T}_d^h$ denote the tree $\mathbb{T}_d$ truncated at generation h. Finally, the notation $\overline{\mathbb{T}}_d^h$ denotes the tree $\mathbb{T}_d^{h+1}$ where all nodes in generation $h+1$ have been identified into one vertex (refer to Figure 1, where such vertex is denoted by W).

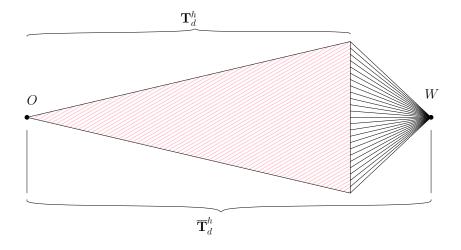[1]A vertex is a cut-point when, if removed, it disconnects the graph.

Figure 1: Sketch of $\overline{\mathbb{T}}_d^h$: the triangle on the left (pink) represents the tree $\mathbb{T}_d^h$, while the one on the right (black) represents the set of edges connecting the h-th generation of $\mathbb{T}_d^h$ to W.

For ease of readability, the reader can refer to Figure 2 for a sketch of G. Start from a vertex O (which we also call *the origin*) and attach two edges to it, that is, identify the starting point of these two edges into O. Call the other two endpoints O_{up} and O_{low}.

For the moment we shall simply say "fix a sufficiently large integer", meaning that we shall be more precise later on by fixing the conditions that such value must satisfy. Fix two sufficiently large integers $D > 2$ and $L > 0$, and identify the root of the graph $\overline{\mathbb{T}}_D^L$ with O_{up}. Similarly, for another large integer $H > 0$ identify the root of $\overline{\mathbb{T}}_2^H$ with O_{low}. Basically, we are "attaching" $\overline{\mathbb{T}}_D^L$ to one edge incident to the origin, and $\overline{\mathbb{T}}_2^H$ to the other one. The last vertex of $\overline{\mathbb{T}}_D^L$ will be denoted by W_{up} and the last vertex of $\overline{\mathbb{T}}_2^H$ will be denoted by W_{low}.

Finally, attach a single edge from W_{up} to a new vertex B and a path of length R (for a large integer R) from W_{low} to B. (Refer to Figure 2.)

Figure 2: Sketch of the tile G and the lengths of its components.

Recall the constants $\mathbf{c_0}, \mathbf{c_1}, \mathbf{c}_{\mathrm{in}}$ and $\mathbf{c}_{\mathrm{out}}$ from Lemma 1.6. Later on, for $d \in \mathbb{N}$, we shall

emphasize this dependence by writing

$\mathbf{c}_{\text{in}}^{(d)} := \mathbf{c}_{\text{in}}^{(d)}(\mathbf{c_0}) :=$ value of $\mathbf{c}_{\text{in}}$ satisfying (1.2) when $\mathbf{c_0}$ is given and the graph is $\mathbb{T}_d$,

$\mathbf{c}_{\text{out}}^{(d)} := \mathbf{c}_{\text{out}}^{(d)}(\mathbf{c_1}) :=$ value of $\mathbf{c}_{\text{out}}$ satisfying (1.1) when $\mathbf{c_1}$ is given and the graph is $\mathbb{T}_d$.

$$(2.1)$$

For the particular case where $d = 1$, we write $\mathbf{c}_{\text{in}}^{(1)}$ and $\mathbf{c}_{\text{out}}^{(1)}$ to denote the values that on a large interval of $\mathbb{Z}$ satisfy (1.2) and (1.1) respectively.

Before proceeding, we introduce some notation. For any fixed $\mu \in (0,1)$, any fixed $d \in \mathbb{N}$ and any infinite tree $\mathbb{T}_d$ define the event

$$\mathcal{E}_{\text{perc}}(\mathbb{T}_d) := \{\text{cluster of non-seeds containing the root percolates on } \mathbb{T}_d\},$$

then set $\mathbf{f}_d(\mu)$ to be

$$\mathbf{f}_d(\mu) := 1 - \mathbb{P}_\mu\left(\mathcal{E}_{\text{perc}}(\mathbb{T}_d)\right). \tag{2.2}$$

By monotonicity of percolation, we know that $\mathbf{f}_d(\mu)$ is increasing with μ. To simplify the notation throughout, for any vertex $x \in V(G)$ and $\mu \in (0,1)$ fixed, set

$$\mathcal{E}_x := \{\text{vertex } x \text{ does not host a seed}\}, \tag{2.3}$$

for which we have

$$\mathbb{P}_\mu(\mathcal{E}_x) = 1 - \mu. \tag{2.4}$$

On $\overline{\mathbb{T}}_d^h$ with the notation borrowed from Figure 1 we set

$$\mathcal{E}_{\text{perc}}(\overline{\mathbb{T}}_d^h) := \{\text{there is a cluster of non-seeds in } \overline{\mathbb{T}}_d^h \text{ connecting } O \text{ to } W\}, \tag{2.5}$$

for which we have

$$\mathbb{P}_\mu\left[\mathcal{E}_{\text{perc}}(\overline{\mathbb{T}}_d^h)\right] \geq \left(1 - \mathbf{f}_d(\mu)\right)(1 - \mu). \tag{2.6}$$

In particular, the first term of the product on the R.H.S. follows from (2.2) and the second is the probability that vertex W does not host a seed. (Recall that in the definition of the model the origin is defined to be free of seeds.)

2.2 Behavior of FPPHE on a tile

Fix $\lambda > 0$ and consider a copy of the tile G as constructed above, and start FPPHE by placing a FPP_1 particle at O, and a seed independently at each vertex with probability μ. From now on, the value of $\lambda > 0$ will be considered fixed and *small enough* for our purposes.

Now we are able to state a result that will be crucial for the whole analysis. The considered graph is $\overline{\mathbb{T}}_d^m$, and the notation that we use is the same as in Figure 1.

Proposition 2.1. *Fix an integer $d \geq 2$, take $\lambda > 0$ small enough, and any $\mu \in (0,1)$. For all $m \in \mathbb{N}$ large enough, consider $\overline{\mathbb{T}}_d^m$ with the following initial setting: place a particle of FPP_1 at o and at every $x \in V(\overline{\mathbb{T}}_d^m) \setminus \{o\}$ place seeds independently with probability μ. Then there is a function $\boldsymbol{\eta}_d = \boldsymbol{\eta}_d(\mu, \lambda) > 0$ so that the following hold.*

(i) For all μ small enough,

$$\mathbb{P}_\mu\left[W \text{ is infected by } \text{FPP}_1\right] \geq \left(1 - \mathbf{f}_d(\mu) - \boldsymbol{\eta}_d\right)(1 - \mu).$$

(ii) If $\mu \in (1/2, 1)$ so that

$$d(1 - \mu) > 1 \quad \text{and} \quad d^2(1 - \mu)^2 \mu^{d-1} < 1, \tag{2.7}$$

then,

$$\mathbb{P}_\mu\left[W \text{ is infected by } \text{FPP}_1\right] \geq \left(1 - \mathbf{f}_d(\mu) - \boldsymbol{\eta}_d\right)(1 - \mu).$$

(iii) For any $\xi > 0$, if λ is small enough and either μ is small enough or d is large enough, then $\boldsymbol{\eta}_d \leq \xi$.

The factor $1 - \mu$ corresponds to the probability that W does not host a seed.

Remark 2.2. Observe that, since $\lambda > 0$ is fixed throughout, the quantity $\boldsymbol{\eta}_d(\mu, \lambda)$ can be seen as a function of μ only. Furthermore, whenever λ and μ are fixed, $\boldsymbol{\eta}_d(\mu, \lambda)$ is decreasing in d. In order to avoid making the notation even more cumbersome, we chose that whenever μ is clear from the context we shall omit its explicit dependence from $\boldsymbol{\eta}_d$.

Since the proof of Proposition 2.1 is quite technical, we postpone it to Section 5, thus we will proceed while assuming the validity of the result. The next statement provides a crucial fact regarding the behavior of the process on a tile G with origin O.

Proposition 2.3. *Fix a large integer D, and let λ be small enough. Then there are values $0 < \mu_{1,low}^{(crit)} < \mu_{1,high}^{(crit)} < 1/2 < \mu_{2,low}^{(crit)} < \mu_{2,high}^{(crit)} < 1$, and values of H, L, R, Φ so that in such a tile G we obtain*

$$\mathbb{P}_\mu \left(B \text{ is infected by } \mathrm{FPP}_1 \right) < \frac{1}{\Phi}, \quad \text{for all } \mu \in \left(\mu_{1,low}^{(crit)}, \mu_{1,high}^{(crit)} \right)$$

and

$$\mathbb{P}_\mu \left(B \text{ is infected by } \mathrm{FPP}_1 \right) > \frac{1}{\Phi}, \quad \text{for all } \mu \in \left(\mu_{2,low}^{(crit)}, \mu_{2,high}^{(crit)} \right).$$

This statement will be proven in Section 4, admitting the result of Proposition 2.1 (which will be proven in Section 5).

2.3 The infinite graph $\mathbb{G}^\infty$

We proceed with the construction that provides the sought counterexample to monotonicity. We view each tile with an orientation, where the vertex O is called the *origin* of the tile, and the vertex B is called the *tail* of the tile. Then $\mathbb{G}^\infty$ is constructed as follows. Fix a large integer Φ and consider an infinite $(\Phi + 1)$-regular tree. Consider an infinite (self-avoiding) path starting at the root, this determines an end of the tree, and thus a specific direction. Now, for each vertex v of the tree, let $u_0, u_1, \ldots, u_\Phi$ be the neighbors of v with u_0 being the unique neighbor that is on the direction of the chosen end. Now we replace each edge (v, u_i) of the tree, for $i \in \{1, 2, \ldots, \Phi\}$, with a copy of the tile, identifying the origin of the tile (i.e., the vertex O) with v and the tail of the tile (i.e., the vertex B) with u_i. For the edge (v, u_0), the orientation of the tile is reversed; i.e., we replace (v, u_0) with a copy of the tile, but identify the origin of the tile with u_0 and the tail of the tile with v.

A graphical representation is shown in Figure 4 (for simplicity we used $\Phi = 2$, although it follows from the computations in Section 4 that Φ is much larger than 2). Note that the graph $\mathbb{G}^\infty$ is quasi transitive.

2.4 Proof of Theorem 1.1

We state our result for FPPHE on the graph $\mathbb{G}^\infty$ below, from which Theorem 1.1 follows directly.

Theorem 2.4. *Consider FPPHE starting from the origin of $\mathbb{G}^\infty$. Fix a large integer D, let λ be small enough, and let the values $\mu_{1,low}^{(crit)}, \mu_{1,high}^{(crit)}, \mu_{2,low}^{(crit)}, \mu_{2,high}^{(crit)}$ be as in Proposition 2.3. Then there are values for H, L, R and Φ so that,*

$$\mathbb{P}_\mu(\mathrm{FPP}_1 \text{ survives}) = 0 \quad \text{for all } \mu \in \left(\mu_{1,low}^{(crit)}, \mu_{1,high}^{(crit)} \right)$$

and

$$\mathbb{P}_\mu(\mathrm{FPP}_1 \text{ survives}) > 0 \quad \text{for all } \mu \in \left(\mu_{2,low}^{(crit)}, \mu_{2,high}^{(crit)} \right).$$

Remark 2.5. We make the following observations.

1. It is immediate that FPP_1 survives if μ is small enough (much smaller than $\mu_{1,\text{low}}^{(\text{crit})}$) since seed-free tiles percolate on $\mathbb{G}^\infty$ for all small enough μ. Moreover, it is immediate that FPP_1 dies out if μ is large enough (much larger than $\mu_{2,\text{high}}^{(\text{crit})}$) so that the origin is blocked by seeds with probability one. Therefore, this reasoning and Theorem 2.4 establish at least three phase transitions on $\mathbb{G}^\infty$: from survival of FPP_1 for small enough μ, then non-survival for some μ in $(0, 1/2)$, then a survival regime for some μ in $(1/2, 1)$ to a final regime of non-survival for all large μ.

2. By further working the construction of the tiles we can make as many phase transitions as we desire. For example, we could have a tile composed of κ copies of the upper and lower parts so that O would have degree 2κ, with each copy of the upper and lower parts of different sizes and different degrees for the corresponding trees. Then by tuning the parameters properly, as we increase μ we could make B be occupied by FPP_λ through the first lower part, then by FPP_1 through the first upper part, then by FPP_λ through the second lower part, then by FPP_1 through the second upper part, and so on and so forth.

3. We choose to fix λ small enough and show non-monotonicity when varying μ. A similar result could be obtained by fixing μ small enough first and then varying λ inside the interval $(0, 1)$. For example, in the lower part of the tile we could split the vertex W_{low} into two vertices (W_{low} and W_{low}'), connect them with a small finite tree, and then create an additional path of length R' connecting O_{low} to W_{low} (see Figure 3). Then, from O_{low} there would be two ways to W_{low}, one through $\overline{\mathbb{T}}_2^H$ which is favorable to FPP_1 by having μ small and $\lambda < 1$, and the new path which will be likely to have a seed in it. Even for not-too-small values of λ, we can make the path of length R' slow enough to let FPP_1 get to W_{low} through $\overline{\mathbb{T}}_2^H$, leading to a fast passage through the small tree from W_{low} to W_{low}' and the vertex B to be occupied by FPP_λ from the lower part through the activation of a seed in the path of length R. On the other direction, by increasing λ we can make the path of length R' to be the fastest way to W_{low}, but since $\lambda < 1$ this makes the way through the small tree from W_{low} to W_{low}' slow enough to let the upper part of the tile be the fastest. This leads to FPP_1 occupying B from the upper part. By tuning the parameters we can then establish the non-monotonicity. For conciseness, we only do the details for the phase transition on μ.

Remark 2.6. In the following, we set the parameters in the following order. First, we choose $\mu_{1,\text{low}}^{(\text{crit})}, \mu_{1,\text{high}}^{(\text{crit})}$ and $\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})}$, then we choose the values of L, R, H, D uniformly over $\mu_1 \in \left(\mu_{1,\text{low}}^{(\text{crit})}, \mu_{1,\text{high}}^{(\text{crit})}\right)$ and $\mu_2 \in \left(\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})}\right)$.

Proof of Theorems 1.1 and 2.4. We now prove Theorem 2.4 (Theorem 1.1 is just a direct consequence of it). The proof of Theorem 2.4 is a consequence of Proposition 2.3 and classical facts about percolation. Proposition 2.3 gives values for the quantities D, L, H, R so that all calculations work for all $\mu_1 \in \left(\mu_{1,\text{low}}^{(\text{crit})}, \mu_{1,\text{high}}^{(\text{crit})}\right)$ and $\mu_2 \in \left(\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})}\right)$. Thus, we can fix μ_1 and μ_2 in the above intervals. We can assume, without loss of generality, that the origin of $\mathbb{G}^\infty$ is a vertex that has been identified with the origin of a tile (i.e., it is not an internal vertex of a tile). The reason why we can assume this is that, if the origin were an internal vertex of a tile, then with positive probability FPP_1 would spread to both the origin of the tile and the tail of the tile (since there is a positive probability that the tile contains no seed). This means that FPP_1 has a positive probability of surviving in this case if and only if it has a positive probability of survival when the origin of $\mathbb{G}^\infty$ is not internal to a tile.
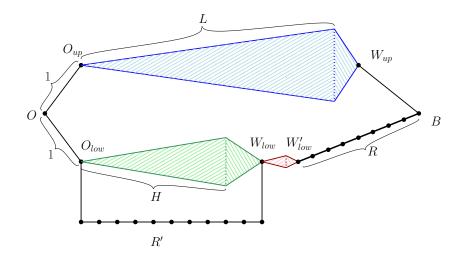
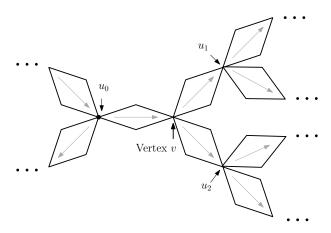Figure 3: The tile G that gives non-monotonicity in λ as in Remark 2.5 part 3.

Figure 4: The construction of $\mathbb{G}^\infty$, every diamond represents a copy of the tile G, while arrows inside each tile point from the origin of the tile to the tail. Considering vertex v as the origin of $\mathbb{G}^\infty$, we obtain that $\vec{\mathbb{G}}^\infty$ (as formally defined in the proof of Theorems 1.1 and 2.4) as the part of the figure to the right of v.

From now on, let o denote the origin of $\mathbb{G}^\infty$, which is identified with the origin of Φ tiles and with the tail of one tile, we denote this last tile G_{parent}. Referring to Figure 4 with the vertex v being o, we obtain that G_{parent} is the tile between u_0 and v. We shall denote the corresponding percolation probability measure by $\mathbb{P}^p$, in order to avoid confusion with the previously defined measures. Now let $\vec{\mathbb{G}}^\infty$ be the subgraph where we remove from $\mathbb{G}^\infty$ all vertices and edges whose geodesic to o intersects G_{parent}; in Figure 4 this is the part of the graph to the right of v. Note that by collapsing each tile of $\vec{\mathbb{G}}^\infty$ to an edge, we obtain a Φ-ary tree rooted at o. Then, by performing (edge) Bernoulli percolation with parameter p on an infinite Φ-ary tree we have (cf., e.g., [15])

$$\mathbb{P}^p\big(\text{cluster containing the root is infinite}\big) > 0 \quad \Leftrightarrow \quad p > \frac{1}{\Phi}.$$

In a similar fashion, consider a percolation process on the graph $\vec{\mathbb{G}}^\infty$ defined as follows. Fix a value $p \in (0,1)$. For every copy of G in $\vec{\mathbb{G}}^\infty$, toss an independent coin which gives

heads with probability p and tails with probability $1 - p$. If the coin gives heads, then declare the corresponding copy of G to be *open*, otherwise declare it *closed*. This process on $\vec{\mathbb{G}}^\infty$ is a percolation process of parameter p on a Φ-ary tree, and thus

$$\mathbb{P}^p(\text{cluster containing vertex } o \text{ is infinite}) > 0 \quad \Leftrightarrow \quad p > \frac{1}{\Phi}. \tag{2.8}$$

From Proposition 2.3 it follows immediately that, for a large integer Φ (which will be specified later in Equation (4.6)) we have

$$\mathbb{P}_{\mu_1}\left(B \text{ is infected by } \mathrm{FPP}_1\right) > \frac{1}{\Phi} \quad \text{and} \quad \mathbb{P}_{\mu_2}\left(B \text{ is infected by } \mathrm{FPP}_1\right) < \frac{1}{\Phi}. \tag{2.9}$$

Therefore, it suffices to realize that the event $\{\mathrm{FPP}_1 \text{ survives }\}$ corresponds to the event $\mathcal{E} := \{\exists \text{ infinite cluster of tiles so that, for each such tile, vertex } B \text{ is infected by } \mathrm{FPP}_1\}$. Hence, by (2.8)

$$\mathbb{P}^p(\mathcal{E}) > 0 \quad \Leftrightarrow \quad p > \frac{1}{\Phi}.$$

Now we turn to analyzing $\mathbb{G}^\infty$. When FPP_1 survives in $\vec{\mathbb{G}}^\infty$, then it also survives on $\mathbb{G}^\infty$. We are then left to show that when FPP_1 dies out almost surely in $\vec{\mathbb{G}}^\infty$ then it also dies out almost surely in $\mathbb{G}^\infty$. To see this, let $o_1, o_2, \ldots$ be a sequence of origins of tiles, starting with $o = o_1$, such that for each $i \geq 1$ there is a tile for which o_{i+1} is its origin and o_i its tail. In other words, $o_1, o_2, \ldots$ is the unique sequence of origins of tiles that are traversed by a path that starts from o and can only traverse tiles from the tail to the origin. Since this sequence is unique and the corresponding tiles form a one-dimensional chain of tiles, almost surely FPP_1 occupies only a finite number of elements in this sequence (for example, it will happen almost surely that there exists a k for which o_k hosts a seed). Let $\ell < \infty$ be such that o_ℓ is occupied by FPP_1 and $o_{\ell+1}$ is occupied by FPP_λ. For FPP_1 to survive, it has to survive by spreading through a path from $o_1, o_2, \ldots, o_\ell$ which first traverses a tile from the origin to the tail. But such paths correspond to spreading of FPP_1 in $\vec{\mathbb{G}}^\infty$. Since FPP_1 dies out almost surely in $\vec{\mathbb{G}}^\infty$, it will also die out when spreading from each one of $o_1, o_2, \ldots, o_\ell$, concluding the proof. $\qquad\square$

3 Analysis of the process on the tile

This section is devoted to the analysis of the process on each tile. In particular, we will state and prove two results (Lemmas 3.1 and 3.2 below) which we will use in the proof of Proposition 2.3. In order to prove the lemmas, we will apply the result from Proposition 2.1. So μ_1 and μ_2 are given as specified in Proposition 2.1. In the following, we shall refer to the *upper part* of G as the subgraph of G induced by the vertices of $\overline{\mathbb{T}}_D^L$ together with O and B. Symmetrically, the *lower part* of G is the subgraph of G induced by all vertices of G except for those of $\overline{\mathbb{T}}_D^L$; cf. Figure 2.

Informally speaking, we study the spread of the process along two main streams, the lower one and the upper one. Since they only communicate at O and at B, they are independent until the moment at which vertex B gets infected. Once this happens, it will be crucial to understand whether B got infected from the upper or from the lower part.

In order to avoid confusion, we define another notation for passage times that are only local to G. More precisely, if the edges of a path $\pi = (e_{\{0,1\}} \ldots e_{\{n-1,n\}})$ are completely contained in $E(G)$ and is such that when exploring the path from vertex 0 to vertex n the first seed is encountered at k, then (similarly to the reasoning in Definition 1.4) we define the quantity

$$\Upsilon^{(1,\lambda)}(\pi) := \sum_{j=0}^{k-1} t_{\{j,j+1\}}^{(1)} + \sum_{j=k}^{n-1} t_{\{j,j+1\}}^{(\lambda)}.$$

Consequently, similarly to what discussed in Definition 1.4, the "mixed" passage time between two vertices $x, y \in V(G)$ along a path completely contained in G is denoted by $\Upsilon^{(1,\lambda)}(x \to y)$. When $k = 0$ or $k = n$ in the above formula, we shall simplify the notation writing $\Upsilon^{(\lambda)}(x \to y)$ or $\Upsilon^{(1)}(x \to y)$, respectively. Note that this is the first passage time between x and y restricted to paths contained in G, and thus in general we would have $\Upsilon^{(1,\lambda)}(x \to y) \geq T^{(1,\lambda)}(x \to y)$.

To avoid confusion, we denote by $\Upsilon_{\mathrm{up}}^{(1,\lambda)}(x \to y)$ the first passage time between x and y restricted to paths in the *upper part* of G, and by $\Upsilon_{\mathrm{low}}^{(1,\lambda)}(x \to y)$ the first passage time from x to y restricted to paths in the *lower part*. Moreover, when we write the event $\{W_{\mathrm{up}}$ is infected by $\mathrm{FPP}_1\}$, we mean that this happens when we restrict to the upper part of the tile, so we do not account for the possibility that W_{up} is infected by B.

We start from the fact that for any $\varepsilon > 0$ fixed and any spread rate $\gamma > 0$ there is a constant $C(\varepsilon, \gamma) > 0$ such that uniformly in $e \in E(G)$ we have $\mathbb{P}\big(t_e^{(\gamma)} < C(\varepsilon, \gamma)\big) > (1 - \varepsilon)$. In what follows, set $\varepsilon > 0$ as an arbitrarily small fixed constant, and we will set

$$\mathfrak{C} := \max\big\{C(\varepsilon, 1), C(\varepsilon, \lambda)\big\}. \tag{3.1}$$

The next two subsections are devoted to prove two results. More specifically, in Lemma 3.1 we show that when μ is small, the process is very likely to be "quick" in the upper part and "slow" in the lower part (where "quick" and "slow" are to be made precise below). Subsequently, in Lemma 3.2 we show that when μ is large (but not too large), it is very likely that the situation is reversed. The proof of Proposition 2.3 will strongly rely on these results.

Lemma 3.1. *Let $\mathfrak{C}$ be as in (3.1). There are critical values $0 < \mu_{1,low}^{(crit)} < \mu_{1,high}^{(crit)} < 1/2$ and large enough values of L, H, R and D so that for any fixed $\varepsilon > 0$ there are $\mathbf{c}_{\mathrm{out}}^{(D)} = \mathbf{c}_{\mathrm{out}}^{(D)}(\varepsilon)$, $\mathbf{c}_{\mathrm{in}}^{(1)} = \mathbf{c}_{\mathrm{in}}^{(1)}(\varepsilon)$, $\mathbf{c}_{\mathrm{in}}^{(2)} = \mathbf{c}_{\mathrm{in}}^{(2)}(\varepsilon)$ such that for all $\mu_1 \in \left(\mu_{1,low}^{(crit)}, \mu_{1,high}^{(crit)}\right)$ the following hold.*

(i) In the lower part,

$$\mathbb{P}_{\mu_1}\left[\Upsilon_{low}^{(1,\lambda)}(O \to B) \leq \mathfrak{C} + \frac{H+1}{\mathbf{c}_{\mathrm{in}}^{(2)}} + \frac{R}{\lambda \mathbf{c}_{\mathrm{in}}^{(1)}}\right]$$
$$\geq (1-\varepsilon)^3(1-\mu_1)^2\left[(1 - \mathbf{f}_2(\mu_1) - \boldsymbol{\eta}_2)(1-\mu_1)^2 - \varepsilon\right],$$

where the factor $\boldsymbol{\eta}_2$ is as in Proposition 2.1 and $\mathbf{f}_2(\mu)$ as in (2.2) when $d = 2$.

(ii) In the upper part,

$$\mathbb{P}_{\mu_1}\left[\Upsilon_{up}^{(1,\lambda)}(O \to B) \geq \frac{L+1}{\mathbf{c}_{\mathrm{out}}^{(D)}}\right] \geq 1 - \varepsilon.$$

Remark The critical value $0 < \mu_{1,\mathrm{high}}^{(\mathrm{crit})} < 1/2$ depends on λ and has to satisfy some technical conditions that will be specified later on (expressed in details in conditions (4.7), (5.37), (5.47), (5.49)). Moreover, since the result holds for all μ_1 small enough, we can simply set $\mu_{1,\mathrm{low}}^{(\mathrm{crit})} := \mu_{1,\mathrm{high}}^{(\mathrm{crit})}/2$. Note that such values are independent of L, H, R, D.

Lemma 3.2. *Let $\mathfrak{C}$ be as in (3.1). There are critical values $1/2 < \mu_{2,low}^{(crit)} < \mu_{2,high}^{(crit)} < 1$ and large enough values of L, H, R and D so that for any fixed $\varepsilon > 0$ there are $\mathbf{c}_{\mathrm{out}}^{(D)} = \mathbf{c}_{\mathrm{out}}^{(D)}(\varepsilon)$, $\mathbf{c}_{\mathrm{in}}^{(1)} = \mathbf{c}_{\mathrm{in}}^{(1)}(\varepsilon)$, $\mathbf{c}_{\mathrm{in}}^{(2)} = \mathbf{c}_{\mathrm{in}}^{(2)}(\varepsilon)$ such that for all $\mu_2 \in \left(\mu_{2,low}^{(crit)}, \mu_{2,high}^{(crit)}\right)$,*

(i) In the lower part,

$$\mathbb{P}_{\mu_2}\left[\Upsilon_{low}^{(1,\lambda)}(O \to B) \geq \frac{H/2}{\mathbf{c}_{\mathrm{out}}^{(2)}} + \frac{H/2+1}{\lambda \mathbf{c}_{\mathrm{out}}^{(2)}} + \frac{R}{\lambda \mathbf{c}_{\mathrm{out}}^{(1)}}\right] \geq (1 - 3\varepsilon).$$

(ii) In the upper part,

$$\mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{up}^{(1,\lambda)}(O \to B) \le 2\mathfrak{C} + \frac{L+1}{\mathbf{c}_{\text{in}}^{(D)}}\right\} \cap \{W_{up} \text{ is infected by } \text{FPP}_1\}\right]$$

$$\ge (1-\varepsilon)^2(1-\mu_2)^2\left[\left(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2)\right)(1-\mu_2)^2 - \varepsilon\right].$$

where $\boldsymbol{\eta}_D$ is as in Proposition 2.1 and $\mathbf{f}_D(\mu)$ as in (2.2) when $d = D$.

3.1 Proof of Lemma 3.1

The aim of this section is to understand the "typical" behavior of the process on the tile G when $\mu = \mu_1$ is small enough. Sections 3.1.1 and 3.1.2 take care of the analysis in the lower and upper part of G respectively.

3.1.1 Proof of Lemma 3.1(i), "lower part" of G

Consider the variable $\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B)$. Fix any pair of vertices x, y in the lower part of G and an initial density of seeds $\mu_1 \in \left(\mu_{1,\text{low}}^{(\text{crit})}, \mu_{1,\text{high}}^{(\text{crit})}\right)$, where $\mu_{1,\text{high}}^{(\text{crit})}$ is a small enough value (cf. Section 5). For this part of the proof consider only the lower part of G, more precisely define the passage times $\{\bar{t}_e^{(\gamma)}\}$ as follows. Let $\bar{t}_e^{(\gamma)} := t_e^{(\gamma)}$ if the edge e is in the lower part of G, and set $\bar{t}_e^{(\gamma)} := \infty$, if e is in the upper part. With respect to the passage times $\{\bar{t}_e^{(\gamma)}\}_{e \in E(G)}$, recalling the definition of $\mathcal{E}_y$ from (2.3), define the event

$$\mathfrak{B}_{\text{low}}(x,y) := \mathcal{E}_y \cap \{\text{FPP}_1 \text{ started at } x \text{ infects } y \text{ before } \text{FPP}_\lambda\}. \tag{3.2}$$

Remark 3.3. Note that the event $\mathfrak{B}_{\text{low}}(x,y)$ does not require x to be seed-less.

Using the fact that O_{low} and W_{low} are cutpoints, under the event $\mathcal{E}_O \cap \mathfrak{B}_{\text{low}}(O, O_{\text{low}}) \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}})$,

$$\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) = \Upsilon_{\text{low}}^{(1,\lambda)}(O \to O_{\text{low}}) + \Upsilon_{\text{low}}^{(1,\lambda)}(O_{\text{low}} \to W_{\text{low}}) + \Upsilon_{\text{low}}^{(1,\lambda)}(W_{\text{low}} \to B).$$

Thus,

$$\mathbb{P}_{\mu_1}\left[\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) \le \mathfrak{C} + \frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}} + \frac{R}{\lambda\mathbf{c}_{\text{in}}^{(1)}}\right]$$

$$\ge \mathbb{P}_{\mu_1}\left[\left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) \le \mathfrak{C} + \frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}} + \frac{R}{\lambda\mathbf{c}_{\text{in}}^{(1)}}\right\} \cap \mathcal{E}_O \cap \mathfrak{B}_{\text{low}}(O, O_{\text{low}}) \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}})\right]$$

$$\ge \mathbb{P}_{\mu_1}\left[\mathcal{E}_O \cap \mathfrak{B}_{\text{low}}(O, O_{\text{low}}) \cap \left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O \to O_{\text{low}}) \le \mathfrak{C}\right\}\right] \times$$

$$\times \mathbb{P}_{\mu_1}\left[\mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}}) \cap \left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O_{\text{low}} \to W_{\text{low}}) \le \frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}}\right\}\right] \times$$

$$\times \mathbb{P}_{\mu_1}\left[\Upsilon_{\text{low}}^{(1,\lambda)}(W_{\text{low}} \to B) \le \frac{R}{\lambda\mathbf{c}_{\text{in}}^{(1)}}\right], \tag{3.3}$$

where we used independence of the passage times in the second bound.

We proceed our analysis by bounding each term in (3.3).

First term Since the graph $\{O, O_{\text{low}}\}$ is actually an edge, then

$$\mathbb{P}_{\mu_1}\left[\mathcal{E}_O \cap \mathfrak{B}_{\text{low}}(O, O_{\text{low}}) \cap \left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O \to O_{\text{low}}) \le \mathfrak{C}\right\}\right]$$

$$= \mathbb{P}_{\mu_1}\left[\left\{\Upsilon_{\text{low}}^{(1)}(O \to O_{\text{low}}) \le \mathfrak{C}\right\} \cap \mathcal{E}_O \cap \mathcal{E}_{O_{\text{low}}}\right] \overset{(2.4),\ (3.1)}{>} (1-\varepsilon)(1-\mu_1)^2.$$

Second term Because of Proposition 2.1, by taking $o \equiv O_{\text{low}}$, $m = H$, $d = 2$ and $W \equiv W_{\text{low}}$, the event $\mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}})$ has probability bounded from below by $(1 - \mathbf{f}_2(\mu_1) - \boldsymbol{\eta}_2)(1 - \mu_1)$. For the following part we need a technical construction which will be developed in detail later on, therefore at this point we present a result whose proof is deferred to Section 5.1, since it requires a comparison with branching random walks that will be described in Section 5. At this point we emphasize that the bound obtained in the following result is independent of λ (provided that λ is chosen to be small).

Claim 3.4. *Suppose that μ_1 and μ_2, as well as L, H, R, D, are as in Lemmas 3.1 and 3.2. Let $\varepsilon > 0$ be a small enough constant so that $\mathbf{c}_{\text{in}}^{(2)} < 1$ and $\mathbf{c}_{\text{in}}^{(D)} < 1$ are small enough. Then there are values $C, \delta, \beta > 0$ independent of H and of $\mathbf{c}_{\text{in}}^{(2)}$ such that*

$$
\mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{low}} \cap \mathfrak{B}_{low}(O_{low}, W_{low}) \cap \left\{\Upsilon_{low}^{(1,\lambda)}(O_{low} \to W_{low}) \leq \frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}}\right\}\right]
$$
$$
\geq (1 - \mathbf{f}_2(\mu_1) - \boldsymbol{\eta}_2)(1 - \mu_1)^2 - e^{-\beta H} - Ce^{-\delta(H+1)/(2\mathbf{c}_{\text{in}}^{(2)})}.
$$

Similarly, there are constants $\overline{C}, \overline{\delta}, \overline{\beta} > 0$ that are independent of L and of $\mathbf{c}_{\text{in}}^{(D)}$, so that

$$
\mathbb{P}_{\mu_2}\left[\mathcal{E}_{O_{up}} \cap \mathfrak{B}_{up}(O_{up}, W_{up}) \cap \left\{\Upsilon_{up}^{(1,\lambda)}(O_{up} \to W_{up}) \leq \frac{L+1}{\mathbf{c}_{\text{in}}^{(D)}}\right\}\right]
$$
$$
\geq (1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2))(1 - \mu_2)^2 - e^{-\overline{\beta} H} - \overline{C}e^{-\overline{\delta}(L+1)/(2\mathbf{c}_{\text{in}}^{(D)})}.
$$

Hence in this context we choose H large enough, so that $e^{-\beta H} + Ce^{-\delta(H+1)/(2\mathbf{c}_{\text{in}}^{(2)})} < \varepsilon$.

Third term By the *geodesic path* connecting W_{low} to B we mean the path of length R as described in Section 2.1. Define the event

$$\mathcal{E}_{\text{seed}} := \{\text{there is a seed on a vertex along the (geodesic) path going from } W_{\text{low}} \text{ to } B\}. \tag{3.4}$$

Since

$$\mathbb{P}_{\mu_1}(\mathcal{E}_{\text{seed}}) = 1 - (1 - \mu_1)^R, \tag{3.5}$$

then, by choosing R large enough uniformly over $\mu_1 \in \left(\mu_{1,\text{low}}^{(\text{crit})}, \mu_{1,\text{high}}^{(\text{crit})}\right)$, we can make the above probability arbitrarily close to 1. We emphasize that if the first seed appears at the beginning of the path, then the passage time of FPP_λ along the path is higher than if it originated towards the end of the path. This can be obtained by a simple coupling argument. In fact, since λ is small, then when adding a seed the passage time for each edge e of the path $(t_e^{(\gamma)})$ is not decreasing. More precisely, if e is preceding the seed, then $t_e^{(\gamma)} = t_e^{(1)}$, whereas if e is after the seed, then $t_e^{(\gamma)} = t_e^{(\lambda)} > t_e^{(1)}$. From this observation together with Lemma 1.6, applied to $\gamma = \lambda < 1$, it follows that for any $\mathbf{c_0}^{(1)} \in (0,1)$ fixed, there is $\mathbf{c}_{\text{in}}^{(1)} < 1$ such that

$$
\mathbb{P}_{\mu_1}\left[\Upsilon_{\text{low}}^{(1,\lambda)}(W_{\text{low}} \to B) \leq \frac{R}{\lambda \mathbf{c}_{\text{in}}^{(1)}}\right] \geq \mathbb{P}_{\mu_1}\left[\left\{\Upsilon_{\text{low}}^{(\lambda)}(W_{\text{low}} \to B) \leq \frac{R}{\lambda \mathbf{c}_{\text{in}}^{(1)}}\right\} \cap \mathcal{E}_{\text{seed}}\right]
$$
$$
\overset{(3.5)}{\geq} \left(1 - e^{-\mathbf{c_0}^{(1)} R}\right)\left(1 - (1 - \mu_1)^R\right). \tag{3.6}
$$

We can now conclude the proof of the first part of the lemma.

Lemma 3.1(i). By plugging all these lower bounds into (3.3) we obtain

$$\mathbb{P}_{\mu_1}\left[\Upsilon_{\text{low}}^{(1)}(O \to B) \le \mathfrak{C} + \frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}} + \frac{R}{\lambda \mathbf{c}_{\text{in}}^{(1)}}\right]$$

$$\ge (1-\varepsilon)(1-\mu_1)^2 \left[(1-\mathbf{f}_2(\mu_1)-\boldsymbol{\eta}_2)(1-\mu_1)^2 - \varepsilon\right]\left[1 - e^{-\mathbf{co}^{(1)}R}\right]\left[1 - (1-\mu_1)^R\right].$$

The statement follows by taking R large enough, uniformly over $\mu_1 \in \left(\mu_{1,\text{low}}^{(\text{crit})}, \mu_{1,\text{high}}^{(\text{crit})}\right)$, so that $e^{-\mathbf{co}^{(1)}R} \le \varepsilon$ and $(1-\mu_1)^R < \varepsilon$. $\qquad\square$

3.1.2 Proof of Lemma 3.1(ii), "upper part" of G

Here we investigate what is likely to happen on the upper part of G. Here $D > 5$ denotes a large integer that will be specified later. Note that the tree structure implies that the geodesics have length exactly $L + 1$. Also, note the following simple coupling argument: for any $\lambda < 1$, if X is an exponential random variable of parameter 1, then X/λ is an exponential random variable of parameter λ. Therefore, the sum of independent exponential random variables of rate λ stochastically dominates the sum of the independent exponential random variables of rate 1. Then, by Lemma 1.6, for any fixed $\mathbf{c}_1^{(D)} > 0$ there is a $\mathbf{c}_{\text{out}}^{(D)} > 1$ so that

$$\mathbb{P}_{\mu_1}\left[\Upsilon_{\text{up}}^{(1,\lambda)}(O \to B) \ge \frac{L+1}{\mathbf{c}_{\text{out}}^{(D)}}\right] \ge \mathbb{P}_{\mu_1}\left[\Upsilon_{\text{up}}^{(1,\lambda)}(O_{\text{up}} \to W_{\text{up}}) \ge \frac{L+1}{\mathbf{c}_{\text{out}}^{(D)}}\right]$$

$$\overset{\text{coupling}}{\ge} \mathbb{P}_{\mu_1}\left[\Upsilon_{\text{up}}^{(1)}(O_{\text{up}} \to W_{\text{up}}) \ge \frac{L+1}{\mathbf{c}_{\text{out}}^{(D)}}\right] \ge 1 - e^{-\mathbf{c}_1^{(D)}(L+1)}. \tag{3.7}$$

Now it suffices to take L so large that $e^{-\mathbf{c}_1^{(D)}(L+1)} \le \varepsilon$ and the proof of the lemma is concluded.

3.2 Proof of Lemma 3.2

As required in Proposition 2.1 (and in the hypotheses of the Lemma), we will assume that D is large enough and $\mu_2 \in \left(\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})}\right)$ satisfies (2.7). Since $D(1-\mu_2) > 1$, then

$$\mathbb{P}_{\mu_2}\left(\mathcal{E}_{\text{perc}}(\mathbb{T}_2)\right) = 0 \quad \text{and} \quad \mathbb{P}_{\mu_2}\left(\mathcal{E}_{\text{perc}}(\mathbb{T}_D)\right) > 0. \tag{3.8}$$

This is clearly not the case for $\mu = \mu_1 < \frac{1}{2}$, and this difference will turn out to be crucial.

3.2.1 Proof of Lemma 3.2(i), "lower part" of G

Consider the connected component containing O_{low} consisting of vertices $\overline{\mathbb{T}}_2^H$ which are not hosting seeds (placed with density μ_2). This set has the same distribution as the first H generations of a *subcritical* Galton-Watson tree (GWT) whose reproduction rule is a Bin $(2, (1-\mu_2))$.

By letting $\mathbf{P}_{\mu_2}$ denote the probability measure defined on the space of Galton-Watson trees with the above-mentioned distribution, then it is a standard fact that

$$\mathbf{P}_{\mu_2}(\text{subcritical GWT survives for more than } H/2 \text{ generations}) \le \left(2(1-\mu_2)\right)^{H/2}.$$

Thus, the event

$$\mathcal{E}_{\text{subcr}}(H) := \left\{\text{cluster of non-seeds on } \overline{\mathbb{T}}_2^H \text{ is a tree of height at most } H/2\right\} \tag{3.9}$$

has probability bounded from below by

$$\mathbf{P}_{\mu_2}\left(\mathcal{E}_{\mathrm{subcr}}(H)\right) \geq 1 - \left(2(1-\mu_2)\right)^{H/2}, \tag{3.10}$$

which can be made as close to 1 as we need, uniformly over all $\mu_2 \in \left(\mu_{2,\mathrm{low}}^{(\mathrm{crit})}, \mu_{2,\mathrm{high}}^{(\mathrm{crit})}\right)$, by taking a sufficiently large value of H.

Recall the definition $\mathfrak{B}_{\mathrm{low}}(x,y)$ from (3.2). We will show that it is very likely that, when restricting the analysis to the lower part, FPP_λ takes a very long time before it can potentially infect W_{low}. In order to proceed, we need an auxiliary result.

Lemma 3.5. *For the quantities defined above we have*

$$\mathbb{P}_{\mu_2}\left(\mathfrak{B}_{low}^c(O_{low}, W_{low}) \cap \left\{\Upsilon_{low}^{(1,\lambda)}(O_{low} \to W_{low}) \geq \frac{H/2}{\mathbf{c}_{\mathrm{out}}^{(2)}} + \frac{H/2+1}{\lambda\mathbf{c}_{\mathrm{out}}^{(2)}}\right\}\right)$$
$$\geq 1 - 2^{H/2+1}e^{-\mathbf{c_1}^{(2)}H/2} - \left(2(1-\mu_2)\right)^{H/2}.$$

Proof. For every $k \in \mathbb{N}$ define $\mathcal{L}_{2,k}$ to be the set of vertices at generation k of $\overline{\mathbb{T}}_2^H$, that is,

$$\mathcal{L}_{2,k} := \left\{x \in V(\overline{\mathbb{T}}_2^H) \ : \ d_{\overline{\mathbb{T}}_2^H}(O_{\mathrm{low}}, x) = k\right\},$$

where $d_{\overline{\mathbb{T}}_2^H}(O_{\mathrm{low}}, x)$ is the distance induced by the graph metric on $\overline{\mathbb{T}}_2^H$. We observe that since $\lambda < 1$, via a simple coupling argument (similar to the one outlined in Section 3.1.2), we see that the probability of the event

$$\mathfrak{B}_{\mathrm{low}}^c(O_{\mathrm{low}}, W_{\mathrm{low}}) \cap \left\{\Upsilon_{\mathrm{low}}^{(1,\lambda)}(O_{\mathrm{low}} \to W_{\mathrm{low}}) \geq \frac{H/2}{\mathbf{c}_{\mathrm{out}}^{(2)}} + \frac{H/2+1}{\lambda\mathbf{c}_{\mathrm{out}}^{(2)}}\right\}$$

is bounded from below by the probability of the event

$$\left\{\Upsilon_{\mathrm{low}}^{(1)}(O_{\mathrm{low}} \to \mathcal{L}_{2,\frac{H}{2}}) \geq \frac{H/2}{\mathbf{c}_{\mathrm{out}}^{(2)}}\right\} \cap \mathcal{E}_{\mathrm{subcr}}(H) \cap \left\{\Upsilon_{\mathrm{low}}^{(\lambda)}(\mathcal{L}_{2,\frac{H}{2}} \to W_{\mathrm{low}}) \geq \frac{H/2+1}{\lambda\mathbf{c}_{\mathrm{out}}^{(2)}}\right\}. \tag{3.11}$$

To see this fact more clearly, observe that the event $\mathcal{E}_{\mathrm{subcr}}(H)$ only depends on the vertices up to generation $H/2$, and that the function $\Upsilon_{\mathrm{low}}^{(1)}$ does not take into account the presence of seeds. Lemma 1.6 yields

$$\mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{low}}^{(1)}(O_{\mathrm{low}} \to \mathcal{L}_{2,\frac{H}{2}}) \geq \frac{H/2}{\mathbf{c}_{\mathrm{out}}^{(2)}}\right\} \cap \mathcal{E}_{\mathrm{subcr}}(H)\right] \geq \left[1 - e^{-\mathbf{c_1}^{(2)}H/2}\right]\left[1 - \left[2(1-\mu_2)\right]^{H/2}\right]. \tag{3.12}$$

Furthermore, by Lemma 1.6 with $\gamma = \lambda$,

$$\mathbb{P}_{\mu_2}\left[\Upsilon_{\mathrm{low}}^{(\lambda)}(\mathcal{L}_{2,\frac{H}{2}} \to W_{\mathrm{low}}) \geq \frac{H/2+1}{\lambda\mathbf{c}_{\mathrm{out}}^{(2)}}\right] \geq 1 - 2^{H/2}e^{-\mathbf{c_1}^{(2)}\left(H/2+1\right)}. \tag{3.13}$$

Thus, by (3.12) and (3.13) and using independence of the passage times before and after level $\mathcal{L}_{2,\frac{H}{2}}$,

$$\mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{low}}^{(1)}(O_{\mathrm{low}} \to \mathcal{L}_{2,\frac{H}{2}}) \geq \frac{H/2}{\mathbf{c}_{\mathrm{out}}^{(2)}}\right\} \cap \mathcal{E}_{\mathrm{subcr}}(H) \cap \left\{\Upsilon_{\mathrm{low}}^{(\lambda)}(\mathcal{L}_{2,\frac{H}{2}} \to W_{\mathrm{low}}) \geq \frac{\frac{H}{2}+1}{\lambda\mathbf{c}_{\mathrm{out}}^{(2)}}\right\}\right]$$

$$= \mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{low}}^{(1)}(O_{\mathrm{low}} \to \mathcal{L}_{2,\frac{H}{2}}) \geq \frac{H/2}{\mathbf{c}_{\mathrm{out}}^{(2)}}\right\} \cap \mathcal{E}_{\mathrm{subcr}}(H)\right]\mathbb{P}_{\mu_2}\left[\Upsilon_{\mathrm{low}}^{(\lambda)}(\mathcal{L}_{2,\frac{H}{2}} \to W_{\mathrm{low}}) \geq \frac{\frac{H}{2}+1}{\lambda\mathbf{c}_{\mathrm{out}}^{(2)}}\right]$$

$$\geq \left(1 - e^{-\mathbf{c_1}^{(2)}H/2} - \left(2(1-\mu_2)\right)^{H/2}\right)\left(1 - 2^{H/2}e^{-\mathbf{c_1}^{(2)}\left(H/2+1\right)}\right)$$

$$\geq 1 - 2^{H/2+1}e^{-\mathbf{c_1}^{(2)}H/2} - \left(2(1-\mu_2)\right)^{H/2},$$

as claimed. □

Furthermore, Lemma 1.6 applied to the path from W_{low} to B when $\gamma = \lambda < 1$ gives

$$\mathbb{P}_{\mu_2}\left[\Upsilon_{\text{low}}^{(\lambda)}(W_{\text{low}} \to B) \geq \frac{R}{\lambda \mathbf{c}_{\text{out}}^{(1)}}\right] \geq 1 - e^{-\mathbf{c_1}^{(1)}R}. \tag{3.14}$$

Now we are ready to proceed with the proof of the first part of Lemma 3.2.

Proof of Lemma 3.2(i). Since W_{low} is a cutpoint we can bound the sought probability as

$$\mathbb{P}_{\mu_2}\left[\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) \geq \frac{H/2}{\mathbf{c}_{\text{out}}^{(2)}} + \frac{H/2+1}{\lambda \mathbf{c}_{\text{out}}^{(2)}} + \frac{R}{\lambda \mathbf{c}_{\text{out}}^{(1)}}\right]$$

$$\geq \mathbb{P}_{\mu_2}\left[\mathfrak{B}_{\text{low}}^c(O_{\text{low}}, W_{\text{low}}) \cap \left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O_{\text{low}} \to W_{\text{low}}) \geq \frac{H/2}{\mathbf{c}_{\text{out}}^{(2)}} + \frac{H/2+1}{\lambda \mathbf{c}_{\text{out}}^{(2)}}\right\} \cap \right.$$

$$\left. \cap \left\{\Upsilon_{\text{low}}^{(\lambda)}(W_{\text{low}} \to B) \geq \frac{R}{\lambda \mathbf{c}_{\text{out}}^{(1)}}\right\}\right]$$

$$= \mathbb{P}_{\mu_2}\left[\mathfrak{B}_{\text{low}}^c(O_{\text{low}}, W_{\text{low}}) \cap \left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O_{\text{low}} \to W_{\text{low}}) \geq \frac{H/2}{\mathbf{c}_{\text{out}}^{(2)}} + \frac{H/2+1}{\lambda \mathbf{c}_{\text{out}}^{(2)}}\right\}\right] \times$$

$$\times \mathbb{P}_{\mu_2}\left[\Upsilon_{\text{low}}^{(\lambda)}(W_{\text{low}} \to B) \geq \frac{R}{\lambda \mathbf{c}_{\text{out}}^{(1)}}\right].$$

By Lemma 3.5 the above is bounded from below by

$$\left(1 - 2^{H/2+1}e^{-\mathbf{c_1}^{(2)}H/2} - \left(2(1-\mu_2)\right)^{H/2}\right) \mathbb{P}_{\mu_2}\left[\Upsilon_{\text{low}}^{(\lambda)}(W_{\text{low}} \to B) \geq \frac{R}{\lambda \mathbf{c}_{\text{out}}^{(1)}}\right]$$

$$\overset{(3.14)}{\geq} \left(1 - 2^{H/2+1}e^{-\mathbf{c_1}^{(2)}H/2} - \left(2(1-\mu_2)\right)^{H/2}\right)\left(1 - e^{-\mathbf{c_1}^{(1)}R}\right).$$

At this point we can choose R and H so large that, for all values of $\mu_2 \in \left(\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})}\right)$ one has $e^{-\mathbf{c_1}^{(1)}R} < \varepsilon$, $\left(2(1-\mu_2)\right)^{H/2} < \varepsilon$ and $2^{H/2+1}e^{-\mathbf{c_1}^{(2)}H/2} < \varepsilon$, hence the statement follows. □

3.2.2 Proof of Lemma 3.2(ii), "upper part" of G

Similarly to the previous section, consider the variable $\Upsilon_{\text{up}}^{(1,\lambda)}(O \to B)$, fix any pair of vertices x, y in the upper part of G and an initial density of seeds μ. For this part of the proof consider only the upper part of G, more precisely define the passage times $\{\bar{\bar{t}}_e^{(\gamma)}\}$ as follows. Set $\bar{\bar{t}}_e^{(\gamma)} := t_e^{(\gamma)}$ if the edge e is in the upper part of G, and set $\bar{\bar{t}}_e^{(\gamma)} := \infty$, if e is in the lower part. With respect to the passage times $\{\bar{\bar{t}}_e^{(\gamma)}\}_{e \in E(G)}$, recalling the definition of $\mathcal{E}_y$ from (2.3), define the event

$$\mathfrak{B}_{\text{up}}(x, y) := \mathcal{E}_y \cap \{\text{FPP}_1 \text{ started at } x \text{ infects } y \text{ before } \text{FPP}_\lambda\}. \tag{3.15}$$

By Proposition 2.1,

$$\mathbb{P}_{\mu_2}\left[\mathfrak{B}_{\text{up}}(O_{\text{up}}, W_{\text{up}})\right] \geq \left(1 - \mathbf{f}_D(\mu_2) - \boldsymbol{\eta}_D\right)(1 - \mu_2). \tag{3.16}$$

Now we want to show that, with high probability, when the initial density of seeds is $\mu_2 \in \left(\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})}\right)$ the time needed by FPP_1 started at O to reach B is smaller than that needed by FPP_λ.

Now recall the definition of $\mathfrak{C}$ from (3.1) and the bounds obtained in Lemma 3.5 and (3.14). At this point we are ready to prove the second bound of Lemma 3.2.

Proof of Lemma 3.2(ii). We can bound the sought probability by

$$
\begin{aligned}
&\mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{up}}^{(1,\lambda)}(O \to B) \leq 2\mathfrak{C} + \frac{L+1}{\mathbf{c}_{\mathrm{in}}^{(D)}}\right\} \cap \{W_{\mathrm{up}} \text{ is infected by } \mathrm{FPP}_1\}\right] \\
&\geq \mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{up}}^{(1,\lambda)}(O \to B) \leq 2\mathfrak{C} + \frac{L+1}{\mathbf{c}_{\mathrm{in}}^{(D)}}\right\} \cap \mathcal{E}_O \cap \mathcal{E}_{O_{\mathrm{up}}} \cap \mathfrak{B}_{\mathrm{up}}(O_{\mathrm{up}}, W_{\mathrm{up}}) \cap \mathcal{E}_{W_{\mathrm{up}}}\right].
\end{aligned}
\tag{3.17}
$$

Using the fact that O_{up} and W_{up} are cutpoints, under the event $\mathcal{E}_O \cap \mathcal{E}_{O_{\mathrm{up}}} \cap \mathfrak{B}_{\mathrm{up}}(O_{\mathrm{up}}, W_{\mathrm{up}})$ we have

$$
\Upsilon_{\mathrm{up}}^{(1,\lambda)}(O \to B) = \Upsilon_{\mathrm{up}}^{(1,\lambda)}(O \to O_{\mathrm{up}}) + \Upsilon_{\mathrm{up}}^{(1,\lambda)}(O_{\mathrm{up}} \to W_{\mathrm{up}}) + \Upsilon_{\mathrm{up}}^{(1,\lambda)}(W_{\mathrm{up}} \to B),
$$

hence the last line in (3.17) is bounded from below by

$$
\begin{aligned}
&\mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{up}}^{(1,\lambda)}(O \to O_{\mathrm{up}}) \leq \mathfrak{C}\right\} \cap \left\{\Upsilon_{\mathrm{up}}^{(1,\lambda)}(O_{\mathrm{up}} \to W_{\mathrm{up}}) \leq \frac{L+1}{\mathbf{c}_{\mathrm{in}}^{(D)}}\right\} \cap \right. \\
&\left. \cap \left\{\Upsilon_{\mathrm{up}}^{(1,\lambda)}(W_{\mathrm{up}} \to B) \leq \mathfrak{C}\right\} \cap \mathcal{E}_O \cap \mathcal{E}_{O_{\mathrm{up}}} \cap \mathfrak{B}_{\mathrm{up}}(O_{\mathrm{up}}, W_{\mathrm{up}}) \cap \mathcal{E}_{W_{\mathrm{up}}}\right].
\end{aligned}
\tag{3.18}
$$

Under the event $\mathcal{E}_O \cap \mathcal{E}_{O_{\mathrm{up}}} \cap \mathfrak{B}_{\mathrm{up}}(O_{\mathrm{up}}, W_{\mathrm{up}}) \cap \mathcal{E}_{W_{\mathrm{up}}}$ we have that

- $\Upsilon_{\mathrm{up}}^{(1,\lambda)}(O \to O_{\mathrm{up}})$ is equal to $\Upsilon_{\mathrm{up}}^{(1)}(O \to O_{\mathrm{up}})$;
- $\Upsilon_{\mathrm{up}}^{(1,\lambda)}(W_{\mathrm{up}} \to B)$ is equal to $\Upsilon_{\mathrm{up}}^{(1)}(W_{\mathrm{up}} \to B)$.

Furthermore, the distributions of such random variables only depend on the passage times of FPP_1, which are independent of the position of the seeds. Therefore, the quantity in (3.18) is bounded from below by

$$
\begin{aligned}
&\mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{up}}^{(1)}(O \to O_{\mathrm{up}}) \leq \mathfrak{C}\right\} \cap \mathcal{E}_O\right] \mathbb{P}_{\mu_2}\left[\Upsilon_{\mathrm{up}}^{(1)}(W_{\mathrm{up}} \to B) \leq \mathfrak{C}\right] \times \\
&\times \mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{up}}^{(1,\lambda)}(O_{\mathrm{up}} \to W_{\mathrm{up}}) \leq \frac{L+1}{\mathbf{c}_{\mathrm{in}}^{(D)}}\right\} \cap \mathcal{E}_{O_{\mathrm{up}}} \cap \mathfrak{B}_{\mathrm{up}}(O_{\mathrm{up}}, W_{\mathrm{up}}) \cap \mathcal{E}_{W_{\mathrm{up}}}\right] \\
&\stackrel{(3.1)}{\geq} (1-\varepsilon)^2(1-\mu_2)^2 \mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{up}}^{(1,\lambda)}(O_{\mathrm{up}} \to W_{\mathrm{up}}) \leq \frac{L+1}{\mathbf{c}_{\mathrm{in}}^{(D)}}\right\} \cap \mathcal{E}_{O_{\mathrm{up}}} \cap \mathfrak{B}_{\mathrm{up}}(O_{\mathrm{up}}, W_{\mathrm{up}})\right]
\end{aligned}
$$

For the last term of the last line, we proceed analogously to the proof of Lemma 3.1. Consider now a sufficiently small value of $\mathbf{c}_{\mathrm{in}}^{(D)}$. By using similar arguments to the ones mentioned in Section 3.1.1, in particular by Claim 3.4, one finds that there are constants $\overline{C}, \overline{\delta}, \overline{\beta} > 0$ that are independent of L and of $\mathbf{c}_{\mathrm{in}}^{(D)}$, so that

$$
\begin{aligned}
&\mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\mathrm{up}}^{(1,\lambda)}(O_{\mathrm{up}} \to W_{\mathrm{up}}) \leq \frac{L+1}{\mathbf{c}_{\mathrm{in}}^{(D)}}\right\} \cap \mathcal{E}_{O_{\mathrm{up}}} \cap \mathfrak{B}_{\mathrm{up}}(O_{\mathrm{up}}, W_{\mathrm{up}})\right] \\
&\geq (1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2))(1-\mu_2)^2 - e^{-\overline{\beta}H} - \overline{C}e^{-\overline{\delta}(L+1)/(2\mathbf{c}_{\mathrm{in}}^{(D)})}.
\end{aligned}
$$

Now it suffices to pick L, H large enough so that $e^{-\overline{\beta}H} + \overline{C}e^{-\overline{\delta}(L+1)/(2\mathbf{c}_{\mathrm{in}}^{(D)})} < \varepsilon$. By putting all these bounds together into (3.17) and using (3.16), we obtain the claim. $\qquad\square$

4 Proof of Proposition 2.3

We start with a technical lemma which we use to appropriately set H, R and L.

Lemma 4.1. *As before, we consider λ to be an arbitrarily small constant fixed beforehand; let D be large. For $\mathfrak{C}$ fixed as in (3.1), for any large integer R, there is a choice of values H, L, so that*

$$\frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}} + \frac{R}{\lambda \mathbf{c}_{\text{in}}^{(1)}} + \mathfrak{C} < \frac{L+1}{\mathbf{c}_{\text{out}}^{(D)}}, \tag{4.1}$$

and

$$\frac{H/2}{\mathbf{c}_{\text{out}}^{(2)}} + \frac{H/2+1}{\lambda \mathbf{c}_{\text{out}}^{(2)}} + \frac{R}{\lambda \mathbf{c}_{\text{out}}^{(1)}} > 2\mathfrak{C} + \frac{L+1}{\mathbf{c}_{\text{in}}^{(D)}} \tag{4.2}$$

hold simultaneously.

Proof. By solving (4.1) and (4.2) for $(L+1)$ and comparing the results, we see that it suffices to prove the following (stronger) condition,

$$\mathbf{c}_{\text{out}}^{(D)} \left(\frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}} + \frac{R}{\lambda \mathbf{c}_{\text{in}}^{(1)}} + 2\mathfrak{C} \right) + 1 \;<\; \mathbf{c}_{\text{in}}^{(D)} \left(\frac{H/2}{\mathbf{c}_{\text{out}}^{(2)}} + \frac{H/2}{\lambda \mathbf{c}_{\text{out}}^{(2)}} + \frac{R}{\lambda \mathbf{c}_{\text{out}}^{(1)}} - 2\mathfrak{C} \right). \tag{4.3}$$

In other words, (4.3) ensures that we can use (4.1) and (4.2) at the same time. The value "+1" on the left-hand side guarantees we can pick a large integer L depending on $\lambda, H, R, \mathfrak{C}, \mathbf{c}_{\text{in}}^{(1)}, \mathbf{c}_{\text{in}}^{(2)}, \mathbf{c}_{\text{in}}^{(D)}, \mathbf{c}_{\text{out}}^{(1)}, \mathbf{c}_{\text{out}}^{(2)}, \mathbf{c}_{\text{out}}^{(D)}$ that satisfies both (4.1) and (4.2). Relation (4.3) is equivalent to

$$\frac{\mathbf{c}_{\text{out}}^{(D)}}{\mathbf{c}_{\text{in}}^{(2)}} + \frac{R}{\lambda} \left(\frac{\mathbf{c}_{\text{out}}^{(D)}}{\mathbf{c}_{\text{in}}^{(1)}} - \frac{\mathbf{c}_{\text{in}}^{(D)}}{\mathbf{c}_{\text{out}}^{(1)}} \right) + 2\mathfrak{C} \left(\mathbf{c}_{\text{out}}^{(D)} + \mathbf{c}_{\text{in}}^{(D)} \right) + 1 < H \left(\frac{\mathbf{c}_{\text{in}}^{(D)}}{2\mathbf{c}_{\text{out}}^{(2)}} \left(1 + \frac{1}{\lambda} \right) - \frac{\mathbf{c}_{\text{out}}^{(D)}}{\mathbf{c}_{\text{in}}^{(2)}} \right). \tag{4.4}$$

Recall from Lemma 1.6 that $\mathbf{c}_{\text{out}}^{(d)} > 1 > \mathbf{c}_{\text{in}}^{(d)}$ for all d. Suppose now that R is a large value chosen in a convenient way. Then we need to check that the coefficient of H is positive, that is,

$$\frac{\mathbf{c}_{\text{in}}^{(D)}}{2\mathbf{c}_{\text{out}}^{(2)}} \left(1 + \frac{1}{\lambda} \right) - \frac{\mathbf{c}_{\text{out}}^{(D)}}{\mathbf{c}_{\text{in}}^{(2)}} > 0,$$

which is guaranteed for all values $0 < \lambda < \lambda_0$ where we have set

$$\lambda_0 := \frac{\mathbf{c}_{\text{in}}^{(D)} \mathbf{c}_{\text{in}}^{(2)}}{2\mathbf{c}_{\text{out}}^{(2)} \mathbf{c}_{\text{out}}^{(D)} - \mathbf{c}_{\text{in}}^{(D)} \mathbf{c}_{\text{in}}^{(2)}} > 0.$$

The proof is now completed since, after fixing R, it suffices to choose H large enough so that (4.4) is satisfied. $\qquad\square$

Now we proceed by showing that for all $\mu_1 \in \left(\mu_{1,\text{low}}^{(\text{crit})}, \mu_{1,\text{high}}^{(\text{crit})} \right)$, it is likely that B is infected by the process spreading in the lower part of the tile, whereas for all $\mu_2 \in \left(\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})} \right)$ it is likely to be the opposite. We recall that, as in Lemma 3.2, the event $\{W_{\text{up}}$ is infected by $\text{FPP}_1\}$ is restricted to the upper part of the tile, so we do not consider the possibility that W_{up} is infected by B.

Lemma 4.2. *Let G be constructed as in Section 2.1. Fix a large integer D and a small enough $\lambda > 0$. Then there are values $0 < \mu_{1,low}^{(crit)} < \mu_{1,high}^{(crit)} < 1/2 < \mu_{2,low}^{(crit)} < \mu_{2,high}^{(crit)} < 1$ satisfying*

$$D \left(1 - \mu_{2,high}^{(crit)} \right) > 1 \quad \text{and} \quad D^2 \left(1 - \mu_{2,low}^{(crit)} \right)^2 \left(\mu_{2,high}^{(crit)} \right)^{D-1} < 1,$$

as well as values H, L and R satisfying Lemma 4.1 together with (3.6) and (3.14), so that

$$\mathbb{P}_{\mu_1}\left(\Upsilon_{low}^{(1,\lambda)}(O \to B) < \Upsilon_{up}^{(1,\lambda)}(O \to B)\right) \geq 1 - \boldsymbol{\eta}_2 - \mathbf{f}_2(\mu_1) - 5\varepsilon - 4\mu_1,$$

$$\mathbb{P}_{\mu_2}\left(\left\{\Upsilon_{low}^{(1,\lambda)}(O \to B) > \Upsilon_{up}^{(1,\lambda)}(O \to B)\right\} \cap \{W_{up} \text{ is infected by } \mathrm{FPP}_1\}\right)$$
$$\geq (1 - 5\varepsilon)(1 - \mu_2)^2\left[(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2))(1 - \mu_2)^2 - \varepsilon\right],$$

where $\boldsymbol{\eta}_2$ and $\boldsymbol{\eta}_D$ are obtained as in Proposition 2.1.

Proof. In order to bound the first formula we use Lemma 3.1 together with relation (4.1) and the fact that FPPHE spreads in the upper and lower part of the tile independently (until B gets infected). We find that

$$\mathbb{P}_{\mu_1}\left(\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) < \Upsilon_{\text{up}}^{(1,\lambda)}(O \to B)\right)$$

$$\geq \mathbb{P}_{\mu_1}\left[\left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) \leq \mathfrak{C} + \frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}} + \frac{R}{\lambda\mathbf{c}_{\text{in}}^{(1)}}\right\} \cap \left\{\Upsilon_{\text{up}}^{(1,\lambda)}(O \to B) \geq \frac{L+1}{\mathbf{c}_{\text{out}}^{(D)}}\right\}\right]$$

$$\geq \left(1-\varepsilon\right)^3\left(1-\mu_1\right)^2\left[(1 - \mathbf{f}_2(\mu_1) - \boldsymbol{\eta}_2)(1-\mu_1)^2 - \varepsilon\right](1-\varepsilon),$$

implying the first part of the claim. Analogously, for the second formula of the statement we use Lemma 3.2, relation (4.2) and again independence of FPPHE in the upper and lower part, finding

$$\mathbb{P}_{\mu_2}\left(\left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) > \Upsilon_{\text{up}}^{(1,\lambda)}(O \to B)\right\} \cap \{W_{\text{up}} \text{ is infected by } \mathrm{FPP}_1\}\right)$$

$$\geq \mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) \geq \frac{H/2}{\mathbf{c}_{\text{out}}^{(2)}} + \frac{H/2+1}{\lambda\mathbf{c}_{\text{out}}^{(2)}} + \frac{R}{\lambda\mathbf{c}_{\text{out}}^{(1)}}\right\} \cap \left\{\Upsilon_{\text{up}}^{(1,\lambda)}(O \to B) \leq 2\mathfrak{C} + \frac{L+1}{\mathbf{c}_{\text{in}}^{(D)}}\right\} \cap\right.$$

$$\left. \cap \{W_{\text{up}} \text{ is infected by } \mathrm{FPP}_1\}\right]$$

$$\geq (1 - 3\varepsilon)(1 - \varepsilon)^2(1 - \mu_2)^2\left[\left(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2)\right)(1 - \mu_2)^2 - \varepsilon\right],$$

which implies the claim. $\qquad\square$

Now recall that λ is small enough and note that, given $\mu_{2,\text{high}}^{(\text{crit})}$, the function $\mathbf{f}_D(\mu_{2,\text{high}}^{(\text{crit})})$ is decreasing in D. Moreover, by Proposition 2.1(iii), the quantity $\left(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_{2,\text{high}}^{(\text{crit})})\right)(1 - \mu_{2,\text{high}}^{(\text{crit})})^2$ can be made positive by choosing D large enough. Hence, let D be so large that, for all $\mu_2 \in \left(\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})}\right)$ one has $\left(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2)\right)(1 - \mu_2)^2 > 0$. Hence we assume that ε is so that, uniformly in $\mu_2 \in \left(\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})}\right)$,

$$\varepsilon < \min\left\{\frac{1}{700}, (1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2))(1 - \mu_2)^2\right\}. \tag{4.5}$$

Since $\left(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2)\right)(1 - \mu_2)^2 - \varepsilon > 0$, define

$$\Phi := \sup_{\mu_2 \in \left(\mu_{2,\text{low}}^{(\text{crit})}, \mu_{2,\text{high}}^{(\text{crit})}\right)}\left\lceil \frac{2D^3}{\frac{99}{100}\left[\left(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2)\right)(1 - \mu_2)^2 - \varepsilon\right]}\right\rceil; \tag{4.6}$$

this quantity will be left constant throughout. Now we are ready to prove Proposition 2.3. Here we shall use that the critical value $\mu_{1,\text{high}}^{(\text{crit})}$ is small enough; more precisely, with the notation above, let $\overline{\mu_1} := \sup\{\mu \in (0,1) : 7\varepsilon + 5\mu + \boldsymbol{\eta}_2 + \mathbf{f}_2(\mu) < 1/\Phi\}$. Then we need

$$0 < \mu_{1,\text{high}}^{(\text{crit})} \leq \overline{\mu_1}. \tag{4.7}$$

Proof of Proposition 2.3. Start by considering $\mu_1 \in \left(\mu_{1,\text{low}}^{(\text{crit})}, \mu_{1,\text{high}}^{(\text{crit})}\right)$ and recall the definition of $\mathcal{E}_y$ from (2.3). Then, since R was so that $(1 - \mu_1)^R < \varepsilon$ (cf. (3.6) and (3.14)).

$$\mathbb{P}_{\mu_1}\left(B \text{ is infected by } \text{FPP}_1\right)$$
$$\leq \mathbb{P}_{\mu_1}\left(\left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) > \Upsilon_{\text{up}}^{(1,\lambda)}(O \to B)\right\} \cap \mathcal{E}_B\right)$$
$$+ \mathbb{P}_{\mu_1}\left(\text{there is no seed on the path to } B\right)$$
$$\overset{\text{Lem. 4.2}}{\leq} \left(1 - (1 - 5\varepsilon - 4\mu_1 - \boldsymbol{\eta}_2 - \mathbf{f}_2(\mu_1))\right)(1 - \mu_1) + (1 - \mu_1)^{R-1}$$
$$< 5\varepsilon + 5\mu_1 + \boldsymbol{\eta}_2 + \mathbf{f}_2(\mu_1) + 2\varepsilon.$$

Using the fact that ε is small (cf. (4.5)), as well as λ, together with (4.7), then, for all $\mu_1 \in \left(\mu_{1,\text{low}}^{(\text{crit})}, \mu_{1,\text{high}}^{(\text{crit})}\right)$ we have $7\varepsilon + 5\mu_1 + \boldsymbol{\eta}_2 + \mathbf{f}_2(\mu_1) < 1/\Phi$, concluding the first part. Analogously,

$$\mathbb{P}_{\mu_2}\left(B \text{ is infected by } \text{FPP}_1\right)$$
$$\geq \mathbb{P}_{\mu_2}\left[\left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O \to B) > \Upsilon_{\text{up}}^{(1,\lambda)}(O \to B)\right\} \cap \mathcal{E}_B \cap \{W_{\text{up}} \text{ is infected by } \text{FPP}_1\}\right]$$
$$\overset{\text{Lem. 4.2}}{\geq} (1 - 5\varepsilon)\left[\left(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2)\right)(1 - \mu_2)^2 - \varepsilon\right](1 - \mu_2)^3$$
$$\overset{(2.7)}{>} (1 - 5\varepsilon)\left[\left(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2)\right)(1 - \mu_2)^2 - \varepsilon\right]\frac{1}{D^3}.$$

By our assumption on ε we have that $1 - 5\varepsilon > 99/100$. By (4.6),

$$(1 - 5\varepsilon)\left[\left(1 - \boldsymbol{\eta}_D - \mathbf{f}_D(\mu_2)\right)(1 - \mu_2)^2 - \varepsilon\right]\frac{1}{D^3} \geq \frac{2}{\Phi} > \frac{1}{\Phi},$$

as claimed. $\qquad\square$

5 Analysis of FPPHE on a finite tree: proof of Proposition 2.1

In this section we analyze FPPHE on the tree $\overline{\mathbb{T}}_d^m$ and prove Proposition 2.1. Since the proof is articulated into several intermediate results, we provide a sketch. The idea is the following. If FPP_λ occupies W, then we can find an index $i \in \{1, \ldots, m\}$ corresponding to a generation, such that there is a *seed* at generation i that is activated by FPP_1 and which belongs to the geodesic that reaches W. Note that i is the generation of the first seed in the geodesic between the root and W.

Once we fix an arbitrarily small constant $\varepsilon_0 > 0$, "Step 1" and "Step 2" below deal with the cases $i \leq \lceil(1 - \varepsilon_0)m\rceil$ and $i > \lceil(1 - \varepsilon_0)m\rceil$ respectively. From now on, in order to avoid cumbersome notation, instead of writing $\lceil(1 - \varepsilon_0)m\rceil$ we only write $(1 - \varepsilon_0)m$.

Step 1. Fix an arbitrarily small constant $\varepsilon_0 > 0$. We start by showing that the probability that there is a seed at some generation $i \leq (1 - \varepsilon_0)m$ that originates a FPP_λ process which will reach level m much quicker than expected, decreases exponentially fast in m. The underlying idea is that if $i \leq (1 - \varepsilon_0)m$, then, in order to reach generation m, FPP_λ has to cross at least $\varepsilon_0 m$ edges, and since m is large, this is unlikely. This step then will show that if FPP_λ does occupy W, then it is more likely that it did so from a seed within distance $\varepsilon_0 m$ of W.

Step 2. Subsequently, we take care of what happens at levels i for $(1 - \varepsilon_0)m \leq i \leq m$. Since the seed is close to W, we cannot argue that the path traversed by FPP_λ from the seed to W has a large delay with respect to what FPP_1 would have had. Note that with probability at least μ the fastest path to W will have a seed in its last vertex before reaching W. Therefore, in order to give an upper bound to the probability that W is occupied by FPP_λ, we will consider what happened on the tree before level $(1 - \varepsilon_0)m$. For simplicity, set $j := m - i$; here is the idea behind this procedure.

(i) We show that for all $j \in \{1, \ldots, \lfloor \varepsilon_0 m \rfloor\}$, it is likely that *all* vertices of type FPP_1 at level j are infected within some time proportional to j.

(ii) Fix $j \in \{1, \ldots, \lfloor \varepsilon_0 m \rfloor\}$. We consider the (unique, if any) geodesic of vertices infected by FPP_1 started at each vertex at level j described in (i), and ended at the first infected vertex at level $m - j$. Note that up to level $m - j - 1$ such line consists of vertices infected by FPP_1, whereas the one at level $m - j$ can be either of type FPP_1 or a seed.

(iii) We use a relation between first passage percolation and a continuous-time branching random walk, as well as concentration results for the latter, to show that all vertices at level $m - j$ described in (ii) are infected within a time that is very concentrated around its expected value (which is of order $m - j$).

(iv) Next, we show that the probability that among the first-infected particles described in (iii) there is a non-seed originating a FPP_1 process that infects level m before any other process does is high.

Before proceeding we make a brief reflection about the steps above. Part (i) is used to get several independent vertices at level j to be occupied by FPP_1, for which the arrival time at level j is within distance of order j of one another. Parts (ii) and (iii) give that each such vertex will give rise to a geodesic path to level $m - j$, whose passage time is very concentrated around a given value of order $m - j$. Call the collection of such geodesic paths $\mathcal{P}$. Since $m - j$ is much bigger than j, the paths in $\mathcal{P}$ all arrive at level $m - j$ within a small time interval. Part (iv) gives that, since $\mathcal{P}$ reaches several vertices at level $m - j$, it is unlikely that too many of them are occupied by seeds. Similarly, since $\mathcal{P}$ is large, it is unlikely that a seed from level $m - j$ that is activated by a non-geodesic path (that is, by a path not in $\mathcal{P}$) arrives at level m faster than the paths in $\mathcal{P}$.

Recall that $d \geq 2$ is the degree of the tree, and consider μ as in the statement of the proposition. In this proof we will deal with the first passage times restricted to $\overline{\mathbb{T}}_d^m$, thus to make this more explicit we shall write $\Upsilon_{\overline{\mathbb{T}}_d^m}^{(1,\lambda)}$. Recall that m is chosen to be large. For $d \geq 2$ and $0 \leq k \leq m$ let

$$\mathcal{L}_k = \mathcal{L}_{d,k} := \{x \in V(\overline{\mathbb{T}}_d^m) \ : \ d_{\overline{\mathbb{T}}_d^m}(o, x) = k\},$$

that is, the set of all vertices of $\overline{\mathbb{T}}_d^m$ at distance k from the root.

Auxiliary facts For $d \geq 2$ and μ as in (2.7), for all levels $i \in \{1, \ldots, m\}$ define the events

$$\begin{aligned} B_i := \{ &\exists \text{ seed in } \mathcal{L}_i \text{ which is infected by } \mathrm{FPP}_1 \text{ and} \\ &\text{gives rise to a } \mathrm{FPP}_\lambda \text{ process that infects } \mathcal{L}_m \text{ first} \}. \end{aligned} \tag{5.1}$$

These are "bad" events, thus we need to show that the probability that there is a level i for which B_i occurs is small. To shorten the notation, sometimes we shall write "vertex x is of type 1 (resp. λ)" as an equivalent formulation of "vertex x is of type FPP_1 (resp. FPP_λ)".

A fundamental tool is provided by [1, Theorem 3] which we recall below. Let $\gamma \in \{1, \lambda\}$. For $d \geq 2$ and each $i, n \in \{1, \ldots, m\}$ with $i < n$, and for every fixed $x \in \mathcal{L}_i$, x of type γ, set

$$M_{i,n}^{(\gamma)}(x) := \inf_{y \in \mathcal{L}_n, \ y \text{ of type } \gamma, \ y \text{ descendant of } x} \{\Upsilon_{\overline{\mathbb{T}}_d^m}^{(\gamma)}(x \to y)\}, \tag{5.2}$$

with the standard convention that $\inf\{\emptyset\} := +\infty$, and thus define

$$M_{i,n}^{(\gamma)} := \inf_{x \in \mathcal{L}_i, \ x \text{ of type } \gamma} M_{i,n}^{(\gamma)}(x).$$

In words, $M_{i,n}^{(\gamma)}$ is the time needed by the quickest among the infections started at any $x \in \mathcal{L}_i$ to reach $\mathcal{L}_n$ when the infection spreads at rate γ. Now we consider a Galton-Watson branching process (where individuals have i.i.d. offspring) whose branching distribution is given by $\text{Bin}(d, (1 - \mu))$. Define $\mathscr{S}$ to be the event

$$\mathscr{S} := \{\text{branching process with distribution given by Bin}(d, (1 - \mu)) \text{ survives}\}.$$

By (2.2),

$$\mathbb{P}_\mu(\mathscr{S}) = 1 - \mathbf{f}_d(\mu). \tag{5.3}$$

Description of a branching random walk in continuous time (BRWCT) We will couple the process FPP_1 inside $\overline{\mathbb{T}}_d^m$ with initial density of seeds μ with a BRWCT on $\mathbb{R}$ as follows.

- Each vertex of type FPP_1 corresponds to an individual in the BRWCT. The origin o plays the role of the starter of the progeny, and for $j \in \{0, \ldots, m - 1\}$ we say that individual $u \in \mathcal{L}_j$ is the *father* of $x \in \mathcal{L}_{j+1}$ if and only if **(a)** u and x are not seeds, and **(b)** $\{u, x\} \in E(\overline{\mathbb{T}}_d^m)$. For $x, y \in V(\overline{\mathbb{T}}_d^m)$, we say that y is a *descendant* of x if there is a line of fathers from y to x.

- The distribution governing the underlying random walk has (independent) increments distributed as exponentials of parameter $\gamma = 1$.

- The offspring distribution is binomial with parameters d and $1 - \mu$.

- The above defines the BRWCT up to generation m; from generation $m + 1$ onwards, we let the BRWCT evolve independently of FPP_1 on $\overline{\mathbb{T}}_d^m$.

Whenever we will write "GWT" we will mean a Galton-Watson Tree having this distribution. Also, from the coupling above, we see that the event that the BRWCT survives does not depend on the passage times of FPP_1, but just on the location of the seeds.

Similarly, we couple the evolution of FPP_λ on $\mathbb{T}_d$ from each seed activated by FPP_1 as independent BRWCT on $\mathbb{R}$ by setting the offspring distribution to be the constant d, and the distribution governing the underlying random walk to have (independent) increments as exponentials of parameter $\gamma = \lambda$. Clearly, individuals here correspond to vertices of type FPP_λ, and each progeny is started by a seed. The theorem below is the concentration result we shall use. The measure $\mathbb{P}$ below is defined on the space of the trajectories of BRWCT on $\mathbb{R}$, where the branching mechanism and the step size are governed by the prescribed distributions.

Theorem 5.1 ([1, Theorem 3, adapted to our setting]). *Consider a supercritical BRWCT on $\mathbb{R}$ with branching distribution $\text{Bin}(d, (1 - \mu))$ and step size $X \sim \text{Exp}(\gamma)$ (for a fixed $\gamma > 0$) coupled with FPPHE as specified above. Then, for each fixed $x \in \mathcal{L}_i$ there exist constants $C_x > 0, \delta_x > 0$ which depend on γ, such that for all $\alpha > 0$ and all $n \geq 1$*

$$\mathbb{P}\left(\left|M_{i,n}^{(\gamma)}(x) - \mathbb{E}\left(M_{i,n}^{(\gamma)}(x) \mid \mathscr{S}\right)\right| > \alpha \mid \mathscr{S}\right) \leq C_x e^{-\delta_x \alpha},$$

where $M_{i,n}^{(\gamma)}(x)$ was defined in (5.2).

Remark 5.2. It follows from the definition that, whenever the branching distribution is $\text{Bin}(d, (1 - \mu))$, the measure $\mathbb{P}$ above can be coupled with $\mathbb{P}_\mu$. From now on, for every measurable event E define $\widetilde{\mathbb{P}}_\mu[E] := \mathbb{P}_\mu[E \mid \mathscr{S}]$, that is, the probability measure $\widetilde{\mathbb{P}}_\mu$ coincides with $\mathbb{P}_\mu$ conditioned on the event of survival $\mathscr{S}$. Consistently, the corresponding expectation will be denoted by $\widetilde{\mathbb{E}}_\mu$.

5.1 Proof of Claim 3.4

Before continuing with the proof of Proposition 2.1, in this short section we use the auxiliary results stated above to establish Claim 3.4, in particular we will only write the proof for the lower part, as it is completely analogous for the upper part.

Proof of Claim 3.4. Start by defining the following auxiliary event

$$\overline{\mathscr{E}} := \{\exists \text{ a path } \pi \text{ from } O_{\text{low}} \text{ to } W_{\text{low}} \text{ that is free of seeds and s.t. } \Upsilon_{\text{low}}^{(1)}(\pi) \leq \frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}}\}.$$

We can bound the sought probability from below as follows,

$$\mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{\text{low}}} \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}}) \cap \left\{\Upsilon_{\text{low}}^{(1,\lambda)}(O_{\text{low}} \to W_{\text{low}}) \leq \frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}}\right\}\right]$$

$$\geq \mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{\text{low}}} \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}}) \cap \overline{\mathscr{E}} \cap \mathscr{S}\right]$$

$$= \mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{\text{low}}} \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}}) \cap \mathscr{S}\right] - \mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{\text{low}}} \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}}) \cap \overline{\mathscr{E}}^c \cap \mathscr{S}\right].$$

Now let $\mathscr{S}_H^{\text{not}}$ denote the event that the GWT survives for $H+1$ generations but it does not survive forever. At this point, the above expression is at least

$$\mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{\text{low}}} \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}}) \cap \mathscr{S}\right] - \mathbb{P}_{\mu_1}\left[\overline{\mathscr{E}}^c \cap \mathscr{S}\right]$$

$$\geq \mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{\text{low}}} \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}}) \cap \mathscr{S}\right] - \mathbb{P}_{\mu_1}\left[\overline{\mathscr{E}}^c \mid \mathscr{S}\right]$$

$$= \mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{\text{low}}} \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}})\right] - \mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{\text{low}}} \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}}) \cap \mathscr{S}^c\right] - \widetilde{\mathbb{P}}_{\mu_1}\left[\overline{\mathscr{E}}^c\right]$$

$$\geq \mathbb{P}_{\mu_1}\left[\mathcal{E}_{O_{\text{low}}} \cap \mathfrak{B}_{\text{low}}(O_{\text{low}}, W_{\text{low}})\right] - \mathbb{P}_{\mu_1}\left[\mathscr{S}_H^{\text{not}}\right] - \widetilde{\mathbb{P}}_{\mu_1}\left[\overline{\mathscr{E}}^c\right].$$

By Proposition 2.1 the first term in the last line is bounded from below by $(1 - \mathbf{f}_2(\mu_1) - \boldsymbol{\eta}_2)(1 - \mu_1)^2$. Regarding the second term, it suffices to recall that a super-critical GWT conditioned on dying out has the same distribution of a *sub-critical* Galton-Watson tree (cf. [2, Section I.12]). Thus, the event $\mathscr{S}_H^{\text{not}}$ is so that there is a constant $\beta > 0$ dependent on μ_1 but independent of H, so that

$$\mathbb{P}_{\mu_1}\left[\mathscr{S}_H^{\text{not}}\right] \leq e^{-\beta H}.$$

In order to bound the third term we proceed as follows. Since the event $\Upsilon_{\text{low}}^{(1)}(\pi) > \frac{H+1}{\mathbf{c}_{\text{in}}^{(2)}}$ must hold for all paths of length $H+1$ starting at O_{low} that are free of seeds, it holds for the one with minimum passage time. By construction, it is easy to see that there is a constant $C' > 0$ (independent of H) so that

$$\widetilde{\mathbb{E}}_\mu\left[M_{0,H+1}^{(1)}\right] \leq C'(H+1).$$

Furthermore, we shall assume that the constant $\mathbf{c}_{\text{in}}^{(2)}$ in the statement is so that $1/\mathbf{c}_{\text{in}}^{(2)} \geq 2C'$. Thus,

$$\widetilde{\mathbb{P}}_{\mu_1}\left[\overline{\mathscr{E}}^c\right] \leq \widetilde{\mathbb{P}}\left(\left|M_{0,H+1}^{(1)}(O_{\text{low}}) - \widetilde{\mathbb{E}}\left(M_{0,H+1}^{(1)}(O_{\text{low}})\right)\right| > \left(\frac{1}{\mathbf{c}_{\text{in}}^{(2)}} - C'\right)(H+1)\right)$$

$$\leq \widetilde{\mathbb{P}}\left(\left|M_{0,H+1}^{(1)}(O_{\text{low}}) - \widetilde{\mathbb{E}}\left(M_{0,H+1}^{(1)}(O_{\text{low}})\right)\right| > \frac{H+1}{2\mathbf{c}_{\text{in}}^{(2)}}\right)$$

$$\overset{\text{Thm 5.1}}{\leq} C_{O_{\text{low}}} e^{-\delta_{O_{\text{low}}}(H+1)/(2\mathbf{c}_{\text{in}}^{(2)})},$$

concluding the proof. $\square$

5.2 Step 1, taking care of levels up to $(1 - \varepsilon_0)m$

We start by showing that it is unlikely that FPP_λ originated from any seed located far away from $\mathcal{L}_m$ reaches level m too quickly. To achieve this we show that the probability of the occurrence of a bad event B_i (cf. (5.1)) at some level much earlier than m decreases exponentially fast.

Now we proceed with our analysis as follows. Fix an arbitrarily small constant $\varepsilon_0 > 0$; for any fixed $i \leq (1 - \varepsilon_0)m$ the event B_i necessarily implies that there must be a vertex $\mathbf{x}_i \in \mathcal{L}_i$ hosting a seed (and whose tree-geodesic to the origin is free of seeds), which starts a FPP_λ process such that

$$M_{i,m}^{(\lambda)}(\mathbf{x}_i) < M_{0,m}^{(1)} - \Upsilon_{\overline{\mathbb{T}}_d^m}^{(1)}(o \to \mathbf{x}_i).$$

For simplicity, let $\mathcal{M}(\mathbf{x}_i)$ denote this event, then

$$B_i \subseteq \left\{ \exists \text{ a seed } \mathbf{x}_i \in \mathcal{L}_i \ : \ \mathcal{M}(\mathbf{x}_i) \text{ is realized} \right\}. \tag{5.4}$$

We shall proceed with bounding the probability of the RHS of the above. We shall use the obvious fact that $\mathcal{M}(\mathbf{x}_i) \Rightarrow M_{i,m}^{(\lambda)}(\mathbf{x}_i) < M_{0,m}^{(1)}$. Fix a large constant $c_1 > 0$, and define the following auxiliary event

$$\mathcal{A}_{(1)} := \left\{ M_{0,m}^{(1)} < (1 + c_1)\widetilde{\mathbb{E}}_\mu \left[M_{0,m}^{(1)} \right] \right\}.$$

The value $M_{0,m}^{(1)}$ is a minimum over all paths starting at the origin as FPP_1 and arriving to $\mathcal{L}_m$ as FPP_1. Thus it is easy to see that there is a constant $C' > 0$ (independent of m) so that

$$\widetilde{\mathbb{E}}_\mu \left[M_{0,m}^{(1)} \right] \leq C'm. \tag{5.5}$$

Similarly, there must be a constant $\mathcal{C}^* > 0$ independent of λ, such that for all $0 \leq i < n \leq m$

$$\mathbb{E}_\mu \left[M_{i,n}^{(\lambda)} \right] \geq \frac{\mathcal{C}^*}{\lambda}(n - i). \tag{5.6}$$

We shall proceed by showing that $\widetilde{\mathbb{P}}_\mu \left[\exists i \leq (1 - \varepsilon_0)m \text{ s.t. } \exists \text{ seed } \mathbf{x}_i \ : \ \mathcal{M}(\mathbf{x}_i) \text{ occurs} \right]$ decreases exponentially fast in m.

Claim 5.3. *Given $\varepsilon_0 > 0$, for all $\lambda > 0$ small enough, there is a constant $\boldsymbol{\alpha_1} = \boldsymbol{\alpha_1}(\lambda, \varepsilon_0) > 0$ independent of m such that*

$$\widetilde{\mathbb{P}}_\mu \left[\exists \, i \leq (1 - \varepsilon_0)m \text{ s.t. } \exists \text{ seed } \mathbf{x}_i \ : \ \mathcal{M}(\mathbf{x}_i) \cap \mathcal{A}_{(1)} \right] \leq e^{-\boldsymbol{\alpha_1} m}.$$

Proof. By definition, the event $\mathcal{M}(\mathbf{x}_i) \cap \mathcal{A}_{(1)}$ implies

$$M_{i,m}^{(\lambda)}(\mathbf{x}_i) < M_{0,m}^{(1)} \leq (1 + c_1)\widetilde{\mathbb{E}}_\mu \left[M_{0,m}^{(1)} \right]. \tag{5.7}$$

Thus, relation (5.7) together with (5.5) imply

$$M_{i,m}^{(\lambda)}(\mathbf{x}_i) \leq (1 + c_1)C'm = (1 + c_1)C'(m - i) + (1 + c_1)C'i$$

$$\overset{i \leq (1-\varepsilon_0)m}{\leq} (1 + c_1)C'(m - i) + (1 + c_1)C'(1 - \varepsilon_0)m$$

$$= (1 + c_1)C'(m - i) + \frac{(1 + c_1)C'}{\varepsilon_0}\varepsilon_0(1 - \varepsilon_0)m \overset{\varepsilon_0 m \leq m-i}{<} C''(m - i),$$

for some constant $C'' > 0$ which depends on ε_0 but is independent of i and m. By (5.6),

$$\widetilde{\mathbb{E}}_\mu \left[M_{i,m}^{(\lambda)}(\mathbf{x}_i) \right] \geq \frac{\mathcal{C}^*}{\lambda}(m - i). \tag{5.8}$$

By putting these facts together we find that

$$M_{i,m}^{(\lambda)}(\mathbf{x}_i) < C''(m-i) < \frac{C''}{\mathcal{C}^*} \lambda \widetilde{\mathbb{E}}_\mu \left[M_{i,m}^{(\lambda)}(\mathbf{x}_i) \right]. \tag{5.9}$$

Since λ is small enough, relations (5.8) and (5.9) ensure that there is a constant $c_2 > 0$ such that uniformly over all $i \le (1-\varepsilon_0)m$,

$$M_{i,m}^{(\lambda)}(\mathbf{x}_i) < \widetilde{\mathbb{E}}_\mu \left[M_{i,m}^{(\lambda)}(\mathbf{x}_i) \right] - c_2 \frac{1}{\lambda}(m-i).$$

In particular, we have that

$$\widetilde{\mathbb{P}}_\mu \left[\mathcal{M}(\mathbf{x}_i) \cap \mathcal{A}_{(1)} \right] \le d^i \max_{\mathbf{x}_i \in \mathcal{L}_i} \widetilde{\mathbb{P}}_\mu \left(\left| M_{i,m}^{(\lambda)}(\mathbf{x}_i) - \widetilde{\mathbb{E}}_\mu \left[M_{i,m}^{(\lambda)}(\mathbf{x}_i) \right] \right| > \frac{c_2}{\lambda}(m-i) \right), \tag{5.10}$$

where in the last step we used a union bound over all vertices of $\mathcal{L}_i$. Thus,

$$\widetilde{\mathbb{P}}_\mu \left(\exists i \le \lceil (1-\varepsilon_0)m \rceil \text{ s.t. } \exists \text{ seed } \mathbf{x}_i \in \mathcal{L}_i \ : \ M_{i,m}^{(\lambda)}(\mathbf{x}_i) < \widetilde{\mathbb{E}}_\mu \left[M_{i,m}^{(\lambda)}(\mathbf{x}_i) \right] - \frac{c_2}{\lambda}(m-i) \right)$$

$$\overset{(5.10)}{\le} \sum_{i=1}^{\lceil (1-\varepsilon_0)m \rceil} d^i \max_{\mathbf{x}_i \in \mathcal{L}_i} \widetilde{\mathbb{P}}_\mu \left(\left| M_{i,m}^{(\lambda)}(\mathbf{x}_i) - \widetilde{\mathbb{E}}_\mu \left[M_{i,m}^{(\lambda)}(\mathbf{x}_i) \right] \right| > \frac{c_2}{\lambda}(m-i) \right)$$

$$\overset{\text{Thm. 5.1}}{\le} \sum_{i=1}^{\lceil (1-\varepsilon_0)m \rceil} C d^i \, e^{-\delta(m-i)c_2/\lambda} \le e^{-\boldsymbol{\alpha_1} m},$$

where the very last inequality follows from the facts that λ is small enough (in order to have $d^m \, e^{-\delta\varepsilon_0 m c_2/\lambda} < 1$ to ensure summability) and $i \le \lceil (1-\varepsilon_0)m \rceil$. □

The following result is of the same fashion.

Claim 5.4. *Given $\varepsilon_0 > 0$, there is a constant $\boldsymbol{\alpha_2} = \boldsymbol{\alpha_2}(\varepsilon_0) > 0$ independent of m so that*

$$\widetilde{\mathbb{P}}_\mu \left[\mathcal{A}_{(1)}^c \right] \le e^{-\boldsymbol{\alpha_2} m}.$$

Proof. By definition of $\mathcal{A}_{(1)}$, we obtain

$$\widetilde{\mathbb{P}}_\mu \left[\mathcal{A}_{(1)}^c \right] \le \widetilde{\mathbb{P}}_\mu \left[\left| M_{0,m}^{(1)} - \widetilde{\mathbb{E}}_\mu \left[M_{0,m}^{(1)} \right] \right| \ge c_1 \widetilde{\mathbb{E}}_\mu \left[M_{0,m}^{(1)} \right] \right] \overset{\text{Thm 5.1}}{\le} C e^{-\delta c_1 \widetilde{\mathbb{E}}_\mu \left[M_{0,m}^{(1)} \right]}.$$

The quantity $M_{0,m}^{(1)}$ is a minimum over all paths starting at the origin as FPP_1 and arriving to $\mathcal{L}_m$ as FPP_1, therefore it is easy to see that there is a constant $\mathcal{C} > 0$ such that

$$\widetilde{\mathbb{E}}_\mu \left[M_{0,m}^{(1)} \right] \ge \mathcal{C}m,$$

implying the statement for a suitable choice of $\boldsymbol{\alpha_2} > 0$. □

Lemma 5.5. *Given $\varepsilon_0 > 0$, for all small enough $\lambda > 0$ there exists a constant $\delta' = \delta'(\lambda, \varepsilon_0) > 0$ independent of m, such that*

$$\widetilde{\mathbb{P}}_\mu \left(\bigcap_{i=1}^{(1-\varepsilon_0)m} B_i^c \right) \ge 1 - e^{-\delta' m}.$$

Proof. For all $i \le (1-\varepsilon_0)m$, for each seed $\mathbf{x}_i \in \mathcal{L}_i$ we have

$$\widetilde{\mathbb{P}}_\mu \left[\mathcal{M}(\mathbf{x}_i) \right] = \widetilde{\mathbb{P}}_\mu \left[\mathcal{M}(\mathbf{x}_i) \cap \mathcal{A}_{(1)} \right] + \widetilde{\mathbb{P}}_\mu \left[\mathcal{M}(\mathbf{x}_i) \cap \mathcal{A}_{(1)}^c \right]. \tag{5.11}$$

The proof is now a simple consequence of the previous two claims, in fact by (5.4)

$$\widetilde{\mathbb{P}}_\mu \left[\bigcup_{i=1}^{(1-\varepsilon_0)m} B_i \right] \leq \widetilde{\mathbb{P}}_\mu \left[\bigcup_{i \leq (1-\varepsilon_0)m} \{\exists \text{ a seed } \mathbf{x}_i \in \mathcal{L}_i \ : \ \mathcal{M}(\mathbf{x}_i) \text{ is realized}\} \right]$$

$$\overset{\text{(5.11) and Claims 5.3, 5.4}}{\leq} e^{-\alpha_1 m} + e^{-\alpha_2 m}.$$

Since α_1, α_2 are independent of m, the statement follows. $\quad\square$

5.3 Step 2, subsequent levels

In this section we will bound the probability of the event

$$\bigcup_{i=(1-\varepsilon_0)m+1}^{m} B_i.$$

Since this part is more delicate, we need to control the probabilities of the B_i's in a more refined way. We proceed by using the BRWCT representation that we defined before. Recall that by "GWT" we mean the Galton-Watson tree with offspring distribution given by $\text{Bin}(d, 1-\mu)$. For any individual u in the GWT define

$$\mathbf{t}_u := \text{birth time of } u, \tag{5.12}$$

where the root o was born at time 0, that is $\mathbf{t}_o = 0$. From the definition of BRWCT we have

$$\mathbf{t}_u \overset{\text{law}}{=} \Upsilon^{(1,\lambda)}_{\overline{\mathbb{T}}_d^m}(o \to u).$$

To make the proof more readable, from now on we set

$$\ell := \lfloor \varepsilon_0 m \rfloor. \tag{5.13}$$

Now for each $j \in \{1, \dots, \ell\}$ define

$$\mathcal{N}_j := \{\text{individuals at } j\text{-th generation of the GWT}\}, \tag{5.14}$$

and let $N_j := |\mathcal{N}_j|$. Consider a large constant $C_1 > 0$ (which will be specified later) and set

$$\mathcal{K}_j := \{u \in \mathcal{N}_j \ : \ \mathbf{t}_u \leq C_1 j\}, \tag{5.15}$$

and let $K_j := |\mathcal{K}_j|$. In words, $\mathcal{K}_j$ is the subset of $\mathcal{N}_j$ of individuals born within time $C_1 j$.

Recall the technical conditions on μ from (2.7) and fix two small constants $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$. With this notation in mind, we define a sequence of "good" events, namely for $j \in \{1, \dots, \ell\}$ set

$$A_j := \left\{ K_j \geq \varepsilon_2 \big(d(1-\mu)\big)^{(1-\varepsilon_1)j} \right\}. \tag{5.16}$$

In order to continue, we need an auxiliary result. Recall that $\widetilde{\mathbb{P}}_\mu(\cdot) = \mathbb{P}(\cdot \mid \mathscr{S})$ is the probability measure conditioned on the event of survival of GWT.

Lemma 5.6. *Assume that d and μ satisfy (2.7) and $\varepsilon_1, \varepsilon_2$ are as above. Then there are two positive constants $c_3 = c_3(\varepsilon_1) < \frac{\varepsilon_1}{2} \log(d(1-\mu))$ and $C_2 = C_2(\varepsilon_1, \varepsilon_2)$ such that for all $1 \leq j \leq \ell$,*

$$\widetilde{\mathbb{P}}_\mu \left(\varepsilon_2 \big(d(1-\mu)\big)^{(1-\varepsilon_1)j} \leq N_j \leq \frac{1}{\varepsilon_2} \big(d(1-\mu)\big)^{(1+\varepsilon_1)j} \right) \geq 1 - C_2 e^{-c_3 j}. \tag{5.17}$$

In particular, C_2 can be made arbitrarily small by choosing ε_2 small enough.

Proof. For all $j \in \{1, \dots, \ell\}$ we have

$$\widetilde{\mathbb{P}}_\mu \left(\frac{N_j}{\left(d(1-\mu)\right)^j} \geq \frac{1}{\varepsilon_2} \big(d(1-\mu)\big)^{\varepsilon_1 j} \right) \leq \frac{\mathbb{P}_\mu \left(\frac{N_j}{\left(d(1-\mu)\right)^j} \geq \frac{1}{\varepsilon_2} \big(d(1-\mu)\big)^{\varepsilon_1 j} \right)}{\mathbb{P}_\mu(\mathscr{S})}$$

$$\leq \frac{1}{\mathbb{P}_\mu(\mathscr{S})} \frac{\mathbb{E}_\mu \left[\frac{N_j}{\left(d(1-\mu)\right)^j} \right] \varepsilon_2}{\big(d(1-\mu)\big)^{\varepsilon_1 j}} = \frac{1}{\mathbb{P}_\mu(\mathscr{S})} \frac{\varepsilon_2}{\big(d(1-\mu)\big)^{\varepsilon_1 j}} \leq \frac{C_2 e^{-c_3 j}}{2}.$$

The above inequalities follow from Markov's inequality, from the fact that $\mathbb{E}_\mu N_j = \big(d(1-\mu)\big)^j$ and that $c_3 < \varepsilon_1 \log(d(1-\mu))$.

To bound the reverse inequality we re-write the sought quantity to apply Markov's inequality,

$$\widetilde{\mathbb{P}}_\mu \left(\frac{N_j}{\left(d(1-\mu)\right)^j} \leq \varepsilon_2 \big(d(1-\mu)\big)^{-\varepsilon_1 j} \right) = \widetilde{\mathbb{P}}_\mu \left(\frac{\left(d(1-\mu)\right)^j}{N_j} \geq \frac{1}{\varepsilon_2} \big(d(1-\mu)\big)^{\varepsilon_1 j} \right)$$

$$\leq \widetilde{\mathbb{E}} \left[\frac{\left(d(1-\mu)\right)^j}{N_j} \right] \varepsilon_2 \big(d(1-\mu)\big)^{-\varepsilon_1 j} \leq \exp \left\{ \left| \log \widetilde{\mathbb{E}} \left[\frac{\left(d(1-\mu)\right)^j}{N_j} \right] \right| \right\} \varepsilon_2 \big(d(1-\mu)\big)^{-\varepsilon_1 j}.$$

$$(5.18)$$

Hence our next aim is to bound the expression $\exp \left\{ \left| \log \widetilde{\mathbb{E}} \left[\frac{\left(d(1-\mu)\right)^j}{N_j} \right] \right| \right\}$. To do so, we employ [16, Corollary 1], which says that under the so-called $L \log L$ condition (which is satisfied here, since the offspring distribution of GWT is binomial) we have that

$$\lim_n \frac{1}{n} \log \widetilde{\mathbb{E}} \left[\frac{1}{N_n} \right] = \max \left\{ \log(p_1), -\log(\mathbb{E}(N_1)) \right\},$$

where $p_1 := \mathbb{P}(N_1 = 1)$. In our specific case this translates into

$$\lim_n \frac{1}{n} \log \widetilde{\mathbb{E}} \left[\frac{1}{N_n} \right] = \max \left\{ \log \big(d(1-\mu)\mu^{d-1}\big), -\log(d(1-\mu)) \right\}.$$

Given that d and μ must satisfy (2.7), we obtain that $\log \big(d(1-\mu)\mu^{d-1}\big) < -\log(d(1-\mu))$, implying that for any fixed $\varepsilon' > 0$, for all n sufficiently large we have

$$\left| \frac{1}{n} \log \widetilde{\mathbb{E}} \left[\frac{1}{N_n} \right] + \log(\mathbb{E}N_1) \right| < \varepsilon'.$$

From now on, for convenience, we set $\varepsilon' := \frac{\varepsilon_1}{2} \log(\mathbb{E}N_1)$. Thus, for all n large enough, the above can be rewritten as

$$\exp \left\{ \left| \log \widetilde{\mathbb{E}} \left[\frac{1}{N_n} (\mathbb{E}N_1)^n \right] \right| \right\} < \exp\{\varepsilon' n\} \stackrel{\varepsilon' = \varepsilon_1 \log(\mathbb{E}N_1)/2}{=} (d(1-\mu))^{\varepsilon_1 n/2}.$$

Hence, there is a constant $C' > 0$ such that uniformly for all $j \geq 1$ we have

$$\exp \left\{ \left| \log \widetilde{\mathbb{E}} \left[\frac{1}{N_j} (\mathbb{E}N_1)^j \right] \right| \right\} < C'(d(1-\mu))^{\varepsilon_1 j/2}. \qquad (5.19)$$

Thus, replacing this quantity in (5.18) we obtain

$$\widetilde{\mathbb{P}}_\mu \left[\frac{N_j}{\left(d(1-\mu)\right)^j} \leq \varepsilon_2 \big(d(1-\mu)\big)^{-\varepsilon_1 j} \right] \stackrel{(5.18),\ (5.19)}{\leq} C'(d(1-\mu))^{\varepsilon_1 j/2} \varepsilon_2 \big(d(1-\mu)\big)^{-\varepsilon_1 j},$$

which equals $C'\varepsilon_2\big(d(1-\mu)\big)^{-\varepsilon_1 j/2}$. Since $c_3 < \frac{\varepsilon_1}{2}\log(d(1-\mu))$ then, for all $j \in \{1, 2, \ldots, \ell\}$,

$$\widetilde{\mathbb{P}}_\mu\left(\frac{N_j}{\big(d(1-\mu)\big)^j} \le \varepsilon_2\big(d(1-\mu)\big)^{-\varepsilon_1 j}\right) \le C'\varepsilon_2\big(d(1-\mu)\big)^{-\varepsilon_1 j/2} \le \frac{C_2 e^{-c_3 j}}{2},$$

finishing the proof. In particular, C_2 can be made arbitrarily small by picking ε_2 small enough. $\qquad\square$

Here we recall an auxiliary result that can be found e.g. in [13], which will allow us to proceed. The following statement and Lemma 1.6 are strongly related. In the result below we denote by $\mathbb{P}$ the probability measure defined on all possible realizations of the sums of independent random variables with the prescribed distribution.

Theorem 5.7. *[13, Thm 5.1(i),(iii)] Let $X := \sum_{k=1}^K X_k$, where $\{X_k\}_k$ is a collection of independent random variables such that $X_k \sim Exp(a_k)$ for some constants $a_k > 0$ and let $a_* := \min_k a_k$. Then the following bounds hold.*

(i) *For all fixed $\delta > 0$ we have $\mathbb{P}\big(X \ge (1+\delta)\mathbb{E}X\big) \le \frac{1}{1+\delta}\exp\big\{-a_*\mathbb{E}X(\delta - \ln(1+\delta))\big\}$.*

(ii) *For all fixed $0 < \delta < 1$ we have $\mathbb{P}\big(X \le (1-\delta)\mathbb{E}X\big) \le \exp\big\{-a_*\mathbb{E}X(-\delta - \ln(1-\delta))\big\}$.*

Claim 5.8. *For every $j \in \{1, \ldots, \ell\}$*

$$\widetilde{\mathbb{P}}_\mu\left[\textit{The entire generation } \mathcal{N}_j \textit{ is born by time } C_1 j\right] \ge 1 - \frac{1}{C_1}e^{-jC_1/2}.$$

Proof. By Theorem 5.7, with high probability all particles of generation j of GWT are born within time $C_1 j$, for some C_1 large enough. Clearly, for all $j \in \{1, \ldots, \ell\}$,

$$\widetilde{\mathbb{P}}_\mu\left[\text{Entire generation } \mathcal{N}_j \text{ is born by time } C_1 j\right]$$
$$\ge 1 - \widetilde{\mathbb{P}}_\mu\big[\exists \text{ individual in } \mathcal{N}_j \text{ born after time } C_1 j\big].$$

Now, let $\{X_k\}_k$ denote an i.i.d. sequence of exponential random variables with parameter 1. By a union bound we obtain that the above probability is bounded from below by

$$1 - d^j\,\mathbb{P}\left[\sum_{k=1}^j X_k \ge C_1 j\right] \overset{\text{Thm 5.7(i)}}{\ge} 1 - d^j \frac{1}{C_1}e^{-j\big(C_1 - 1 - \ln(C_1)\big)}$$
$$= 1 - \frac{1}{C_1}e^{-j\big(C_1 - 1 - \ln(C_1) - \ln(d)\big)} \overset{C_1 \text{ large}}{\ge} 1 - \frac{1}{C_1}e^{-jC_1/2}.$$

This concludes the proof. $\qquad\square$

The next result proves that for all $j \in \{1, \ldots, \ell\}$, it is likely that at level j there are *a lot* of vertices of type FPP_1 which have been infected by time $C_1 j$.

Lemma 5.9. *Given any small constant $\varepsilon_3 > 0$, we can choose $\varepsilon_1 > 0$ in (5.16) such that*

$$\widetilde{\mathbb{P}}_\mu\big(\exists j \in \{1, \ldots, \ell\} \textit{ s.t. } A_j^c \textit{ occurs}\big) \le \varepsilon_3.$$

Proof. The probability $\widetilde{\mathbb{P}}_\mu\big(K_j \ge \varepsilon_2(d(1-\mu))^{(1-\varepsilon_1)j}\big)$ is bounded from below by

$$\widetilde{\mathbb{P}}_\mu\left[\{\text{Entire } \mathcal{N}_j \text{ born by time } C_1 j\} \cap \left\{N_j \ge \varepsilon_2(d(1-\mu))^{(1-\varepsilon_1)j}\right\}\right]$$
$$\ge 1 - \widetilde{\mathbb{P}}_\mu[\{\text{Entire } \mathcal{N}_j \text{ born by time } C_1 j\}^c] - \mathbb{P}\left[N_j < \varepsilon_2(d(1-\mu))^{(1-\varepsilon_1)j}\right].$$

The last term of the sum is taken care of by Lemma 5.6, while for the middle term we rely on Claim 5.8. Using these facts we obtain that for every $j \in \{1, \ldots, \ell\}$

$$\widetilde{\mathbb{P}}_\mu(A_j^c) \le \frac{1}{C_1}e^{-jC_1/2} + C_2 e^{-c_3 j}.$$

Hence it suffices to pick $C_1 = C_1(d, \varepsilon_3)$ large enough and $\varepsilon_2 = \varepsilon_2(\varepsilon_3)$, $\varepsilon_1 = \varepsilon_1(\varepsilon_3)$ so that

$$\sum_{j=1}^{\ell} \frac{1}{C_1} e^{-jC_1/2} \leq \frac{\varepsilon_3}{2} \quad \text{and} \quad \sum_{j=1}^{\ell} C_2 e^{-c_3 j} \leq \frac{\varepsilon_3}{2}, \tag{5.20}$$

in order to obtain $\sum_{j=1}^{\ell} \widetilde{\mathbb{P}}_\mu(A_j^c) \leq \varepsilon_3$, which is the claim. $\qquad\square$

For a B_i to be realized, there must be a first seed (at level $\mathcal{L}_i$) initiating the FPP_λ process which infects level $\mathcal{L}_m$ before any other process does. From now on we set

$\mathbf{s} :=$ first seed initiating the FPP_λ process which infects the first vertex on $\mathcal{L}_m$. (5.21)

In other words, $\cup_{i=m-\ell}^m B_i$ is realized only if there is a seed $\mathbf{s} \in \cup_{j=0}^\ell \mathcal{L}_{m-j}$ such that

$$\min_{y \in \mathcal{L}_m} \Upsilon_{\overline{\mathbb{T}}_d^m}^{(1,\lambda)}(o \to y) = \Upsilon_{\overline{\mathbb{T}}_d^m}^{(1)}(o \to \mathbf{s}) + \min_{y \in \mathcal{L}_m} \Upsilon_{\overline{\mathbb{T}}_d^m}^{(\lambda)}(\mathbf{s} \to y).$$

Remark 5.10. Note that $j = 0$ means that the first infected vertex on $\mathcal{L}_m$ is a seed, hence from now on we shall consider $j \geq 1$.

Recall that for any vertex $y \in V(\overline{\mathbb{T}}_d^m) \setminus \{o\}$ the *father* of y is the unique vertex y^- on the (unique) geodesic line connecting o to y such that

$$d_{\mathbb{T}_d}(o, y) = d_{\mathbb{T}_d}(o, y^-) + 1.$$

For all $j \in \{1, \ldots, \ell\}$ fixed, for every vertex $v \in \mathcal{N}_j$ with birth time $\mathbf{t}_v \leq C_1 j$, let the FPP_1 process started at v continue until it reaches level $\mathcal{L}_{m-j}$ for the first time. More precisely, define the set

$$\mathcal{S}(m-j) := \{y \in \mathcal{L}_{m-j} \ : \ y^- \in \mathcal{N}_{m-j-1}\}, \tag{5.22}$$

that is, $\mathcal{S}(m-j)$ is the set of vertices $y \in \mathcal{L}_{m-j}$ whose father is of type FPP_1. Thus, for every $v \in \mathcal{N}_j$ we look for a vertex $h(v) \in \mathcal{S}(m-j)$, that corresponds to the first vertex that has been infected by FPP_1 and has v as ancestor. (Note that $h(v)$ can be of type FPP_1 or a seed.) Furthermore, it will be convenient to set

$$h(\mathcal{N}_j) := \cup_{v \in \mathcal{N}_j} h(v), \tag{5.23}$$

keeping in mind that $h(\mathcal{N}_j) \subseteq \mathcal{S}(m-j)$; cf. Figure 5.

The next result shows that with high probability all vertices $h(v)$ are born relatively close (in time) to each other.

Lemma 5.11. *For any large enough constant C_3', we can take $C_3 = C_3(C_3')$ so large that*

$$\widetilde{\mathbb{P}}_\mu \left[\sup_{v, v' \in \mathcal{N}_j} \left| \mathbf{t}_{h(v)} - \mathbf{t}_{h(v')} \right| \leq (2C_3 + C_1)j \right] \geq 1 - 2e^{-C_3' j} - \frac{1}{C_1} e^{-jC_1/2}.$$

Proof. For simplicity, denote

$$\tau(v, h(v)) := \mathbf{t}_{h(v)} - \mathbf{t}_v, \tag{5.24}$$

that is the time needed by FPP_1 started at v to advance up to infecting $\mathcal{S}(m-j)$ for the first time. Note that by definition we have that the variables $\{\tau(v, h(v))\}_{v \in \mathcal{N}_j}$ are independent. Furthermore, since the process advances as FPP_1 for $m - 2j$ steps, we also have that $\{\tau(v, h(v))\}_{v \in \mathcal{N}_j}$ are identically distributed. In particular,

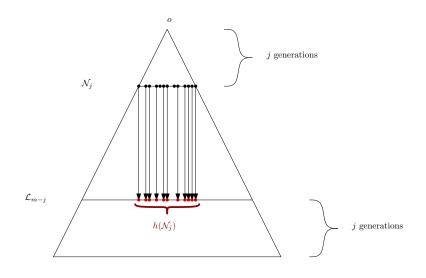$$\text{for all } v \in \mathcal{N}_j, \quad \tau(v, h(v)) \stackrel{\text{law}}{=} M_{0, m-2j}^{(1)}. \tag{5.25}$$

Figure 5: Sketch of the construction of $h(\mathcal{N}_j)$.

We can exploit this fact by fixing a large constant $C_3 > 0$ and applying Theorem 5.1, obtaining

$$\widetilde{\mathbb{P}}_\mu \left[\left| \tau\big(v, h(v)\big) - \widetilde{\mathbb{E}}_\mu \big[\tau\big(v, h(v)\big) \big] \right| > C_3 j \right] \le C e^{-\delta C_3 j} \le e^{-C_3' j}, \tag{5.26}$$

for a suitable (and still large, when C_3 is large enough) constant $C_3' > 0$. Furthermore,

$$\left| \mathbf{t}_{h(v)} - \mathbf{t}_{h(v')} \right| = \left| \mathbf{t}_{h(v)} - \mathbf{t}_v + \mathbf{t}_v - \mathbf{t}_{h(v')} + \mathbf{t}_{v'} - \mathbf{t}_{v'} \right| \le \left| \mathbf{t}_{h(v)} - \mathbf{t}_v - (\mathbf{t}_{h(v')} - \mathbf{t}_{v'}) \right| + \left| \mathbf{t}_v - \mathbf{t}_{v'} \right|$$
$$\overset{(5.24)}{=} \left| \tau\big(v, h(v)\big) - \tau\big(v', h(v')\big) \right| + \left| \mathbf{t}_v - \mathbf{t}_{v'} \right|.$$

From (5.25) it follows that $\widetilde{\mathbb{E}}_\mu \big[\tau\big(v, h(v)\big) \big] = \widetilde{\mathbb{E}}_\mu \big[\tau\big(v', h(v')\big) \big]$, for all $v, v' \in \mathcal{N}_j$. Thus, the above gives

$$\left| \mathbf{t}_{h(v)} - \mathbf{t}_{h(v')} \right| \le \left| \tau\big(v, h(v)\big) - \tau\big(v', h(v')\big) \right| + \left| \mathbf{t}_v - \mathbf{t}_{v'} \right|$$
$$\le \left| \tau\big(v, h(v)\big) - \widetilde{\mathbb{E}}_\mu \big[\tau\big(v, h(v)\big) \big] \right| + \left| \tau\big(v', h(v')\big) - \widetilde{\mathbb{E}}_\mu \big[\tau\big(v', h(v')\big) \big] \right| + \left| \mathbf{t}_v - \mathbf{t}_{v'} \right|.$$

By putting this fact together with Claim 5.8 and (5.26) we obtain that for every $v, v' \in \mathcal{N}_j$,

$$\widetilde{\mathbb{P}}_\mu \left[\left| \mathbf{t}_{h(v)} - \mathbf{t}_{h(v')} \right| \le \big(2C_3 + C_1 \big) j \right]$$
$$\ge \widetilde{\mathbb{P}}_\mu \Big[\left| \tau\big(v, h(v)\big) - \widetilde{\mathbb{E}}_\mu \big[\tau\big(v, h(v)\big) \big] \right| +$$
$$+ \left| \tau\big(v', h(v')\big) - \widetilde{\mathbb{E}}_\mu \big[\tau\big(v', h(v')\big) \big] \right| + \left| \mathbf{t}_v - \mathbf{t}_{v'} \right| \le \big(2C_3 + C_1 \big) j \Big]$$
$$\ge 1 - 2 e^{-C_3' j} - \frac{1}{C_1} e^{-j C_1 / 2},$$

yielding the claim. $\qquad\square$

Using the notation in Lemma 5.11, define the constant

$$C_4 := 2C_3 + C_1, \tag{5.27}$$

furthermore, for all $j \in \{1, \dots, \ell\}$ define the following event

$$\mathcal{Z}_j := \left\{ \sup_{v, v' \in \mathcal{N}_j} \left| \mathbf{t}_{h(v)} - \mathbf{t}_{h(v')} \right| \le C_4 j \right\}. \tag{5.28}$$

For $\gamma \in \{\lambda, 1\}$ recall the definition of $M_{i,n}^{(\gamma)}(x)$ from (5.2). For all $j \in \{1, 2, \ldots, \ell\}$, whenever a vertex $\mathbf{x}_{m-j} \in \mathcal{L}_{m-j}$ hosts a seed, all variables $M_{m-j,m}^{(\lambda)}(\mathbf{x}_{m-j})$ are independent and have the same distribution. Thus introduce $M_\lambda(j)$ as an independent copy of $M_{m-j,m}^{(\lambda)}(\mathbf{x}_{m-j})$, that is,

$$M_\lambda(j) \sim M_{m-j,m}^{(\lambda)}(\mathbf{x}_{m-j}).$$

Similarly, let $M_1(j)$ be an independent copy of $M_{m-j,m}^{(1)}(x_{m-j})$ for some (any) vertex $x_{m-j} \in \mathcal{N}_{m-j}$, that is,

$$M_1(j) \sim M_{m-j,m}^{(1)}(x_{m-j}).$$

(If the corresponding GWT is so that the subtree started at x_{m-j} dies out before reaching level m, then we set $M_{m-j,m}^{(1)}(x_{m-j}) := +\infty$.) As a consequence of relation (5.6) we see that for any large fixed constant $\mathbf{K}$, if λ is small enough, we obtain that for all $j \in \{1, 2, \ldots, \ell\}$

$$\mathbb{E}_\mu [M_\lambda(j)] > \mathbf{K}\, \widetilde{\mathbb{E}}_\mu [M_1(j)]. \tag{5.29}$$

Recall $\mathcal{C}^*$ from (5.6). Since λ is small enough,

$$\frac{\mathcal{C}^*}{\lambda} > 2C_4,$$

with C_4 satisfying (5.27). Subsequently, for $j \in \{1, \ldots, \ell\}$ define $\mathbf{t}_{\text{first}}(m-j) := \min_{x \in \mathcal{L}_{m-j}} \mathbf{t}_x$, and consider the following (random) time interval

$$\mathbf{T}_\lambda(j) := \left[\mathbf{t}_{\text{first}}(m-j),\ \mathbf{t}_{\text{first}}(m-j) + \frac{\mathcal{C}^* j}{2\lambda} \right].$$

Our next aim is to use these tools to define auxiliary events that, for each $j \in \{1, \ldots, \ell\}$, contain B_{m-j} and subsequently bound (from above) their probability.

Recall the definition of $\mathbf{s}$ from (5.21), and the definition of $\mathcal{S}(m-j)$ from (5.22). We now look at all particles in $\mathcal{S}(m-j)$ and order them by their birth time and subsequently, for all $j \in \{1, \ldots, \ell\}$ and for all $1 \leq i \leq |\mathcal{S}(m-j)|$ we set

$$E_i(j) := \big\{\text{the } i\text{-th particle coincides with } \mathbf{s} \text{ and initiates a } \mathrm{FPP}_\lambda$$
$$\text{process reaching } \mathcal{L}_m \text{ within the time interval } \mathbf{T}_\lambda(j)\big\}. \tag{5.30}$$

Then define

$$E(j) := \bigcup_{i=1}^{|\mathcal{S}(m-j)|} E_i(j). \tag{5.31}$$

Recall the definition of $h(\mathcal{N}_j)$ from (5.23). This time consider $i \in h(\mathcal{N}_j)$, and set

$$U_i(j) := \big\{\text{particle } i \text{ is } not \text{ a seed, and the } \mathrm{FPP}_1 \text{ process}$$
$$\text{started from it infects } \mathcal{L}_m \text{ within } \mathbf{T}_\lambda(j)\big\}. \tag{5.32}$$

Subsequently, define

$$U(j) := \bigcup_{i \in h(\mathcal{N}_j)} U_i(j). \tag{5.33}$$

Note that the events $E(j)$ and $U(j)$ are *not* mutually exclusive. Each event $E(j)$ can be seen as follows: for some $i \in \{1, \ldots, |\mathcal{S}(m-j)|\}$ the i-th particle is a seed $\mathbf{s}$ initiating a FPP_λ process which is atypically quick. On the other hand, $U(j)$ states that in the set of all individuals in $h(\mathcal{N}_j) \subset \mathcal{S}(m-j) \subset \mathcal{L}_{m-j}$, there is one that starts a FPP_1 process

that infects level $\mathcal{L}_m$ within a reasonably large time. (Recall that for small values of λ we have (5.29).) Now for all $1 \leq i \leq |\mathcal{S}(m-j)|$ we define

$$F_i^{(j)} := E_i(j) \cap \Big\{ \forall r \in \{1, \ldots, i-1\}, \text{ the process originated from the } r\text{-th particle}$$

$$\text{reaches } \mathcal{L}_m \text{ after the process originated by the } i\text{-th particle} \Big\}. \tag{5.34}$$

Roughly speaking, the event $F_i^{(j)}$ implies that the i-th particle of $\mathcal{S}(m-j)$ is a seed initiating a FPP_λ process that is quicker than what we expect, and at the same time all previously born particles originate processes (FPP_1 or FPP_λ) that are slower than that. A graphical representation can be seen in Figure 6, where the horizontal segments represent time-intervals.
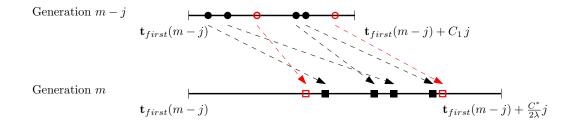

Figure 6: Empty (red) dots are seeds and full (black) dots are non-seeds. Here the first individual to be born at the m-th level is a descendant of a seed at level $m-j$. Dashed arrows represent the quickest line of descendants for each particle, and the first descendant is represented as a square. Empty (red) squares are descendants of seeds and full (black) squares are descendants of non-seeds. Note that among the descendants of the non-seeds we only consider those of type FPP_1, as further seeds are treated at the corresponding level.

We proceed by showing that $E_i(j)$ is unlikely.

Claim 5.12. *There is large constant $\overline{\mathbf{C}}_\lambda > 0$ such that for all fixed $j \in \{1, \ldots, \ell\}$ and uniformly over all $1 \leq i \leq |\mathcal{S}(m-j)|$,*

$$\mathbb{P}_\mu(E_i(j)) \leq e^{-\overline{\mathbf{C}}_\lambda j}. \tag{5.35}$$

Furthermore, $\overline{\mathbf{C}}_\lambda$ can be made large by picking λ sufficiently small.

Proof. Recall relation (5.6); to make the notation less cumbersome, for this proof we set

$$\mathbf{C}_\lambda := \frac{\mathcal{C}^*}{\lambda}.$$

From Theorem 5.1 we see that for all $j \in \{1, \ldots, \ell\}$ and uniformly in $1 \leq i \leq |\mathcal{S}(m-j)|$,

$$\mathbb{P}_\mu(E_i(j)) \leq \mathbb{P}_\mu \left[|M_\lambda(j) - \mathbb{E}(M_\lambda(j))| > \frac{\mathbf{C}_\lambda j}{2} \right] \leq C e^{-\delta \mathbf{C}_\lambda j/2}.$$

Thus, there is a large constant $\overline{\mathbf{C}}_\lambda = \overline{\mathbf{C}}_\lambda(C, \delta, \mathbf{C}_\lambda)$ so that uniformly over $i \in \{1, \ldots, |\mathcal{S}(m-j)|\}$

$$\mathbb{P}_\mu(E_i(j)) \leq C e^{-\delta \mathbf{C}_\lambda j/2} \leq e^{-\overline{\mathbf{C}}_\lambda j}, \tag{5.36}$$

proving the statement. $\qquad\square$

The next result is where we need the assumption that either d is arbitrary and μ is small, or $\mu > 1/2$ is fixed and d is *large enough*. In the first case we start by defining μ' so small that

$$\mathbb{P}_{\mu'}(\mathscr{S}) > 1/2, \tag{5.37}$$

and consider only $\mu < \mu'$.

Lemma 5.13. *For any fixed constant* $\mathbf{C} > 0$ *the following occurs. Suppose that* $d \geq 2$ *is arbitrary and* μ *is small enough, or* $\mu \in (1/2, 1)$ *is fixed and* d *is large enough and satisfies* (2.7). *Then for all fixed* $j \in \{1, \ldots, \ell\}$,

$$\widetilde{\mathbb{P}}_\mu(U^c(j)) \leq e^{-\mathbf{C}j}. \tag{5.38}$$

Proof. Start by fixing the constant $\mathbf{C} > 0$. Now we want to show that $|h(\mathcal{N}_j)|$ is likely to be large, so that $\widetilde{\mathbb{P}}_\mu(U^c(j))$ satisfies the sought bound. For this proof denote

$$h_j := |h(\mathcal{N}_j)|.$$

Using Lemma 5.6, we see that with probability at least $1 - C_2 e^{-c_3 j}$, the variable h_j stochastically dominates a Binomial random variable

$$Y_j \sim \text{Bin}\left(\max\left\{1, \left\lfloor \varepsilon_2 \big(d(1-\mu)\big)^{(1-\varepsilon_1)j} \right\rfloor\right\}, \mathbb{P}_\mu(\mathscr{S})\right).$$

In particular, since we can choose ε_2 as small as we please, we can fix it so small that

$$C_2 e^{-c_3 j} \leq \frac{e^{-\mathbf{C}j}}{3}. \tag{5.39}$$

We proceed with a case distinction, following the two conditions in the statement.

Case (i): $d \geq 2$ is arbitrary and μ is small enough. Note that the case $Y_j \sim \text{Bin}(1, \mathbb{P}_\mu(\mathscr{S}))$ is a Bernoulli random variable. Recall that $\mu < \mu'$ (cf. (5.37)), then for convenience we pick a large constant $\tilde{C} = \tilde{C}(\mathbf{C})$ so that

$$\frac{1}{\mathbb{P}_\mu(\mathscr{S})} e^{-\tilde{C}j/4} \leq \frac{e^{-\mathbf{C}j}}{3}. \tag{5.40}$$

First we prove the claim for all j such that

$$\tilde{C}j \leq \frac{1}{2}\mathbb{E}_\mu Y_j = \frac{1}{2}\left\lfloor \varepsilon_2 \big(d(1-\mu)\big)^{(1-\varepsilon_1)j} \right\rfloor \mathbb{P}_\mu(\mathscr{S}). \tag{5.41}$$

Here we simply apply a Chernoff bound and the definition of stochastic domination,

$$\widetilde{\mathbb{P}}_\mu\left(h_j \leq \tilde{C}j\right) \leq \frac{\mathbb{P}_\mu\left(Y_j \leq \frac{1}{2}\mathbb{E}_\mu Y_j\right)}{\mathbb{P}_\mu(\mathscr{S})} \leq \frac{1}{\mathbb{P}_\mu(\mathscr{S})} e^{-\mathbb{E}_\mu Y_j/8} \overset{(5.41)}{\leq} \frac{1}{\mathbb{P}_\mu(\mathscr{S})} e^{-\tilde{C}j/4}.$$

At this point our choice of $\tilde{C}$ as in (5.40) gives us that for all j satisfying (5.41),

$$\widetilde{\mathbb{P}}_\mu\left(h_j \leq \tilde{C}j\right) \leq \frac{e^{-\mathbf{C}j}}{3}. \tag{5.42}$$

Whenever $\tilde{C}j > \frac{1}{2}\mathbb{E}_\mu Y_j$ we perform a different approach, using Chebyshev's inequality. Let

$$j_0 = j_0(\varepsilon_1, \varepsilon_2, d) := \max\left\{j \in \{1, \ldots, \ell\} \; : \; \tilde{C}j \geq \frac{1}{2}\varepsilon_2 d^{(1-\varepsilon_1)j/2}\right\}. \tag{5.43}$$

By definition, under our assumptions, all j's for which (5.41) fails are smaller or equal than j_0. Letting

$$\mathscr{S}_j := \{\text{GWT survives for } j \text{ generations}\},$$

then we observe that, given N_j, h_j is distributed as a binomial random variable $h_j \sim \mathrm{Bin}\left(N_j, \mathbb{P}_\mu(\mathscr{S}_j)\right)$, thus

$$\mathbb{E}_\mu h_j = \mathbb{E}_\mu\left[\mathbb{E}_\mu[h_j \mid N_j]\right] = \mathbb{E}_\mu[N_j \mathbb{P}_\mu(\mathscr{S}_j)] = \left(d(1-\mu)\right)^j \mathbb{P}_\mu(\mathscr{S}_j). \tag{5.44}$$

This holds for all j, and hence for all $j \leq j_0$. Since d and μ must satisfy (2.7), there are two constants $0 < \underline{\kappa} < \overline{\kappa} < 1$ so that, on the one hand,

$$\underline{\kappa} d(1-\mu) > 1; \tag{5.45}$$

and on the other hand, for all $j \leq j_0$ $\overline{\kappa}\mathbb{P}_\mu(\mathscr{S}_j) > \underline{\kappa}$. Then,

$$\begin{aligned}
\widetilde{\mathbb{P}}_\mu\left(h_j \leq \underline{\kappa}\left(d(1-\mu)\right)^j\right) &\leq \widetilde{\mathbb{P}}_\mu\left(h_j \leq \overline{\kappa}\left(d(1-\mu)\right)^j \mathbb{P}_\mu(\mathscr{S}_j)\right) \\
&\leq \widetilde{\mathbb{P}}_\mu\left(|h_j - \mathbb{E}_\mu h_j| \geq (1-\overline{\kappa})\mathbb{E}_\mu h_j\right) \leq \frac{1}{\mathbb{P}_\mu(\mathscr{S})}\frac{\mathrm{Var}_\mu(h_j)}{\left((1-\overline{\kappa})\mathbb{E}_\mu h_j\right)^2}.
\end{aligned} \tag{5.46}$$

To evaluate the variance of h_j we reason like above, namely,

$$\begin{aligned}
\mathbb{E}_\mu[(h_j - \mathbb{E}_\mu h_j)^2] &= \mathbb{E}_\mu[\mathbb{E}_\mu[(h_j - \mathbb{E}_\mu h_j)^2 \mid N_j]] = \mathbb{E}[N_j \mathbb{P}_\mu(\mathscr{S}_j)(1 - \mathbb{P}_\mu(\mathscr{S}_j))] \\
&= \left(d(1-\mu)\right)^j \mathbb{P}_\mu(\mathscr{S}_j)(1 - \mathbb{P}_\mu(\mathscr{S}_j)).
\end{aligned}$$

Using these facts in (5.46) we obtain

$$\widetilde{\mathbb{P}}_\mu\left(|h_j - \mathbb{E}_\mu h_j| > (1-\overline{\kappa})\mathbb{E}_\mu h_j\right) \leq \frac{(1 - \mathbb{P}_\mu(\mathscr{S}_j))}{(1-\overline{\kappa})^2 \left(d(1-\mu)\right)^j \mathbb{P}_\mu(\mathscr{S}_j)\mathbb{P}_\mu(\mathscr{S})}.$$

Now, since for all $j \geq 1$ we have $\mathbb{P}_\mu(\mathscr{S}_j) \geq \mathbb{P}_\mu(\mathscr{S})$, then for all $j \geq 1$ one has

$$1 - \mathbb{P}_\mu(\mathscr{S}_j) \leq 1 - \mathbb{P}_\mu(\mathscr{S}).$$

Hence, Chebyshev's inequality for all $j \in \{1, \ldots, j_0\}$ reduces to

$$\widetilde{\mathbb{P}}_\mu\left(|h_j - \mathbb{E}_\mu h_j| > (1-\overline{\kappa})\mathbb{E}_\mu h_j\right) \leq \frac{(1 - \mathbb{P}_\mu(\mathscr{S}))}{(1-\overline{\kappa})^2 \left(d(1-\mu)\right)^j \mathbb{P}_\mu(\mathscr{S})^2}.$$

Since we are assuming that $\mu < \mu'$ (defined in (5.37)) then $\mathbb{P}_\mu(\mathscr{S}) > 1/2$, which yields

$$\widetilde{\mathbb{P}}_\mu\left(|h_j - \mathbb{E}_\mu h_j| > (1-\overline{\kappa})\mathbb{E}_\mu h_j\right) \leq \frac{4(1 - \mathbb{P}_\mu(\mathscr{S}))}{(1-\overline{\kappa})^2 \left(d(1-\mu)\right)^j}.$$

Now define $\mu'' := \mu''(\mathbf{C}, \overline{\kappa}) > 0$ so that, uniformly over all $j \geq 1$ one has

$$\frac{2}{(1-\overline{\kappa})^2 \left(2(1-\mu'')\right)^j} \leq \frac{e^{-\mathbf{C}j}}{3}. \tag{5.47}$$

Hence we can restrict our attention to $\mu < \mu''$ so that, uniformly over all $j \in \{1, \ldots, j_0\}$ one has

$$\widetilde{\mathbb{P}}_\mu\left(h_j \leq \underline{\kappa}\left(d(1-\mu)\right)^j\right) \leq \frac{4(1 - \mathbb{P}_\mu(\mathscr{S}))}{(1-\overline{\kappa})^2 \left(d(1-\mu)\right)^j} \leq \frac{e^{-\mathbf{C}j}}{3}. \tag{5.48}$$

Now we find an upper bound for $\widetilde{\mathbb{P}}_\mu(U_i^c(j))$, for all $i \in h(\mathcal{N}_j)$. By definition we have

$$\begin{aligned}
\widetilde{\mathbb{P}}_\mu(U_i^c(j)) &\leq \widetilde{\mathbb{P}}_\mu\left[\left|M_1(j) - \widetilde{\mathbb{E}}_\mu[M_1(j)]\right| > \frac{\mathbb{E}(M_\lambda(j))}{2} - \widetilde{\mathbb{E}}_\mu[M_1(j)]\right] + \widetilde{\mathbb{P}}_\mu\left(\text{particle } i \text{ is seed}\right) \\
&\overset{(5.29)}{\leq} \widetilde{\mathbb{P}}_\mu\left(\left|M_1(j) - \widetilde{\mathbb{E}}_\mu[M_1(j)]\right| > \left(\frac{\mathbf{K}}{2} - 1\right)\widetilde{\mathbb{E}}_\mu[M_1(j)]\right) + \mu \overset{\text{Thm 5.1}}{\leq} Ce^{-\delta\mathbf{K}\widetilde{\mathbb{E}}_\mu[M_1(j)]/3} + \mu.
\end{aligned}$$

At this point fix a value $\mu''' = \mu'''(\lambda, C, \delta, \mathbf{C}, j_0) > 0$ so small that

$$Ce^{-\delta\mathbf{K}\widetilde{\mathbb{E}}_{\mu'''}[M_1(1)]/3} + \mu''' \leq \frac{e^{-\mathbf{C}j_0}}{3}. \tag{5.49}$$

(This quantity is well defined because j_0 is independent of μ, cf. (5.43).) Thus, for all $\mu < \mu'''$ and for all λ small enough (that is for all $\mathbf{K}$ large enough), one has

$$\widetilde{\mathbb{P}}_{\mu}(U_i^c(j)) \leq Ce^{-\delta\mathbf{K}\widetilde{\mathbb{E}}_{\mu}[M_1(j)]/3} + \mu \leq Ce^{-\delta\mathbf{K}\widetilde{\mathbb{E}}_{\mu}[M_1(1)]/3} + \mu \leq \frac{e^{-\mathbf{C}j_0}}{3}. \tag{5.50}$$

As a consequence of this, for all $j \in \{1, \ldots, j_0\}$ and $\mu < \min\{\mu', \mu'', \mu'''\}$ (defined in (5.37), (5.47), (5.49) respectively) we get

$$\widetilde{\mathbb{P}}_{\mu}(U^c(j)) \leq \widetilde{\mathbb{P}}_{\mu}\left(\bigcap_{i \in h(\mathcal{N}_j)} U_i^c(j)\right) \leq \widetilde{\mathbb{P}}_{\mu}\left(U_1^c(j)\right)^{\underline{\kappa}\left(d(1-\mu)\right)^j} + \widetilde{\mathbb{P}}_{\mu}\left(h_j \leq \underline{\kappa}\left(d(1-\mu)\right)^j\right) + $$

$$\overset{(5.50),\,(5.45),\,(5.48)}{\leq} \frac{e^{-\mathbf{C}j_0}}{3} + \frac{e^{-\mathbf{C}j}}{3} \overset{j \leq j_0}{\leq} e^{-\mathbf{C}j}.$$

Similarly, for all $j \in \{j_0, \ldots, \ell\}$ and $\mu < \min\{\mu', \mu'', \mu'''\}$ we get

$$\widetilde{\mathbb{P}}_{\mu}(U^c(j)) \leq \widetilde{\mathbb{P}}_{\mu}\left(\bigcap_{i \in h(\mathcal{N}_j)} U_i^c(j)\right)$$

$$\leq \widetilde{\mathbb{P}}_{\mu}\left(U_1^c(j)\right)^{\tilde{C}j} + \widetilde{\mathbb{P}}_{\mu}\left(h_j \leq \tilde{C}j\right) + \widetilde{\mathbb{P}}_{\mu}\left(N_j \leq \varepsilon_2\left(d(1-\mu)\right)^{(1-\varepsilon_1)j}\right)$$

$$\overset{\text{Lemma 5.6}}{\leq} \widetilde{\mathbb{P}}_{\mu}\left(U_1^c(j)\right)^{\tilde{C}j} + \widetilde{\mathbb{P}}_{\mu}\left(h_j \leq \tilde{C}j\right) + C_2 e^{-c_3 j}$$

$$\overset{(5.50),\,(5.45),\,(5.42),\,(5.39)}{\leq} \frac{e^{-\mathbf{C}j_0\tilde{C}j}}{3} + \frac{e^{-\mathbf{C}j}}{3} + \frac{e^{-\mathbf{C}j}}{3} \leq e^{-\mathbf{C}j},$$

as claimed.

Case (ii): $\mu \in (1/2, 1)$ is fixed and d is large enough. In this case, when $\mu \in (1/2, 1)$ is fixed, $\tilde{C} > 0$ can be chosen again as in (5.40). Thus, once $\varepsilon_1, \varepsilon_2$ have been set, we can pick d large enough, so that condition (5.41) holds. Hence the proof for this case follows from a Chernoff bound exactly as performed before, yielding the claim. This concludes the proof. $\quad\square$

We proceed with the goal of bounding $\widetilde{\mathbb{P}}_{\mu}(B_{m-j})$. For all j, by applying the inclusion-exclusion principle, we obtain the following bound

$$\widetilde{\mathbb{P}}_{\mu}(B_{m-j}) \leq \widetilde{\mathbb{P}}_{\mu}\left[B_{m-j} \cap \left(\bigcap_{i=1}^{\exp\{\overline{\mathbf{C}}_\lambda j/2\}} E_i^c(j) \cap U(j)\right)\right] + \widetilde{\mathbb{P}}_{\mu}\left[\bigcup_{i=1}^{\exp\{\overline{\mathbf{C}}_\lambda j/2\}} E_i(j) \cup U^c(j)\right]. \tag{5.51}$$

Recall the definition of $\mathcal{Z}_j$ from (5.28), A_j from (5.16), and the definition of the events $E(j)$ and $U(j)$ from (5.31) and (5.33). Consider the case where all the following assumptions are satisfied, for all $j \in \{1, \ldots, \ell\}$.

(i) There are at least $\varepsilon_2\left(d(1-\mu)\right)^{(1-\varepsilon_1)j}$ individuals of type FPP_1 alive at generation j which are born by time $C_1 j$ (that is, A_j holds).

(ii) All their first descendants at generation $m - j$ which are of type FPP_1 or seeds (if there are any) are born within time $C_4 j$ of each other (that is, $\mathcal{Z}_j$ holds).

(iii) All particles $i \in \mathcal{S}(m-j)$ such that $i > \exp\{\overline{\mathbf{C}}_\lambda j/2\}$ are either not seeds, or seeds initiating a FPP_λ process which is not reaching $\mathcal{L}_m$ within the time interval $\mathbf{T}_\lambda(j)$, or yet there is a particle r with $r < i$ originating a process (FPP_1 or FPP_λ) reaching $\mathcal{L}_m$ before the process originated from the i-th particle, that is, $\left(F_i^{(j)}\right)^c$ holds for all $i > \exp\{\overline{\mathbf{C}}_\lambda j/2\}$. (Note that if $|\mathcal{S}(m-j)| \leq \exp\{\overline{\mathbf{C}}_\lambda j/2\}$ this assumption is empty.)

Then, necessarily at least one of the following events must hold. Possibly, at level $\mathcal{L}_{m-j}$ there are no seeds that originate a FPP_λ process infecting $\mathcal{L}_m$ first; or there is at least a particle $i \leq \exp\{\overline{\mathbf{C}}_\lambda j/2\}$ originating a FPP_1 or an FPP_λ process reaching $\mathcal{L}_m$ within the time interval $\mathbf{T}_\lambda(j)$. (This latter possibility is ensured by assumption (iii) above.) Hence, by considering the complementary events, for every $j \in \{1, \ldots, \ell\}$

$$B_{m-j} \cap \left(\bigcap_{i=1}^{\exp\{\overline{\mathbf{C}}_\lambda j/2\}} E_i^c(j) \cap U(j) \right) \subseteq \left\{ \bigcup_{i > \exp\{\overline{\mathbf{C}}_\lambda j/2\}} F_i^{(j)} \right\} \cup \mathcal{Z}_j^c \cup A_j^c. \qquad (5.52)$$

Hence, relations (5.51) and (5.52) yield

$$\widetilde{\mathbb{P}}_\mu(B_{m-j}) \leq \left[\sum_{i > e^{\overline{\mathbf{C}}_\lambda j/2}} \widetilde{\mathbb{P}}_\mu(F_i^{(j)}) + \widetilde{\mathbb{P}}_\mu(\mathcal{Z}_j^c) + \widetilde{\mathbb{P}}_\mu(A_j^c) \right] + \widetilde{\mathbb{P}}_\mu(U^c(j)) + \sum_{i=1}^{e^{\overline{\mathbf{C}}_\lambda j/2}} \widetilde{\mathbb{P}}_\mu(E_i(j))$$

$$\overset{(5.35)}{\leq} \sum_{i > \exp\{\overline{\mathbf{C}}_\lambda j/2\}} \widetilde{\mathbb{P}}_\mu(F_i^{(j)}) + \widetilde{\mathbb{P}}_\mu(\mathcal{Z}_j^c) + \widetilde{\mathbb{P}}_\mu(A_j^c) + \widetilde{\mathbb{P}}_\mu(U^c(j)) + e^{\overline{\mathbf{C}}_\lambda j/2} e^{-\overline{\mathbf{C}}_\lambda j}$$

$$\overset{(5.38)}{\leq} \sum_{i > \exp\{\overline{\mathbf{C}}_\lambda j/2\}} \widetilde{\mathbb{P}}_\mu(F_i^{(j)}) + \widetilde{\mathbb{P}}_\mu(\mathcal{Z}_j^c) + \widetilde{\mathbb{P}}_\mu(A_j^c) + e^{-\mathbf{C}j} + e^{-\overline{\mathbf{C}}_\lambda j/2}.$$

$$(5.53)$$

Our next aim is to show that the quantity $\sum_{i > \exp\{\overline{\mathbf{C}}_\lambda j/2\}} \widetilde{\mathbb{P}}_\mu(F_i^{(j)})$ decreases exponentially fast in j, which we do in the next lemma.

Lemma 5.14. *With the above definitions, there is a constant $C_5 > 0$ which does not depend on λ, such that for all $j \in \{1, \ldots, \ell\}$ we have*

$$\sum_{i > \exp\{\overline{\mathbf{C}}_\lambda j/2\}} \widetilde{\mathbb{P}}_\mu(F_i^{(j)}) \leq C_5 e^{-\overline{\mathbf{C}}_\lambda j/2},$$

Proof. We define two auxiliary events as follows. For all $k \in \mathbb{N}$ define

$$\mathcal{H}_k(i) := \big\{ \forall r \in \{1, \ldots, i-1\} \text{ with } r\text{-th particle being a non-seed, the process}$$
$$\text{originated from the } r\text{-th particle reaches } \mathcal{L}_m \text{ in time } \geq k \big\};$$

$$\mathcal{Y}_k^{(\lambda)}(i) := \big\{ i\text{-th particle is a seed initiating a } \mathrm{FPP}_\lambda \text{ process reaching } \mathcal{L}_m \text{ in time } \leq k \big\}.$$

(Note that the above events also depend on j, but we omitted this dependence to simplify the notation.) Recall that $\mathscr{S}$ is the event that GWT survives indefinitely and let

$$\mathbf{Q} := \mathbb{P}_\mu(\{o \text{ is a non-seed}\} \cap \mathscr{S}),$$

then fix any $r \leq i - 1$ and let $\{X_l\}_l$ denote an i.i.d. sequence of exponential random variables with rate 1. We have

$\mathbb{P}(\{r \text{ is not a seed}\} \cap \{r \text{ starts a } \mathrm{FPP}_1 \text{ process infecting } \mathcal{L}_m \text{ in time } \leq k\})$

$$\geq \mathbb{P}(\{r \text{ is not a seed}\} \cap \{r = \text{root of a GWT surviving for } j \text{ generations}\}) \mathbb{P}\left[\sum_{l=1}^{j} X_l \leq k \right]$$

$$\geq \mathbf{Q}\, \mathbb{P}\left(\sum_{l=1}^{j} X_l \leq k \right).$$

In particular, according to the relative sizes of k and j, we shall make a case distinction, namely, fix $1 \leq j \leq |\mathcal{S}(m-j)|$ and consider the following two situations. Recall $\mathcal{C}^*$ from (5.6) and for simplicity set $\mathbf{C}_\lambda := \frac{\mathcal{C}^*}{\lambda}$.

(i) For all $1 \leq k \leq \frac{\mathbf{C}_\lambda j}{2}$, we see that

$$\mathbf{Q}\, \mathbb{P}\left(\sum_{l=1}^{j} X_l \leq k\right) = \mathbf{Q}\, \mathbb{P}\left[\mathrm{Poi}(k) \geq j\right] \geq \mathbf{Q}\, \mathbb{P}\left[\mathrm{Poi}(k) = k+j\right]$$

$$\geq \mathbf{Q} e^{-k} \frac{k^{(k+j)}}{(k+j)!} \geq \mathbf{Q}\, \frac{e^{-k-\frac{1}{12(k+j)}}}{\sqrt{2\pi(k+j)}}\left(\frac{ke}{k+j}\right)^{(k+j)},$$

in the last inequality we used a standard refinement of Stirling's approximation. Set $C_6 := \mathbf{Q}\frac{e^{-\frac{1}{12}}}{\sqrt{2\pi}}$, then

$$\mathbf{Q}\, \frac{e^{-k-\frac{1}{12(k+j)}}}{\sqrt{2\pi(k+j)}}\left(\frac{ke}{k+j}\right)^{(k+j)} \overset{k,j\geq 1}{\geq} C_6 \frac{e^j}{\sqrt{k+j}}\left(\frac{k}{k+j}\right)^{(k+j)}$$

$$\overset{k\leq \mathbf{C}_\lambda j/2}{\geq} C_6 \frac{e^j}{\sqrt{\frac{\mathbf{C}_\lambda}{2}j + j}}\left(\frac{k}{\frac{\mathbf{C}_\lambda}{2}j + j}\right)^{(\frac{\mathbf{C}_\lambda}{2}j+j)} \geq C_7 \frac{e^{j/2}}{\sqrt{\mathbf{C}_\lambda}}\left(\frac{k}{\left(\frac{\mathbf{C}_\lambda}{2}+1\right)j}\right)^{(\frac{\mathbf{C}_\lambda}{2}+1)j},$$

for some suitable constant $C_7 = C_7(C_6) > 0$. To ease the notation, set

$$\mathbf{C}_\lambda^* := \frac{\mathbf{C}_\lambda}{2} + 1, \tag{5.54}$$

which, for all $1 \leq k \leq \frac{\mathbf{C}_\lambda j}{2}$ gives

$$\widetilde{\mathbb{P}}_\mu(\mathcal{H}_k(i)) \leq \left(1 - \frac{C_7}{\sqrt{\mathbf{C}_\lambda}}e^{j/2}\left(\frac{k}{\mathbf{C}_\lambda^* j}\right)^{\mathbf{C}_\lambda^* j}\right)^i \leq \exp\left\{-\frac{C_7}{\sqrt{\mathbf{C}_\lambda}}i\, e^{j/2}\left(\frac{k}{\mathbf{C}_\lambda^* j}\right)^{\mathbf{C}_\lambda^* j}\right\}. \tag{5.55}$$

(ii) For all $k > \frac{\mathbf{C}_\lambda j}{2}$ we use Theorem 5.7(i) in the following way. First, we note that for a sequence of i.i.d. exponential variables of parameter 1 we have $\mathbb{E}\left(\sum_{l=1}^{j} X_l\right) = j$. Then, since k is much larger than j (for λ small enough) we can write

$$\mathbb{P}\left(\sum_{l=1}^{j} X_l > k\right) = \mathbb{P}\left(\sum_{l=1}^{j} X_l > \frac{k}{j}j\right) \overset{\mathrm{Thm\ 5.7}(i)}{\leq} \frac{j}{k}e^{-j\left(\frac{k}{j}-1-\ln\left(\frac{k}{j}\right)\right)} \leq \frac{j}{k}e^{-k/2},$$

where in the last inequality we have used that λ is small enough (and thus k is large). Hence, for all $k > \frac{\mathbf{C}_\lambda j}{2}$ we obtain

$$\mathbb{P}(\mathcal{H}_k(i)) \leq \left(\frac{j}{k}e^{-k/2}\right)^i \leq \left(\frac{2}{\mathbf{C}_\lambda}\right)^i e^{-i\, k/2}. \tag{5.56}$$

Using again Theorem 5.1 we obtain that for all $k \leq \frac{\mathbb{E}(M_\lambda(j))}{2}$

$$\mathbb{P}\left(\mathcal{Y}_k^{(\lambda)}(i)\right) = \mathbb{P}\left(M_\lambda(j) \leq k\right) \leq \mathbb{P}\left(|M_\lambda(j) - \mathbb{E}\left(M_\lambda(j)\right)| > \mathbb{E}(M_\lambda(j)) - k\right)$$

$$\leq \mathbb{P}\left(|M_\lambda(j) - \mathbb{E}\left(M_\lambda(j)\right)| > \frac{\mathbf{C}_\lambda j}{2}\right) \leq C e^{-\delta \mathbf{C}_\lambda j/2}.$$

Hence, as we did previously (cf. (5.36)) we find that for all $k \leq \frac{\mathbf{C}_\lambda j}{2}$ and all i,

$$\mathbb{P}\left(\mathcal{Y}_k^{(\lambda)}(i)\right) \leq e^{-\overline{\mathbf{C}}_\lambda j}. \tag{5.57}$$

From the definition of $F_i^{(j)}$ it follows that

$$\widetilde{\mathbb{P}}_\mu\left(F_i^{(j)}\right) \leq \widetilde{\mathbb{P}}_\mu\left(\cup_{k\in\mathbb{N}}\left(\mathcal{H}_k(i)\cap\mathcal{Y}_k^{(\lambda)}(i)\right)\right),$$

thus, since processes born from different particles (at the same level) are independent we have $\widetilde{\mathbb{P}}_\mu\left(\mathcal{H}_k(i)\cap\mathcal{Y}_k^{(\lambda)}(i)\right) = \widetilde{\mathbb{P}}_\mu\left(\mathcal{H}_k(i)\right)\mathbb{P}\left(\mathcal{Y}_k^{(\lambda)}(i)\right)$. Hence,

$$\widetilde{\mathbb{P}}_\mu\left(F_i^{(j)}\right) \leq \sum_{k\geq 1}\widetilde{\mathbb{P}}_\mu\left(\mathcal{H}_k(i)\cap\mathcal{Y}_k^{(\lambda)}(i)\right) \leq \sum_{k=1}^{\mathbf{C}_\lambda j/2}\widetilde{\mathbb{P}}_\mu\left(\mathcal{H}_k(i)\right)\mathbb{P}\left(\mathcal{Y}_k^{(\lambda)}(i)\right) + \sum_{k\geq\mathbf{C}_\lambda j/2+1}\widetilde{\mathbb{P}}_\mu\left(\mathcal{H}_k(i)\right)$$

$$\overset{(5.57)}{\leq} e^{-\overline{\mathbf{C}}_\lambda j}\sum_{k=1}^{\mathbf{C}_\lambda j/2}\widetilde{\mathbb{P}}_\mu\left(\mathcal{H}_k(i)\right) + \sum_{k\geq\mathbf{C}_\lambda j/2+1}\widetilde{\mathbb{P}}_\mu\left(\mathcal{H}_k(i)\right).$$

Now we apply relations (5.55) and (5.56), which give that the above is bounded by

$$e^{-\overline{\mathbf{C}}_\lambda j}\frac{\mathbf{C}_\lambda j}{2}\exp\left\{-C_7\, i\, e^{j/2}\left(\frac{k}{\mathbf{C}_\lambda^* j}\right)^{\mathbf{C}_\lambda^* j}\right\} + \sum_{k>\mathbf{C}_\lambda j/2}\left(\frac{2}{\mathbf{C}_\lambda}\right)^i e^{-i\,k/2}$$

$$\leq e^{-\overline{\mathbf{C}}_\lambda j}\mathbf{C}_\lambda j\exp\left\{-C_7\, i\, e^{j/2}\left(\frac{1}{\mathbf{C}_\lambda^* j}\right)^{\mathbf{C}_\lambda^* j}\right\} + 2\left(\frac{2}{\mathbf{C}_\lambda}\right)^i e^{-i\,\mathbf{C}_\lambda j/4},$$

where we have used that $\mathbf{C}_\lambda$ is large enough. Now it follows easily that the above quantity is summable in i, giving at most

$$e^{-\overline{\mathbf{C}}_\lambda j}\mathbf{C}_\lambda j\,\frac{e^{-C_7\, e^{j/2}\left(\frac{1}{\mathbf{C}_\lambda^* j}\right)^{\mathbf{C}_\lambda^* j}}}{1 - e^{-C_7\, e^{j/2}\left(\frac{1}{\mathbf{C}_\lambda^* j}\right)^{\mathbf{C}_\lambda^* j}}} + \frac{4}{\mathbf{C}_\lambda}\frac{e^{-\mathbf{C}_\lambda j/4}}{1 - e^{-\mathbf{C}_\lambda j/4}}$$

$$\leq C_8 e^{-\overline{\mathbf{C}}_\lambda j/2}e^{-C_7\, e^{j/2}\left(\frac{1}{\mathbf{C}_\lambda^* j}\right)^{\mathbf{C}_\lambda^* j}} + \frac{8}{\mathbf{C}_\lambda}e^{-\mathbf{C}_\lambda j/4}$$

$$\leq C_8 e^{-\overline{\mathbf{C}}_\lambda j/2} + \frac{8}{\mathbf{C}_\lambda}e^{-\mathbf{C}_\lambda j/4} \leq C_8 e^{-\overline{\mathbf{C}}_\lambda j/2} + 8e^{-\delta\mathbf{C}_\lambda j/4},$$

for a suitable constant $C_8 > 0$, independent of λ. The statement follows from the definition of $\overline{\mathbf{C}}_\lambda$ as in (5.36) and by picking $C_5 := C_8 + 8$. $\qquad\square$

At this point we are ready to conclude the proof of Proposition 2.1.

Proof of Proposition 2.1. For any large (fixed) value of m we have

$$\mathbb{P}_\mu\left[W \text{ reached first by } \mathrm{FPP}_\lambda\right] \leq \mathbb{P}_\mu\left[\bigcup_{j=1}^m B_j \cup \mathscr{S}^c\right] \leq \mathbb{P}_\mu\left[\bigcup_{j=1}^m B_j\right] + \mathbb{P}_\mu\left[\mathscr{S}^c\right]$$

$$\overset{(5.3)}{\leq} \mathbb{P}_\mu\left[\bigcup_{j=1}^m B_j\right] + \mathbf{f}_d(\mu) \overset{\text{Lemma 5.5}}{\leq} e^{-\delta'm} + \mathbb{P}_\mu\left[\bigcup_{j=m-\ell}^m B_j\right] + \mathbf{f}_d(\mu).$$

Now, in order to bound $\mathbb{P}_\mu\left[\bigcup_{j=m-\ell}^m B_j\right]$ we use 5.53. By putting together the results obtained in Lemmas 5.9, 5.11 and 5.14 we obtain

$$\mathbb{P}_\mu\left[\bigcup_{j=m-\ell}^m B_j\right] \leq \varepsilon_3 + \sum_{j=1}^\ell\left[2e^{-C_3'j} + \frac{1}{C_1}e^{-jC_1/2} + e^{-\mathbf{C}j} + e^{-\overline{\mathbf{C}}_\lambda j/2} + C_5 e^{-\overline{\mathbf{C}}_\lambda j/4}\right]$$

$$\leq \varepsilon_3 + \sum_{j=1}^\ell\left[2e^{-C_3'j} + \frac{1}{C_1}e^{-jC_1/2} + e^{-\mathbf{C}j} + (C_5+1)e^{-\overline{\mathbf{C}}_\lambda j/4}\right].$$

Recall the definitions of $\mathbf{C}$ and $\overline{\mathbf{C}}_\lambda$ from Claim 5.12 and Lemma 5.13, and in particular recall that they could be picked as large as we wanted, when λ is small and whenever $\mu < \min\{\mu', \mu'', \mu'''\}$ (defined in (5.37), (5.47), (5.49) respectively) or d large enough. Now denote by

$$\boldsymbol{\omega} := \min\{C_3', C_1/2, \mathbf{C}, \overline{\mathbf{C}}_\lambda/4\},$$

hence

$$\mathbb{P}_\mu\left[\bigcup_{i=m-\ell}^{m} B_i\right] \leq \varepsilon_3 + \left(2 + \frac{1}{C_1} + 1 + (C_5 + 1)\right)\sum_{j=1}^{\ell} e^{-\boldsymbol{\omega} j}.$$

Set $C_9 := 5 + C_5$, since λ is chosen to be small enough, then $\boldsymbol{\omega}$ is extremely large, hence

$$\mathbb{P}_\mu\left[\bigcup_{i=m-\ell}^{m} B_i\right] \leq \varepsilon_3 + 2C_9 e^{-\boldsymbol{\omega}}.$$

Since m was chosen large enough, we let

$$\boldsymbol{\eta}_d = \boldsymbol{\eta}_d(\delta', \mu, d, \lambda, \varepsilon_3, C_9, \boldsymbol{\omega}) := e^{-\delta' m} + \varepsilon_3 + 2C_9 e^{-\boldsymbol{\omega}}.$$

Moreover, ε_3 can be made arbitrarily small and $\boldsymbol{\omega}$ arbitrarily large, provided $\lambda > 0$ is small enough and either μ is small enough, or d is large enough. This establishes the desired property of $\boldsymbol{\eta}_d$ in Proposition 2.1. To conclude, we note that

$$\mathbb{P}_\mu\Big[W \text{ is infected by } \mathrm{FPP}_\lambda\Big]$$
$$= \mathbb{P}_\mu\Big[W \text{ reached first by } \mathrm{FPP}_1 \cap \{W \text{ is seed}\}\Big] + \mathbb{P}_\mu\Big[W \text{ reached first by } \mathrm{FPP}_\lambda\Big]$$
$$\leq (1 - \boldsymbol{\eta}_d - \mathbf{f}_d(\mu))\,\mu + (\boldsymbol{\eta}_d + \mathbf{f}_d(\mu)) \leq \mu + (1-\mu)(\boldsymbol{\eta}_d + \mathbf{f}_d(\mu)).$$

Thus, $\mathbb{P}_\mu\Big[W \text{ is infected by } \mathrm{FPP}_1\Big] \geq (1-\mu)(1 - \boldsymbol{\eta}_d - \mathbf{f}_d(\mu))$, as claimed. $\qquad\square$

References

[1] Louigi Addario-Berry and Bruce Reed, *Minima in branching random walks*, Ann. Probab. **37** (2009), no. 3, 1044–1079. MR2537549

[2] Krishna B. Athreya and Peter E. Ney, *Branching processes*, Springer-Verlag, New York-Heidelberg, 1972, Die Grundlehren der mathematischen Wissenschaften, Band 196. MR0373040

[3] Elisabetta Candellero and Alexandre Stauffer, *Coexistence of competing first passage percolation on hyperbolic graphs*, Ann. Inst. Henri Poincaré Probab. Stat. **57** (2021), no. 4, 2128–2164. MR4330845

[4] Duncan Dauvergne and Allan Sly, *Spread of infections in a heterogeneous moving population*, Arxiv preprint: https://arxiv.org/abs/2105.11947. MR4634337

[5] Maria Deijfen and Olle Häggström, *Nonmonotonic coexistence regions for the two-type Richardson model on graphs*, Electron. J. Probab. **11** (2006), no. 13, 331–344. MR2217820

[6] Thomas Finn, *Topics in random growth models*, 2021, Thesis (Ph.D.)–University of Bath.

[7] Thomas Finn and Alexandre Stauffer, *Non-equilibrium multi-scale analysis and coexistence in competing first passage percolation*, J.E.M.S. (2022). MR4721033

[8] Olivier Garet and Régine Marchand, *Coexistence in two-type first-passage percolation models*, Ann. Appl. Probab. **15** (2005), no. 1A, 298–330. MR2115045

[9] Olivier Garet and Régine Marchand, *First-passage competition with different speeds: positive density for both species is impossible*, Electronic Journal of Probability **13** (2008), no. none, 2118 – 2159. MR2461538

[10] Olle Häggström and Robin Pemantle, *First passage percolation and a model for competing spatial growth*, Journal of Applied Probability **35** (1998), no. 3, 683–692. MR1659548

[11] Olle Häggström and Robin Pemantle, *Absence of mutual unbounded growth for almost all parameter values in the two-type Richardson model*, Stochastic Process. Appl. **90** (2000), no. 2, 207–222. MR1794536

[12] Christopher Hoffman, *Coexistence for richardson type competing spatial growth models*, Ann. Appl. Probab. **15** (2005), no. 1B, 739–747. MR2114988

[13] Svante Janson, *Tail bounds for sums of geometric and exponential variables*, Statist. Probab. Lett. **135** (2018), 1–6. MR3758253

[14] Thomas M. Liggett, *Stochastic interacting systems: contact, voter and exclusion processes*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 324, Springer-Verlag, Berlin, 1999. MR1717346

[15] Russell Lyons and Yuval Peres, *Probability on trees and networks*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 42, Cambridge University Press, New York, 2016. MR3616205

[16] Peter. E. Ney and Anand. N. Vidyashankar, *Harmonic moments and large deviation rates for supercritical branching processes*, Ann. Appl. Probab. **13** (2003), no. 2, 475–489. MR1970272

[17] Vladas Sidoravicius and Alexandre Stauffer, *Multi-particle diffusion limited aggregation*, Invent. Math. **218** (2019), no. 2, 491–571. MR4011705