Accessibility Percolation with Crossing Valleys on $n$-ary Trees

Frank Duque · Alejandro Roldán-Correa · Leon A. Valencia

Received: 25 January 2018 / Accepted: 8 January 2019 / Published online: 19 January 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
In this paper, we study a variation of the accessibility percolation model. This work is also motivated by evolutionary biology and evolutionary computation. Consider a tree whose vertices are labeled with random numbers. We study the probability of having a monotone subsequence of a path from the root to a leaf, where any $k$ consecutive vertices in the path contain at least one vertex of the subsequence. An $n$-ary tree, with height $h$, is a tree whose vertices at distance at most $h - 1$ to the root have $n$ children. For the case of $n$-ary trees, we prove that, as $h$ tends to infinity, the probability of having such subsequence: tends to 1, if $n$ grows significantly faster than $\sqrt[3]{h/(ek)}$; and tends to 0, if $n$ grows significantly slower than $\sqrt[3]{h/(ek)}$.

Keywords Percolation · Dynamics of evolution · Fitness landscape

Mathematics Subject Classification 60K35 · 60C05 · 92D15

1 Introduction
Let $T$ be a rooted tree (whose edges are directed from fathers to children). Assume that the vertices of $T$ are labeled with independent and identically distributed continuous random variables. Further, given a vertex $v \in T$, we denote by $w(v)$ its label. Let $P = v_0 \rightarrow v_1 \rightarrow \cdots \rightarrow v_h$ be a path in $T$ from the root to a leaf. We say that $P$ is a $k$-accessible if there is a subsequence $S := v_{r(0)}, v_{r(1)}, v_{r(2)}, \ldots, v_{r(t)}$ of the vertices in $P$ such that:

$$w(v_{r(0)}) < w(v_{r(1)}) < w(v_{r(2)}) < \cdots < w(v_{r(t)}),$$

1 Instituto de Matemáticas UNAM and Instituto de Física UASLP, Mexico, Mexico
2 Instituto de Matemáticas, Universidad de Antioquia, Medellín, Colombia
Fig. 1 An example of a labelled tree without 1-accessible paths, and for which the only 2-accessible paths are labeled as: 53,99,68,4,71; and 53,65,13,78,26,91. Each $k$ consecutive vertices of $P$ contain at least one vertex in $S$, $v_0 = v_{r(0)}$ and $v_h = v_{r(t)}$. In Fig. 1 we illustrate a labelled tree and its 2-accessible paths. We denote by $\theta_k(T)$ the probability of having a $k$-accessible path in $T$.

1.1 Motivation

Fitness landscapes are by construction a static concept that assigns fitness values to points of an underlying configuration space [11]. This concept was first proposed in 1932 by Sewall Wright [16] (as a mapping from a set of genotypes to fitness). Since then, fitness landscapes have attracted particular interest in evolutionary biology and evolutionary computation, because it offers an approach to conceptualize and visualize how an evolving population may change over time, in various population genetics models [9].

For the case of a haploid asexual population on a given fitness landscape, framed by the ‘strong selection, weak mutation’ (SSWM) regime, a mutations path is considered selectively accessible if the fitness values encountered along it are monotonically increasing [3,13,14]. Motivated by the concept of selective accessibility, a new kind of percolation was introduced by Nowak and Krug in [8], in which they studied the probability of having a monotonically increasing path in a graph whose vertices have been labelled with random numbers [1,2,6,8,10].

The model of accessibility percolation was introduced by Nowak and Krug [8] as follows. Imagine a population of some life form endowed with the same genetic type (genotype). If a mutation occurs, a new genotype is created which can die out or replace the old one. Provided that natural selection is sufficiently strong, the latter only happens if the new genotype has larger fitness. As a consequence, on a longer timescale, the genotype of the population takes a path through the space of genotypes along which the fitness is monotonically increasing [4].

The hypercube is a graph whose vertices are all possible $N$-tuples in $\{0, 1\}^N$ (for some positive integer $N$), where two vertices are connected by an undirected edge, if the number of coordinates at which they differ is one. In many basic mathematical models of genetic mutations, the genotype sequence space is represented by the hypercube: each genome is represented as a node of the hypercube and each mutation involves the flipping of a single bit from 0 (the “wild” state) to 1 (the “mutant” state) [3,5,13]. An $f$-ary (complete $f$-ary) tree with height $h$ is a rooted tree (whose edges are directed from parents to children) whose
vertices at distance at most $h - 1$ to the root, have $f$ children. Being a simpler case and motivated by the hypercube case, Nowak and Krug in [8] study the problem of determining the growth of $f$ as a function of $h$ for having accessibility percolation in $f$-ary trees.

According to the ruggedness or smoothness of the landscape, there are some topological features (as peaks, valleys, ridges, etc.) that block the selectively accessible mutations paths towards the highest peak in the landscape. The evolutionary process that may allow escaping those topological features is known as valley crossing, which is not allowed in the selectively accessible mutation paths. However, in natural populations of sufficient size, a number of double mutants is present at all times, and the crossing valleys can be relatively facile [3,12,15]; the SSWM assumption may therefore seem very restrictive.

A simple way of allowing peaks and valleys on the accessibility percolation model, consists in “allowing holes” in the monotonicity of the path that the genotype population takes. In this paper, as a variation of the concept of accessibility percolation introduced by Nowak and Krug in [8], we introduce the concept of $k$-accessibility percolation in which crossing small valleys is allowed. Imagine a population of some organisms endowed with the same genotype. If a mutation occurs, a new genotype is created. The genotype die, when in the small valleys is allowed. According to the ruggedness or smoothness of the landscape, there are some topological features (as peaks, valleys, ridges, etc.) that block the selectively accessible mutations paths.

### 1.2 Previous Results

Given a function $f: \mathbb{Z}_+ \rightarrow \mathbb{R}_+$, we denote by $T_h(f)$ the $\lfloor f(h) \rfloor$-ary tree with height $h$ and we denote by $T_f$ the sequence of trees $(T_h(f))_{h \in \mathbb{Z}_+}$. We denote by $\theta_k(T_f) := \lim_{h \to \infty} \theta_k(T_h(f))$ provided the limit exists. Further, when $\theta_k(T_f) > 0$ we say that there is $k$-percolation in $T_f$.

In the first work on accessibility percolation, Nowak and Krug [8] studied the problem of determining the growth of $f$, with respect to $h$, such that there is 1-accessibility percolation on $T_f$. Nowak and Krug prove that: if $f$ is a super-linear function, then there is 1-accessibility percolation in $T_f$; and if $f$ is a sub-linear function, then there is no 1-accessibility percolation in $T_f$. They also studied the linear case, $f(h) = ah$, establishing the existence of a threshold between $e^{-1}$ and 1 on the scaling constant factor; there is 1-accessibility percolation in $T_f$ when $\alpha > 1$, and there is no 1-accessibility percolation in $T_f$ when $\alpha < e^{-1}$. See Theorem 1.1

**Theorem 1.1** (Nowak–Krug [8]) $\theta_1(T_f) > 0$, if $f(h) \geq |ah|$ for $h$ large enough and $\alpha > 1$. $\theta_1(T_f) = 0$, if $f(h) \leq |ah|$ for $h$ large enough and $\alpha \leq e^{-1}$.

Continuing with the work of Nowak and Krug, Roberts and Zhao [10] determined that $\alpha = e^{-1}$ is a threshold on the scaling constant factor for 1-accessibility percolation on $T_f$, see Theorems 1.2 and 1.3.

**Theorem 1.2** (Roberts–Zhao [10])

$$
\theta_1(T_f) = \begin{cases} 
1, & \text{if } f(h) \geq |ah| \text{ for } h \text{ large enough and } \alpha > e^{-1}, \\
0, & \text{if } f(h) \leq |ah| \text{ for } h \text{ large enough and } \alpha \leq e^{-1}.
\end{cases}
$$

**Theorem 1.3** (Roberts–Zhao [10]) If $f(h) = \left(\frac{1+\beta_h}{e}\right)h$ where $\beta_h \to 0$ as $h \to \infty$, then

$$
\theta_1(T_f) = \begin{cases} 
1, & \text{if } h\beta_h / \log h \to \infty, \\
0, & \text{if } \log h - 2h\beta_h \to \infty.
\end{cases}
$$

Accessibility percolation has also been studied, recently, for the case when the underlying graph is the hypercube; see [1,2,5–7].
1.3 Our Results

In this paper we determine the growth of $f$, as a function of $h$, for which there is $k$-accessibility percolation on $T_f$. Our main result is the following.

**Theorem 1.4** Let $T_f$ be the sequence of $n$-ary trees $\{T_h(f)\}_{h \in \mathbb{Z}_+}$, $k \geq 2$, $c_1 > 0$ and $c_2 > 1/k$. Then,

$$\theta_k(T_f) = \begin{cases} 
1, & \text{if } f(h) \geq \sqrt{h/(ek)} + c_1 \text{ for } h \text{ large enough,} \\
0, & \text{if } f(h) \leq \sqrt{h/(ek)} - c_2 \text{ for } h \text{ large enough.}
\end{cases} (1.1)$$

2 Preliminaries

Before proceeding with the proof of Theorem 1.4, we introduce the concept of $k$-transitive closure and state an equivalent version of Theorem 1.3.

We define the $k$-transitive closure of a graph $G$, $G^k$, as the graph obtained from it, by adding new edges from each vertex $u$ to each vertex $v$, with the property that $G$ does not already contain the directed edge from $u$ to $v$ but does contain a directed path from $u$ to $v$ with length at most $k$.

Although the concept of $k$-accessibility percolation was only introduced for trees, it can easily be extended to directed graphs where there is a single vertex distinguished as the source, some vertices distinguished as sinks, and all of the maximal paths are directed paths from the source to some sink. We name the graphs that satisfy the previous statement as monotone graphs. For the case of monotone graphs, we say that a path in it is 1-accessible if it starts in the source, ends at some sink and its vertices have increasing labels. Similar to the case of trees, we denote by $\theta_1(G)$ the probability of having a 1-accessible path in $G$.

In Fig. 2 we illustrate the 2-transitive closure of the graph depicted in Fig. 1 and its 1-accessible paths. Note that the 1-accessible paths in the graph depicted in Fig. 2 correspond to the 2-accessible paths in the graph depicted in Fig. 1.

**Remark 2.1** Let $T$ be a rooted tree with height $h$ and $T^k$ be its $k$-transitive closure. Then $\theta_1(T^k) = \theta_k(T)$.

![Fig. 2](image-url) An illustration of the 2-transitive closure of the graph depicted in Fig. 1, the 1-accessible paths in such graph are labeled as: 53,68,71; and 53,65,78,91
Proof Let $T$ be a rooted tree with height $h$ and $T^k$ be its $k$-transitive closure. If $P$ is a $k$-accessible path in $T$, then there is a subsequence $S$ of the vertices in $P$ with increasing labels, which contains at least one vertex in each $k$ consecutive vertices in $P$ and that contains the root and a leaf of $T$; therefore $S$ is a 1-accessible path in $T^k$. On the other hand, if there is a 1-accessible path in $T^k$, the vertices in such path define a subsequence $S$ of the vertices in some path $P$ in $T$ that make $P$ a $k$-accessible path. \[ \square \]

We denote by $\omega(g)$ any function $t$ such that $\lim_{h \to \infty} t(h)/g(h) = \infty$, and we denote by $o(g)$ any function $t$ such that $\lim_{h \to \infty} g(h)/t(h) = \infty$ (for a more formal definition of $\omega$ and $o$, the reader may change $\lim_{h \to \infty}$ by $\liminf_{h \to \infty}$). Note that $\omega(1)$ denotes a function $t$ such that $\lim_{h \to \infty} t(h) = \infty$, and $\omega(\log(h))$ denotes a function $t$ such that, for any constant $C > 0$, $t(h) \geq C \log(h)$ for $h$ large enough.

For proving Theorem 1.4 we use the following proposition, that it is an equivalent version of Theorem 1.3.

**Proposition 2.1**

$$\theta_1(T_f) = \begin{cases} 
1, & \text{if } f(h) = \frac{h}{e} + \omega(\log(h)), \\
0, & \text{if } f(h) = \frac{h}{e} + \frac{\log(h)}{2e} - \omega(1). 
\end{cases}$$

**Proof** Let $f(h)$, $g(h)$ and $\beta_h$ be such that

$$f(h) = \frac{h}{e} + g(h) \log(h) = \frac{1 + \beta_h}{e} h.$$ 

As $\beta_h = g(h) \frac{e \log(h)}{h}$, it follows that, $h \beta_h / \log h$ goes to infinity when $h$ goes to infinity, if and only if $g(h) = \omega(1)$. Therefore, by Theorem 1.3, if $f(h) = \frac{h}{e} + \omega(\log(h))$, then $f(h) = \frac{h}{e} + \omega(1) \log(h)$ and $\theta_1(T_f) = 1$.

Let $f(h)$, $g(h)$ and $\beta_h$ be such that

$$f(h) = \frac{h}{e} + \frac{\log(h)}{2e} - g(h) = \frac{1 + \beta_h}{e} h.$$ 

As $\beta_h = \frac{\log(h) - 2e g(h)}{2e}$, it follows that, $\log h - 2h \beta_h$ goes to infinity when $h$ goes to infinity, if and only if $g(h) = \omega(1)$. Therefore, by Theorem 1.3, if $f(h) = \frac{h}{e} + \frac{\log(h)}{2e} - \omega(1)$, then $\theta_1(T_f) = 0$. \[ \square \]

**3 Proof of Theorem 1.4**

Now we proceed to prove Theorem 1.4. The proof is divided in two steps to cover the two cases stated in (1.1).

**Step 1**

Let $T_f$ be the sequence of $n$-ary trees $\{T_h(f)\}_{h \in \mathbb{Z}_{+}}$. Let $k \geq 2$ and $c > 0$, in this section we assume

$$f(h) \geq \sqrt{\frac{k}{ek}} + c$$

and we prove that $\theta_k(T_f) = 1$. \[ \square \]
Let $T^k_f$ be the sequence of graphs $\{T^k_h(f)\}_{h \in \mathbb{Z}^+}$, where $T^k_h(f)$ is the $k$-transitive closure of $T_h(f)$. Let $g(h) = h/e + \omega(\log(h))$, $T^k = T^k_h(f)$ and $T' = T_{[h/k]}(g)$. By Remark 2.1, provided the limit exists, we have
\[
\theta_k(T_f) = \lim_{h \to \infty} \theta_k(T_h(f)) = \lim_{h \to \infty} \theta_1(T^k_h(f)).
\]
By Proposition 2.1
\[
\lim_{h \to \infty} \theta_1(T_h(g)) = \theta_1(T_g) = 1.
\]
To prove that $\theta_k(T_f) = 1$, it is sufficient to show that $\theta_1(T^k) \geq \theta_1(T')$.

Let $G$ be the subgraph of $T^k$, obtained by removing those vertices in $T$, whose distances from the root are not a multiple of $k$. Note that $G$ is a tree contained in $T^k$ with height $[h/k]$. Also note that, for $h$ large enough,
\[
\deg_G(v) \geq \left(\frac{\sqrt{h}/(ek)}{c} + c\right)^k \geq h/(ek) + ck(h/(ek))^{(k-1)/k} \geq h/(ek) + \omega(\log(h)) \geq g([h/k]).
\]
Thus the non-leaves vertices of $G$ have degree at least $g([h/k])$. Therefore, for $h$ large enough we have $T' \subseteq G \subseteq T^k$. Further, the probability of having at least one 1-accessible path in $T'$ is a lower bound of the probability of having at least an 1-accessible path in $T^k$.

**Step 2**

Let $T_f$ be the sequence of $n$-ary trees $\{T_h(f)\}_{h \in \mathbb{Z}^+}$. Let $k \geq 2$ and $c > 1/k$, in this section we assume $f(h) \leq \frac{\sqrt{h}/(ek)}{c} - c$ and we prove that $\theta_k(T_f) = 0$.

Let $T^k_f$ be the sequence of graphs $\{T^k_h(f)\}_{h \in \mathbb{Z}^+}$, where $T^k_h(f)$ is the $k$-transitive closure of $T_h(f)$. Let $g(h) = \frac{h}{e}$, $T = T_h(f)$, $T^k = T^k_h(f)$ and $T' = T_{[h/k]}(g)$. By Remark 2.1, provided the limit exists,
\[
\theta_k(T_f) = \lim_{h \to \infty} \theta_k(T_h(f)) = \lim_{h \to \infty} \theta_1(T^k_h(f)).
\]
By Proposition 2.1
\[
\lim_{h \to \infty} \theta_1(T_h(g)) = \theta_1(T_g) = 0.
\]
To prove that $\theta_k(T_f) = 0$, it is sufficient to show that $\theta_1(T^k) \leq \theta_1(T')$. For this, we define a tree $H^k$ such that
\[
\theta_1(T^k) \leq \theta_1(H^k) \leq \theta_1(T').
\]

Denote by $H^k$ a tree whose vertices are in correspondence with the paths in $T^k$ which starts in the root. As an outline of the construction of $H^k$ note that, any path in $T^k$ is characterized by the following two features. The first feature is the vertex $v$ in which the path ends. Suppose that $v$ is at distance $l$ to the root. The second feature is the information about the levels in which it has no vertices; it can be represented by a subset $s$ of $\{1, 2, \ldots, l-1\}$ that does not contain $k$ consecutive numbers. Therefore the vertices in $H^k$ will be determined by a vertex $v$ in $T^k$ and a $s$ subset of $\{1, 2, \ldots, l-1\}$ which does not contain $k$ consecutive numbers. About the edges in $H^k$, two vertices will be adjacent in $H^k$ if: one of the corresponding paths in $T^k$ contains the other corresponding path, and those corresponding only differ in one edge. As an example of this outline, consider the paths $v_0, v_1, v_3$ and $v_0, v_1, v_3, v_5$ in the graph $T^2$. 

\(\copyright\) Springer
Fig. 3 An example of the subgraphs of $T$ (a), $T^2$ (b) and $H^2$ (c), respectively, corresponding to six consecutive vertices in the path $v_0, v_1, v_2, v_3, v_4, v_5$ depicted Fig. 1

illustrated in Fig. 3b; in $H^2$, illustrated in Fig. 3c, those paths are represented by $v_3^{(2)}$ and $v_5^{(2,4)}$, respectively, and they are adjacent.

We define $H^k$ formally as follows.

Let $A_l$ be the set whose elements are the subsets $s$ of $\{1, 2, \ldots, l-1\}$ that do not contain $k$ consecutive numbers. Let $H^k = (V, E)$ be the graph whose vertices and edges are defined as follows. See Fig. 3.

The set of vertices in $H^k$ is defined as $V := \{v^s : v \in T \text{ and } s \in A_l, \text{ where } l \text{ is the distance of } v \text{ to the root}\}$.

The edges in $H^k$ are defined recursively as follows. Let $v$ be a vertex of $T$ at distance $l$ to the root and let $s \in A_l$. We say that $w^{s'}$ is a son of $v^s$ if

- $w$ is a son of $v$ in $T$ and $s = s'$, or
- $w$ is a descendant of $v$ at a distance $j$, $1 < j \leq k$, in $T$ and $s' = s \cup \{l + 1, l + 2, \ldots, l + j - 1\}$.

Given a vertex $v$ in $T^k$, we claim that: the paths in $T^k$ from the root to $v$, the elements in $A_l$, and the paths in $H^k$ from the root to $v^s$ (for some $s \in A_l$), are in one-to-one correspondence. For each path $P$ in $T^k$ from the root to $v$, there exist one and only one $s$ in $A_l$ whose elements are the indices $i$ such that $v_i$ is not in $P$. For each $s$ in $A_l$, there is one and only one path $P$ that starts in the root and ends in $v$; such path contains the vertices $v_i$ such that $i$ is not in $s$.

For each path in $H^k$ from the root to $v^s$, obviously there is one and only one $s \in A_l$. Now, suppose that $s$ is in $A_l$, then there is one and only one path $P$ in $H^k$ from the root to $v^s$; such path contains the vertices $v_i^{s(i)}$, where $i$ is not in $s$ and $s(i) = s \cap \{1, 2, \ldots, i - 1\}$. As an example consider: In the graph $G^2$ for the graph $G$ in Fig. 1, the path $v_0, v_1, v_3, v_5$ in $A_5$, the set $\{2, 4\}$; and in $H^2$, the path $v_0^\phi, v_1^\phi, v_3^{[2]}, v_5^{[2,4]}$ (See Fig. 3).

Assume that the vertices in $H^k$ are labelled with independent and identically distributed random variables with the same distribution as the labels in $T$.

**Lemma 3.1** $\theta_1(H^k) \geq \theta_1(T^k)$. 

\[123\]
Fig. 4 An example of a sequence of graphs in which the graph depicted in Fig. 3c is obtained from the graph in Fig. 3b. It illustrates how to obtain the sequence of graphs used in Lemma 3.1

**Proof** In what follows, we define a sequence of graphs $H_0, H_1, \ldots, H_t$, such that $H_0 = T_k$, $H_t = H_k$ and $\theta_1(H_i) \leq \theta_1(H_{i+1})$, for $i = 0, 1, \ldots, t-1$. The graph $H_0$ is obtained from $T_k$ by changing each vertex $v$ for a vertex $v^\phi$, see Fig. 4a. Given some $H_i$, we say that a vertex $v^j_i \in H_i$ is divisible if it satisfies both of the following properties: (see Fig. 5a)

- The subgraph of $H_i$, induced by $v^j_i$ and its descendants vertices, is a directed tree. We denote such tree as $T(v^j_i)$.
- There are at least two edges that start in an ancestor of $v^j_i$ and end at $v^j_i$. From such edges we denote by $e(v^j_i)$ the edge that starts in the ancestor $x = v^j_i'$ of $v^j_i$ such that $v^j_j'$ is the nearest to the root in $T$.

Whenever $H_i$ has at least one divisible vertex, we define $H_{i+1}$ as follows. Let $v^j_i$ be a divisible vertex in $H_i$. Further, let $v_0, v_1, \ldots, v_j$ be a path in $T$ from the root to $v_j$, then $v^{s(0)}_0, v^{s(1)}_1, \ldots, v^{s(j)}_j$ is a path in $H_i$ from the root to $v^j_i$. Furthermore, let $x = v^{s(j)}_{j'}$ be
the vertex where \( e(v^i_j) \) starts. Let \( Q \) be the indices of the vertices that \( e(v^i_j) \) jumps, i.e. \( Q \coloneqq \{j + 1, j + 2, \ldots, j - 1\} \). Let \( T^Q \) be a copy of \( T(v^i_j) \) where each vertex \( u^i_j \) is changed by \( u^i_j \cup Q \). Thus, \( H_{i+1} \) is defined as the graph obtained from \( H_i \) by removing \( e(v^i_j) \), adding \( T^Q \), and adding an edge \( e'(v^i_j) \) that starts in \( x \) and ends at the root of \( T^Q \). See Fig. 4.

The reader may notice that, if \( H_i \) does not have divisible vertices, then \( H_i = H_k \). It remains to prove that \( \theta_1(H_i) \leq \theta_1(H_{i+1}) \).

Given a vertex \( y \) in a labeled directed graph \( G \), we denote by \( [y \twoheadrightarrow G] \) the event of having a path in \( G \) that starts in \( y \), ends at a sink of \( G \) and has increasing labels; we denote by \( [y \not\twoheadrightarrow G] \) its complement. For the case when \( y \) is the root it is denoted by \( 0 \).

Consider in the graph in Fig. 5a, the event of having an 1-accessible path that contains the edge \( e(v^i_j) \). Note that, such event occurs, if and only if, there is an increasing path from 0 to \( x \), \( w(x) < w(v^i_j) \) and \( [v^i_j \twoheadrightarrow T(v^i_j)] \). Similarly, consider the event in Fig. 5b, the event of having an 1-accessible path that contains the edge \( e'(v^i_j) \). Note that, such event occurs, if and only if, there is an increasing path from 0 to \( x \), \( w(x) < w(v^i_j \cup Q) \) and \( [v^i_j \cup Q \twoheadrightarrow T^Q] \).

With the notation introduced in the construction of \( H_{i+1} \), define \( \Gamma_1 \) has the event of having an 1-accessible path in \( H_i \) that contains the edge \( e(v^i_j) \), and define \( \Gamma_2 \) has the event of having an 1-accessible path in \( H_{i+1} \) that contains the edge \( e'(v^i_j) \).

This lemma follows from the following facts:

- If \( A_1 \) is the event of having an increasing path from 0 to \( x \) and \( w(x) < w(v^i_j) \), and \( A_2 \) is the event of having an increasing path from 0 to \( x \) and \( w(x) < w(v^i_j \cup Q) \), then \( \Gamma_1 \) is the event of having both \( A_1 \) and \( [v^i_j \twoheadrightarrow T(v^i_j)] \), and then \( \Gamma_2 \) is the event of having both \( A_2 \) and \( [v^i_j \cup Q \twoheadrightarrow T^Q] \). Therefore, as \( \mathbf{P}(A_1) = \mathbf{P}(A_2) \) and \( \mathbf{P}([v^i_j \twoheadrightarrow T(v^i_j)]|A_1) = \mathbf{P}([v^i_j \cup Q \twoheadrightarrow T^Q]|A_2) \) then \( \mathbf{P}(\Gamma_1) = \mathbf{P}(\Gamma_2) \).
- As the only edge that joins \( T^Q \) with the vertices in \( H_i \) is \( e'(v^i_j) \), and \( H_i \) and \( H_{i+1} \) only differ in \( e(v^i_j) \), \( e'(v^i_j) \) and \( T^Q \), then \( H_i - e(v^i_j) = H_{i+1} - \{e'(v^i_j), T^Q\} \). Therefore, as there are no paths in \( H_{i+1} - e'(v^i_j) \) from the root to a leaf in \( T^Q \), we have
By definition we require to prove that: the probability of having a 1-accessible path in $H$.

Let $[0 \sim x]$ be the event of having an increasing path from 0 to $x$. Then

$$
\mathbf{P}([0 \sim H - e(v)] \mid \Gamma_1) \geq \mathbf{P}([0 \sim H - e(v)] [0 \sim x], w(x) < w(v_j^*)
$$

$$
= \mathbf{P}([0 \sim H_{i+1} - \{e'(v_j^*), T^Q\}] [0 \sim x], w(x) < w(v_j^*)
$$

$$
= \mathbf{P}([0 \sim H_{i+1} - \{e'(v_j^*), T^Q\}] \mid \Gamma_2)
$$

$$
= \mathbf{P}([0 \sim H_{i+1} - e'(v_j^*)] \mid \Gamma_2)
$$

Let $\theta_1(H_{i+1}) - \theta_1(H_i) \geq 0$.

$$
\theta_1(H_{i+1}) - \theta_1(H_i) = \mathbf{P}([0 \sim H_{i+1}] - \mathbf{P}([0 \sim H_i])
$$

$$
= \mathbf{P}([0 \sim H_{i+1}] \cap [0 \sim H_{i+1} - e'(v)] - \mathbf{P}([0 \sim H_i] \cap [0 \sim H_i - e(v)])
$$

$$
+ \mathbf{P}([0 \sim H_{i+1}] \cap [0 \sim H_{i+1} - e'(v)] - \mathbf{P}([0 \sim H_i] \cap [0 \sim H_i - e(v)])
$$

$$
= \mathbf{P}([0 \sim H_{i+1} - e'(v)]) - \mathbf{P}([0 \sim H_i - e(v)])
$$

$$
+ \mathbf{P}([0 \sim H_{i+1} - e'(v)]) - \mathbf{P}([0 \sim H_i - e(v)])
$$

$$
= \mathbf{P}(\Gamma_2) - \mathbf{P}(\Gamma_2 \cap [0 \sim H_{i+1} - e'(v)])
$$

$$
= \mathbf{P}(\Gamma_1) \left[ \mathbf{P}([0 \sim H_i - e(v)] \mid \Gamma_1) - \mathbf{P}([0 \sim H_{i+1} - e'(v)] \mid \Gamma_2) \right] \geq 0. \quad \square
$$

Lemma 3.2 $\theta_1(T') \geq \theta_1(H^k)$.

Proof By definition we require to prove that: the probability of having a 1-accessible path in $T'$, is an upper bound for the probability of having a 1-accessible path in $H^k$.

Let $H' \subseteq H^k$ be the tree induced by the vertices of $H^k$ at distance at most $h/k$ to the root. Notice that $\theta_1(H') \geq \theta_1(H^k)$. Also, notice that, if $v$ is a vertex of $H'$ at distance at most $h/k - 1$ to the root, then (for $h$ large)

$$
\deg_{H'}(v) = \sum_{j=1}^{k} \lfloor f(h) \rfloor^j \leq \sum_{j=1}^{k} f(h)^j = \frac{f(h)^{k+1} - 1}{f(h) - 1}.
$$

We claim that $\deg_{H'}(v) \leq g([h/k])$ (recall that $g$ is defined as $g(h) = \frac{h}{e}$), from which $H' \subseteq T'$ and, as both trees have the same height, then $\theta_1(T') \geq \theta_1(H')$.

Now, to conclude this proof, we show that the claim holds. Let $A = \sqrt[3]{h/(ek)}$. It is sufficient to prove that

$$
\frac{(A - c)^{k+1} - 1}{(A - c) - 1} \leq A^k.
$$

But it follows from that

$$
(A - c)^{k+1} - 1 = A^{k+1} - (k + 1)cA^k + o(A^{k-1})
$$

and

$$
A^k (A - c - 1) = A^{k+1} - (c + 1)A^k.
$$
Acknowledgements The authors are thankful to Ricardo Restrepo and Daya K. Nagar for helpful discussions. Thanks are due to the anonymous referees for their careful reading, criticism and suggestions which helped us to considerably improve the paper. Research Partially supported by FORDECYT 265667 (Mexico). Research Partially supported by Universidad de Antioquia (Colombia).

References

1. Berestycki, J., Brunet, E., Shi, Z.: The number of accessible paths in the hypercube. Bernoulli 22(2), 653–680 (2016)
2. Berestycki, J., Brunet, É., Shi, Z.: Accessibility percolation with backsteps. ALEA Lat. Am. J. Probab. Math. Stat 14, 45–62 (2017)
3. Franke, J., Klözer, A., de Visser, J.A.G.M., Krug, J.: Evolutionary accessibility of mutational pathways. PLOS Comput. Biol. 7(8), 1–9 (2011)
4. Gillespie, J.H.: Molecular evolution over the mutational landscape. Evolution 38(5), 1116–1129 (1984)
5. Hegarty, P., Martinsson, A.: On the existence of accessible paths in various models of fitness landscapes. Ann. Appl. Probab. 24(4), 1375–1395 (2014)
6. Li, L.: Phase transition for accessibility percolation on hypercubes. J. Theor. Probab. 31, 2072–2111 (2018)
7. Martinsson, A.: Accessibility percolation and first-passage site percolation on the unoriented binary hypercube. arXiv:1501.02206 (2015)
8. Nowak, S., Krug, J.: Accessibility percolation on n-trees. EPL (Europhys. Lett.) 101(6), 66004 (2013)
9. Richter, H., Engelbrecht, A.: Recent advances in the theory and application of fitness landscapes. Springer, New York (2014)
10. Roberts, M., Zhao, L.: Increasing paths in regular trees. Electron. Commun. Probab. 18(87), 1–10 (2013)
11. Stadler, P.F., Stephens, C.R.: Landscapes and effective fitness. Comment Theor. Biol. 8((4–5)), 389–431 (2003)
12. Weinreich, D.M., Chao, L.: Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. Evolution 59(6), 1175–1182 (2005)
13. Weinreich, D.M., Watson, R.A., Chao, L.: Perspective: sign epistasis and genetic constraint on evolutionary trajectories. Evolution 59, 1165–1174 (2005)
14. Weinreich, D.M., Delaney, N.F., DePristo, M.A., Hartl, D.L.: Darwinian evolution can follow only very few mutational paths to fitter proteins. Science 312(5770), 111–114 (2006)
15. Weissman, D.W., Desai, M.M., Fisher, D.S., Feldman, M.W.: The rate at which asexual populations cross fitness valleys. Theor. Popul. Biol. 75, 286–300 (2009)
16. Wright, S.: The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proc. Sixth Int. Congress Genet. 1, 356–366 (1932)