Accessibility percolation and first-passage percolation on the hypercube

Anders Martinsson

Department of Mathematical Sciences,
Chalmers University of Technology and University of Gothenburg

Abstract

In this thesis, we consider two percolation models on the n-dimensional binary hypercube, known as accessibility percolation and first-passage percolation. First-passage percolation randomly assigns non-negative weights, called passage times, to the edges of a graph and considers the minimal total weight of a path between given end-points. This quantity is called the first-passage time. Accessibility percolation is a biologically inspired model which has appeared in the mathematical literature only recently. Here, the vertices of a graph are randomly assigned heights, or fitnesses, and a path is considered accessible if strictly ascending. We let $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$ denote the all zeroes and all ones vertices respectively.

A natural simplification of both models is the restriction to oriented paths, i.e. paths that only flip 0:s to 1:s. Paper I considers the existence of such accessible paths between $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$ for fitnesses assigned according to the so-called House-of-Cards and Rough Mount Fuji models. In Paper II we consider the first-passage time between $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$ in the case of independent standard exponential passage times. It is previously known that, in the oriented case, this quantity tends to 1 in probability as $n \to \infty$. We show that without this restriction, the limit is instead $\ln\left(1+\sqrt{2}\right) \approx 0.88$. By adapting ideas in Paper II to unoriented accessibility percolation, we are able to determine a limiting probability for the existence of accessible paths from $\hat{\mathbf{0}}$ to the global fitness maximum. This is presented in Paper III.

Key words and phrases: hypercube, percolation, accessible path, house of cards, rough mount Fuji, first-passage percolation, Richardson's model, branching translation process

List of Papers

This thesis consists of the following papers:

- Peter Hegarty and Anders Martinsson,
  "*On the existence of accessible paths in various models of fitness landscapes*",
  in *Ann. Appl. Probab.* **24** (2014), no. 4, 1375-1395.

- Anders Martinsson,
  "*Unoriented first-passage percolation on the n-cube*",
  Submitted

- Anders Martinsson,
  "*Accessibility percolation and first-passage site percolation on the unoriented binary hypercube*",
  Preprint

Acknowledgements

My deepest gratitude goes to my advisor, Peter Hegarty, for his constant support and numerous insightful comments and discussions, which have been vital in the making of this thesis. May you have that which is good! I also wish to thank my co-advisor, Johan Wästlund, as well as Jeff Steif and Johan Tykesson for valuable input on my research.

Thanks to Joachim Krug for introducing me to the topic of accessibility percolation, and to him and Stefan Nowak for helpful discussions on the topic.

I would further like to extend my gratitude to my former and current colleagues at the department for contributing to this inspiring and welcoming working atmosphere, to the many of you I have had as teachers throughout my years at Chalmers, and to the friends I have met here.

Lastly, many thanks to my family and friends outside of the department for helping to keep me sane.

Anders Martinsson
Göteborg, December 10, 2014

Contents

Chapter 1

Introduction

Percolation theory is a branch of probability theory which is roughly speaking concerned with connectivity in a random medium. It encompasses many simple, seemingly innocent random models whose rigorous investigations have turned out highly non-trivial. Since its introduction in 1957 by Broadbent and Hammersley, this field has gained much popularity, both for its mathematical appeal and for its relevance in various applications.

The classical example of percolation is the following model for a porous stone submerged in a bucket of water: Assume that the bulk of the stone contains pores placed periodically in the shape of a d-dimensional integer lattice, $\mathbb{Z}^d$, for some integer $d \geq 2$[1]. We shall here consider $\mathbb{Z}^d$ as a graph by connecting two lattice points by an edge if their Euclidean distance is 1. For each edge, i.e. for each pair of adjacent pores, we flip a coin which turns up heads with probability p, and tails with probability $1 - p$. If heads, then the edge is considered *open*, meaning that the corresponding pores are connected and water can flow between them. Otherwise, the edge is considered *closed* and no water can pass.

Depending on p, we can get two qualitatively different behaviors of this graph. Either all components formed by the open edges are finite, or there is an infinite region connected by open edges, which we interpret as meaning that that water cannot, respectively can, seep into the stone.

[1] For a 3-dimensional stone (as most real stones are) one should reasonably take $d = 3$, though the problem is mathematically interesting for any $d \geq 2$ and the most studied case is $d = 2$.

The model described above is often referred to as *Bernoulli percolation*, *bond percolation*, or simply *percolation* [14]. Here we defined it specifically on the integer lattice, but we can analogously define Bernoulli percolation on any graph, though whether or not a certain question is meaningful may depend on this choice.

In this thesis, we will consider two percolation models, not in the integer lattice, but instead when the underlying graph is the *n-dimensional binary hypercube*, $\mathbb{Q}_n$, for large n. This is the graph whose vertices are the length n binary strings, and where two vertices are connected by an edge if they differ at exactly one position. This graph is sometimes also referred to as the Hamming cube. We shall specifically be concerned with what is known as *accessibility percolation* and *first-passage percolation* on this graph. First-passage percolation is one of the more well-known variations of the bond percolation model above. Here all edges are considered open, but for each edge there is an associated random time, or cost, required to traverse it. Accessibility percolation is a simple biologically inspired model for connectivity in an evolutionary landscape, which has gained some recent attention. This model randomly assigns heights to the vertices of a graph and only allows traversing edges if the new vertex is higher.

Similarly to Bernoulli percolation, the classical setting for first-passage percolation is on the integer lattice. Fill and Pemantle [11] proposed first-passage percolation on the hypercube as a way to capture the high-dimensional limiting behavior of the lattice case. Because the hypercube has an inherent geometry, it is arguably also more in the spirit of the classical setting than for example trees, which is popular alternative to the lattice. In contrast, the hypercube is the natural choice in accessibility percolation. This dates back to ideas about the topography of evolutionary landscapes by Wright in 1932.

Below, we let $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$ denote the all zeroes and all ones vertices respectively in the hypercube. For any two length n binary strings, their *Hamming distance* is the number of locations where the strings differ. The *oriented binary n-dimensional hypercube*, $\mathbb{Q}_n^{\mathcal{O}}$, is the digraph obtained by orienting each edge of $\mathbb{Q}_n$ towards the vertex that has the extra one. Hence, traversing an edge in $\mathbb{Q}_n$ corresponds to flipping a bit in the binary string, and traversing an edge in $\mathbb{Q}_n^{\mathcal{O}}$ corresponds to changing a '0' to a '1'.

Chapter 2

Accessibility Percolation

In Papers I and III we will consider a problem with motivation from biological evolution. This subject is based on rather old ideas from the biology literature, but has only appeared in its current form in the last few years. Due to its similarity to percolation, Nowak and Krug named this problem accessibility percolation in [21]. Here the underlying structure is thought of as the possible genotypes of a species. The idea is to try to capture the notion of whether or not a certain advantageous genotype is selectively accessible from some initial position.

Work on accessibility percolation prior to Paper I, e.g. [7, 12], consists mostly of simulations. Paper I is the first article in the mathematical literature on this subject, and the first to prove significant rigorous results. Since this paper, a number of articles on the subject have appeared in the mathematical literature. Besides Papers I and III, there are also [4, 5] for accessibility percolation on the hypercube, and [3, 21, 23] for accessibility percolation on rooted regular trees.

2.1 Biological motivation

Suppose we want to model the evolution of a population of haploid organisms, reproducing asexually without recombination. A common assumption in describing the set of possible genotypes for such organisms is that their genome contains n loci (i.e. positions in the genome where specific genes are stored) that are susceptible to point mutations, and that for each locus there are two possible states (or alleles). One typically thinks of n as very large.

We shall designate the initially most prevalent allele at each locus its *wild state* and the less prevalent as the *mutated state*. By denoting the wild state by a '0' and the mutated one by a '1', the possible genotypes can be encoded as binary strings of length n. Here it is useful to give this space a graph structure by considering the possible genotypes as vertices and letting two vertices share an edge if it is possible for a single mutation to change one genotype into the other, i.e. if the Hamming distance between their corresponding binary strings is 1. We note that this graph is precisely the n-dimensional binary hypercube [17, 26].

An important concept in evolutionary theory is the notion of a fitness landscape, as first introduced by Wright in 1932. Assume that the reproductive success of an organism can be measured by a real number $\omega(v)$, uniquely defined by its genotype v. Wright's fitness landscape is the idea to visualize this relationship between genotypes and reproductive success by considering the possible genotypes as positions in a "hilly landscape", where the height of each point is given by its fitness. Hence, the evolutionary process can be seen as a kind of random walk on this landscape with a bias towards moving uphill.

In a remarkable recent development, a number of experimental studies have begun to map local regions of empirical fitness landscapes by constructing all or a large subset of possible combinations of up to 9 mutations, and measuring the reproductive success of each. Surveys of this data are given in [24, 25].

Since the relationship between genotype and fitness is complicated and still largely unknown, it makes sense to try to model the fitnesses by a random distribution. A number of models have been suggested to this end. One such model that has gained popularity in the biological community is Kauffman's NK model [17, 18]. Unfortunately, from a mathematical point of view, this model is very complex and the only rigorous results in the setting we are interested in seems to be the ones given in [10, 20] for the distribution of local fitness maxima. We shall not delve further into the NK model in this thesis, and instead restrict our attention to the simpler House-of-Cards and Rough Mount Fuji models as described below.

A question of long history in evolutionary theory is how accessible real fitness landscapes are to the evolutionary process. Here there are two conflicting intuitions. On the one hand, as Wright's illustration would seem to indicate, the landscape may consist of many small hills isolated by valleys of considerably lower fitness, which would mean that evolution would be restricted to a relatively small set of genotypes near a local optimum. On the other hand, the topography of the hypercube is very different to that of the plane, and the sheer number of different paths between two vertices in the hypercube would seem to indicate that some paths are bound to be nice, meaning that a large portion of the viable states should, at least in principle, be accessible [12].

Following in the footsteps of [13, 17, 26] we shall here take a rather strict definition of evolutionary accessibility by assuming the strong selection/weak mutation (SSWM) regime: The population has one predominant genotype, which may change over time. Occasionally a point mutation generates an organism of a genotype adjacent to the currently predominant one. If the new organism is fitter it has a chance to overtake the population. If it is not fitter, then its lineage will die out before it has a chance to mutate again. Hence, given an initial predominant genotype $\hat{\mathbf{0}}$, we see that it is possible for a genotype v to become the predominant one if and only if there is a path from $\hat{\mathbf{0}}$ to v in $\mathbb{Q}_n$ along which fitness is strictly increasing.

2.2 Mathematical formulation

Combining the ideas described in the preceding section, the question of accessibility in fitness landscapes boils down to a very concrete mathematical model: We define a fitness landscape as a graph $G = (V, E)$ together with a fitness function $\omega : V \to \mathbb{R}$, which is generated according to some random distribution. We say that a path $v_0 \to v_1 \to \ldots \to v_l$ in G is *accessible* if

$$\omega(v_0) < \omega(v_1) < \ldots < \omega(v_l). \tag{2.1}$$

For $u, v \in V$ we say that v *is accessible from* u if there exists an accessible path from u to v. Of primary concern is which vertices are accessible from some given starting point. When G is the hypercube, we will always take this starting point to be $\hat{\mathbf{0}}$. In particular, is the global fitness maximum accessible?

From the biological perspective, the most natural choice of G is arguably $\mathbb{Q}_n$ by the reasoning given in the previous section. However, the first papers on accessibility percolation on the hypercube, [7,12], Paper I and [4], instead consider this problem on the oriented hypercube $\mathbb{Q}_n^{\mathcal{O}}$. The interpretation of replacing $\mathbb{Q}_n$ by $\mathbb{Q}_n^{\mathcal{O}}$ is that "mutational reversions", i.e. changing a gene back to its wild state, become disallowed. In [12], this is motivated by stating that as mutations are assumed rare, the paths without reversions are the most feasible ones because they require the least number of mutations to reach the target vertex. It should however be noted that the oriented hypercube is combinatorially a much simpler graph than the unoriented one.

For a probabilist, the most natural choice of distribution of fitnesses is independent and identically distributed of, say, some continuous distribution (to avoid ties). By convention, we will use $U(0,1)$, the uniform distribution on the interval $[0,1]$. This way to assign fitnesses is often called Kingman's *House-of-Cards model* [19]. Kauffman and Levin [17] instead refer to this as an uncorrelated landscape. For accessibility percolation we will however modify this slightly by choosing some vertex $\hat{\boldsymbol{v}}$, which we assign the fitness 1. As accessibility percolation only considers the order of fitnesses, this is equivalent to conditioning on $\hat{\boldsymbol{v}}$ being the global fitness maximum. In particular, if $\hat{\boldsymbol{v}}$ is chosen uniformly at random, this is equivalent to the original formulation of the House-of-Cards model, with $\hat{\boldsymbol{v}}$ denoting the global fitness maximum. When considering this model on the oriented hypercube, one usually takes $\hat{\boldsymbol{v}} = \hat{\mathbf{1}}$.

Another simple way to assign fitnesses which has been proposed in the literature is the *Rough Mount Fuji model* [1]. It includes two parameters, a fixed probability distribution η, and a positive number θ, called the *drift*, which may depend on the dimension n. For each $v \in \mathbb{Q}_n$, or $\mathbb{Q}_n^{\mathcal{O}}$, one lets

$$\omega(v) = \theta \cdot d(v, \hat{\mathbf{0}}) + \eta(v). \tag{2.2}$$

In other words, one first assigns a fitness to each node at random, according to η, and independent of all other nodes. Then the fitness of each node is shifted upwards by a fixed multiple of the Hamming distance from $\hat{\mathbf{0}}$.

It probably comes as no surprise that the House-of-Cards model is not considered very realistic. House-of-Cards landscapes are extremely rugged with no correlation between adjacent genotypes. The fitness of a genotype one mutation away from the global fitness maximum is no better on average than the fitness of a random genotype (in fact, it is slightly worse). Due to its mathematical simplicity it is often still considered, but presented as a null model.

The Rough Mount Fuji model is a natural way to construct a smoother landscape and introduce an "arrow of evolution", since the drift factor implies that successive $0 \to 1$ mutations will tend to increase fitness. Despite its simplicity, comparisons to empirical fitness landscapes in [24] found that the Rough Mount Fuji model "captures the features of the experimental landscapes remarkably well".

Chapter 3

First-Passage Percolation

In appended Paper II, we will consider a variation of percolation known as first-passage percolation on the n-dimensional binary hypercube.

First-passage percolation is a random process on a graph $G = (V, E)$ proposed by Hammersley and Welsh in 1965. A closely related model, the Eden growth model, was proposed by Eden in 1961. In first-passage percolation, the edges of G are assigned weights $\{\tau_e\}_{e \in E}$, called *passage times*, according to non-negative, independent and identically distributed random variables. We define the *passage time of a path* Γ by $T(\Gamma) = \sum_{e \in \Gamma} \tau_e$. For two vertices $u, v \in V$, we define the *first-passage time from* u *to* v by

$$T(u, v) = \inf \{T(\Gamma) : \Gamma \text{ from } u \text{ to } v\} . \tag{3.1}$$

Equivalently $T(u, v)$ denotes the shortest distance from u to v with respect to the edge weights. We call a path from u to v that attains the infimum in (3.1) a *geodesic* between the two points.

Often when considering first-passage percolation, one does not consider $T(u, v)$ directly, but instead use it to define a random model for the spread of some property throughout the graph. For simplicity, we will follow the same convention as for Bernoulli percolation and assume that we are modeling the spread of water through some random medium (i.e. the property considered is "wetness"). When seen in this light, first-passage percolation can be described as the following modification of the original percolation model: We assume that water can flow between any pair of adjacent pores in the stone, but flowing across a certain edge e takes some amount of time τ_e determined by the random structure of the stone. Hence the question is no longer if it is possible for water to spread through the structure, but rather how fast it happens. Let us make this a bit more precise. Let v_0 be some vertex in G. Suppose at time 0 we connect v_0 to a water source, and that water then flows from v_0 along all possible paths at the speeds dictated by their respective passage times. It follows that the time at which water first reaches a vertex v is given by $T(v_0, v)$. Hence, the set of vertices that is reached within time t is given by

$$B_t = \{v \in V : T(v_0, v) \leq t\} . \tag{3.2}$$

We will refer to this set as the *wet region* at time t.

One interesting special case of first-passage percolation is when passage times are exponentially distributed with mean one. In this case, it turns out that $\{B_t\}_{t\geq 0}$ is Markovian. This is an implication of the so-called memory-less property of the exponential distribution. More precisely, $\{B_t\}_{t\geq 0}$ can equivalently be described as the following continuous-time Markov chain: Initially we have $B = \{v_0\}$, and for each $v \in V \setminus B$, the transition $B \to B \cup \{v\}$ occurs at rate $|N(v) \cap B|$, where $N(v)$ denotes the neighborhood of v. This Markov chain is known as Richardson's model [9, 22]. Instead of modeling the spread of water, this process is often interpreted as the spread of an infection, where B_t denotes the set of infected vertices at time t. Informally, the transition rates can be interpreted as meaning that, if you are infected, you sneeze on each of your neighbors according to independent Poisson processes with rate 1. If you get sneezed on, you get sick.

Similarly to Bernoulli percolation, the classical setting for first-passage percolation is the lattice case, i.e. when $G = \mathbb{Z}^d$ for some $d \geq 2$. A central result in this setting is the Shape Theorem which is due to Cox and Durrett [8], see also [16, 22]. For any distribution of edge weights τ, let B_t denote the corresponding wet region starting at the origin. We define a smoothed version of B_t given by

$$\overline{B}_t = B_t + [-0.5, 0.5]^d . \tag{3.3}$$

The Shape Theorem states that under some regularity conditions on τ, there exists a compact convex set $B^* \subseteq \mathbb{R}^d$ with non-empty interior such that for any $\varepsilon > 0$ we almost surely have

$$(1 - \varepsilon)\, B^* \subseteq \frac{1}{t}\overline{B}_t \subseteq (1 + \varepsilon)\, B^* \tag{3.4}$$

for sufficiently large t. The natural interpretation of this statement is that the wet region grows linearly in t, and when scaled appropriately approaches a limit shape B^*.

There is a lot of literature on the subject of first-passage percolation on $\mathbb{Z}^d$, but many aspects of this process are still poorly understood. The limit shape is an important example of this. We know from the Shape Theorem that B^* is always convex and must have the obvious symmetries inherited from $\mathbb{Z}^d$, but beyond that very little is known about the geometry of this set. I will not attempt to give an overview of what is known for the integer lattice, partly because I do not know it sufficiently well. The reader is instead referred to either of [15, 16].

3.1 First-passage percolation on the hypercube

We now turn to the case where G is the n-dimensional hypercube. Here we will always assume standard exponential passage times, hence the wet region spreads according to Richardson's model.

For the hypercube, first-passage percolation has a slightly different flavor than on the lattice. Since $\mathbb{Q}_n$ is a finite graph for each n, the behavior of the wet region as $t \to \infty$ is trivial; for sufficiently large t the wet region is simply the entire cube. Instead, we will consider the limiting behavior of first-passage percolation on $\mathbb{Q}_n$ as $n \to \infty$. We will for simplicity always assume that B_t starts from $\hat{\mathbf{0}}$, though it can be noted that, by symmetry, the choice of starting vertex does not matter.

We will consider two natural quantities for first-passage percolation on the hypercube. Firstly, the first-passage time between antipodal vertices (e.g. $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$). Equivalently, this is the random time when $\hat{\mathbf{1}}$ becomes infected. We will denote this time by T_n rather than $T(\hat{\mathbf{0}}, \hat{\mathbf{1}})$ to emphasize the dependence of the dimension of the hypercube. Secondly, the *covering time of* $\mathbb{Q}_n$, which is the random time at which the entire graph becomes infected. We will denote this by C_n.

One would intuitively expect the vertex antipodal to the source to be the hardest to reach. More precisely, one would expect that $T(\hat{\mathbf{0}}, \hat{\mathbf{1}})$ is stochastically larger than $T(\hat{\mathbf{0}}, v)$ for any other vertex v. Interestingly, this seems to be an open problem. This is not to say that $\hat{\mathbf{1}}$ is likely to become infected last however, and indeed this turns out not to be true, see Theorem 3.1 below.

A third and more tractable quantity, the *oriented first-passage time from* $\hat{\mathbf{0}}$ *to* $\hat{\mathbf{1}}$ also appears in the literature. This is the first-passage time from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$, but with respect to the oriented hypercube, $\mathbb{Q}_n^{\mathcal{O}}$. We will denote this by $T_n^{\mathcal{O}}$. We will couple $T_n^{\mathcal{O}}$ to T_n and C_n by using the same edge passage times for $\mathbb{Q}_n^{\mathcal{O}}$ as for $\mathbb{Q}_n$.

The literature on first-passage percolation on the hypercube prior to Paper II consists of two papers, [11] by Fill and Pemantle (1993) and [6] by Bollobás and Kohayakawa (1997). The result for this process in [11] is summarized in the following theorem. To simplify notation, we say that a sequence of random variables X_n lies asymptotically between two constants $a < b$ if for each $\varepsilon > 0$ we have $a - \varepsilon \leq X_n \leq b + \varepsilon$ with probability tending to 1 as $n \to \infty$.

Theorem 3.1. *(Fill, Pemantle)*

1. $T_n^{\mathcal{O}} \to 1$ *in probability as* $n \to \infty$.
2. T_n *lies asymptotically between* $\ln\left(1 + \sqrt{2}\right) \approx 0.88$ *and* 1
3. C_n *lies asymptotically between* $\frac{1}{2}\ln\left(2 + \sqrt{5}\right) + \ln 2 \approx 1.41$ *and* $3 + 2\ln\left(4 + 2\sqrt{2}\right) \approx 14.0$.

This theorem deserves a few comments. It is straight-forward to show that $T_n^{\mathcal{O}}$ must be at least $1 - o(1)$ with probability tending to 1 as $n \to \infty$ by considering the expected number of oriented paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ with passage time at most t. This had previously been observed by Aldous in [2]. Using a second moment analysis, Fill and Pemantle were able to prove a corresponding upper bound, implying that $T_n^{\mathcal{O}}$ tends to 1 in probability as $n \to \infty$. As $T_n \leq T_n^{\mathcal{O}}$, this also implies that T_n is at most $1 + o(1)$ with probability tending to 1 as $n \to \infty$. They remark that they doubt this upper bound for T_n is sharp, but state that they do not know how to improve it. Prior to Paper II, this seems to be the best known upper bound on T_n.

The lower bound on T_n is due to Durrett, even though he is not officially one of the authors of the paper. The main idea of the proof is to consider a certain branching process, which Durrett calls the branching translation process, or BTP, whose individuals live on the vertices of the hypercube. It is a simple observation that this process is an upper bound on Richardson's model in the sense that, when coupled appropriately, the vertices get occupied by individuals in the BTP before they become infected. This means that T_n is bounded from below by the time the first individual at $\hat{\mathbf{1}}$ is born. Durrett shows that this time converges to $\sinh^{-1}(1) = \ln\left(1+\sqrt{2}\right)$ in probability as $n \to \infty$.

In [6] Bollobás and Kohayakawa improved the upper bound on C_n, by relating it to T_n.

Theorem 3.2. *(Bollobás, Kohayakawa) Let*

$$T_\infty = \inf\left\{t \geq 0 : \mathbb{P}\left(T_n \leq t\right) \to 1 \text{ as } n \to \infty\right\}. \tag{3.5}$$

Then, with probability tending to 1 *as* $n \to \infty$, *we have* $C_n \leq T_\infty + \ln 2 + o(1)$. *In particular by the result of Fill and Pemantle, we have* $T_\infty \leq 1$, *hence asymptotically almost surely* $C_n \leq 1.69\cdots + o(1)$.

Actually, their result is more general but also considerably more technical. Essentially what it states is as follows: Take any two vertices u, v in $\mathbb{Q}_n$. Then, disregarding possible bottlenecks from u to $N(u)$ and from $N(v)$ to v, it is extremely unlikely that $T(u,v)$ is larger than $T_\infty + o(1)$. Applying this to the cover time, it turns out that the worst-case bottlenecks that are likely to occur are those where every edge incident to the end-point has passage time at least $\ln 2 - o(1)$, hence the additional term of $\ln 2$ in the upper bound of C_n.

Bollobás and Kohayakawa referred to the quantity T_∞ as simply the first-passage percolation time. Note that if T_n has a limit in probability as $n \to \infty$, then it must converge to T_∞. Indeed, they conjectured that this was the case, hence the notation T_∞.

Prior to Paper II, it has been an open problem how to improve these results. Fill and Pemantle state that in the same way that the upper bound for $T_n^{\mathcal{O}}$ is harder to prove than the lower one, it is reasonable to expect that the same holds true for T_n. Durrett's branching process is the only method to yield significant results for T_n beyond those obtained from oriented first-passage percolation. However, beyond the fact that this process dominates Richardson's model, the connection between these processes is rather vague, and it is hence not clear how this process could be used to prove an upper bound on T_n.

Chapter 4

Summary of Appended Papers

4.1 Paper I: On the existence of accessible paths in various models of fitness landscapes

(coauthored with Peter Hegarty)

In this paper, we consider accessibility percolation on $\mathbb{Q}_n^{\mathcal{O}}$ as $n \to \infty$, in the cases where fitnesses are assigned according to the House-of-Cards or the Rough Mount Fuji models.

Let us first consider the House-of-Cards case, i.e. fitnesses are assigned according to $\omega(\hat{\mathbf{1}}) = 1$ and independently $\omega(v) \sim U(0,1)$ for all other vertices. We let X denote the number of accessible paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$.

As there are $n!$ oriented paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$, and each path is accessible if and only if the n random fitnesses along the path are in ascending order, we see that

$$\mathbb{E}X = 1. \tag{4.1}$$

At first glance, this may seem to imply a positive probability of accessible paths existing. However, a much clearer picture of what occurs is obtained by conditioning on the fitness of the starting vertex. Following the convention in [4, 5], for $\alpha \in [0,1]$ we let $\mathbb{P}^{\alpha}$ and $\mathbb{E}^{\alpha}$ denote conditional probability and expectation given

$\omega(\hat{\mathbf{0}}) = \alpha$. In the paper, we refer to this conditional distribution as α-constrained House-of-Cards. We see that

$$\mathbb{E}^\alpha X = n(1-\alpha)^{n-1}. \tag{4.2}$$

We can interpret (4.2) in the following way: There is a critical threshold for α at $\frac{\ln n}{n} + O\left(\frac{1}{n}\right)$ such that above this threshold, the expected number of accessible paths tends to 0 as $n \to \infty$. Hence, asymptotically almost surely there are no accessible paths when α is above this value. Below the threshold, we have that the expected number of paths tends to infinity as $n \to \infty$. Strictly speaking, this by itself does not tell us anything about the probability that such paths exist (other than that it cannot be exactly zero), but a natural guess is that the probability should tend to one in this case.

The first main result of Paper I shows that this analysis "tells the truth" about the behavior of X:

Theorem 4.1. *For any sequence* $\{\varepsilon_n\}_{n=1}^\infty$ *such that* $n\varepsilon_n \to \infty$, *as* $n \to \infty$ *we have*

$$\mathbb{P}^{\frac{\ln n}{n}+\varepsilon_n}(X \geq 1) \to 0 \tag{4.3}$$

$$\mathbb{P}^{\frac{\ln n}{n}-\varepsilon_n}(X \geq 1) \to 1. \tag{4.4}$$

Furthermore,

$$\mathbb{P}(X \geq 1) \sim \frac{\ln n}{n}. \tag{4.5}$$

In the case where $\alpha = O\left(\frac{1}{n}\right)$, this theorem has later been strengthened by Berestycki, Brunet and Shi [4] who determine a limit distribution for the number of accessible paths.

In order to state our result when fitnesses are assigned according to the Rough Mount Fuji model we require a definition. For any function $f : \mathbb{R} \to \mathbb{R}$, we define the *support* of f, denoted $\mathrm{Supp}(f)$, as the set of points at which f is non-zero[1], i.e.: $\mathrm{Supp}(f) = \{x : f(x) \neq 0\}$. We say that f has *connected support* if $\mathrm{Supp}(f)$ is a connected subset of $\mathbb{R}$. Our result is as follows:

Theorem 4.2. *Let* η *be any probability distribution whose p.d.f. is continuous on its support and whose support is connected. Let* θ_n *be any strictly positive function of* n *such that* $n\theta_n \to \infty$ *as* $n \to \infty$. *Then in the model* (2.2), $\mathbb{P}(X \geq 1)$ *tends to one as* $n \to \infty$.

[1] Sometimes in the mathematical literature, the support of a function is defined to be the closure of this set.

4.2 Paper II: Unoriented first-passage percolation on the n-cube

In this paper, we consider first-passage percolation on $\mathbb{Q}_n$ with independent standard exponential passage times on the edges. Of primary concern is the quantity T_n, the first-passage time from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ in $\mathbb{Q}_n$.

Our main result is that T_n converges to $\vartheta := \ln\left(1+\sqrt{2}\right) \approx 0.88$ in L^p-norm as $n \to \infty$ for any $1 \leq p < \infty$, proving that the lower bound by Fill and Pemantle is sharp. More precisely, we show that for any fixed $1 \leq p < \infty$, the L^p-norm of $T_n - \vartheta$ is $\Theta\left(\frac{1}{n}\right)$. In particular, this means that T_n has mean $\vartheta + O\left(\frac{1}{n}\right)$ and standard deviation of order $\frac{1}{n}$. A direct implication is that the first-passage percolation time T_∞ in Theorem 3.2 is ϑ, which improves the best known upper bound on the covering time of the hypercube to $\vartheta + \ln 2 \approx 1.57$.

We further give a characterization of the geodesic from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$, Γ_n, i.e. the path from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ that attains the minimal passage time. As a corollary, we get that the number of edges in Γ_n is concentrated around $\sqrt{2}\ln\left(1+\sqrt{2}\right)n$.

A key idea of our proof is to consider a subset of Durrett's branching process, which we call the set of uncontested particles. We show that Richardson's model is stochastically sandwiched between the full BTP and this subset. Using this relation, we derive an explicit lower bound on the probability that $\hat{\mathbf{1}}$ is infected at time t.

4.3 Paper III: Accessibility percolation and first-passage site percolation on the unoriented binary hypercube

The third paper considers accessibility percolation on the unoriented hypercube. Suppose the vertices of $\mathbb{Q}_n$ are assigned fitnesses according to the House-of-Cards model, i.e. we pick some vertex $\hat{\boldsymbol{v}}$ which we assign the fitness 1, and for all other vertices we independently let $\omega(v) \sim U(0,1)$. Let X denote the number of accessible paths from $\hat{\mathbf{0}}$ to $\hat{\boldsymbol{v}}$.

We will here focus on the case where $\hat{\boldsymbol{v}}$ is picked uniformly at random among the vertices of $\mathbb{Q}_n$, though the full result in Paper III is more general. The problem is then equivalent to assigning all fitnesses i.i.d. $U(0,1)$ and asking for the number of accessible paths from $\hat{\mathbf{0}}$ to the global fitness maximum.

A natural first step in determining the behavior of X is to try to mimic the analysis for the oriented hypercube in (4.1) and (4.2). However, while these derivations were basically one-liners in the oriented case, the increased complexity introduced by allowing flips from 1:s to 0:s makes this procedure much more complicated. In a recent paper by Berestycki, Brunet and Shi [5], it was shown that for any $\alpha \in [0,1]$

$$(\mathbb{E}^\alpha X)^{1/n} \to \sqrt{\frac{1}{2}\sinh(2-2\alpha)} \text{ as } n \to \infty. \tag{4.6}$$

As a consequence, there is a critical value of $\alpha^* = 1 - \frac{1}{2}\sinh^{-1}(2) = 1 - \frac{1}{2}\left(2+\sqrt{5}\right) \approx 0.28$ such that $\mathbb{E}^\alpha X$ tends to 0 for $\alpha > \alpha^*$, and tends to ∞ for $\alpha < \alpha^*$. This implies that probability of accessible paths tends to 0 as $n \to \infty$ when $\alpha > \alpha*$.

The first main result of Paper III shows that, for any $\alpha > \alpha^*$, we have $\mathbb{P}^\alpha(X \geq 1) \to 1$ as $n \to \infty$, as conjectured by Berestycki, Brunet and Shi. Hence, we have $\mathbb{P}(X \geq 1) \to 1 - \frac{1}{2}\left(2+\sqrt{5}\right)$ as $n \to \infty$. As a second result, we show that accessibility percolation with House-of-Cards fitnesses for any graph of digraph can be equivalently formulated in terms of first-passage percolation with $U(0,1)$ passage times on the vertices rather than edges, as in the usual formulation of first-passage percolation. Implications for earlier results on accessibility percolation are discussed. The reader is referred to Theorem 1.4 in appended Paper III for details.

Bibliography

[1] T. Aita, H. Uchiyama, T. Inaoka, M. Nakajima, T. Kokubo, and Y. Husimi, *Analysis of a local fitness landscape with a model of the rough Mt. Fuji-type landscape: application to prolyl endopeptidase and thermolysin.*, Biopolymers **54** (2000), 64-79.

[2] David Aldous, *Probability approximations via the Poisson clumping heuristic*, Applied Mathematical Sciences, vol. 77, Springer-Verlag, New York, 1989. MR **969362 (90k:**60004)

[3] Xinxin Chen, *Increasing paths on N-ary trees*, available at `arXiv:1403.0843[math.PR]`.

[4] Julien Berestycki, Éric Brunet, and Zhan Shi, *The number of accessible paths in the hypercube*, available at `arXiv:1304.0246[math.PR]`.

[5] ———, *Accessibility percolation with backsteps*, available at `arXiv:1401.6894[math.PR]`.

[6] Béla Bollobás and Yoshiharu Kohayakawa, *On Richardson's model on the hypercube*, Combinatorics, geometry and probability (Cambridge, 1993), Cambridge Univ. Press, Cambridge, 1997, pp. 129–137. MR **1476439 (98j:**05122)

[7] Maurício Carneiro and Daniel L. Hartl, *Colloquium papers: Adaptive landscapes and protein evolution*, Proc. Natl. Acad. Sci. **107 Suppl 1** (2010), 1747–1751, DOI 10.1073/pnas.0906192106.

[8] J. Theodore Cox and Richard Durrett, *Some limit theorems for percolation processes with necessary and sufficient conditions*, Ann. Probab. **9** (1981), no. 4, 583–603. MR **624685 (82k:**60208)

[9] Richard Durrett, *Lecture notes on particle systems and percolation*, The Wadsworth & Brooks/Cole Statistics/Probability Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1988. MR **940469 (89k:**60157)

[10] Richard Durrett and Vlada Limic, *Rigorous results for the NK model*, Ann. Probab. **31** (2003), no. 4, 1713–1753, DOI 10.1214/aop/1068646364. MR **2016598 (2005a:**60067)

[11] James Allen Fill and Robin Pemantle, *Percolation, first-passage percolation and covering times for Richardson's model on the n-cube*, Ann. Appl. Probab. **3** (1993), no. 2, 593–629. MR **1221168 (94h:**60145)

[12] Jasper Franke, Alexander Klözer, J. Arjan G. M. de Visser, and Joachim Krug, *Evolutionary accessibility of mutational pathways*, PLoS Comput. Biol. **7** (2011), no. 8, e1002134, 9, DOI 10.1371/journal.pcbi.1002134. MR 2845072

[13] John H. Gillespie, *Molecular evolution over the mutational landscape*, Evolution **38** (1984), no. 5, 1116–1129.

[14] Geoffrey Grimmett, *Percolation*, 2nd ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 321, Springer-Verlag, Berlin, 1999. MR **1707339 (2001a:**60114)

[15] C. Douglas Howard, *Models of first-passage percolation*, Probability on discrete structures, Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004, pp. 125–173. MR **2023652 (2005b:**60258)

[16] Harry Kesten, *Aspects of first passage percolation*, École d'été de probabilités de Saint-Flour, XIV—1984, Lecture Notes in Math., vol. 1180, Springer, Berlin, 1986, pp. 125–264, DOI 10.1007/BFb0074919, (to appear in print). MR **876084 (88h:**60201)

[17] Stuart Kauffman and Simon Levin, *Towards a general theory of adaptive walks on rugged landscapes*, Journal of Theoretical Biology **128** (1987), no. 1, 11-45, DOI 10.1016/s0022-5193(87)80029-2.

[18] Stuart Kauffman and Edward Weinberger, *The NK Model of rugged fitness landscapes and its application to the maturation of the immune response*, J. Theor. Biol. **141** (1989), no. 2, 211–245.

[19] John Kingman, *A simple model for the balance between selection and mutation*, J. Appl. Probability **15** (1978), no. 1, 1–12. MR 0465272 (57 #5177)

[20] Vlada Limic and Robin Pemantle, *More rigorous results on the Kauffman-Levin model of evolution*, Ann. Probab. **32** (2004), no. 3A, 2149–2178, DOI 10.1214/009117904000000081. MR **2073188 (2005b:**92032)

[21] Stefan Nowak and Joachim Krug, *Accessibility percolation on n-trees*, Europhys. Lett. **101** (2013), no. 6.

[22] Daniel Richardson, *Random growth in a tessellation*, Proc. Cambridge Philos. Soc. **74** (1973), 515–528, DOI 10.1017/S0305004100077288.

[23] Matthew Roberts and Lee Zhuo Zhao, *Increasing paths in regular trees*, Electron. Commun. Probab. **18** (2013), no. 87, 1-10, DOI 10.1214/ECP.v18-2784.

[24] Ivan G. Szendro, Martijn F. Schenk, Jasper Franke, Joachim Krug, and J. Arjan G. M. de Visser, *Quantitative analyses of empirical fitness landscapes*, J. Stat. Mech., posted on 2013, DOI 10.1088/1742-5468/2013/01/P01005, (to appear in print).

[25] Daniel M. Weinreich, Yinghong Lan, C. Scott Wylie, and Robert B. Heckendorn, *Should evolutionary geneticists worry about higher-order epistasis?*, Current Opinion in Genetics & Development **23** (2013), no. 6, 700–707, DOI 10.1016/j.gde.2013.10.007.

[26] Daniel M. Weinreich, Richard A. Watson, and Lin Chao, *Perspective: Sign epistasis and genetic constraint on evolutionary trajectories*, Evolution **59** (2005), 1165–1174.

Paper I

On the existence of accessible paths in various models of fitness landscapes

Peter Hegarty and Anders Martinsson

Abstract

We present rigorous mathematical analyses of a number of well-known mathematical models for genetic mutations. In these models, the genome is represented by a vertex of the n-dimensional binary hypercube, for some n, a mutation involves the flipping of a single bit, and each vertex is assigned a real number, called its fitness, according to some rules. Our main concern is with the issue of existence of (selectively) accessible paths, that is, monotonic paths in the hypercube along which fitness is always increasing. Our main results resolve open questions about three such models, which in the biophysics literature are known as House of Cards (HoC), Constrained House of Cards (CHoC) and Rough Mount Fuji (RMF). We prove that the probability of there being at least one accessible path from the all-zeroes node $\boldsymbol{v^0}$ to the all-ones node $\boldsymbol{v^1}$ tends respectively to 0, 1 and 1, as n tends to infinity. A crucial idea is the introduction of a generalisation of the CHoC model, in which the fitness of $\boldsymbol{v^0}$ is set to some $\alpha = \alpha_n \in [0, 1]$. We prove that there is a very sharp threshold at $\alpha_n = \frac{\ln n}{n}$ for the existence of accessible paths from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$. As a corollary we prove significant concentration, for α below the threshold, of the number of accessible paths about the expected value (the precise statement is technical, see Corollary 1.4). In the case of RMF, we prove that the probability of accessible paths from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$ existing tends to 1 provided the drift parameter $\theta = \theta_n$ satisfies $n\theta_n \to \infty$, and for any fitness distribution which is continuous on its support and whose support is connected.

Notation

Throughout this paper, $\mathbb{Q}_n$ will denote the *directed* n-dimensional binary hypercube. This is the directed graph whose nodes are all binary strings of length n, with an edge between any pair of nodes that differ in exactly one bit, the edge being always directed toward the node with the greater number of ones.

Let $g, h : \mathbb{N} \to \mathbb{R}_+$ be any two functions. We will employ the following notations throughout, all of which are quite standard:

(i) $g(n) \sim h(n)$ means that $\lim_{n\to\infty} \frac{g(n)}{h(n)} = 1$.

(ii) $g(n) \lesssim h(n)$ means that $\limsup_{n\to\infty} \frac{g(n)}{h(n)} \leq 1$.

(iii) $g(n) \gtrsim h(n)$ means that $h(n) \lesssim g(n)$.

(iv) $g(n) = O(h(n))$ means that $\limsup_{n\to\infty} \frac{g(n)}{h(n)} < \infty$.

(v) $g(n) = \Omega(h(n))$ means that $h(n) = O(g(n))$.

(vi) $g(n) = \Theta(h(n))$ means that both $g(n) = O(h(n))$ and $h(n) = O(g(n))$ hold.

(vii) $g(n) = o(h(n))$ means that $\lim_{n\to\infty} \frac{g(n)}{h(n)} = 0$.

Now suppose instead that $(g(n))_{n=1}^{\infty}, (h(n))_{n=1}^{\infty}$ are two sequences of random variables. We write $g(n) \sim h(n)$ if, for all $\varepsilon_1, \varepsilon_2 > 0$ and n sufficiently large,

$$\mathbb{P}\left(1 - \varepsilon_1 < \frac{g(n)}{h(n)} < 1 + \varepsilon_1\right) > 1 - \varepsilon_2. \tag{0.1}$$

Similarly, we write $g(n) \gtrsim h(n)$ if, for all $\varepsilon_1, \varepsilon_2 > 0$ and n sufficiently large,

$$\mathbb{P}\left(\frac{g(n)}{h(n)} > 1 - \varepsilon_1\right) > 1 - \varepsilon_2. \tag{0.2}$$

1 Introduction

In many basic mathematical models of genetic mutations, the genome is represented as a node of the directed n-dimensional binary hypercube $\mathbb{Q}_n$ and each mutation involves the flipping of a single bit from 0 (the "wild" state) to 1 (the "mutant" state), hence displacement along an edge of $\mathbb{Q}_n$. Each node $v \in \mathbb{Q}_n$ is assigned a real number $f(v)$, called its *fitness*. The fitness of a node is not a constant, but is drawn from some probability distribution specified by the model. This distribution may vary from node to node in more or less complicated ways, depending on the model. Basically, however, evolution is considered as favoring mutational pathways which, on average, lead to higher fitness. A fundamental concept in this regard is the following (see [16], [15], [7]):

Definition 1.1. Let $f : \mathbb{Q}_n \to \mathbb{R}$ be a fitness function. A *(selectively) accessible path* in $\mathbb{Q}_n$ is a path

$$v_0 \to v_1 \to \cdots \to v_{k-1} \to v_k, \tag{1.1}$$

such that $f(v_i) > f(v_{i-1})$ for $i = 1, \ldots, k$.

Let $\boldsymbol{v^0} = (0, 0, \ldots, 0)$, $\boldsymbol{v^1} = (1, 1, \ldots, 1)$ denote the all-zeroes and all-ones vertices in $\mathbb{Q}_n$. A basic question in such models is whether accessible paths from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$ exist or not with high probability. For the remainder of this paper, unless explicitly stated otherwise, the words "accessible path" will always refer to such a path which starts at $\boldsymbol{v^0}$ and ends at $\boldsymbol{v^1}$. In fact, it will only be in the proof of Proposition 2.18 that we will need to consider accessible paths with other start- and endpoints.

We shall be concerned below with the following three well-known models, in which no rigorous answer has previously been given to the question of whether or not accessible paths exist with high probabilty.

Model 1: Unconstrained House of Cards (HoC)

This model is originally attributed to Kingman [10]. In the form we consider below, it was first studied by Kauffman and Levin [9]. We set $f(\boldsymbol{v}^1) := 1$ and, for every other node $v \in \mathbb{Q}_n$, independently let $f(v) \sim \mathrm{U}(0, 1)$, the uniform distribution on the interval $[0, 1]$.

Model 2: Constrained House of Cards (CHoC)

This variant seems to have been considered only more recently, see for example [11] and [3]. The only difference from Model 1 is that we fix $f(\boldsymbol{v^0}) := 0$.

Model 3: Rough Mount Fuji (RMF)

This model was first proposed in [1], see also [8]. For each $v \in \mathbb{Q}_n$ one lets

$$f(v) = \theta \cdot d(v, \boldsymbol{v^0}) + \eta(v), \tag{1.2}$$

where $\theta = \theta_n$ is a positive number, called the *drift*, $d(\cdot, \cdot)$ denotes Hamming distance and the $\eta(v)$ are independent random variables of some fixed distribution. In other words, one first assigns a fitness to each node at random, according to η, and independent of all other nodes. Then the fitness of each node is shifted upwards by a fixed multiple of the Hamming distance from $\boldsymbol{v^0}$.

Before proceeding, it is worth noting that the above models are also of interest in physics in the context of so-called *spin glasses* [12]. In this setting, each node of $\mathbb{Q}_n$ represents a point in the state space of all possible configurations of spins in a disordered magnet. The analogue of fitness is in this case energy, or more precisely "energy times -1". Accessible paths (not necessarily from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$) correspond to trajectories in which energy decreases monotonically, and which are therefore easily accessible even at zero temperature. The HoC model appears in the spin glass context as Derrida's Random Energy Model (REM), and the RMF-model is a REM in an external magnetic field. For further discussion of the connection between fitness landscapes and spin glasses, see [6].

In all three models, the basic random variable of interest is the number $X = X(n)$ of accessible paths. One thinks of $\boldsymbol{v^0}$ as the starting point of some evolutionary process, and $\boldsymbol{v^1}$ as the desirable endpoint. The HoC model is often referred to as a "null model" for evolution, since the fitnesses of all nodes other than $\boldsymbol{v^1}$ are assigned at random and independently of one another. No mechanism is prescribed which might push an evolutionary process in any particular direction. The CHoC model is not much better, though it does specify that the starting point is a global fitness minimum. The RMF model is a very natural, and simple, way to introduce an "arrow of evolution", since the drift factor implies that successive $0 \to 1$ mutations will tend to increase fitness.

It seems intuitively obvious that the number X of accessible paths should, on average, be much higher in RMF than in HoC, with the CHoC model lying somewhere in between. One should be a little careful here, since in RMF, the node $\boldsymbol{v^1}$ is not assumed to be a global fitness maximum. Nevertheless, it is easy to verify that $\mathbb{E}[X] = 1$ in HoC, $\mathbb{E}[X] = n$ in CHoC, whereas in many situations $\mathbb{E}[X]$ grows super-exponentially with n in RMF: see [7], along with Propositions 2.1 and 3.1 below. Of more interest, however, is the quantity $P = P(n)$, which is the probability of there being at least one accessible path, i.e.: $P = \mathbb{P}(X > 0)$. The idea here is that, as long as *some* accessible path exists, then evolution will eventually find it. The quantity P has been simulated in the biophysics literature. In [7] it was conjectured explicitly that $P \to 0$ in the HoC model, and that $P \to 1$ in the RMF model, when η is a normal distribution and θ is any positive constant. In [3], the CHoC model was simulated for $n \leq 13$, and the authors conjecture, if somewhat implicitly, that P is monotonic decreasing in n and approaches a limiting value close to 0.7. In [7], simulations were continued up to $n = 19$ and these indicated clearly that P was not, after all, monotonic decreasing. The authors abstain from making any explicit conjecture about the limiting behaviour of P in CHoC.

Our main results below resolve all these issues. A crucial idea is to consider the following slight generalisation of the CHoC model:

Model 4: α-Constrained House of Cards (α-HoC)

Let $\alpha \in [0, 1]$. In this model, fitnesses are assigned as in the CHoC model, with the exception that we set $f(\boldsymbol{v}^{\mathbf{0}}) := \alpha$. Hence, CHoC is the case $\alpha = 0$.

For $\alpha \in [0, 1]$, let $P(n, \alpha)$ denote the probability of there being an accessible path in the α-HoC model. To simplify notation below, we define $P(n, \alpha) = P(n, 0)$ for $\alpha < 0$ and $P(n, \alpha) = P(n, 1)$ for $\alpha > 1$. Note that $P(n, \alpha)$ decreases as α increases. Our first main result is the following:

Theorem 1.2. *Let $\varepsilon = \varepsilon_n > 0$. If $n\varepsilon_n \to \infty$ then*

$$\lim_{n\to\infty} P\left(n, \frac{\ln n}{n} - \varepsilon_n\right) = 1 \tag{1.3}$$

and

$$\lim_{n\to\infty} P\left(n, \frac{\ln n}{n} + \varepsilon_n\right) = 0. \tag{1.4}$$

It follows immediately that $P \to 1$ in the CHoC model and that $P(n, \alpha) \to 0$ for any strictly positive constant α. The above result says a lot more, however. It shows that there is a very sharp threshold at $\alpha = \alpha_n = \frac{\ln n}{n}$ for the existence of accessible paths in the α-HoC model. Theorem 1.2 will be proven in Section 2. We have the following immediate corollary for HoC:

Corollary 1.3. *Let X denote the number of accessible paths in the HoC model. Then*

$$\mathbb{P}\left(X > 0\right) \sim \frac{\ln n}{n}. \tag{1.5}$$

Proof. As $P(n, \alpha)$ is decreasing in α we know that, for any $\alpha \in [0, 1]$, $\mathbb{P}\left(X > 0\right) \geq \alpha P(n, \alpha)$. Picking $\alpha = \frac{\ln n}{n} - \varepsilon_n$ where $n\varepsilon_n$ tends to infinity sufficiently slowly, it follows from Theorem 1.2 that $\mathbb{P}\left(X > 0\right) \gtrsim \frac{\ln n}{n}$.

To get the upper bound, let $\alpha = \frac{\ln n}{n}$. Now, if the hypercube has accessible paths, then either $\boldsymbol{v}^{\mathbf{0}}$ has fitness at most α or there is an accessible path where all nodes involved have fitness at least α. Obviously the former event occurs with probability α. Concerning the latter, if

$$\boldsymbol{v}^{\mathbf{0}} \rightarrow v_1 \rightarrow \cdots \rightarrow v_{n-1} \rightarrow \boldsymbol{v}^{\mathbf{1}} \tag{1.6}$$

is any path, then the probability of all nodes along it having fitness at least α is $(1-\alpha)^n$. The probability of fitness being increasing along the path is $1/n!$. Since there are $n!$ possible paths of the form (1.6), it follows from a union bound that

$$\mathbb{P}(X > 0) \leq \alpha + n! \frac{(1-\alpha)^n}{n!} \leq \frac{\ln n}{n} + \frac{1}{n}. \tag{1.7}$$

■

Another Corollary of Theorem 1.2 concerns the distribution of the number of accessible paths in α-HoC for $\alpha = \frac{\ln n}{n} - \varepsilon_n$, where $n\varepsilon_n \rightarrow \infty$. It is straightforward to show that the expected number of paths in α-HoC is $n(1-\alpha)^{n-1}$ (see Proposition 2.1), which, for this choice of α, is $\sim e^{n\varepsilon_n}$. We have the following result:

Corollary 1.4. *Let X denote the number of accessible paths in α-HoC for $\alpha = \frac{\ln n}{n} - \varepsilon_n$ where $n\varepsilon_n \rightarrow \infty$. If $w_n \rightarrow \infty$ then*

$$\lim_{n\rightarrow\infty} \mathbb{P}\left(\frac{1}{w_n}\mathbb{E}[X] \leq X \leq w_n \mathbb{E}[X]\right) = 1. \tag{1.8}$$

Corollary 1.4 will be proven in Subsection 2.5.

Our second main result concerns the RMF model. For any function $f : \mathbb{R} \rightarrow \mathbb{R}$, recall that the *support* of f, denoted $\mathrm{Supp}(f)$, is the set of points at which f is non-zero[1], i.e.: $\mathrm{Supp}(f) = \{x : f(x) \neq 0\}$.

[1]Sometimes in the mathematical literature, the support of a function is defined to be the closure of this set.

We say that f has *connected support* if $\mathrm{Supp}(f)$ is a connected subset of $\mathbb{R}$. Our result is the following:

Theorem 1.5. *Let η be any probability distribution whose p.d.f. is continuous on its support and whose support is connected. Let θ_n be any strictly positive function of n such that $n\theta_n \to \infty$ as $n \to \infty$. Then in the model* (1.2)*, $P(n)$ tends to one as $n \to \infty$.*

This result is proven in Section 3. The proof follows similar lines to that of Theorem 1.2, but the analysis is somewhat simpler.

Remark 1.6. More generally, the proof of Theorem 1.5 presented in this article holds for any distribution η that satisfies, with notation taken from Section 3, $\kappa_{\eta,\delta} = \inf_{I \subseteq I_\delta} \frac{1}{l(I)} \int_I \eta(x)\,dx > 0$ for any $\delta \in (0, 1)$. This condition essentially states that η is not allowed to have "isolated modes". For instance, it is satisfied for any unimodal distribution.

2 Results for the HoC models

For each path i from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$ let X_i be the indicator function of the event that i is accessible, and let $X = \sum_i X_i$ denote the number of accessible paths from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$. Furthermore, given a path i from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$ in the n-dimensional hypercube, let $T(n, k)$ denote the number of paths from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$ that intersect i in exactly $k - 1$ interior nodes (by symmetry, this is independent of i).

Proposition 2.1. *Let X denote the number of accessible paths in the α-HoC model. Then*

$$\mathbb{E}[X] = n(1 - \alpha)^{n-1}. \tag{2.1}$$

Proof. There are $n!$ paths through the hypercube. A path is accessible if all $n-1$ interior nodes have fitness at least α and the fitness of the interior nodes is increasing along the path. This occurs with probability $(1 - \alpha)^{n-1}/(n - 1)!$. ■

Note that for $\alpha = \frac{\ln n}{n} + \varepsilon_n$, the Proposition implies that the expected number of accessible paths tends to 0 for any sequence ε_n satisfying $n\varepsilon_n \to \infty$. This directly implies equation (1.4). Similarly, for $\alpha = \frac{\ln n}{n} - \varepsilon_n$ where $n\varepsilon_n \to \infty$, the expected number of paths tends to infinity.

To show the remaining part of Theorem 1.2, that the probability of there being at least one accessible path tends to 1 in the case $\alpha = \frac{\ln n}{n} - \varepsilon_n$, we will begin by showing that the probability is at least $\frac{1}{4} - o(1)$ by the second moment method. In subsection 2.4 we will then provide a proof that the probability must tend to 1.

Lemma 2.2. *Let X be a random variable with finite expected value and finite and non-zero second moment. Then*

$$\mathbb{P}(X \neq 0) \geq \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}. \tag{2.2}$$

Proof. Let $1_{X\neq 0}$ denote the indicator function of $X \neq 0$. Then, by the Cauchy-Schwartz inequality, $\mathbb{E}[X]^2 = \mathbb{E}[1_{X\neq 0}X]^2 \leq \mathbb{E}[1_{X\neq 0}^2] \cdot \mathbb{E}[X^2] = \mathbb{P}(X \neq 0) \cdot \mathbb{E}[X^2]$. ■

See also Exercise 4.8.1 in [2].

Proposition 2.3. *Let i and j be paths with exactly $k-1$ interior nodes in common. Then*

$$\mathbb{E}[X_i X_j] \leq \frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!}, \tag{2.3}$$

where equality holds if the nodes where i and j differ are consecutive along the paths, i.e. if i and j diverge at most once. Furthermore,

$$\mathbb{E}[X^2] \leq \sum_{k=1}^{n} n!\, T(n,k) \frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!}. \tag{2.4}$$

Proof. The event that i and j are both accessible occurs if all $2n-k-1$ interior nodes have fitness at least α and the fitnesses of the interior nodes are ordered in such a way that fitness increases along both paths.

Conditioned on the event that all interior nodes have fitness at least α, all possible ways in which the fitnesses of the interior nodes can be ordered are equally likely. This implies that the probability that both paths are accessible is $(1-\alpha)^{2n-k-1}/(2n-k-1)!$ times the number of ways to order the fitnesses of the interior nodes such that fitness increases along both paths.

To count the number of ways this can be done we color the numbers $1, \ldots, 2n-k-1$ in the following way: The number l is colored gray if the interior node with the l:th smallest fitness is contained in both paths, red if it is only contained in i and blue if only in j. Note that i and j uniquely determine which numbers must be gray for a valid order, and that any coloring corresponds to at most one order.

Clearly, any coloring corresponding to a valid order colors half of the non-gray numbers red and half blue, which implies that there can be at most $\binom{2n-2k}{n-k}$ such orders. Furthermore, if i and j diverge at most once, one can always construct a valid order from such a coloring, so in this case there are exactly $\binom{2n-2k}{n-k}$ such orders.

As the number of ordered pairs of paths that intersect in exactly $k-1$ interior nodes is $n!T(n,k)$, (2.4) follows from this estimate. ■

2.1 Useful formulas for $T(n,k)$

The numbers $T(n,k)$ already appear in the mathematical literature. The usual terminology is that $T(n,k)$ is the number of permutations of $\{1,2,\ldots,n\}$ with k *components*, where the number of components of a permutation $\pi_1\pi_2\cdots\pi_n$ is defined as the number of choices for $1 \le s \le n$ such that $\pi_1\pi_2\cdots\pi_s$ is a permutation of $\{1,2,\ldots,s\}$. In terms of paths in $\mathbb{Q}_n$, we can represent each path from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$ by a permutation $\pi_1\pi_2\cdots\pi_n$ of $\{1,2,\ldots,n\}$ where π_s denotes which coordinate to increase in step s. If we let i be the path represented by the identity permutation, then a path j, represented by $\pi_1\pi_2\cdots\pi_n$, intersects i in step $s \ge 1$ if and only if $\pi_1\pi_2\cdots\pi_s$ is a permutation of $\{1,2,\ldots,s\}$. This means that, if $\pi_1\pi_2\cdots\pi_n$ has k components, then i and j intersect in $k-1$ interior nodes (the k:th component corresponds to $s=n$). We can thus consider a component as an interval $[s,t]$ where i and j intersect in steps s and t, but at no step in between.

An alternative formulation is that $T(n,k)$ is the number of permutations of $\{1,2,\ldots,n\}$ with $k-1$ *global descents*. A global descent in a permutation $\pi_1\pi_2\cdots\pi_n$ of $\{1,2,\ldots,n\}$ is a number $t \in [1,n-1]$ such that $\pi_i > \pi_j$ for all $i \le t$ and $j > t$. There is a simple $1-1$ correspondence between permutations with k components and those with $k-1$ global descents obtained by reading a permutation backwards. In other words, $\pi_1\pi_2\cdots\pi_n$ has $k-1$ global descents if and only if $\pi_n\pi_{n-1}\cdots\pi_1$ has k components.

There is a database of the numbers $T(n,k)$ for small n and k, see [14]. The book of Comtet [5] referred to at this link contains a couple of exercises and an implicit recursion formula for $T(n,k)$. Comtet has also performed a detailed asymptotic analysis of the numbers $T(n,1)$ in [4]. Permutations with one component (i.e.: no global descents) are variously referred to as *connected, indecomposable, irreducible.* These seem to crop up quite a lot, see [13]. However, estimates of the numbers $T(n,k)$ for general n and k like those in Propositions 2.9 and 2.11 below do not appear to have been obtained before.

Proposition 2.4. *$T(n,1)$ is uniquely defined by*

$$n! = \sum_{k=1}^{n} T(k,1)(n-k)!. \tag{2.5}$$

Proof. Given a path i through $\mathbb{Q}_n$, the number of paths j that intersect i for the first time in step k is $T(k,1)(n-k)!$. As any path through $\mathbb{Q}_n$ intersects i for the first time after between 1 and n steps, the Proposition follows. ∎

Proposition 2.5.

$$n!\left(1-O\left(\frac{1}{n}\right)\right) \leq T(n,1) \leq n! \tag{2.6}$$

Proof. By definition, $T(n,1) \leq n!$. Using this, Proposition 2.4 implies that $T(n,1)$ is at least $n! - \sum_{k=1}^{n-1} k!(n-k)! = n! - O\left((n-1)!\right)$. ∎

Proposition 2.6.

$$T(n,k) = \sum_{\substack{s_1,\dots s_k \geq 1 \\ s_1+\dots s_k = n}} T(s_1,1)\cdot \ldots \cdot T(s_k,1) \tag{2.7}$$

Proof. Given a path i, the number of paths that intersect i for the first time after s_1 steps, for the second time after s_2 more steps and so on up to the last time (at $\boldsymbol{v}^{\mathbf{1}}$) after n steps is $T(s_1,1)\cdot \ldots T(s_{k-1},1)\cdot T(n-s_1-\cdots-s_{k-1},1)$. Let $s_k = n - s_1 - \cdots - s_{k-1}$. $T(n,k)$ is obtained by summing over all possible values of $s_1,\ldots s_k$. ∎

Proposition 2.7. *For $k \geq 2$, $T(n,k)$ satisfies*

$$T(n,k) = \sum_{s=1}^{n-k+1} T(s,1)T(n-s,k-1). \tag{2.8}$$

Proof. It follows by induction that this sum equals the right hand side in (2.7). ■

2.2 Upper bounds for $T(n,k)$

Proposition 2.8. *For any $n \geq k \geq 1$,*

$$T(n,k) \leq k \sum \left((n - \sum_{j=1}^{k-1} s_j)! \prod_{j=1}^{k-1} s_j! \right) \tag{2.9}$$

where the first sum goes over all $(k-1)$-tuples of integers $s_1, \ldots s_{k-1}$ such that $s_j \geq 1$ for all j and $\max_j s_j \leq n - \sum_j s_j$.

Proof. Consider the formula for $T(n,k)$ in Proposition 2.6. By symmetry, $T(n,k)$ is at most k times the contribution from terms where $s_j \leq s_k$ for $j = 1, \ldots, k-1$. The Proposition follows by applying $T(s,1) \leq s!$. ■

Proposition 2.9. *There is a positive constant c such that for all $n \geq k \geq 1$,*

$$T(n,k) \leq k(n-k+1)!\, e^{c(k-1)/(n-k+1)}. \tag{2.10}$$

Proof. We use Proposition 2.8 and make the following approximations:

- substitute $(n - \sum_j s_j)!$ by $\beta^{n-\sum_j s_j}$ where $\beta = ((n-k+1)!)^{1/(n-k+1)}$. It follows from log-convexity of $l!$ that $\beta^l \geq l!$ for any $0 \leq l \leq n-k+1$.
- let all s_j go from 1 to $\lfloor (n-k+1)/2 + 1 \rfloor$.

This yields

$$T(n,k) \leq k(n-k+1)! \left(\sum_{s=1}^{\lfloor (n-k+1)/2+1 \rfloor} s! \beta^{1-s} \right)^{k-1}. \tag{2.11}$$

We now claim that the sum in the above expression is always less than $1 + c/(n-k+1)$ for sufficiently large c. Indeed

$$\begin{aligned}
&\sum_{s=1}^{\lfloor (n-k+1)/2+1 \rfloor} s!\beta^{1-s} \\
&= 1 + 2\beta^{-1} + \beta^{-1} \sum_{t=1}^{\lfloor (n-k+1)/2-1 \rfloor} t!(t+1)(t+2)\beta^{-t} \\
&\leq 1 + 2\beta^{-1} + e\beta^{-1} \sum_{t=1}^{\lfloor (n-k+1)/2-1 \rfloor} \sqrt{t}(t+1)(t+2) \cdot \\
&\quad \cdot \left(\frac{n-k+1}{2e}\right)^t \left(\frac{n-k+1}{e}\right)^{-t} \\
&\leq 1 + 2\beta^{-1} + e\beta^{-1} \sum_{t=1}^{\infty} \sqrt{t}(t+1)(t+2)2^{-t} \\
&\leq 1 + c(n-k+1)^{-1}.
\end{aligned}$$

Here we have used that $(n-k+1)/e \leq \beta \leq (n-k+1)$ and that $n! \leq en^{n+1/2}e^{-n}$, which follows from standard estimates of factorials.

The Proposition now follows from this result together with (2.11). ■

Proposition 2.10. *For any fixed l there is a constant $C_l > 0$ such that*

$$T(n, n-l) \leq C_l n^l \tag{2.12}$$

for all $n \geq 1$.

Proof. We may, without loss of generality, assume that $n \geq 2l$.

Recall the formula for $T(n, n-l)$ in Proposition 2.6. As $s_1, \ldots, s_{n-l} \geq 1$ and $s_1 + \cdots + s_{n-l} = n$ it is easy to see that all but at most l variables are equal to 1. This implies that $T(n, n-l)$ is at most $\binom{n-l}{l}$ times the contribution from all terms where $s_{l+1} = \cdots = s_{n-l} = 1$. Using $T(1,1) = 1$, we get

$$T(n, n-l) \leq \binom{n-l}{l} \sum_{\substack{s_1, \ldots s_l \geq 1 \\ s_1 + \ldots s_l = 2l}} T(s_1, 1) \cdot \ldots \cdot T(s_l, 1) \leq C_l n^l. \tag{2.13}$$

■

Proposition 2.11. *For sufficiently large c, we have*

$$T(n, n-l) \leq c(l+1)\left(\frac{n+2l}{5}\right)^l. \tag{2.14}$$

Proof. Let

$$S(n, n-l) = (l+1)\left(\frac{n+2l}{5}\right)^l \tag{2.15}$$

i.e.

$$S(n, k) = (n-k+1)\left(\frac{3n-2k}{5}\right)^{n-k}. \tag{2.16}$$

We will begin by showing that $S(n, k)$ satisfies

$$S(n, k) \geq \sum_{i=1}^{n-k+1} i!S(n-i, k-1) \tag{2.17}$$

for $k > 1$ and sufficiently large $n-k$. Here we have

$$\begin{aligned}
&\sum_{i=1}^{n-k+1} i!S(n-i, k-1) \\
&\quad= \sum_{i=1}^{n-k+1} i!(n-k+2-i)\left(\frac{3n-2k-3i+2}{5}\right)^{n-k-i+1} \\
&\quad\leq (n-k+1)\left(\frac{3n-2k-1}{5}\right)^{n-k} \\
&\qquad\quad+ \sum_{i=2}^{n-k+1} i!(n-k+1)\left(\frac{3n-2k}{5}\right)^{n-k-i+1} \\
&\quad= S(n, k)\left(\left(1-\frac{1}{3n-2k}\right)^{n-k} + \sum_{i=2}^{n-k+1} i!\left(\frac{3n-2k}{5}\right)^{-i+1}\right),
\end{aligned}$$

where

$$\begin{aligned}
\left(1-\frac{1}{3n-2k}\right)^{n-k} &\leq \exp\left(-\frac{n-k}{3n-2k}\right) \\
&\leq \exp\left(-\frac{n-k}{3n}\right) \\
&\leq \max\left(\frac{1}{2}, 1-\frac{n-k}{6n}\right),
\end{aligned}$$

and

$$
\begin{aligned}
&\sum_{i=2}^{n-k+1} i! \left(\frac{3n-2k}{5}\right)^{-i+1} \\
&\leq \frac{10}{3n-2k} + \frac{5}{3n-2k} \sum_{j=1}^{n-k-1} j!(j+1)(j+2) \left(\frac{3n-2k}{5}\right)^{-j} \\
&\leq \frac{10}{3n-2k} + \frac{5e}{3n-2k} \sum_{j=1}^{\infty} \sqrt{j}(j+1)(j+2) \left(\frac{n-k}{e}\right)^{j} \left(\frac{3n-2k}{5}\right)^{-j} \\
&\leq \frac{1}{n} \left(10 + 5e \sum_{j=1}^{\infty} \sqrt{j}(j+1)(j+2) \left(\frac{5}{3e}\right)^{j}\right) \\
&= \frac{C}{n}.
\end{aligned}
$$

It follows directly that (2.17) holds for $k > 1$ and $n - k \geq 6C$.

Now, if we can choose c so that $T(n,k) \leq cS(n,k)$ for $k = 1$ and for $n - k < 6C$, the Proposition will follow from Proposition 2.7 by induction on k. Hence it suffices to show the Proposition for these two cases.

For $k = 1$, the inequality holds for sufficiently large c by the fact that

$$
\begin{aligned}
\frac{T(n,1)}{S(n,1)} &\leq \frac{n!}{n\left(\frac{3n-2}{5}\right)^{n-1}} \\
&\leq e\sqrt{n}\left(\frac{n}{e}\right)^{n} \frac{1}{n\left(\frac{3n-2}{5}\right)^{n-1}} \\
&= \frac{3e}{5}\sqrt{n}\left(\frac{5}{3e}\right)^{n}\left(1 - \frac{2}{3n}\right)^{-n+1} \\
&\to 0 \text{ as } n \to \infty.
\end{aligned}
$$

For $n - k < 6C$, just apply Proposition 2.10. ■

2.3 Computing $\mathbb{E}[X^2]$

Pick $\delta > 0$ sufficiently small. We divide the sum in (2.4) into the contribution from $k \le (1-\delta)n$ and that from $k > (1-\delta)n$.

$$\begin{aligned}\sum_{k=1}^{n} n!T(n,k)\frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!}\\ &= \sum_{k=1}^{(1-\delta)n} n!T(n,k)\frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!}\\ &\quad + \sum_{l=0}^{\delta n} n!T(n,n-l)\frac{\binom{2l}{l}(1-\alpha)^{n+l-1}}{(n+l-1)!}\\ &:= S_1 + S_2.\end{aligned} \tag{2.18}$$

Proposition 2.12. *For k constant and $\alpha = o(1)$*

$$n!T(n,k)\frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!} \sim k2^{1-k}n^2(1-\alpha)^{2n}. \tag{2.19}$$

Proof. A simple lower bound on $T(n,k)$ is the number of permutations with k components where all but one component contains exactly one element. For sufficiently large n this is given by $kT(n-k+1,1)$, which by Proposition 2.5 is $\sim k(n-k+1)!$. Furthermore, from Proposition 2.9 we know that $T(n,k)$ is most $(1+o(1))\,k(n-k+1)!$. Hence for constant k, $T(n,k) \sim k(n-k+1)!$. The Proposition now follows from standard estimates of factorials. ■

Proposition 2.13. *Let $\alpha = o(1)$. For any $0 < \delta < 1$, we have $S_1 \sim 4n^2(1-\alpha)^{2n}$.*

Proof. From Proposition 2.9 it follows that there is a constant C_δ such that $T(n,k) \le C_\delta k(n-k+1)!$ whenever $k \le (1-\delta)n$. Using this we have

$$n!T(n,k)\frac{\binom{2n-2k}{n-k}}{(2n-k-1)!} \le C_\delta n!k(n-k+1)!\frac{\binom{2n-2k}{n-k}}{(2n-k-1)!} \tag{2.20}$$

for all $k \le (1-\delta)n$. Now by extensive use of Stirling's formula there is a constant $C > 0$ such that:

$$\begin{aligned}
&C_\delta n! k(n-k+1)! \frac{\binom{2n-2k}{n-k}}{(2n-k-1)!} \\
&\leq C_\delta C k \sqrt{n} \left(\frac{n}{e}\right)^n \sqrt{n-k} \left(\frac{n-k}{e}\right)^{n-k} \cdot \\
&\quad \cdot (n-k+1) \frac{\frac{4^{n-k}}{\sqrt{n-k}}(2n-k)}{\sqrt{2n-k}\left(\frac{2n-k}{e}\right)^{2n-k}} \\
&= C_\delta C k(n-k+1)\sqrt{n(2n-k)} 2^{-k} \cdot \\
&\quad \cdot \left(\left(1-\frac{k}{n}\right)^{\frac{n}{k}-1}\left(1-\frac{k}{2n}\right)^{-\frac{2n}{k}+1}\right)^k,
\end{aligned}$$

where

$$\begin{aligned}
\left(1-\frac{k}{n}\right)^{\frac{n}{k}-1}\left(1-\frac{k}{2n}\right)^{-\frac{2n}{k}+1} &\leq \left(1-\frac{k}{2n}\right)^{\frac{2n}{k}-2}\left(1-\frac{k}{2n}\right)^{-\frac{2n}{k}+1} \\
&= \left(1-\frac{k}{2n}\right)^{-1} \\
&\leq \left(1-\frac{1-\delta}{2}\right)^{-1} \\
&= \frac{2}{1+\delta}.
\end{aligned}$$

This means that, for all $\delta > 0$, there exists a constant C'_δ such that, for $k \leq (1-\delta)n$ and sufficiently large n, we have

$$n! T(n,k) \frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!} \leq C'_\delta n^2 (1-\alpha)^{2n} k \left(1+\delta\right)^{-k} \left(1-\alpha\right)^{-k}. \tag{2.21}$$

Since $\sum k(1+\delta)^{-k}(1-\alpha)^{-k}$ converges for sufficiently small α we have shown that $S_1 = O\left(n^2(1-\alpha)^{2n}\right)$. Furthermore, if we assume that n is sufficiently large so that $(1+\delta)(1-\alpha) \geq (1+\frac{\delta}{2})$, then as the terms in the sum

$$\sum_{k=1}^{(1-\delta)n} \frac{1}{n^2(1-\alpha)^{2n}} n! T(n,k) \frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!} \tag{2.22}$$

are dominated by the terms in

$$\sum_{k=1}^{\infty} C'_{\delta} k \left(1+\frac{\delta}{2}\right)^{-k} \tag{2.23}$$

which converges, it follows by dominated convergence together with Proposition 2.12 that

$$\sum_{k=1}^{(1-\delta)n} \frac{1}{n^2(1-\alpha)^{2n}} n!T(n,k)\frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!} \longrightarrow \sum_{k=1}^{\infty} k2^{1-k} = 4$$

as $n \to \infty$. ■

Proposition 2.14. *For sufficiently small $\delta > 0$ and $\alpha = o(1)$, we have $S_2 = O\left(n(1-\alpha)^n\right)$.*

Proof. Using Proposition 2.11 there is a constant C such that this sum is bounded by

$$\begin{aligned}
&\sum_{l=0}^{\delta n} n!T(n,n-l)\frac{\binom{2l}{l}(1-\alpha)^{n+l-1}}{(n+l-1)!} \\
&\quad \le C\sum_{l=0}^{\delta n} n!(l+1)\left(\frac{n+2l}{5}\right)^l \frac{\binom{2l}{l}(1-\alpha)^{n+l-1}}{(n+l-1)!} \\
&\quad \le C(1-\alpha)^{n-1}\sum_{l=0}^{\delta n} n^{1-l}(l+1)\left(\frac{n+2l}{5}\right)^l 4^l \\
&\quad \le Cn(1-\alpha)^{n-1}\sum_{l=0}^{\infty}(l+1)\left(\frac{4(1+2\delta)}{5}\right)^l
\end{aligned}$$

where the last sum clearly converges for sufficiently small δ.

■

Proposition 2.15. *Let X be the number of accessible paths in the α-HoC model where $\alpha = \frac{\ln n}{n} - \varepsilon_n$ where $n\varepsilon_n \to \infty$. Then*

$$\mathbb{E}[X^2] \sim 4n^2(1-\alpha)^{2n}. \tag{2.24}$$

Proof. From Proposition 2.3 together with Propositions 2.13 and 2.14 we know that

$$\mathbb{E}[X^2] \le (4+o(1))\, n^2(1-\alpha)^{2n} + O\left(n(1-\alpha)^n\right), \tag{2.25}$$

where one can show that $n(1-\alpha)^n = o\left(n^2(1-\alpha)^{2n}\right)$, provided $n\varepsilon_n \to \infty$.

To derive a tight lower bound for $\mathbb{E}[X^2]$, consider the sum of $\mathbb{E}[X_i X_j]$ over all pairs of paths whose number of common interior nodes, $k-1$, is at most $\frac{n}{2}-1$ and that diverge at most once. Expressed in terms of components of permutations, for a fixed i and k, the number of paths j that satisfy this equals the number of permutations with k components, where all but one component contains exactly one element. This can clearly be done in $kT(n-k+1,1) \sim k(n-k+1)!$ ways.

By Proposition 2.3 this yields

$$\mathbb{E}[X^2] \geq \sum_{k=1}^{n/2} n!kT(n-k+1,1)\frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!}. \tag{2.26}$$

Proceeding in a manner similar to the proof of Proposition 2.13, we get that

$$\sum_{k=1}^{n/2} n!kT(n-k+1,1)\frac{\binom{2n-2k}{n-k}(1-\alpha)^{2n-k-1}}{(2n-k-1)!} \sim 4n^2(1-\alpha)^{2n} \tag{2.27}$$

which concludes the proof. ■

From this proof we can observe that almost all of the contribution to $\mathbb{E}[X^2]$ comes from pairs of paths we considered in the lower bound. This implies the following:

Corollary 2.16. *Assume $\alpha = \frac{\ln n}{n} - \varepsilon_n$ where $n\varepsilon_n \to \infty$. For any $0 < \delta < 1$, the contribution to $\mathbb{E}[X^2]$ from all pairs of paths that either share more than $(1-\delta)n$ common nodes or that diverge more than once is $o\left(n^2(1-\alpha)^{2n}\right)$.*

2.4 Proof of Theorem 1.2

Let X as above denote the number of accessible paths in α-HoC, where $\alpha = \frac{\ln n}{n} - \varepsilon_n$, $0 \leq \varepsilon_n \leq \frac{\ln n}{n}$ and $n\varepsilon_n \to \infty$. Applying Lemma 2.2 to X and using the expressions for $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$ from Propositions 2.1 and 2.15 respectively, yields the lower bound

$$\liminf_{n\to\infty} P\left(n, \frac{\ln n}{n} - \varepsilon_n\right) \geq \frac{1}{4}. \tag{2.28}$$

In this subsection, we will prove that this probability can be "bootstrapped" up to 1, proving the remaining part of Theorem 1.2.

Lemma 2.17. *Let* $0 \le a \le 1-b \le 1$ *and let* $f : \mathbb{Q}_n \to \mathbb{R}$ *be a fitness function whose values are generated independently according to*

$$f(v) = \begin{cases} a & \text{if } v = \boldsymbol{v^0} \\ 1-b & \text{if } v = \boldsymbol{v^1} \\ \sim \mathrm{U}(0,1) & \text{otherwise.} \end{cases} \tag{2.29}$$

Then the probability of accessible paths with respect to f *equals* $P(n, a+b)$.

Proof. Define the function $g : \mathbb{Q}_n \to \mathbb{R}$ by setting $g(v) = f(v)+b$ if $f(v) \le 1-b$ and $g(v) = f(v)-1+b$ otherwise. Then $g(\boldsymbol{v^0}) = a+b$, $g(\boldsymbol{v^1}) = 1$ and $g(v) \sim \mathrm{U}(0,1)$ independently for all other v, so g is distributed as in α-HoC with $\alpha = a+b$. As this transformation only constitutes a translation for any node on an accessible path, we see that a path is accessible with respect to f if and only if it is so with respect to g. ■

Proposition 2.18. *Assume there is a positive constant* C *such that* $\liminf_{n\to\infty} P(n, \frac{\ln n}{n} - \varepsilon_n) \ge C$ *whenever* $0 \le \varepsilon_n \le \frac{\ln n}{n}$ *is a sequence satisfying* $n\varepsilon_n \to \infty$. *Then, the same inequality holds if* C *is replaced by* $1-(1-C)(1-\frac{C}{2})$.

Proof. Let $\alpha = \frac{\ln n}{n} - \varepsilon_n$. We wish to pick four nodes, a_1, a_2, b_1, b_2, satisfying the following conditions:

(i) $d(a_1, \boldsymbol{v^0}) = d(a_2, \boldsymbol{v^0}) = 1$, and a_1, a_2 each has fitness in the range $[\alpha, \alpha + \varepsilon_n/3]$.

(ii) $d(b_1, \boldsymbol{v^1}) = d(b_2, \boldsymbol{v^1}) = 1$ and b_1, b_2 each has fitness at least $1 - \varepsilon_n/3$.

(iii) none of the four pairs (a_i, b_j) are antipodal (in the undirected hypercube).

By (i), the number of possibilities for each a_i is binomially distributed with parameters $\mathrm{Bin}(n, \varepsilon_n/3)$. Then, by (ii) and (iii), the number of options for each b_j is distributed as $\mathrm{Bin}(n-2, \varepsilon_n/3)$. Since $n\varepsilon_n/3 \to \infty$, it follows that it is possible to choose four nodes satisfying (i)-(iii) with probability $1 - o_n(1)$.

Condition on the fitness of all vertices v with $d(v, \boldsymbol{v^0}) = 1$ or $d(v, \boldsymbol{v^1}) = 1$. Let H_1 and H_2 be the induced subgraphs consisting of all nodes on paths from a_1 to b_1 and from a_2 to b_2 respectively and let H_2' be the induced subgraph consisting of all nodes on paths between a_2 and b_2 that does not intersect H_1 in any vertex. Then H_1 and H_2 are isomorphic to $\mathbb{Q}_{n-2}$. Note that any accessible path from a_1 to b_1 or a_2 to b_2 can be extended to an accessible path from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$.

Let us denote the probability of accessible paths through the respective induced subgraphs by p_{H_1}, p_{H_2} and $p_{H_2'}$. By construction, H_1 and H_2' are vertex disjoint, so the events of accessible paths through the two subgraphs are independent. By Lemma 2.17, $p_{H_1} = P(n-2, f(a_1)+1-f(b_1)) \geq P(n-2, \alpha + \frac{2\varepsilon_n}{3})$. It is straightforward to show that this is still below the threshold, which implies that $p_{H_1} \geq C - o_n(1)$.

To estimate $p_{H_2'}$, we note that a path in H_2 from a_2 to b_2 is contained in H_2' if and only if it "flips the bit that is 1 in a_1 after that which is 0 in b_1". In the cases where there is an accessible path through H_2, let γ be chosen uniformly among all such paths. Then, by symmetry, we know that it flips the two bits corresponding to a_1 and b_1 in the allowed order, and is thus contained in H_2', with probability $\frac{1}{2}$. Hence $p_{H_2'} \geq \frac{1}{2} p_{H_2} = \frac{1}{2} p_{H_1}$.

As the events of accessible paths through H_1 and H_2' are independent, we get $P(n, \alpha) \geq 1 - (1 - p_{H_1})(1 - p_{H_2'}) - o_n(1) \geq 1 - (1 - C)(1 - \frac{C}{2}) - o_n(1)$ and the Proposition follows. ■

Now to conclude the proof of Theorem 1.2. By equation (2.28) and repeated use of Proposition 2.18 we can construct a sequence $\{C_k\}_{k=0}^{\infty}$ such that $C_k \to 1$ and $\liminf_{n\to\infty} P(n, \alpha) \geq C_k$ for all k. Hence we must have $\liminf_{n\to\infty} P(n, \alpha) = 1$.

2.5 Proof of Corollary 1.4

Similar to the proof of Theorem 1.2, that of Corollary 1.4 will use an alternative formulation of the α-HoC model. A key observation is that if one generates fitnesses according to α-HoC but then removes interior vertices independently with some probability δ, then this results in a model equivalent to α'-HoC for some $\alpha' > \alpha$. The intuition is that if α is far below the threshold $\frac{\ln n}{n}$, then not only is there an accessible path with probability $1 - o_n(1)$, but even if we remove a sufficient amount of vertices so that most paths become forbidden, we will still be below the threshold and so will still have accessible paths with probability $1-o_n(1)$. This intuitively requires the original number of accessible paths to be large. Interestingly, this argument only requires the first equation in Theorem 1.2 even though the Corollary itself is a stronger form of that statement.

This idea is formalized in the following Lemmas:

Lemma 2.19. *Let $\alpha, \delta \in [0,1]$. Consider the fitness model that first assigns fitnesses as in α-HoC, but then independently removes each vertex in $\mathbb{Q}_n \setminus \{\boldsymbol{v^0}, \boldsymbol{v^1}\}$ with probability δ. Then the probability of accessible paths using only the remaining vertices is $P(n, 1-(1-\alpha)(1-\delta))$.*

Proof. Let $\alpha' = 1-(1-\alpha)(1-\delta)$. We compare the model described above with α'-HoC.

Let us make the slight modification to α'-HoC and the above model that we additionally consider any vertex removed if it is less fit than $\boldsymbol{v^0}$. As no such node can be part of an accessible path, this will not change accessibility in either model. We see that these formulations are equivalent up to a translation and scaling, so they will have the same distribution of accessible paths. ■

Lemma 2.20. *Let Ω be a finite universal set and let R be a random subset of Ω given by $\mathbb{P}(r \in R) = p_r$, these events being mutually independent over $r \in \Omega$. Let $\{A_i\}_{i \in I}$ be subsets of Ω, I a finite index set. Let B_i be the event $A_i \subseteq R$. Then*

$$\prod_{i \in I} \mathbb{P}(\bar{B}_i) \leq \mathbb{P}\left(\bigwedge_{i \in I} \bar{B}_i\right). \tag{2.30}$$

This inequality is commonly used as a lower bound in Janson's inequality. See for instance Theorem 8.1.1 in [2].

Proof of Corollary 1.4. The upper bound is simply Markov's inequality. We now turn to the lower bound. To simplify calculations we may, without loss of generality, assume that $w_n = o(n\varepsilon_n)$ and that $1 \leq w_n \leq e^{n\varepsilon_n}$ for all n.

Let $\delta_n = \varepsilon_n - \frac{\ln w_n}{n}$ and let Y denote the number of intact accessible paths using the same fitness function as for X but after removing each node except $\boldsymbol{v^0}$ and $\boldsymbol{v^1}$ independently with probability δ_n. By assumption, we know that $0 \leq \delta_n \leq \varepsilon_n \leq \frac{\ln n}{n}$, so δ_n is always a valid probability.

Using Lemma 2.19 we see that $\mathbb{P}(Y > 0) = P(n, \alpha'_n)$ where $\alpha'_n = 1 - (1-\alpha)(1-\delta_n) = \frac{\ln n}{n} - \frac{o(1)+\ln w_n}{n}$. As $o(1) + \ln w_n \to \infty$ as $n \to \infty$ it follows from Theorem 1.2 that $\lim_{n\to\infty} \mathbb{P}(Y = 0) = 0$.

Condition on the set of accessible paths before removing vertices. Let I be the set of accessible paths, R the random set of non-removed vertices and B_i the event that path $i \in I$ only consist of non-removed vertices. Then we are in the setting of Lemma 2.20. As the probability that each accessible path remains intact is $(1-\delta_n)^{n-1}$, averaging conditioned on X we get the inequality

$$\mathbb{P}(Y = 0 \mid X) \geq \left(1 - (1-\delta_n)^{n-1}\right)^X . \tag{2.31}$$

But since $\lim_{n\to\infty} \mathbb{P}(Y = 0) = 0$ and $\left(1 - (1-\delta_n)^{n-1}\right)^X = e^{-(1+o(1))e^{-n\delta_n}X}$ it follows that $e^{-n\delta_n}X$ must tend to infinity in probability. To conclude the proof we note that $e^{-n\delta_n}X = \frac{X}{e^{n\varepsilon_n}/w_n} \sim \frac{X}{\mathbb{E}[X]/w_n}$.

■

Remark 2.21. Note that Proposition 2.15 implies that $\mathrm{Var}(X) \sim 3\mathbb{E}[X]^2$ for α in this regime, so no significant improvement on Corollary 1.4 can be made by a naive application of Chebyshev's inequality.

3 Results for the RMF model

Let $n \in \mathbb{N}$ and let $\varepsilon = \varepsilon_n$ be some strictly positive function. Consider the n-dimensional hypercube in which $\boldsymbol{v^0}$ and $\boldsymbol{v^1}$ are present, and where every other vertex is present with probability ε_n, independently of all other vertices. Let $Y = Y_{n,\varepsilon_n}$ denote the number of accessible paths from $\boldsymbol{v^0}$ to $\boldsymbol{v^1}$, where in this model a path is accessible if Hamming distance from $\boldsymbol{v^0}$ is strictly increasing and all vertices along the path are present. The following proposition may be well-known, as it can be interpreted in the context of site percolation on the directed hypercube. However, we were not able to locate a suitable reference.

Proposition 3.1. *(i)* $\mathbb{E}[Y] = n! \cdot \varepsilon_n^{n-1}$.
(ii) Let $n \to \infty$ and suppose that $n\varepsilon_n \to \infty$. Then Var$(Y) = o(\mathbb{E}[Y]^2)$, and hence

$$Y \sim \mathbb{E}[Y] \sim \frac{\sqrt{2\pi n}}{\varepsilon_n} \left(\frac{n\varepsilon_n}{e}\right)^n . \tag{3.1}$$

Proof. There are $n!$ possible paths in the n-hypercube. Each path contains $n-1$ interior vertices, each of which is present with probability ε_n. This proves (i). Set $\mu = \mu_n := n!\varepsilon_n^{n-1}$. Now suppose $n\varepsilon_n \to \infty$. Let Y_i be the indicator of the event that the i:th increasing path is accessible, where the paths have been ordered in any way. Fix any path i_0. Then, by a standard second moment estimate (see Section 2),

$$\mathrm{Var}(Y) \leq \mu + n! \cdot \sum_{j \sim i_0} \mathbb{E}(Y_{i_0} Y_j), \tag{3.2}$$

where the sum is taken over all paths j which intersect the path i_0 in at least one interior vertex. Let k be the number of intersection points. This leaves $T(n, k+1)$ possibilities for the path j. The paths i_0 and j contain a total of $2n-2-k$ different interior vertices, hence the probability of both being present is ε_n^{2n-2-k}. Hence

$$\mathrm{Var}(Y) \leq \mu + n! \cdot \sum_{k=2}^{n} T(n,k)\varepsilon_n^{2n-1-k} \leq \mu + \mu^2 \cdot \sum_{k=2}^{n} \frac{T(n,k)}{n!\varepsilon_n^{k-1}}. \tag{3.3}$$

Hence, since $\mu \to \infty$ when $n\varepsilon_n \to \infty$, it suffices to show that

$$\sum_{k=2}^{n} \frac{T(n,k)}{n!\varepsilon_n^{k-1}} = o(1). \tag{3.4}$$

We now follow the same strategy as in Section 2, but the analysis here is much simpler. Let $\delta \in (0,1)$. We divide the sum in (3.4) into two parts, one for $k \leq (1-\delta)n$ and the other for $k > (1-\delta)n$. From Proposition 2.9 and Lebesgue's dominated convergence theorem, it follows easily that, for any $\delta > 0$, the sum over terms $k \leq (1-\delta)n$ is bounded by $(1+o_n(1))\sum_{k=2}^{\infty} \frac{k}{(n\varepsilon_n)^{k-1}} = O(\frac{1}{n\varepsilon_n}) = o(1)$, provided $n\varepsilon_n \to \infty$. Similarly, from Proposition 2.11 it follows that the sum over terms $k > (1-\delta)n$ is bounded by

$$\frac{c}{\mu} \sum_{l=0}^{\delta n} (l+1) \left(\frac{1+2\delta}{5} \cdot n\varepsilon_n \right)^l, \tag{3.5}$$

where c is an absolute constant. Since $n\varepsilon_n \to \infty$, the sum in (3.5) is bounded by $1+o(1)$ times the last term, and hence is $O((n\varepsilon_n)^{\delta n})$, which is in turn $o(\mu)$. This proves (3.4) and completes the proof of the proposition. ■

We now turn to the RMF model and prove Theorem 1.5.

We shall abuse notation and also use η to denote the p.d.f. of the probability distribution under consideration. So suppose η has connected support and is continuous there. Let $\delta > 0$ be given. Then there exists a bounded, closed interval $I = I_\delta \subseteq \mathrm{Supp}(\eta)$ such that $\int_{I_\delta} \eta(x)\, dx > 1 - \delta$. The quantity $c_{\eta,\delta} = \min_{x \in I_\delta} \eta(x)$ exists, is non-zero and, obviously, depends only on η and δ. Now let $n \in \mathbb{N}$ and $\theta = \theta_n > 0$ be given. Without loss of generality, we may assume that the interval I_δ has length $l(I_\delta) > \theta_n/2$ (in fact any multiple $c\theta_n$, where $0 < c < 1$, would do in the argument that follows). By definition of I_δ, with probability at least $(1-\delta)^2$ each of $\eta(\boldsymbol{v^0})$ and $\eta(\boldsymbol{v^1})$ lie in I_δ. Let X_{δ,n,θ_n} be the number of accessible paths in the n-hypercube, where fitnesses are assigned as in (1.2), and conditioning on the fact that both $\eta(\boldsymbol{v^0})$ and $\eta(\boldsymbol{v^1})$ lie in I_δ. We claim that, if n is sufficiently large, then X_{δ,n,θ_n} stochastically dominates the random variable Y_{n,ε_n} in Proposition 3.1, where $\varepsilon_n = c_{\eta,\delta} \cdot \frac{\theta_n}{2}$.

To see this, first note that, as long as $l(I_\delta) > \theta_n/2$ then, for any point $x \in I_\delta$, there will be an interval I_x of length at least $\theta_n/2$, which contains x and lies entirely within I_δ. By assumption, any such interval captures at least $c_{\eta,\delta} \cdot \frac{\theta_n}{2}$ of the distribution η. For any adjacent pair (v, v') of vertices in the hypercube such that $d(v', \boldsymbol{v^0}) = d(v, \boldsymbol{v^0}) + 1$, if $\eta(v') > \eta(v) - \theta_n$, then v' is accessible from v. Assuming $\eta(\boldsymbol{v^0}) \in I_\delta$, it follows that we can choose, for each layer i in the hypercube, an interval $I_i \subseteq I_\delta$ of length $\theta_n/2$ such that any path

$$\boldsymbol{v^0} \to v_1 \to v_2 \cdots \to v_{n-1} \tag{3.6}$$

for which $\eta(v_i) \in I_i$ for all $i = 1, \ldots, n-1$, is accessible. If n is sufficiently large, we can also ensure that the interval I_{n-1} contains $\eta(\boldsymbol{v^1})$, so that any viable path (3.6) can definitely be continued to $\boldsymbol{v^1}$. The stochastic domination of Y_{n,ε_n} by X_{δ,n,θ_n} now follows. Then one just needs to apply Proposition 3.1 and Theorem 1.5 follows immediately.

Remark 3.2. Suppose $\mathrm{Supp}(\eta)$ is also bounded, and that θ is a constant, independent of n. Let

$$C_{\eta,\theta} := \min_{l(I)=\theta/2,\ I \subseteq \mathrm{Supp}(\eta)} \int_I \eta(x)\, dx, \tag{3.7}$$

where I denotes a closed interval. Then this minimum exists and is non-zero. It follows from Proposition 3.1 and the argument above that the number $X = X(n)$ of accessible paths in this case satisfies

$$X \gtrsim n! \cdot C_{\eta,\theta}^{n-1}, \tag{3.8}$$

The point is that $C_{\eta,\theta} \in (0,1]$ is a constant depending only on η and θ.

Acknowledgements

We thank Joachim Krug for making us aware of the problems studied here, and both he and Stefan Nowak for helpful discussions. We thank both referees for their very careful reading of the manuscript.

References

[1] Aita T., Uchiyama H., Inaoka T., Nakajima M., Kokubo T. and Husimi Y. (2000). Analysis of a local fitness landscape with a model of the rough Mt. Fuji-type landscape: application to prolyl endopeptidase and thermolysin. *Biopolymers* **54**, No.1, 64–79.

[2] Alon N. and Spencer J. (2008). *The Probabilistic Method*, 3rd ed. Wiley, Hoboken NJ. MR2437651

[3] Carneiro M. and Hartl D.L. (2010). Adaptive landscapes and protein evolution. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 1747–1751.

[4] Comtet L. (1972). Sur les coefficients de l'inverse de la série formelle $\sum n!t^n$ (French). *C. R. Acad. Sci. Paris Sér. A-B* **275**, A569–A572. MR0307928

[5] Comtet L. (1974). *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, Reidel, Dordrecht. MR0460128

[6] Franke J. and Krug J. (2012). Evolutionary accessibility in tunably rugged fitness landscapes. *J. Stat. Phys.* **148**, No.4, 705–722.

[7] Franke J., Klözer A., de Visser J.A.G.M. and Krug J. (2011). Evolutionary accessibility of mutational pathways. *PLoS Comput. Biol.* **7**, No.8, e1002134, 9pp. MR2845072

[8] Franke J., Wergen G. and Krug J. (2010). Records and sequences of records from random variables with a linear trend. *J. Stat. Mech. Theory Exp.* **10**, P10013, 21pp. MR2800498

[9] Kauffman S. and Levin S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *J. Theoret. Biol.* **128**, No.1, 11–45. MR0907587

[10] Kingman J.F.C. (1978). A simple model for the balance between selection and mutation. *J. Appl. Probability* **15**, No.1, 1–12. MR0465272

[11] Klözer A. (2008). *NK Fitness Landscapes*, Diplomarbeit Universität zu Köln.

[12] Mézard M., Parisi G. and Virasoro M.A. (1987). *Spin Glass Theory and Beyond*, World Scientific Publishing, Singapore. MR1026102

[13] The Online Encyclopedia of Integer Sequences, Sequence #*A*003319. `http://oeis.org/A003319`

[14] The Online Encyclopedia of Integer Sequences, Sequence #*A*059438. `http://oeis.org/A059438`

[15] Weinreich D.M., Delaney N.F., DePristo M.A. and Hartl D.M. (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**, 111–114.

[16] Weinreich D.M., Watson R.A. and Chao L. (2005). Perspective: Sign epistasis and genetic constraints on evolutionary trajectories. *Evolution* **59**, No.6, 1165–1174.

Paper II

Unoriented first-passage percolation on the n-cube

Anders Martinsson

Abstract

The n-dimensional binary hypercube is the graph whose vertices are the binary n-tuples $\{0,1\}^n$ and where two vertices are connected by an edge if they differ at exactly one coordinate. We prove that if the edges are assigned independent mean 1 exponential costs, the minimum length T_n of a path from $(0,0,\ldots,0)$ to $(1,1,\ldots,1)$ converges in probability to $\ln(1+\sqrt{2}) \approx 0.881$. It has previously been shown by Fill and Pemantle (1993) that this so-called first-passage time asymptotically almost surely satisfies $\ln(1+\sqrt{2}) - o(1) \leq T_n \leq 1+o(1)$, and has been conjectured to converge in probability by Bollobás and Kohayakawa (1997). A key idea of our proof is to consider a lower bound on Richardson's model, closely related to the branching process used in the article by Fill and Pemantle to obtain the bound $T_n \geq \ln\left(1+\sqrt{2}\right) - o(1)$. We derive an explicit lower bound on the probability that a vertex is infected at a given time. This result is formulated for a general graph and may be applicable in a more general setting.

1 Introduction

The n-dimensional binary hypercube $\mathbb{Q}_n$ is the graph with vertex set $\{0,1\}^n$ where two vertices share an edge if they differ at exactly one coordinate. We let $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$ denote the all zeroes and all ones vertices respectively. For any vertex $v \in \mathbb{Q}_n$, we let $|v|$ denote the number of coordinates of v that are 1. A path $v_0 \to v_1 \to \cdots \to v_k$ in $\mathbb{Q}_n$ is called *oriented* if $|v_i|$ is strictly increasing along the path.

First-passage percolation is a random process on a graph G, which was introduced by Hammersley and Welsh. In this process, each edge e in the graph is assigned a random variable W_e called the *passage time* of e. In this paper, the passage times will always be independent exponentially distributed random variables with expected value 1. The usual way in which this process is described is that there exists some vertex $v_0 \in G$ which is assigned some property, usually either that it is infected (v_0 is the source of some disease) or wet (v_0 is connected to a water source), which then spreads throughout the graph. The passage time of an edge corresponds to the time it takes for an infection to spread in any direction along the edge, that is, when a vertex v gets infected the infection spreads to each neighbor w after $W_{\{v,w\}}$ time, assuming w is not already infected at that time. More concretely, we can let the edge weights generate a metric on G. For a path γ in G we define the *passage time* of γ as the sum of passage times of the edges along γ. Moreover, for any two vertices $v, w \in G$, we say that the *first-passage time* from v to w, denoted by $d_W(v, w)$, is the infimum of passage times over all paths from v to w in G. Then for any $v \in G$, the time at which v is infected is given by $d_W(v_0, v)$.

An alternative way to formulate first-passage percolation with independent exponentially distributed passage times is to consider the process $\{R(\cdot, t)\}_{t\geq 0}$, where for each $t \geq 0$, $R(v, t)$ is the map from the vertex set of G to $\{0, 1\}$ given by

$$R(v,t) = \begin{cases} 1 & \text{if } d_W(v_0, v) \leq t \\ 0 & \text{otherwise,} \end{cases} \tag{1.1}$$

that is, $R(v, t)$ is the indicator function for the event that v is infected at time t. When the edge passage times are independent exponentially distributed with mean one, the memory-less property implies that the process $\{R(\cdot, t)\}_{t\geq 0}$ is Markovian, and its distribution is given by the initial condition $R(\cdot, 0) = \delta_{v_0,\cdot}$ together with the transitions $\{R(\cdot) \to R(\cdot) + \delta_{v,\cdot}\}$ at rate equal to the number of infected neighbors of v if v is healthy, and 0 if v is infected, see [1]. Here $\delta_{\cdot,\cdot}$ denotes the Kronecker delta function. This Markov process is known as Richardson's model.

First-passage percolation and Richardson's model on the hypercube have previously been studied by Fill and Pemantle [2], and later by Bollobás and Kohayakawa [3]. For Richardson's model we always assume that the original infected vertex is $\hat{\mathbf{0}}$, though by transitivity of the hypercube it is clear that the analogous statements hold for any starting vertex. The quantities considered in these articles of most relevance to this paper are the first-passage time from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$, which we denote by T_n, the *oriented first-passage time* from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$, and the *covering time*. Note that, in terms of Richardson's model, T_n is the time until the vertex furthest from the starting point gets infected. The oriented first-passage time is a simplified version of the first-passage time, first proposed by Aldous [5], where the minimum is only taken over all oriented paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$. The covering time is the random amount of time in Richardson's model on $\mathbb{Q}_n$ until all vertices are infected or, equivalently, $\max_{v \in \mathbb{Q}_n} d_W(\hat{\mathbf{0}}, v)$, the maximum first-passage time from $\hat{\mathbf{0}}$ to any other vertex in $\mathbb{Q}_n$.

In case of oriented first-passage percolation, it was shown by Fill and Pemantle that the oriented first-passage time from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ converges to 1 in probability as $n \to \infty$. The fact that $1 - o(1)$ is an asymptotic almost sure lower bound had already been observed by Aldous in [5], and can be shown in a straight-forward manner by considering the expected number of oriented paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ with passage time at most t. The argument by Fill and Pemantle for the upper bound is essentially a second moment analysis on the number of such paths, though as they remark, a direct application of the second moment method can only show that the probability that the oriented first-passage time is at most $1 + \varepsilon$ is bounded away from 0. To circumvent this, they consider a "variance reduction trick", which effectively means that they consider a slightly different random variable.

For the unoriented first-passage time from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$, Fill and Pemantle showed that, as $n \to \infty$, we have

$$\ln\left(1 + \sqrt{2}\right) - o(1) \leq T_n \leq 1 + o(1) \tag{1.2}$$

with probability $1-o(1)$. The upper bound follows directly from the oriented first-passage time. They remark that they doubt the upper bound is sharp, but state that they do not know how to improve it. Prior to this article, this seems to be the best known upper bound on T_n. For the lower bound, Fill and Pemantle relayed an argument by Durrett. In this argument we consider a random process on $\mathbb{Q}_n$, which Durrett calls a branching translation process (BTP). We will postpone the definition of this process to the next section, but the essential difference to Richardson's model is that we allow each site to contain multiple instances of the infection at the same time. Durrett argues that this process stochastically dominates Richardson's model in the sense that it is possible to couple the models such that the infected vertices in Richardson's model are always a subset of the so-called occupied vertices in the BTP. He proves that the time at which $\hat{\mathbf{1}}$ becomes occupied tends to $\ln\left(1+\sqrt{2}\right)$ in probability as $n \to \infty$. As BTP stochastically dominates Richardson's model, this directly implies that $T_n \geq \ln\left(1+\sqrt{2}\right) - o(1) = 0.881\cdots - o(1)$ with probability $1 - o(1)$.

Bollobás and Kohayakawa [3] showed that many global first-passage percolation properties on $\mathbb{Q}_n$, such as the covering time and the graph diameter with respect to $d_W(\cdot,\cdot)$, can be bounded from above in terms of T_n. They defined the quantity

$$T_\infty = \inf\left\{t \in \mathbb{R} | \mathbb{P}\left(T_n \leq t\right) \to 1 \text{ as } n \to \infty\right\}. \tag{1.3}$$

Their main result is that asymptotically almost surely the covering time is at most $T_\infty + \ln 2 + o(1)$ and the graph diameter is at most $T_\infty + 2\ln 2 + o(1)$. Note that it follows from the results by Fill and Pemantle that $\ln\left(1+\sqrt{2}\right) \leq T_\infty \leq 1$. Furthermore, it is easy to see that if T_n converges in probability as $n \to \infty$, then it must converge to T_∞. In fact, Bollobás and Kohayakawa explicitly conjectured that this is the case, and they consequently referred to T_∞ as simply the first-passage percolation time between two antipodal vertices in $\mathbb{Q}_n$. While their article does not prove that T_n converges in probability, the ideas do have some implications for T_n. For instance, with some small modifications of their proof it follows that if T_n converges in distribution, then the limit must be concentrated on one point, meaning that T_n converges in probability.

Besides first-passage percolation, percolation on the hypercube with restriction to oriented paths has also been considered in regards to Bernoulli percolation by Fill and Pemantle (in the same article), and, more recently, accessibility percolation[1] by Hegarty and the author in [4]. Common for these three cases of oriented percolation is that the proofs are based on second moment analyses. Arguably, this is made possible by the relatively simple combinatorial properties of oriented paths. We have $n!$ oriented paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ in $\mathbb{Q}_n$, all of length n and all equivalent up to permutation of coordinates. Perhaps more importantly, one can derive good estimates on the number of pairs of oriented paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ that

[1] The name accessibility percolation is not mentioned in the cited article. The term was coined by Joachim Krug and Stefan Nowak after its writing.

intersect a given number of times, something which is made possible by the natural representation of oriented paths as permutations. In contrast, general paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ do not seem to have a similar representation in any meaningful way, and in any case, there is certainly a lot more variation between general paths than oriented such. Hence, it seems that these type of ideas from oriented percolation on the hypercube cannot be transferred to unoriented percolation.

The most promising approach to improve the result by Fill and Pemantle for T_n seems to be the BTP. Comparing the BTP to path-counting arguments, on the hypercube the former has the advantage that a number of relevant quantities, such as moment estimates, can be expressed by explicit analytical expressions, hence circumventing the problem of counting paths. However, beyond the fact that the BTP stochastically dominates Richardson's model, the relation between the two models is fairly subtle. It is therefore not immediately clear how proving anything about the BTP could imply upper bounds on the first-passage time.

In this article, we propose a way to do precisely this. A central idea of our approach is to consider a subprocess of the BTP with two important properties: Firstly, Richardson's model is stochastically sandwiched between the full BTP and this subprocess, and secondly, it is possible to derive an explicit lower bound on the probability that a vertex is occupied at a given time in this subprocess, expressed in tractable quantities for the BTP. Applying these ideas to the hypercube, we are able to resolve the problem of determining the limit of T_n. This result is summarized in the following Theorem, which is the main result of this paper:

Theorem 1.1. *Let T_n denote the first-passage time from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ in $\mathbb{Q}_n$ with exponentially distributed edge costs with mean 1. For any $1 \leq p < \infty$ we have $\|T_n - \ln\left(1+\sqrt{2}\right)\|_p = \Theta\left(\frac{1}{n}\right)$. In particular, we have $\mathbb{E}T_n = \ln\left(1+\sqrt{2}\right) + O\left(\frac{1}{n}\right)$ and $\operatorname{Var}(T_n) = \Theta\left(\frac{1}{n^2}\right)$.*

A direct consequence of this result is that $T_\infty = \ln\left(1+\sqrt{2}\right)$, which in particular improves the best known upper bound on the covering time to $\ln\left(1+\sqrt{2}\right) + \ln 2 + o(1) = 1.574\cdots + o(1)$. One can compare this with the best known lower bound $\frac{1}{2}\ln\left(2+\sqrt{5}\right) + \ln 2 - o(1) = 1.414\cdots - o(1)$, as shown by Fill and Pemantle.

Given this result for T_n, the question naturally arises how the path from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ with the smallest first-passage time typically behaves. In particular, how long is this path (here length means the number of edges along the path), and how are the "backsteps" distributed along it. Let us denote this path by Γ_n. This question may also be interesting from the point of view of accessibility percolation. Though strictly speaking not part of the mathematical formulation of accessibility percolation, shorter paths are considered more biologically feasible. Hence, an important question for unoriented accessibility percolation on the hypercube is how much longer typical accessible paths are in this case compared to oriented accessibility percolation.

We propose the following way to describe the asymptotic properties of Γ_n: Run a simple random walk on $\mathbb{Q}_n$ starting at $\hat{\mathbf{0}}$ with rate n for $\ln\left(1+\sqrt{2}\right)$ time, and condition on the event that the walk stops at $\hat{\mathbf{1}}$. Let σ_n denote the traversed path.

Theorem 1.2. *Any asymptotically almost sure property of σ_n is also an asymptotically almost sure property of Γ_n. In particular, the length of Γ_n is asymptotically almost surely $\sqrt{2}\ln\left(1+\sqrt{2}\right)n \pm o(n)$.*

In applying Theorem 1.2, it is helpful to note that each coordinate of a simple random walk on $\mathbb{Q}_n$ with rate n is an independent simple random walk on $\{0,1\}$ with rate 1.

The remainder of the paper will be structured in the following way: In Section 2 we define the BTP and describe our stochastical sandwiching of Richardson's model. At the end of this section, we give an outline of the proof of Theorem 1.1. This proof is divided into three steps, which are shown in Sections 3, 4 and 5 respectively. Lastly, in Section 6 we give a short proof of Theorem 1.2 based on ideas from the preceding section.

2 Richardson's model, the BTP, and uncontested particles

We first give an overview of the technique used by Durrett to obtain the lower bound on T_n in [2]. To accommodate Theorem 2.2 below, we present this technique in terms of a general graph G rather than just the hypercube. We remark that though Durrett only defined the branching translation process for the hypercube, the process can be extended to a general graph unambiguously. We let v_0 denote a fixed vertex in G. For simplicity, we will assume that G is finite, connected and simple.

The *branching translation process* (BTP), as introduced by Durrett, is a branching process on G defined in the following way: At time 0 we place a particle at v_0. After this, each existing particle generates offspring independently at rate equal to the degree of the vertex it is placed at. Each offspring is then placed with uniform probability at any neighboring vertex. Equivalently, each existing particle generates offspring at each neighboring vertex independently with rate 1. For a fixed G and fixed location of the first particle $v_0 \in G$, we let $Z(v,t)$ denote the number of particles at vertex v at time t in the BTP (originating at v_0) and define $m(v,t) = \mathbb{E}Z(v,t)$. One can observe that $\{Z(\cdot,t)\}_{t\geq 0}$ is a Markov process with the initial value $Z(v,0) = \delta_{v,v_0}$ and where, for each vertex v, the transition $\{Z(\cdot) \to Z(\cdot) + \delta_{\cdot,v}\}$ occurs at rate $\sum_{w\in N(v)} Z(w)$ where $N(v)$ denotes the neighborhood of v. It can be noted that in [2], the BTP was formally defined as this Markov process. However, this way to describe the states contains an insufficient amount of information for our applications since there is no way to discern ancestry. We will return to the problem of formally defining the state space of the BTP in Section 3. For now, the reader not satisfied with the informal definition of the BTP given here is free to consider any state space in which the particles can be individually identified and for each particle except the first, it is possible to determine its parent.

Below, we will use the terms ancestor and descendant of a particle to denote the natural partial order of particles generated by the BTP. For convenience, we use the convention that a particle is both an ancestor and a descendant of itself. We will sometimes write $x \geq y$ to denote that x is a descendant of y, and $x \leq y$ to denote that x is an ancestor of y. The terms parent and child are defined in the natural way. In order to indicate the location of a child of a particle x, we will sometimes use the term e-child of x to denote a child of x which at the time of its birth was displaced along an edge e. We define the *ancestral line* of a particle x as the ordered set of all ancestors of x (including x itself). If σ is the path obtained by following the locations of the vertices along the ancestral line of a particle x, then we say that the ancestral line of x *follows* σ, and we say that the ancestral line of x is *simple* if this path is simple. In certain parts of our proof we will need to consider BTPs where the location of the initial particle can vary. In that case, we will refer to a BTP where the original particle is placed at v as the BTP *originating at* v.

As pointed out in [2], the BTP stochastically dominates Richardson's model in the sense that, for a common starting vertex v_0, the models can be coupled in such a way that $R(v,t) \leq Z(v,t)$ for all $v \in G$ and $t \geq 0$. This is clear from a comparison of the transition rates of Z and R. However, for our applications we need to consider this relation more closely. To this end, we imagine that we partition the particles in the BTP into two sets, which we call the set of *alive* particles and the set of *ghosts*. We stress that the state of a particle is decided at the time of its birth, and is then never changed. The original particle is placed in the set of alive particles. After this, whenever a new particle is born it is placed in the set of ghosts if its parent is a ghost or if its location is already occupied by an alive particle, and placed in the set of alive particles otherwise. Clearly, the subprocess of the BTP consisting of all alive particles initially contains one particle, located at v_0, and it is straightforward to see that the rate at which alive particles are born at a given vertex v equals the number of adjacent vertices that contain alive particles if v does not currently contain an alive particle, and 0 if it does. As this is the same transition rate as for the corresponding transition in Richardson's model, we can consider Richardson's model as the subprocess of the BTP consisting of all alive particles. In a sense, for an observer not able to see the ghosts, the BTP will look like Richardson's model. Hence, with this coupling, the time at which a vertex gets infected is equal to one of the arrival times at the corresponding vertex in the full BTP, though not necessarily the first. We may here note that as at most one particle can be alive at each vertex, we can interpret $R(v,t)$ as the number of alive particles at v at time t.

A simplified version of the proof of the lower bound on T_n in [2] can now be summarized as follows: Consider a BTP on $\mathbb{Q}_n$ originating at $\hat{\mathbf{0}}$. Since the BTP dominates Richardson's model it suffices to show that with probability $1-o(1)$, no particle occupies $\hat{\mathbf{1}}$ at time $\ln\left(1+\sqrt{2}\right)-\varepsilon$ for all $\varepsilon > 0$ fixed. This is shown by a first moment method. It follows from standard methods in the theory of continuous-time Markov chains that $m(v,t)$ is the unique solution to the initial value problem

$$\begin{aligned} \frac{d}{dt} m(v,t) &= \sum_{w \in N(v)} m(w,t), \ t > 0 \\ m(v,0) &= \delta_{v,v_0}. \end{aligned} \tag{2.1}$$

In the case where $G = \mathbb{Q}_n$ and $v_0 = \hat{\mathbf{0}}$, it is straightforward to check that the solution to (2.1) is

$$m(v,t) = (\sinh t)^{|v|} (\cosh t)^{n-|v|} \tag{2.2}$$

and hence $m(\hat{\mathbf{1}}, t) = (\sinh t)^n$. Clearly, this tends to 0 as $n \to \infty$ for any $t < \sinh^{-1} 1 = \ln\left(1 + \sqrt{2}\right)$, as desired.

Inspired by the coupling between Richardson's model and the BTP as above, we introduce the notion of a particle being uncontested. For a particle x in a BTP, we let $c(x)$ denote the number of pairs of distinct particles y, z such that

- y is an ancestor of x
- y and z occupy the same vertex
- z was born before y.

Note that, according to our definition of ancestor, it is allowed for y to be equal to x. We let $a(x)$ denote the number of such pairs where z is an ancestor of x, and let $b(x)$ denote the number of pairs where z is not an ancestor of x. Clearly $a(x) + b(x) = c(x)$. We say that a particle x is *uncontested* if $c(x) = 0$.

Lemma 2.1. *We have the following properties:*

i) $a(x) = 0$ *if and only if the ancestral line of* x *is simple*

ii) if a particle is uncontested, then it is the first particle to be born at its location

iii) if a particle is uncontested, then it is alive.

Proof. i) This is obvious. *ii)* If some particle z was born before x at a vertex, then the pair (x, z) is counted in $c(x)$. *iii)* For any ghost x in the BTP, there must exist an earliest ancestor y which is a ghost. As the original particle is, by definition, alive, y must have a parent in the BTP. As the parent of y is alive but y is a ghost, the vertex occupied by y must already have been occupied by some alive particle z at the time of birth of y. The pair (y, z) is then counted in $c(x)$. ■

The third property is of particular interest as it allows us to express a lower bound on Richardson's model in terms of the BTP. Letting $Z_k(v, t)$ denote the number of particles x at vertex v at time t such that $c(x) = k$, we conclude that

$$Z_0 \overset{d}{\leq} \text{Richardson's model} \overset{d}{\leq} Z, \tag{2.3}$$

and with the proposed coupling between BTP and Richardson's model above we even have $Z_0 \leq R \leq Z$. However, it should be noted that, unlike Z and R, there is no reason why $Z_0(v)$ could not remain 0 forever. In fact, with the exception of the case where G is a chain of length 1, this occurs with positive probability. In order to see this, one can observe that if the first particle to arrive at a vertex is contested, which occurs with positive probability, then this particle will prevent all subsequent particles from being uncontested. On the other hand, in the event that $Z_0(v)$ is eventually non-zero, it follows from the second and third properties in Lemma 2.1 that the uncontested particle must have been the first particle at v and that this particle must have been alive. Hence, either $Z_0(v)$ remains 0 forever, or the time of the first arrival at v coincides in all three models.

2.1 Outline of proof of Theorem 1.1

For each vertex v and $t \geq 0$, we define $A(v, t)$ and $B(v, t)$ as the expected value of $\sum_x a(x)$ and $\sum_x b(x)$ respectively, where the sums goes over all particles at vertex v at time t in the BTP. We similarly define $S(v, t)$ as the expected number of particles at vertex v at time t with simple ancestral lines, that is, the expected number of particles x at v at time t such that $a(x) = 0$. The core of finding upper bounds on the first-passage time using the BTP is the following theorem, which will be shown in Section 3:

Theorem 2.2. *Let G be a finite connected simple graph. Consider the BTP on G originating at v_0, and let $Z_0(v, t)$, $B(v, t)$ and $S(v, t)$*

be as above. Then, for any vertex v and $t \geq 0$ we have

$$\mathbb{P}\left(Z_0(v,t) > 0\right) \geq S(v,t)e^{-\frac{B(v,t)}{S(v,t)}}. \tag{2.4}$$

In essence, Theorem 2.2 states that if, at a time t, the expected number of particles with simple ancestral line at v in the BTP is bounded away from 0, and if $B(v,t)$ is bounded, then with probability bounded away from 0 there is a particle at v at this time such that $a(x) = b(x) = 0$. Using the relation between the BTP and Richardson's model in (2.3), this immediately implies a lower bound on the probability that the first-passage time from v_0 to v in G is at most t. We remark that while the left-hand side of (2.4) certainly is increasing in t, the right-hand side is generally not, and instead typically attains a maximum for t such that $m(v,t) \approx 1$.

We now apply this result to the hypercube. We let $G = \mathbb{Q}_n$, $v_0 = \hat{\mathbf{0}}$ and $t = \vartheta := \ln\left(1+\sqrt{2}\right)$. In this case, the quantities $A(\hat{\mathbf{1}},\vartheta)$ and $B\left(\hat{\mathbf{1}},\vartheta\right)$ can be expressed analytically in a similar manner to the variance calculations for the BTP in [2]. This will be done in Section 4. The result of this can be summarized as follows:

Proposition 2.3. *For $\vartheta = \ln\left(1+\sqrt{2}\right)$, we have*

$$A\left(\hat{\mathbf{1}},\vartheta\right) = \frac{\vartheta}{\sqrt{2}} + o(1) = 0.623\cdots + o(1) \tag{2.5}$$

$$B\left(\hat{\mathbf{1}},\vartheta\right) = \vartheta + \frac{1}{3-2\sqrt{2}} + o(1) = 6.709\cdots + o(1). \tag{2.6}$$

In order to bound $S(\hat{\mathbf{1}},\vartheta)$, we observe that $A(v,t)$ is an upper bound on the expected number of particles at v at time t whose ancestral lines are not simple. This follows directly from the definition of $A(v,t)$ as $a(x)$ is an upper bound on the indicator function for the event that $a(x)$ is non-zero. We conclude that

$$m(v,t) - A(v,t) \leq S(v,t) \leq m(v,t), \tag{2.7}$$

and in particular, $1 - \frac{\vartheta}{\sqrt{2}} - o(1) = 0.376\cdots - o(1) \leq S(\hat{\mathbf{1}},\vartheta) \leq 1$.

Plugging these values into Theorem 2.2 we conclude the following:

Corollary 2.4. *Let T_n denote the first-passage time from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ in $\mathbb{Q}_n$ and let $\vartheta = \ln\left(1+\sqrt{2}\right)$. There exists a constant $\varepsilon > 0$ such that $\mathbb{P}(T_n \leq \vartheta) \geq \varepsilon$ for all n, and in particular* $\liminf_{n\to\infty} \mathbb{P}(T_n \leq \vartheta) \geq 6.9 \cdot 10^{-9}$.

Proof. The asymptotic lower bound on $\mathbb{P}(T_n \leq \vartheta)$ is obtained directly from Theorem 2.2 and Proposition 2.3. From this, the uniform bound follows by the observation that $\mathbb{P}(T_n \leq \vartheta)$ is non-zero for all n. ■

For our applications, we will need a more technical version of Corollary 2.4, Proposition 5.3, but other than that we are done with the BTP given this result. It may seem like Corollary 2.4 is far from our claimed result of convergence in L^p-norm, but given this result there are in fact a number of different ways to show that T_n converges to ϑ at least in probability, using the self-similar structure of the hypercube. One could for instance apply the ideas by Bollobás and Kohayakawa in [3]. In this paper we will instead apply a bootstrapping argument similar to one given in [4], which has the benefit of letting us get good bounds on the L^p-norms of $T_n - \vartheta$. This will be shown in Section 5, completing the proof of Theorem 1.1.

3 Proof of Theorem 2.2

Before proceeding with the proof, we need to discuss the parametrization of the BTP more carefully. For a BTP originating at a vertex v, a particle is identified by a finite sequence $\{e_1t_1e_2t_2 \ldots e_kt_k\}$ where $e_1, e_2, \ldots, e_k$ are edges forming a path that starts at v and $t_1, t_2, \ldots, t_k$ are positive real numbers. The original particle is identified by $\{\}$, the empty sequence. For any other particle x, $e_1e_2 \ldots e_k$ denotes the edges along the path followed by the ancestral line of x, and if $x_0, x_1, \ldots, x_k = x$ are the ancestors of x in ascending order, then for each $1 \leq i \leq k$, we have t_i equal to the time from the birth of x_{i-1} to the birth of x_i. It is easy to see that such a sequence uniquely defines the location and birth time of x. In particular, as, almost surely, no two particles are born at exactly the same time, this means that this representation is unique for each particle in the BTP. Note that this means that the parent of $x = \{e_1t_1e_2t_2 \ldots e_kt_k\}$ is $\{e_1t_1e_2t_2 \ldots e_{k-1}t_{k-1}\}$. More generally, the ancestors of x are the prefixes of x of even length. By a BTP originating at a vertex v we formally mean a random set of particles, which is interpreted as the set of all particles that will ever be born in the BTP, and, of course, whose distribution is given according to the transition rates as described above. We remark that this means that the event that a particle $x = \{e_1 t_1 e_2 t_2 \dots e_k t_k\}$ exists is interpreted as the event that the original particle has a e_1-child at time t_1, that this child has an e_2-child at time $t_1 + t_2$ and so on.

Below will use $\oplus$ to denote concatenation of sequences. For instance, if y is a child of x, born a time t after its parent and displaced along the edge e, then we may write $y = x \oplus \{et\}$. For a sequence a and a set of sequences B, we define $a \oplus B = \{a \oplus b | b \in B\}$.

It is easy to see that, in a BTP, each vertex can at most contain one uncontested particle, see for instance property *ii)* in Lemma 2.1. This means that the probability that a vertex v contains an uncontested particle at time t is equal to the expected number of such particles. Hence the conclusion of Theorem 2.2 basically states that among the particles at v at time t such that $a(x) = 0$, the probability that $b(x) = 0$ is on average at least $\exp\left(-\frac{B(v,t)}{S(v,t)}\right)$. In principle, it is possible to show this by considering the conditional distributions of $b(x)$ given the event that the particle x exists in the BTP. However, it is not formally possible by the usual definitions of conditional expectation and conditional distribution to condition on the event that a particle exists in the BTP since the event occurs with probability 0 and the particle itself is not the output of some well-defined random variable. In order to solve this problem, we need some ideas from Palm theory, and, in particular, the following special case of the Slivnyak-Mecke formula. The proof of this can be found in various text books on point processes. See for instance Corollary 3.2.3 in [6].

Theorem 3.1. *(Slivnyak-Mecke formula) Let* $\mathbf{T}$ *be a Poisson point process on the positive part of the real line with with constant intensity* 1*. Let* G *be a function mapping pairs* (T, t) *where* T *is a discrete subset of* $\mathbb{R}_+$ *and* $t \in T$ *to non-negative real numbers. Then*

$$\mathbb{E} \sum_{t \in \mathbf{T}} G(\mathbf{T}, t) = \int_0^\infty \mathbb{E} G\left(\mathbf{T} \cup \{t\}, t\right) \, dt. \tag{3.1}$$

If instead of a Poisson process on $\mathbb{R}_+$, we imagine $\mathbf{T}$ being a random subset of a finite, or even countable set, then we clearly have

$$\mathbb{E} \sum_{t \in \mathbf{T}} G(\mathbf{T}, t) = \sum_t \mathbb{P}\left(t \in \mathbf{T}\right) \mathbb{E}\left[G(\mathbf{T}, t) | t \in \mathbf{T}\right]. \tag{3.2}$$

By the standard way to translate this statement, if $\mathbf{T}$ is a Poisson process on $\mathbb{R}_+$ with constant intensity 1, then we would expect the sum over t to translate to an integral and $\mathbb{P}(t \in \mathbf{T})$ to dt, the Lebesgue measure on $\mathbb{R}_+$. Hence, the theorem states that if $\mathbf{T}$ is a Poisson process as above, then we should translate $\mathbb{E}[G(\mathbf{T}, t) | t \in \mathbf{T}]$ to $\mathbb{E}G(\mathbf{T} \cup \{t\}, t)$, and so we may interpret $\mathbf{T} \cup \{t\}$ as the conditional distribution of $\mathbf{T}$ given $t \in \mathbf{T}$.

The following lemma proves a corresponding result for the BTP. In a similar manner as above, we may interpret the lemma as that, conditioned on the event that a particle $x^{z_1,\ldots,z_l}$ exists in the BTP $\mathbf{X}_0$, the conditional distribution of the process is given by $\mathbf{X}^{z_1,\ldots,z_l}$, where $x^{z_1,\ldots,z_l}$ and $\mathbf{X}^{z_1,\ldots,z_l}$ are as defined below. This result may be well-known from the properties of more general processes.

Lemma 3.2. *Let σ be a path of length $l \geq 1$. We denote the vertices along the path $v_0, \ldots, v_l$ and the edges $\sigma_1, \ldots, \sigma_l$. Let $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_l$ be independent branching translation processes where $\mathbf{X}_i$ for $0 \leq i \leq l$ is a BTP originating at vertex v_i. Let f be a function taking pairs (X, x), X a realization of a BTP and x a particle in X, to non-negative real numbers. Let $V_\sigma = V_\sigma(X)$ denote the set of particles at vertex v_l (no matter when they are born) whose ancestral line follows σ. Then, we have*

$$\mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} f(\mathbf{X}_0, x) = \int_0^\infty \cdots \int_0^\infty \mathbb{E} f(\mathbf{X}^{z_1,\ldots,z_l}, x^{z_1,\ldots,z_l})\, dz_1 \ldots dz_l \tag{3.3}$$

where

$$\begin{aligned} \mathbf{X}^{z_1,\ldots,z_l} = \mathbf{X}_0 &\cup (\{\sigma_1 z_1\} \oplus \mathbf{X}_1) \cup (\{\sigma_1 z_1 \sigma_2 z_2\} \oplus \mathbf{X}_2) \\ &\cup \cdots \cup (\{\sigma_1 z_1 \sigma_2 z_2 \ldots \sigma_l z_l\} \oplus \mathbf{X}_l) \end{aligned} \tag{3.4}$$

and $x^{z_1,\ldots,z_l} = \{\sigma_1 z_1 \sigma_2 z_2 \ldots \sigma_l z_l\}$.

Proof. For a vertex v and an edge e we let $e \ni v$ denote that v is one of the end points of e. For each edge $e \ni v_0$, we let $\mathbf{T}_e$ denote the set of birth times of the e-children of the original particle in $\mathbf{X}_0$. Clearly, $\mathbf{T}_e$ for $e \ni v_0$ are independent Poisson processes on $\mathbb{R}_+$ with constant intensity 1.

A central property of the BTP is that, after a particle is born, the set of its descendants is itself distributed as a BTP. Furthermore, this subprocess is then independent of the behavior of any other particle. Hence we can express $\mathbf{X}_0$ recursively by

$$\mathbf{X}_0 = \bigcup_{e \ni v_0} \bigcup_{t_i \in \mathbf{T}_e} \{e t_i\} \oplus \mathbf{Y}_{e,i} \tag{3.5}$$

where for each $e \ni v_0$ and each $i = 1, 2, \ldots,$ we have $\mathbf{Y}_{e,i}$ independently distributed as a BTP originating at the vertex opposite to v_0 along e. For any discrete set $T \subset \mathbb{R}_+$ we let $\mathbf{X}_0(T)$ denote the random variable obtained by replacing $\mathbf{T}_{\sigma_1}$ by T in (3.5). Then $\mathbf{X}_0(\cdot)$ is a random function independent of $\mathbf{T}_{\sigma_1}$, and we have $\mathbf{X_0} = \mathbf{X_0}(\mathbf{T}_{\sigma_1})$. Note that, by independence, $\mathbf{X}_0(T)$ is a version of the conditional distribution of $\mathbf{X}_0$ given $\mathbf{T}_{\sigma_1} = T$.

For each T as above and $t \in T$, we define

$$F(T) = \mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0(T))} f(\mathbf{X}_0(T), x) \tag{3.6}$$

$$F(T, t) = \mathbb{E} \sum_{\substack{x \in V_\sigma(\mathbf{X}_0(T)) \\ x \geq \{\sigma_1 t\}}} f(\mathbf{X}_0(T), x). \tag{3.7}$$

It is clear from the definition that, for any fixed T, we have $F(T) = \sum_{t \in T} F(T, t)$. Furthermore, as $\mathbf{T}_{\sigma_1}$ and $\mathbf{X}_0(\cdot)$ are independent we have $\mathbb{E}F(\mathbf{T}_{\sigma_1}) = \mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} f(\mathbf{X}_0, x)$. Hence by the Slivnyak-Mecke formula we have

$$\mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} f(\mathbf{X}_0, x) = \mathbb{E} \sum_{t \in \mathbf{T}_{\sigma_1}} F(\mathbf{T}_{\sigma_1}, t) = \int_0^\infty \mathbb{E}F(\mathbf{T}_{\sigma_1} \cup \{z_1\}, z_1)\, dz_1. \tag{3.8}$$

By independence of $\mathbf{X}_0(\cdot)$ and $\mathbf{T}_{\sigma_1} \cup \{z_1\}$ we can conclude that

$$\begin{aligned} &\mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} f(\mathbf{X}_0, x) \\ &\quad = \int_0^\infty \mathbb{E} \sum_{\substack{x \in V_\sigma(\mathbf{X}_0(\mathbf{T}_{\sigma_1} \cup \{z_1\})) \\ x \geq \{\sigma_1 z_1\}}} f(\mathbf{X}_0(\mathbf{T}_{\sigma_1} \cup \{z_1\}), x)\, dz_1. \end{aligned} \tag{3.9}$$

Let us now consider the random process $\mathbf{X}_0(\mathbf{T}_{\sigma_1} \cup \{z_1\})$. We can interpret the expression for $\mathbf{X}_0$ in (3.5) and the subsequent definition of $\mathbf{X}_0(T)$ as that these processes are generated by first determining the birth time for each child of the original particle, and then for each child independently generating a BTP which determines its descendants. When seen in this light, it is clear that the only difference between $\mathbf{X}_0$ and $\mathbf{X}_0(\mathbf{T}_{\sigma_1} \cup \{z_1\})$ is that the latter has an additional particle in generation 1. Hence, $\mathbf{X}_0(\mathbf{T}_{\sigma_1} \cup \{z_1\})$ has the same distribution as $\mathbf{X}_0 \cup (\{\sigma_1 z_1\} \oplus \mathbf{X}_1\})$, and so we can replace $\mathbf{X}_0(\mathbf{T}_{\sigma_1} \cup \{z_1\})$ in (3.9) by this other random process.

Letting $\tilde{\sigma} = \{\sigma_2, \sigma_3, \ldots, \sigma_l\}$, we note that the subset of elements in $V_\sigma(\mathbf{X}_0 \cup (\{\sigma_1 z_1\} \oplus \mathbf{X}_1\}))$ that are descendants of $\{\sigma_1 z_1\}$ is precisely the set $\{\sigma_1 z_1\} \oplus V_{\tilde{\sigma}}(\mathbf{X}_1)$. Hence (3.9) simplifies to

$$\mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} f(\mathbf{X}_0, x) = \int_0^\infty \mathbb{E} \sum_{x \in V_{\tilde{\sigma}}(\mathbf{X}_1)} f\left(\mathbf{X}_0 \cup (\{\sigma_1 z_1\} \oplus \mathbf{X}_1\}), \{\sigma_1 z_1\} \oplus x\right) dz_1. \tag{3.10}$$

The lemma follows by induction. If $l = 1$, then the only particle in $V_{\tilde{\sigma}}(\mathbf{X}_1)$ is $\{\}$, the original particle in $\mathbf{X}_1$, and so equation (3.10) simplifies to

$$\mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} f(\mathbf{X}_0, x) = \int_0^\infty \mathbb{E} f(\mathbf{X}^{z_1}, \{\sigma_1 z_1\})\, dz_1 \tag{3.11}$$

as desired.

Now assume $l > 1$. By the induction hypothesis we have for any non-negative function $\tilde{f}$

$$\mathbb{E} \sum_{x \in V_{\tilde{\sigma}}(\mathbf{X}_1)} \tilde{f}(\mathbf{X}_1, x) = \int_0^\infty \cdots \int_0^\infty \mathbb{E} \tilde{f}(\tilde{\mathbf{X}}^{z_2,\ldots,z_l}, \tilde{x}^{z_2,\ldots,z_l})\, dz_2 \ldots dz_l, \tag{3.12}$$

where

$$\begin{aligned} \tilde{\mathbf{X}}^{z_2,\ldots,z_l} &= \mathbf{X}_1 \cup (\{\sigma_2 z_2\} \oplus \mathbf{X}_2) \cup (\{\sigma_2 z_2 \sigma_3 z_3\} \oplus \mathbf{X}_3) \\ &\quad \cup \cdots \cup (\{\sigma_2 z_2 \sigma_3 z_3 \ldots \sigma_l z_l\} \oplus \mathbf{X}_l) \end{aligned} \tag{3.13}$$

and $\tilde{x}^{z_2,\ldots,z_l} = \{\sigma_2 z_2 \sigma_3 z_3 \ldots \sigma_l z_l\}$.

Let us consider the expression

$$\mathbb{E} \sum_{x \in V_{\tilde{\sigma}}(\mathbf{X}_1)} f\left(\mathbf{X}_0 \cup (\{\sigma_1 z_1\} \oplus \mathbf{X}_1\}), \{\sigma_1 z_1\} \oplus x\right), \tag{3.14}$$

the integrand on the right-hand side of equation (3.10). If we fix $z_1 > 0$ and condition on $\mathbf{X}_0 = X_0$, then $f\big(\mathbf{X}_0 \cup (\{\sigma_1 z_1\} \oplus \mathbf{X}_1\}), \{\sigma_1 z_1\} \oplus x\big)$ is a function of $\mathbf{X}_1$ and x only. By the induction hypothesis,

$$\mathbb{E} \sum_{x \in V_{\tilde{\sigma}}(\mathbf{X}_1)} f\left(X_0 \cup (\{\sigma_1 z_1\} \oplus \mathbf{X}_1\}), \{\sigma_1 z_1\} \oplus x\right) = \int_0^\infty \cdots \int_0^\infty$$
$$\mathbb{E} f(X_0 \cup \{\sigma_1 z_1\} \oplus \tilde{\mathbf{X}}^{z_2,\ldots,z_l}, \{\sigma_1 z_1\} \oplus \tilde{x}^{z_2,\ldots,z_l})\, dz_2 \ldots dz_l.$$

Hence, by integrating this expression over z_1 and $\mathbf{X}_0$ we conclude that

$$\mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} f(\mathbf{X}_0, x) = \int_0^\infty \dots \int_0^\infty$$
$$\mathbb{E} f(\mathbf{X}_0 \cup \{\sigma_1 z_1\} \oplus \tilde{\mathbf{X}}^{z_2,\dots,z_l}, \{\sigma_1 z_1\} \oplus \tilde{x}^{z_2,\dots,z_l})\, dz_1 \dots dz_l,$$

where clearly $\mathbf{X}^{z_1,\dots,z_l} = \mathbf{X}_0 \cup \{\sigma_1 z_1\} \tilde{\mathbf{X}}^{z_2,\dots,z_l}$ and $x^{z_1,\dots,z_l} = \{\sigma_1 z_1\} \oplus \tilde{x}^{z_2,\dots,z_l}$. ∎

Lemma 3.3. *Let* $\mathbf{X}$ *be a BTP originating at a vertex* v*. Let* φ *be an indicator function defined over the set of potential particles. If* $\varphi(\{\}) = 0$*, then*

$$\mathbb{P}\left(\varphi(x) = 0 \, \forall x \in \mathbf{X}\right) \geq \exp\left(-\mathbb{E} \sum_{x \in \mathbf{X}} \varphi(x)\right). \tag{3.15}$$

Proof. For any particle $x \in \mathbf{X}$, let $\psi(x)$ be the indicator function for the event that $\varphi(y) = 1$ for at least one descendant y of x. Clearly, we have $\sum_{x \text{ in gen } 1} \psi(x) \leq \sum_{x \in \mathbf{X}} \varphi(x)$. Furthermore, $\sum_{x \text{ in gen } 1} \psi(x) = 0$ if and only if $\sum_{x \in \mathbf{X}} \varphi(x) = 0$.

Let d denote the degree of the vertex v. Then the particles in generation one are born according to a Poisson process on $\mathbb{R}_+^d$. Conditioned on the particles in generation one, the random variables $\psi(x)$ for each such particle x are independent, and are one with probability only depending on the location and birth time of x. Hence, by the random selection property of a Poisson process, the particles in generation one that satisfy $\psi(x) = 1$ are also born according to a Poisson process, and, in particular, the number of such particles is Poisson distributed.

We conclude that the probability that $\varphi(x) = 0$ for all $x \in \mathbf{X}$ is $e^{-\mathbb{E}\sum_{x \text{ in gen } 1} \psi(x)}$, which is at least $e^{-\mathbb{E}\sum_{x \in \mathbf{X}} \varphi(x)}$. ∎

Proof of Theorem 2.2. For any path σ from v_0 to v, let $S_\sigma(v,t)$ and $B_\sigma(v,t)$ denote the contributions to $S(v,t)$ and $B(v,t)$ respectively from particles whose ancestral lines follow σ. Similarly, we define $P(v,t) = \mathbb{E} Z_0(v,t)$ and $P_\sigma(v,t)$ the contribution to $P(v,t)$ from particles whose ancestral lines follow σ. As no two particles at the same vertex can both be uncontested, $Z_0(v,t)$ can only assume the values 0 and 1, so $P(v,t)$ is indeed the probability that $Z_0(v,t)$ is non-zero.

We start by considering the case where σ is a non-simple path. As $S(v,t)$ is the expected number of particles at vertex v at time t whose ancestral line follows a simple path, it is clear that the contribution to $S(v,t)$ from any non-simple path is zero. Similarly, if the ancestral line of a particle follows a non-simple path, then the particle cannot be uncontested. Hence for any non-simple path σ we have $S_\sigma(v,t) = P_\sigma(v,t) = 0$, and trivially $B_\sigma(v,t) \geq 0$.

Let us now fix σ, a simple path from v_0 to v. We denote the length of σ by l. For any realization X of $\mathbf{X}_0$ and $x \in X$, let $T(X,x)$ denote the birth time of x. Then it follows from Lemma 3.2 that

$$\begin{aligned} &S_\sigma(v,t) \\ &= \mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} \mathbb{1}_{T(\mathbf{X}_0,x) \leq t} = \int_0^\infty \dots \int_0^\infty \mathbb{1}_{z_1+\dots+z_l \leq t}\, dz_1 \dots dz_l. \end{aligned} \tag{3.16}$$

In order to express B_σ and P_σ in a similar manner, we need to revise our notation. Strictly speaking, $b(x)$ is a function not only of a particle, but also of the realization of the BTP. Following the convention we have used earlier in this section, we will now denote this quantity by $b(\mathbf{X}_0, x)$. Using this notation we have

$$B_\sigma(v,t) = \mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} \mathbb{1}_{T(\mathbf{X}_0,x) \leq t} b(\mathbf{X}_0, x) \tag{3.17}$$

$$P_\sigma(v,t) = \mathbb{E} \sum_{x \in V_\sigma(\mathbf{X}_0)} \mathbb{1}_{T(\mathbf{X}_0,x) \leq t} \mathbb{1}_{b(\mathbf{X}_0,x)=0}. \tag{3.18}$$

Hence, again by Lemma 3.2

$$B_\sigma(v,t) = \int_0^\infty \dots \int_0^\infty \mathbb{1}_{z_1+\dots+z_l \leq t} \mathbb{E}\left[b(\mathbf{X}^{z_1,\dots,z_l}, x^{z_1,\dots,z_l})\right] dz_1 \dots dz_l \tag{3.19}$$

and

$$\begin{aligned} &P_\sigma(v,t) \\ &= \int_0^\infty \dots \int_0^\infty \mathbb{1}_{z_1+\dots+z_l \leq t} \mathbb{P}\left(b(\mathbf{X}^{z_1,\dots,z_l}, x^{z_1,\dots,z_l}) = 0\right) dz_1 \dots dz_l. \end{aligned} \tag{3.20}$$

Fix $z_1, \dots, z_l > 0$ such that $z_1 + \dots + z_l \leq t$ and consider the random variable $b(\mathbf{X}^{z_1,\dots,z_l}, x^{z_1,\dots,z_l})$. As σ is a simple path, it follows that $b(\mathbf{X}^{z_1,\dots,z_l}, x^{z_1,\dots,z_l})$ is equal to the number of particles $x \in \mathbf{X}^{z_1,\dots,z_l}$ such that, for some $0 \le i \le l$, x is born at the vertex v_i before time $\sum_{k=1}^{i} z_k$. This means that for appropriate indicator functions $\varphi_0, \dots, \varphi_l$ we have

$$b(\mathbf{X}^{z_1,\dots,z_l}, x^{z_1,\dots,z_l}) = \sum_{i=0}^{l} \sum_{x \in \mathbf{X}_i} \varphi_i(x). \tag{3.21}$$

As the original particles in $\mathbf{X}_0, \dots, \mathbf{X}_l$ correspond to ancestors of $x^{z_1,\dots,z_l}$, these are never counted in b and hence the corresponding indicator functions are always zero. Furthermore, as $\mathbf{X}_0, \dots, \mathbf{X}_l$ are independent processes, we have by Lemma 3.3

$$\begin{aligned} \mathbb{P}\left(b(\mathbf{X}^{z_1,\dots,z_l}, x^{z_1,\dots,z_l}) = 0\right) &= \prod_{i=0}^{l} \mathbb{P}\left(\varphi_i(x) = 0 \, \forall x \in \mathbf{X}_i\right) \\ &\ge \prod_{i=0}^{l} \exp\left(-\mathbb{E} \sum_{x \in \mathbf{X}_i} \varphi_i(x)\right) \\ &= \exp\left(-\mathbb{E} b(\mathbf{X}^{z_1,\dots,z_l}, x^{z_1,\dots,z_l})\right) \end{aligned} \tag{3.22}$$

Hence, by (3.20),

$$\begin{aligned} &P_\sigma(v,t) \\ &\ge \int_0^\infty \dots \int_0^\infty \mathbb{1}_{z_1+\dots+z_l \le t} \exp\left(-\mathbb{E} b(\mathbf{X}^{z_1,\dots,z_l}, x^{z_1,\dots,z_l})\right) dz_1 \dots dz_l. \end{aligned} \tag{3.23}$$

Let $r_0 \in \mathbb{R}$ be fixed. By convexity we have $e^{-r} \ge e^{-r_0}(1 + r_0) - e^{-r_0} r$. Applying this inequality to the integrand in (3.23) and comparing to (3.16) and (3.19) we get, for any simple path σ,

$$P_\sigma(v,t) \ge e^{-r_0}(1 + r_0) S_\sigma(v,t) - e^{-r_0} B_\sigma(v,t). \tag{3.24}$$

As remarked, for non-simple paths σ we have $P_\sigma = S_\sigma = 0$ and $B_\sigma \ge 0$, so clearly (3.24) holds for all paths σ from v_0 to v. Summing this inequality over all such paths σ, we get

$$P(v,t) \ge e^{-r_0}(1 + r_0) S(v,t) - e^{-r_0} B(v,t). \tag{3.25}$$

It is easy to verify that the right-hand side is maximized by $r_0 = \frac{B(v,t)}{S(v,t)}$, which yields the inequality $P(v,t) \ge S(v,t) e^{-\frac{B(v,t)}{S(v,t)}}$ as desired. ■

4 Proof of Proposition 2.3

Throughout this section we assume that the underlying graph in the BTP is $\mathbb{Q}_n$, and, unless stated otherwise, the BTP is assumed to originate at $\hat{\mathbf{0}}$. We will accordingly let $m(v,t)$ denote the expected number of particles at v at time t for a BTP originating at $\hat{\mathbf{0}}$, as given by (2.2). In order to simplify notation, we will interpret the vertices of $\mathbb{Q}_n$ as the elements of the additive group $\mathbb{Z}_2^n$, the n-fold group product of $\mathbb{Z}_2$, and we let $e_1, e_2, \ldots, e_n \in \mathbb{Z}_2^n$ denote the standard basis. We note that for any fixed vertex $w \in \mathbb{Q}_n$, the map $v \mapsto v - w$ is a graph isomorphism taking w to $\hat{\mathbf{0}}$. Hence, for a BTP originating at w, the expected number of particles at v at time t is given by $m(v-w,t)$. While addition and subtraction are equivalent in $\mathbb{Z}_2^n$, we will sometimes make a formal distinction between them in order to indicate direction.

Lemma 4.1. *For any $t > 0$ and $v \in \mathbb{Q}_n$ we have*

$$\frac{d^2}{dt^2} m(v,t) = \sum_{i=1}^{n} \sum_{j=1}^{n} m(v + e_i + e_j, t) \tag{4.1}$$

and

$$\frac{1}{2}\frac{d^2}{dt^2} m(v,t)^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} m(v+e_i+e_j,t)m(v,t) + m(v+e_i,t)m(v+e_j,t). \tag{4.2}$$

Proof. Recall that $m(v,t)$ satisfies

$$\frac{d}{dt} m(v,t) = \sum_{i=1}^{n} m(v + e_i, t). \tag{4.3}$$

This directly implies that

$$\begin{aligned}
\frac{d^2}{dt^2} m(v,t) &= \frac{d}{dt} \sum_{i=1}^{n} m(v + e_i, t) \\
&= \sum_{i=1}^{n} \frac{d}{dt} m(v + e_i, t) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} m(v + e_i + e_j, t).
\end{aligned}$$

The second equation now follows from $\frac{1}{2}\frac{d^2}{dt^2}m(v,t)^2 = m''(v,t)m(v,t) + m'(v,t)m'(v,t)$. ■

Lemma 4.2. *Let $s, t \geq 0$ and $v \in \mathbb{Q}_n$. Then*

$$\sum_{w \in \mathbb{Q}_n} m(w,s)m(v+w,t) = m(v,s+t). \tag{4.4}$$

Proof. If we condition on the state of the BTP at time s, then, at subsequent times, the process can be described as a superposition of independent branching processes, originating from each particle alive at time s. For each such process originating from a particle at vertex w, we have, by symmetry of $\mathbb{Q}_n$, that the expected number of particles at vertex v at time $t+s$ is $m(v+w,t)$. Hence

$$\mathbb{E}\left[Z(v,s+t)|Z(s)\right] = \sum_{w \in \mathbb{Q}_n} Z(w,s)m(v+w,t). \tag{4.5}$$

The lemma follows by taking the expected value of this expression. ■

We now turn to the problem of expressing $A(\hat{\mathbf{1}},u)$ and $B(\hat{\mathbf{1}},u)$ in terms of $m(v,t)$. Fix $u > 0$ and let $\mathbf{X}$ be a BTP on $\mathbb{Q}_n$ originating at $\hat{\mathbf{0}}$. Let $\mathcal{T}$ denote the random set of triples of particles (x,y,z) in $\mathbf{X}$ such that

- x is located at $\hat{\mathbf{1}}$ at time u
- y is an ancestor of x
- y and z occupy the same vertex
- z was born before y.

We furthermore partition this set into $\mathcal{T}_a$, the set of all such triples where y is a descendant of z, and $\mathcal{T}_b$, the set of all such triples where y is not a descendant of z. For any x at $\hat{\mathbf{1}}$ at time u in $\mathbf{X}$, it is clear that $c(x)$ gives the number of triples in $\mathcal{T}$ where the first element is x. Hence by summing $c(x)$ over all particles at $\hat{\mathbf{1}}$ at time u we obtain the size of $\mathcal{T}$. Similarly we see that $\sum_x a(x)$ and $\sum_x b(x)$ where x goes over all particles x at $\hat{\mathbf{1}}$ at time u gives the size of $\mathcal{T}_a$ and $\mathcal{T}_b$ respectively. Hence $A(\hat{\mathbf{1}},u) = \mathbb{E}\,|\mathcal{T}_a|$, $B(\hat{\mathbf{1}},u) = \mathbb{E}\,|\mathcal{T}_b|$ and $A(\hat{\mathbf{1}},u) + B(\hat{\mathbf{1}},u) = \mathbb{E}\,|\mathcal{T}|$.

In the following proposition, we derive explicit expressions for $A(\hat{\mathbf{1}}, u)$ and $B(\hat{\mathbf{1}}, u)$ by counting the expected number of elements in $\mathcal{T}_a$ and $\mathcal{T}$ respectively. Our argument is reminiscent of the second moment calculation for $Z(\hat{\mathbf{1}}, u)$ by Durrett in [2].

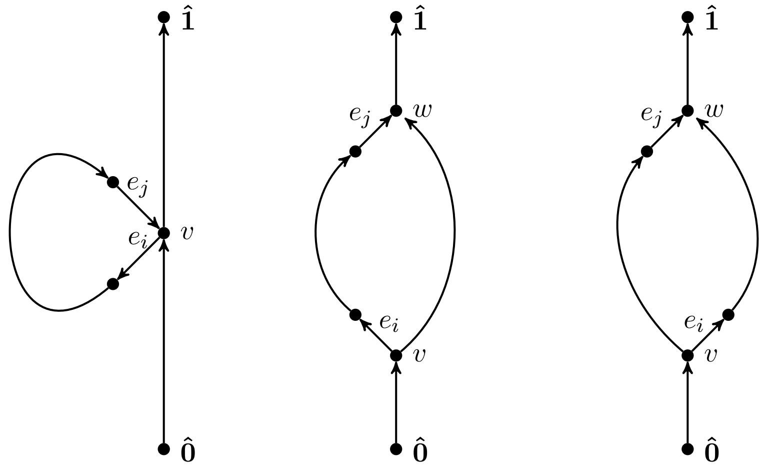
Proposition 4.3. *For any $u > 0$ we have*

$$A(\hat{\mathbf{1}}, u) = \sum_{v \in \mathbb{Q}_n} \sum_{i=1}^{n} \sum_{j=1}^{n} \int_0^\infty \int_0^\infty \mathbb{1}_{s+t \le u} m(v, s) \cdot \\ \cdot\, m(\hat{\mathbf{1}} - v, u - s - t) m(e_j + e_i, t)\, ds\, dt \tag{4.6}$$

$$A(\hat{\mathbf{1}}, u) + B(\hat{\mathbf{1}}, u) = \sum_{v \in \mathbb{Q}_n} \sum_{w \in \mathbb{Q}_n} \sum_{i=1}^{n} \sum_{j=1}^{n} \int_0^\infty \int_0^\infty \mathbb{1}_{s+t \le u} m(v, s) \cdot \\ \cdot\, m(\hat{\mathbf{1}} - w, u - s - t) \Big(m(w - v, t) m(w - v - e_i + e_j, t) \\ + m(w - v - e_i, t) m(w - v + e_j, t) \Big)\, ds\, dt. \tag{4.7}$$

Proof. Let us start by considering $A(\hat{\mathbf{1}}, u)$. For any $(x, y, z) \in \mathcal{T}_a$ there are well-defined particles c, the particle subsequent to z in the ancestral line of x, and p, the parent of y. We note that y is not a child of z as then y and z would not be located at the same vertex, hence c must be an ancestor of p. This means that the for each triple (x, y, z), the particles (x, y, z, c, p) must be related as illustrated in Graph 1 of Figure 1.

Fix $v \in \mathbb{Q}_n$, $1 \le i, j \le n$, and infinitesimal time intervals $(s, s + ds]$ and $(s+t, s+t+dt]$ where $0 \le s < s+t < u$. We now count the expected number of such quintuples of particles where the common location of y and z is v, the location of c is $v + e_i$, the location of p is $v - e_j$, c is born during $(s, s+ds]$ and y is born during $(s+t, s+t+dt]$. A particle is a potential z if it is located at v at time s. For each potential z, a potential c is a child of z born at $v + e_i$ during the time interval $(s, s + ds]$. For each pair of a potential z and c, a particle is a potential p if it is a descendant of c located at $v - e_j$ at time $s + t$. For each potential triple (z, c, p), a particle is a potential y if it is a child of p born at v during $(s + t, s + t + dt]$. Lastly, for each potential quadruple (z, c, p, y) each particle x at $\hat{\mathbf{1}}$ at time u which is a descendant of y forms a triple in $\mathcal{T}_a$. By computing the expected number of potential particles in each step, we see that



Graph 1. Graph 2. Graph 3.

Figure 1: Illustration of the possible configurations of ancestral lines of elements in $\mathcal{T}_a$ and $\mathcal{T}$ respectively. Graph 1 shows the configuration of elements in $\mathcal{T}_a$. Here z is located at v, c is a child of z at $v + e_i$, p is a descendant of c at $v - e_j$, y is a child of p at v, and x a descendant of y at $\hat{\mathbf{1}}$. The possible configurations corresponding to elements in $\mathcal{T}$ are shown in Graphs 2 and 3. After the ancestral lines of x and z split, the unique ancestors of x and z are given by the left-most and right-most paths respectively. In both configurations, the last common ancestor of x and z, l, is located at v, the first particle which is an ancestor of precisely one of x and z, c, is located at $v + e_i$, the parent of y, p, is located at $w - e_j$, y and z are located at w, and x is located at $\hat{\mathbf{1}}$.

the expected number of elements in $\mathcal{T}_a$ corresponding to fixed v, i, j and fixed time intervals $(s, s+ds]$ and $(s+t, s+t+dt]$ is

$$m(v,s)\,ds\,m(-e_j - e_i, t)\,dt\,m(v, u-s-t). \qquad (4.8)$$

Equation (4.6) follows by integrating over all $s, t > 0$ such that $s + t < u$ and summing over all $v \in \mathbb{Q}_n$ and all $1 \leq i, j \leq n$.

We now turn to the formula for $\mathbb{E}\,|\mathcal{T}|$. For each triple $(x, y, z) \in \mathcal{T}$ we define the particles l, the last common ancestor of x and z, c the first particle which is an ancestor of precisely one of x and z, and p the parent of y. Note that c must be a child of l. Similar to the case of $\mathcal{T}_a$, we note that we cannot have $c = y$. In order to see this, we assume that $c = y$. As c is the first particle to be an ancestor of precisely one of x and z, but z is older than y it follows that z must be an ancestor of x, and hence $l = z$. But then, $y = c$ and $z = l$ are located at adjacent vertices, which is a contradiction.

In order to count the elements in $\mathcal{T}$, we need to consider two cases depending on whether c is an ancestor of x or of z. In the former case, as $c \neq y$, c must be an ancestor of p and so the particles x, y, z, l, c and p must be related as illustrated in Graph 2 of Figure 1. Similarly, it is clear that in the latter case, the particles must be related as illustrated in Graph 3 in Figure 1.

We now fix $v, w \in \mathbb{Q}_n$, $1 \leq i, j \leq n$ and time intervals $(s, s+ds]$ and $(s+t, s+t+dt]$ where $0 \leq s < s+t < u$, and consider the elements in $\mathcal{T}$ where l is located at v, c is located at $v + e_i$, p is located at $w - e_j$, y and z are located at w, c is born during $(s, s+ds]$ and y is born during $(s+t, s+t+dt]$. We start by counting the triples where c is an ancestor of x. Here, a particle is a potential l if it is located at v at time s. For each potential l, a particle is a corresponding potential c if it is a child of l born at $v + e_i$ during $(s, s+ds]$. Hence the expected number of pairs of potential l:s and c:s is $m(v,s)\,ds$. For each pair of a potential l and c, we see that if one conditions on the BTP at the time of birth of c, the corresponding potential triples (p, y, x) originates from c whereas the potential z:s originate from l. Hence the potential triples (p, y, x) occur independently of the potential z:s. Furthermore, for each pair of a potential l and c, we see that the expected number of potential (p, y, x) is $m(w-e_j-v-e_i, t)\,dt\,m(\hat{\mathbf{1}}-w, u-s-t)$, and the expected number of potential z:s is $m(w-v, t)$. Combining this, we see that the expected number of elements in $\mathcal{T}$ corresponding to fixed v, w,

i, j, fixed time intervals as above and where c is an ancestor of x is

$$m(v,s)\,ds\,m(w-v,t)m(w-e_j-v-e_i,t)\,dt\,m(\hat{\mathbf{1}}-w,u-s-t). \quad (4.9)$$

Proceeding in a similar manner for the case where c is an ancestor of z we see that the expected number of corresponding elements in $\mathcal{T}$ is

$$m(v,s)\,ds\,m(w-e_j-v,t)m(w-v-e_i,t)\,dt\,m(\hat{\mathbf{1}}-w,u-s-t). \quad (4.10)$$

The proposition follows by summing these expressions over all $v, w \in \mathbb{Q}_n$, all $1 \leq i, j \leq n$ and integrating over all $s, t > 0$ such that $s + t < u$. ■

Remark 4.4. In the proof of Proposition 4.3, the only crucial property of the underlying graph is that it should not contain loops (if the graph does contain loops our counting argument may miss elements in $\mathcal{T}_a$ and $\mathcal{T}$). Hence this can directly be generalized to any loop-free graph by replacing the sums over i and j by sums over the corresponding neighborhoods.

Proposition 4.5. *For $\vartheta = \ln\left(1+\sqrt{2}\right)$, we have $A(\hat{\mathbf{1}},\vartheta) \to \frac{\vartheta}{\sqrt{2}}$ as $n \to \infty$.*

Proof. By reordering the sums and integrals in (4.6) we have

$$\begin{aligned} A(\hat{\mathbf{1}},\vartheta) = \int_0^\infty \int_0^\infty \mathbb{1}_{s+t\leq\vartheta} \sum_{v\in\mathbb{Q}_n} m(v,s)m(\hat{\mathbf{1}}-v,\vartheta-s-t)\cdot \\ \cdot \sum_{i=1}^n \sum_{j=1}^n m(e_j-e_i,t)\,ds\,dt. \end{aligned} \quad (4.11)$$

Applying Lemmas 4.1 and 4.2, the right-hand side simplifies to

$$\begin{aligned} &\int_0^\infty \int_0^\infty \mathbb{1}_{s+t\leq\vartheta} m(\hat{\mathbf{1}},\vartheta-t)\frac{d^2}{dt^2}m(\hat{\mathbf{0}},t)\,ds\,dt \\ &\quad = \int_0^\vartheta (\vartheta-t)m(\hat{\mathbf{1}},\vartheta-t)\frac{d^2}{dt^2}m(\hat{\mathbf{0}},t)\,dt, \end{aligned} \quad (4.12)$$

and by plugging in the analytical formula (2.2) for $m(v,t)$ we get

$$\begin{aligned} A(\hat{\mathbf{1}},\vartheta) &= \int_0^{\vartheta} (\vartheta - t)\,(\sinh(\vartheta - t))^n \,\frac{d^2}{dt^2}\,(\cosh t)^n \; dt \\ &= \int_0^{\vartheta} (\vartheta - t)\,(\sinh(\vartheta - t))^n \left(n + n(n-1)\,(\tanh t)^2\right) (\cosh t)^n \; dt \\ &= \int_0^{\vartheta} (\vartheta - t) \left(n + n(n-1)\,(\tanh t)^2\right) e^{nf(t)}\, dt, \end{aligned} \tag{4.13}$$

where $f(t) := \ln\left(\sinh(\vartheta - t)\cosh t\right)$.

What follows is a textbook application of the Lebesgue dominated convergence theorem. We begin examining the function f. The first and second derivatives of f are given by

$$f'(t) = -\coth(\vartheta - t) + \tanh t \tag{4.14}$$

$$f''(t) = -\operatorname{csch}(\vartheta - t)^2 + \operatorname{sech}(t)^2. \tag{4.15}$$

As $\operatorname{sech} t \leq 1$ for all $t \in \mathbb{R}$ and $\operatorname{csch} t \geq 1$ for $0 < t < \vartheta$, it follows that $f''(t) < 0$ for $0 < t < \vartheta$. Hence f is concave in this interval, so in particular $f(t) \leq f(0) + f'(0)\,t = -\sqrt{2}\,t$. Furthermore, we have $\tanh t \leq Ct$ for some appropriate $C > 0$.

Substituting t by $z = nt$ in (4.13), we obtain

$$A(\hat{\mathbf{1}},\vartheta) = \int_0^{\infty} \mathbb{1}_{z \leq n\vartheta} \left(\vartheta - \frac{z}{n}\right) \left(1 + (n-1)\tanh\left(\frac{z}{n}\right)^2\right) e^{nf\left(\frac{z}{n}\right)}\, dz. \tag{4.16}$$

It is clear that the integrand is bounded for all n by $\vartheta\left(1 + Cz^2\right) e^{-\sqrt{2}\,s}$, which is integrable over $[0,\infty)$. Hence, by dominated convergence, it follows that

$$A(\hat{\mathbf{1}},\vartheta) \to \int_0^{\infty} \vartheta e^{-\sqrt{2}\,z}\, dz = \frac{\vartheta}{\sqrt{2}} \text{ as } n \to \infty. \tag{4.17}$$

■

Proposition 4.6. *For $\vartheta = \ln\left(1 + \sqrt{2}\right)$ we have*

$$A(\hat{\mathbf{1}},\vartheta) + B(\hat{\mathbf{1}},\vartheta) \to \frac{\vartheta e^{\vartheta}}{\sqrt{2}} + \frac{1}{3 - 2\sqrt{2}} \quad \text{as } n \to \infty. \tag{4.18}$$

Hence, as $n \to \infty$ we have $B(\hat{\mathbf{1}},\vartheta) \to \vartheta + \frac{1}{3-2\sqrt{2}}$.

Proof. By reordering the sums in (4.6) and applying Lemma 4.1 we see that $A(\hat{\mathbf{1}},\vartheta)+B(\hat{\mathbf{1}},\vartheta)$ can be expressed as

$$\begin{aligned}\frac{1}{2}\sum_{v\in\mathbb{Q}_n}\sum_{w\in\mathbb{Q}_n}\int_0^\infty\int_0^\infty \mathbb{1}_{s+t\le\vartheta}m(v,s)m(\hat{\mathbf{1}}-w,\vartheta-s-t)\cdot\\ \cdot\frac{d^2}{dt^2}m(w-v,t)^2\,ds\,dt.\end{aligned}\tag{4.19}$$

Letting $\Delta=w-v$, this sum can be rewritten as

$$\begin{aligned}\frac{1}{2}\sum_{v\in\mathbb{Q}_n}\sum_{\Delta\in\mathbb{Q}_n}\int_0^\infty\int_0^\infty \mathbb{1}_{s+t\le\vartheta}m(v,s)\cdot\\ \cdot m(\hat{\mathbf{1}}-\Delta+v,\vartheta-s-t)\frac{d^2}{dt^2}m(\Delta,t)^2\,ds\,dt,\end{aligned}\tag{4.20}$$

which by Lemma 4.2 simplifies to

$$\frac{1}{2}\int_0^\vartheta(\vartheta-t)\sum_{\Delta\in\mathbb{Q}_n}m(\hat{\mathbf{1}}-\Delta,\vartheta-t)\frac{d^2}{dt^2}m(\Delta,t)^2\,dt.\tag{4.21}$$

To evaluate the sum in the above integral we use a small trick. Let us replace $\vartheta-t$ in this sum by z which we consider as a variable not depending on t. Then

$$\sum_{\Delta\in\mathbb{Q}_n}m(\hat{\mathbf{1}}-\Delta,z)\frac{d^2}{dt^2}m(\Delta,t)^2=\frac{\partial^2}{\partial t^2}\sum_{\Delta\in\mathbb{Q}_n}m(\hat{\mathbf{1}}-\Delta,z)m(\Delta,t)^2.$$

By grouping all terms with $|\Delta|=k$ we get

$$\begin{aligned}&\sum_{\Delta\in\mathbb{Q}_n}m(\hat{\mathbf{1}}-\Delta,z)m(\Delta,t)^2\\ &\quad=\sum_{k=0}^n\binom{n}{k}(\sinh z)^k(\cosh z)^{n-k}(\sinh t)^{2n-2k}(\cosh t)^{2k}\\ &\quad=\sum_{k=0}^n\binom{n}{k}\left(\sinh z\,(\cosh t)^2\right)^k\left(\cosh z\,(\sinh t)^2\right)^{n-k}\\ &\quad=\left(\sinh z\,(\cosh t)^2+\cosh z\,(\sinh t)^2\right)^n\\ &\quad=\left(\frac{1}{2}e^z\cosh 2t-\frac{1}{2}e^{-z}\right)^n.\end{aligned}$$

Note that $\frac{1}{2}e^z \cosh 2t - \frac{1}{2}e^{-z} > 0$ for any $t, z \geq 0$. Hence

$$\sum_{\Delta \in \mathbb{Q}_n} m(\hat{\mathbf{1}} - \Delta, z) \frac{d^2}{dt^2} m(\Delta, t)^2 = \frac{\partial^2}{\partial t^2} \left(\frac{1}{2} e^z \cosh 2t - \frac{1}{2} e^{-z} \right)^n$$

$$= 2ne^z \cosh t \left(\frac{1}{2} e^z \cosh 2t - \frac{1}{2} e^{-z} \right)^{n-1}$$

$$+ n(n-1)e^{2z} (\sinh 2t)^2 \left(\frac{1}{2} e^z \cosh 2t - \frac{1}{2} e^{-z} \right)^{n-2}.$$

Letting

$$f(t) = \ln \left(\frac{1}{2} e^{\vartheta - t} \cosh 2t - \frac{1}{2} e^{-\vartheta + t} \right) \tag{4.22}$$

$$g(t) = 2e^{\vartheta - t} \cosh t \, e^{-f(t)} \tag{4.23}$$

$$h(t) = e^{2\vartheta - 2t} (\sinh t)^2 \, e^{-2f(t)} \tag{4.24}$$

we can write

$$A(\hat{\mathbf{1}}, \vartheta) + B(\hat{\mathbf{1}}, \vartheta) = \frac{1}{2} \int_0^{\vartheta} (\vartheta - t) \left(n \, g(t) + n(n-1) \, h(t) \right) e^{nf(t)} \, dt. \tag{4.25}$$

One can check that $f(0) = f(\vartheta) = 0$, $f\left(\frac{1}{2}\right) < -\frac{1}{5}$, and that f has derivatives

$$f'(t) = -1 + 2 \frac{\sinh 2t - e^{-2\vartheta + 2t}}{\cosh 2t - e^{-2\vartheta + 2t}} \tag{4.26}$$

and

$$f''(t) = 4 \frac{1 - 2e^{-2\vartheta}}{\left(\cosh 2t - e^{-2\vartheta + 2t}\right)^2}. \tag{4.27}$$

Note that $\frac{1}{2}e^{\vartheta - t} \cosh 2t - \frac{1}{2}e^{-\vartheta + t} = \sinh(\vartheta - t) (\cosh t)^2 + \cosh(\vartheta - t) (\sinh t)^2 > 0$ for $t \in [0, \vartheta]$. Hence it follows that $f(t)$ is convex. Furthermore, for $0 \leq t \leq \vartheta$, $g(t)$ and $h(t)$ are non-negative bounded functions and $h(t) = O\left(t^2\right)$.

To evaluate the integral in equation (4.25), we divide it into two integrals, one over the interval $\left[0, \frac{1}{2}\right]$, and one over $\left[\frac{1}{2}, \vartheta\right]$, that is into the two integrals

$$\begin{aligned} &\int_0^{\frac{1}{2}} (\vartheta - t) \left(n \, g(t) + n(n-1) \, h(t) \right) e^{nf(t)} \, dt \\ &\quad = \int_0^{\frac{n}{2}} \left(\vartheta - \frac{z}{n} \right) \left(g\left(\frac{z}{n}\right) + (n-1) \, h\left(\frac{z}{n}\right) \right) e^{nf\left(\frac{z}{n}\right)} \, dz \end{aligned} \tag{4.28}$$

and

$$\int_{\frac{1}{2}}^{\vartheta} (\vartheta - t)\,(n\,g(t) + n(n-1)\,h(t))\,e^{nf(t)}\,dt$$
$$= \int_0^{(\vartheta-\frac{1}{2})n} z\left(\frac{1}{n}\,g\left(\vartheta - \frac{z}{n}\right) + \frac{n-1}{n}\,h\left(\vartheta - \frac{z}{n}\right)\right) e^{nf\left(\vartheta - \frac{z}{n}\right)}\,dz. \tag{4.29}$$

Now, using the convexity of $f(t)$ it is a standard calculation to show that the integrands of these expressions are uniformly dominated by
$C\left(1+t^2\right)e^{-\lambda t}$ and $Cte^{-\lambda t}$ respectively, for appropriate positive constants λ and C. Hence, by the Lebesgue dominated convergence theorem, these integrals converge to

$$\int_0^{\infty} 2\vartheta e^{\vartheta + f'(0)z}\,dz = \frac{2\vartheta e^{\vartheta}}{-f'(0)} = \sqrt{2}\,\vartheta e^{\vartheta} \tag{4.30}$$

and

$$\int_0^{\infty} z\,(\sinh 2\vartheta)^2\,e^{-f'(\vartheta)z}\,dz = \frac{8}{f'(\vartheta)^2} = \frac{2}{3-2\sqrt{2}} \tag{4.31}$$

respectively, as $n \to \infty$. We conclude that

$$A(\hat{\mathbf{1}},\vartheta) + B(\hat{\mathbf{1}},\vartheta) \to \frac{1}{2}\left(\sqrt{2}\,\vartheta e^{\vartheta} + \frac{2}{3-2\sqrt{2}}\right) \quad \text{as } n \to \infty. \tag{4.32}$$

■

5 Proof of Theorem 1.1

In order to bound $\|T_n - \vartheta\|_p$ it is natural to treat the problems of bounding $T_n - \vartheta$ from above and below separately. To this end, we let T_n^+ and T_n^- denote the positive and negative part of $T_n - \vartheta$ respectively, that is, T_n^+ is the maximum of $T_n - \vartheta$ and 0 and T_n^- is the maximum of $\vartheta - T_n$ and 0. Hence, we can bound $\|T_n - \vartheta\|_p$ by $\|T_n^+\|_p + \|T_n^-\|_p$. We will begin by proving two simple propositions. The first shows that the variance of T_n and the L^p-norm of $T_n - \vartheta$ for any $1 \le p < \infty$ are $\Omega\left(\frac{1}{n}\right)$. The second proposition uses the lower bound on T_n obtained by Durrett to prove that $\|T_n^-\|_p = O\left(\frac{1}{n}\right)$. The remaining part of the section will be dedicated to bounding $\|T_n^+\|_p$.

Proposition 5.1. *T_n has fluctuations of order at least $\frac{1}{n}$.*

Proof. We can write T_n in terms of Richardson's model as the time until the first neighbor of $\hat{\mathbf{0}}$ gets infected plus the time from this event until $\hat{\mathbf{1}}$ gets infected. It is easy to see that these are independent, and the former is exponentially distributed with mean $\frac{1}{n}$. ■

Proposition 5.2. *Let $1 \le p < \infty$ be fixed. Then $\|T_n^-\|_p = O\left(\frac{1}{n}\right)$.*

Proof. We have

$$\mathbb{E}\left[(T_n^-)^p\right] = \mathbb{E}\int_0^\infty \mathbb{1}_{t \le T_n^-}\, p\, t^{p-1}\, dt = \int_0^\infty p\, t^{p-1}\mathbb{P}\left(T_n \le \vartheta - t\right)\, dt. \tag{5.1}$$

To bound this, we use that $\mathbb{P}\left(T_n \le \vartheta - t\right) \le m(\hat{\mathbf{1}}, \vartheta - t)$ $= (\sinh(\vartheta - t))^n$ for any $t \le \vartheta$ and $\mathbb{P}\left(T_n \le \vartheta - t\right) = 0$ for $t > \vartheta$ (naturally T_n is always non-negative). It is straightforward to show that $\ln\sinh(\vartheta - t) \le -\sqrt{2}\, t$ for any $0 \le t \le \vartheta$. Using this, we conclude that

$$\mathbb{E}\left[(T_n^-)^p\right] \le \int_0^\vartheta p t^{p-1} e^{-\sqrt{2}\, nt}\, dt = O\left(\frac{1}{n^p}\right). \tag{5.2}$$

■

We now turn to the upper bound on T_n. Assume $n \ge 4$. Let $\{W_e\}_{e \in E(\mathbb{Q}_n)}$ be a collection of independent exponentially distributed random variables with expected value 1, denoting the passage times of the edges in $\mathbb{Q}_n$. For any vertex v adjacent to $\hat{\mathbf{0}}$ we will use W_v to denote the passage time of the edge between $\hat{\mathbf{0}}$ and v. Similarly, for any v adjacent to $\hat{\mathbf{1}}$, W_v denotes the passage time of the edge between v and $\hat{\mathbf{1}}$.

Condition on the weights of all edges connected to either $\hat{\mathbf{0}}$ or $\hat{\mathbf{1}}$. We pick vertices a_1 and a_2 adjacent to $\hat{\mathbf{0}}$ such that W_{a_1} and W_{a_2} have the smallest and second smallest edge weights respectively among all edges adjacent to $\hat{\mathbf{0}}$. Among all $n-2$ neighboring vertices of $\hat{\mathbf{1}}$ which are not antipodal to a_1 or a_2 we then pick b_1 and b_2 such that W_{b_1} and W_{b_2} have the smallest and second smallest values. Then W_{a_1}, $W_{a_2} - W_{a_1}$, W_{b_1} and $W_{b_2} - W_{b_1}$ are independent exponentially distributed random variables with respective expected values $\frac{1}{n}$, $\frac{1}{n-1}$, $\frac{1}{n-2}$ and $\frac{1}{n-3}$.

As a_1 and a_2 are adjacent to $\hat{\mathbf{0}}$ and b_1 and b_2 are adjacent to $\hat{\mathbf{1}}$, there is exactly one coordinate in each of a_1 and a_2 which is 1, and exactly one coordinate in b_1 and b_2 which is 0. Let the locations of these coordinates in a_1, a_2, b_1 and b_2 be denoted by i, j, k and l respectively. Note that the requirement on a_1, a_2, b_1 and b_2 not to be antipodal means that i, j, k and l are all distinct. We define H_1 as the induced subgraph of $\mathbb{Q}_n$ consisting of all vertices $v \in \mathbb{Q}_n$ such that the i:th coordinate is 1 and the k:th coordinate is 0. We similarly define H_2 as the induced subgraph of $\mathbb{Q}_n$ consisting of all vertices $v \in \mathbb{Q}_n$ such that the j:th coordinate is 1 and the l:th coordinate is 0. We furthermore define H_2' as the induced subgraph of $\mathbb{Q}_n$ whose vertex set is given by $H_2 \setminus H_1$. Note that H_1 and H_2' are vertex disjoint and hence also edge disjoint.

The idea to bound T_n is essentially to express it in terms of the minimum of the first-passage time from a_1 to b_1 in H_1 and the first-passage time from a_2 to b_2 in H_2', where the passage times for the edges are taken from $\{W_e\}_{e\in E(\mathbb{Q}_n)}$. As H_1 and H_2 are both isomorphic to $\mathbb{Q}_{n-2}$, where a_1 and b_1 are antipodal in H_1 and a_2 and b_2 are antipodal in H_2, Corollary 2.4 implies that the corresponding first-passage times in each of H_1 and H_2 are at most ϑ with probability bounded away from 0. However, for our proof it is not needed to make this connection. Rather, we will make use of the slightly stronger statement that the same holds true for H_2'. The following proposition is a consequence of Corollary 2.4. We postpone the proof of this to the end of the section.

Proposition 5.3. *There exists a constant $\varepsilon_2 > 0$ such that for all $n \geq 4$, with probability at least ε_2 the first-passage time in H_2' from a_2 to b_2 is at most ϑ.*

Now, let ξ denote the indicator function for the event that the first-passage time from a_2 to b_2 in H_2' is at most ϑ. As H_1 is isomorphic to $\mathbb{Q}_{n-2}$ it is clear that the first-passage time from a_1 to b_1 in H_1 is distributed as T_{n-2}, and so we may couple T_{n-2} to $\{W_e\}_{e\in E(\mathbb{Q}_n)}$ such that T_{n-2} denotes this quantity. Note that this means that ξ and T_{n-2} are independent random variables. With this coupling it is clear that $T_n \leq W_{a_1} + W_{b_1} + T_{n-2}$ as this is the passage time of the path that traverses the edge from $\hat{\mathbf{0}}$ to a_1, then follows the path to b_1 in H_1 with minimal passage time and lastly traverses the edge from b_1 to $\hat{\mathbf{1}}$. Furthermore, if $\xi = 1$ we similarly see that $T_n \leq W_{a_2} + W_{b_2} + \vartheta$. Combining these bounds we see that for any $n \geq 4$ we have

$$T_n \leq \xi\left(W_{a_2} + W_{b_2} + \vartheta\right) + (1-\xi)\left(W_{a_1} + W_{b_1} + T_{n-2}\right). \qquad (5.3)$$

We may interpret this inequality as follows. We flip a coin ξ. If the coin turns up heads then T_n is bounded by ϑ plus a small penalty. If the coin turns up heads, then we can bound T_n by a small penalty plus T_{n-2}, where T_{n-2} is independent of ξ. Assuming n is sufficiently large, we can then repeat this process on T_{n-2} and so on until one coin turns up heads. As each coin toss ends up heads with probability at least $\varepsilon_2 > 0$, this is likely to occur after $O(1)$ steps. Hence the total penalty before this occurs is likely to be small.

We now employ (5.3) to bound the L^p-norm of T_n^+. By subtracting ϑ and taking the positive part of both sides we get

$$T_n^+ \leq \xi\left(W_{a_2} + W_{b_2}\right) + (1-\xi)\left(W_{a_1} + W_{b_1} + T_{n-2}^+\right). \qquad (5.4)$$

As $W_{a_2} \geq W_{a_1}$ and $W_{b_2} \geq W_{b_1}$ we can replace $\xi\left(W_{a_2} + W_{b_2}\right) + (1-\xi)\left(W_{a_1} + W_{b_1}\right)$ in the right-hand side of (5.4) by $W_{a_2} + W_{b_2}$. Taking the L^p-norm of both sides we obtain the inequality

$$\|T_n^+\|_p \leq \|W_{a_2} + W_{b_2}\|_p + \|(1-\xi)T_{n-2}^+\|_p. \qquad (5.5)$$

For each fixed p, it is straightforward to show that $\|W_{a_2} + W_{b_2}\|_p = O\left(\frac{1}{n}\right)$. Furthermore, as ξ and T_{n-2}^+ are independent we have $\|(1-\xi)T_{n-2}^+\|_p = \|(1-\xi)\|_p \|T_{n-2}^+\|_p \leq (1-\varepsilon_2)^{\frac{1}{p}} \|T_{n-2}^+\|_p$. Hence, for any fixed p we have the inequality

$$\|T_n^+\|_p \leq O\left(\frac{1}{n}\right) + (1-\varepsilon_2)^{\frac{1}{p}} \|T_{n-2}^+\|_p. \qquad (5.6)$$

As $(1-\varepsilon_2)^{\frac{1}{p}} < 1$ it follows that we must have $\|T_n^+\|_p = O\left(\frac{1}{n}\right)$. Combining this with the corresponding bound on $\|T_n^-\|_p$ from Proposition 5.2, we have $\|T_n - \vartheta\|_p = O\left(\frac{1}{n}\right)$, as desired.

■

It only remains to prove Proposition 5.3.

In the following argument, we will identify H_2 with $\mathbb{Q}_{n-2}$ by simply disregarding the two coordinates of the vertices in H_2 which are fixed. Hence we will consider a_2 and b_2 to be the all zeroes and all ones vertices in $\mathbb{Q}_{n-2}$ respectively. When seen in this light, is clear that H_2' is the induced subgraph if H_2 consisting of all vertices where either the i':th coordinate is 1 or the k':th coordinate is 0 for some $i' \neq k'$.

It makes sense to think of H_2' as half a hypercube. For instance, exactly half of the oriented paths from a_2 to b_2 in H_2 are contained in H_2', namely those that move in direction i' before direction k'. Now, the paths from a_2 to b_2 in H_2 which are relevant for the early arrivals in the BTP are extremely unlikely to be oriented, but they are not too far from being oriented either. Our approach to showing Proposition 5.3 is essentially to show that H_2' is a sufficiently large subset of H_2 that when considering a BTP on H_2 originating at a_2, if there is an uncontested particle at b_2 at time ϑ, then with probability bounded away from 0, its ancestral line is contained in H_2'.

In order to show this, we need a property of the BTP which was hinted at briefly in [2]. Let $\mathbf{X}$ denote a BTP on $\mathbb{Q}_n$ originating at $\hat{\mathbf{0}}$. For any set of paths A in $\mathbb{Q}_n$, let $X_t(A)$ denote the expected number of particles in the BTP at time t whose ancestral line follows some path in A. Let $\{y(t)\}_{t\geq 0}$ denote a simple random walk on $\mathbb{Q}_n$ starting at $\hat{\mathbf{0}}$ with rate n, and for each $t \geq 0$ let σ_t denote the path that the random walk has followed up to time t.

Lemma 5.4. *Let S denote the set of paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ in $\mathbb{Q}_n$. For any $S' \subseteq S$ and for any $t \geq 0$ we have*

$$\frac{X_t(S')}{X_t(S)} = \mathbb{P}\left(\sigma_t \in S' \middle| y(t) = \hat{\mathbf{1}}\right). \tag{5.7}$$

Proof. Let σ be any fixed path from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ and let l denote the length of σ. By applying Lemma 3.2, we get

$$\begin{aligned} X_t(\{\sigma\}) &= \mathbb{E} \sum_{x \in V_\sigma(\mathbf{X})} \mathbb{1}_{T(\mathbf{X},x)\leq t} \\ &= \int_0^\infty \dots \int_0^\infty \mathbb{1}_{z_1+\dots+z_l \leq t}\, dz_1 \dots dz_l = \frac{t^l}{l!}, \end{aligned} \tag{5.8}$$

where $T(\mathbf{X}, x)$ denotes the birth time of x. In comparison, it is straightforward to see that $\mathbb{P}(\sigma_t = \sigma) = e^{-nt}\frac{t^l}{l!}$. It follows that, for any set of paths A, we have $X_t(A) = e^{nt}\mathbb{P}(\sigma_t \in A)$, and so in particular

$$\frac{X_t(S')}{X_t(S)} = \frac{\mathbb{P}(\sigma_t \in S')}{\mathbb{P}(\sigma_t \in S)} = \mathbb{P}\left(\sigma_t \in S' \middle| y(t) = \hat{\mathbf{1}}\right), \tag{5.9}$$

as desired. ■

Lemma 5.5. *Let* $\mathbf{X}$ *be a BTP on* $\mathbb{Q}_n$ *originating at* $\hat{\mathbf{0}}$*. Then with probability* $1 - o(1)$*, all particles at* $\hat{\mathbf{1}}$ *at time* ϑ *have ancestral lines of length* $\sqrt{2}\,\vartheta n \pm o(n)$*.*

Proof. We apply Lemma 5.4 with $t = \vartheta$. As $X_u(S) = m(\hat{\mathbf{1}}, \vartheta) = 1$ we see that it suffices to show that the number of steps performed by $\{y(t)\}_{t\geq 0}$ up to time ϑ, conditioned on the event that $y(\vartheta) = \hat{\mathbf{1}}$, is concentrated around $\sqrt{2}\,\vartheta n$.

In order to show this, we note that if $y(t) = (y_1(t), \ldots, y_n(t))$ is a simple random walk on $\mathbb{Q}_n$ with rate n, then each coordinate, $y_i(t)$, is an independent simple random walk on $\{0,1\}$ with rate one. Hence, conditioned on the event that $y(\vartheta) = \hat{\mathbf{1}}$, each coordinate $y_i(t)$ is an independent simple random walk on $\{0,1\}$ conditioned on the event that $y_i(\vartheta) = 1$. It is easy to see that the expected number of steps taken by such a process up to time ϑ is

$$\frac{e^{-\vartheta}}{e^{-\vartheta}} \frac{\vartheta + \frac{\vartheta^3}{2!} + \frac{\vartheta^5}{4!} + \ldots}{\vartheta + \frac{\vartheta^3}{3!} + \frac{\vartheta^5}{5!} + \ldots} = \vartheta \coth \vartheta = \sqrt{2}\,\vartheta. \tag{5.10}$$

The lemma follows by the law of large numbers. ■

Proof of Proposition 5.3. Consider the BTP:s $\mathbf{X}$ and $\mathbf{X}'$ on H_2 and H_2' respectively, both originating at a_2. We may couple these processes such that $\mathbf{X}'$ consists of all particles in $\mathbf{X}$ whose ancestral lines are contained in H_2'. Note that any particle in $\mathbf{X}'$ is uncontested in $\mathbf{X}'$ if it is uncontested in $\mathbf{X}$.

As H_2 is graph isomorphic to $\mathbb{Q}_{n-2}$, we know from Corollary 2.4 that, with probability bounded away from zero, there exists an uncontested particle in $\mathbf{X}$ at b_2 at time ϑ. Furthermore, by Lemma 5.5 we know that if such a particle exists, then with probability $1 - o(1)$ the length of its ancestral line is at most $1.25(n-2)$.

Let us now condition on the event that there exists an uncontested particle x in $\mathbf{X}$ at $\hat{\mathbf{1}}$ at time ϑ whose ancestral line is of length at most $1.25(n-2)$. As a path from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ must traverse edges in each of the $n-2$ directions of $\mathbb{Q}_{n-2}$ an odd number of times, this bound on the length of the ancestral line implies that there are at least $\frac{7}{8}(n-2)$ directions in which the path followed by the ancestral line of x only traverses one edge. By the symmetry of the hypercube, the distribution of this path must be invariant under permutation of coordinates. Hence, with probability $\approx \frac{49}{128}$, this path only traverses one edge in direction i' and one in direction k', and traverses the edge in direction i' before that in direction k'. Hence with probability bounded away from 0, this path is contained in H_2'.

We conclude that with probability bounded away from zero, there exists an uncontested particle at $\hat{\mathbf{1}}$ at time ϑ in $\mathbf{X}'$. The proposition follows from the fact that Richardson's model stochastically dominates the set of uncontested particles in a BTP. ■

6 Proof of Theorem 1.2

In the following proof we adopt the notation $X_t(A)$, $\{y(t)\}_{t\geq 0}$ and σ_t from the previous section. Hence, σ_n in the statement of Theorem 1.2 will here be denoted by σ_ϑ conditioned on $y(\vartheta) = \hat{\mathbf{1}}$. For any set of paths A in $\mathbb{Q}_n$ we let $Z_t(A)$ denote the expected number of simple paths in A starting at $\hat{\mathbf{0}}$ with passage time at most t. As Γ_n must be a simple path, it follows from the union bound that for any $c \in \mathbb{R}$ and any set A of paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ in $\mathbb{Q}_n$, we have

$$\mathbb{P}\left(\Gamma_n \in A\right) \leq Z_{\vartheta+\frac{c}{n}}(A) + \mathbb{P}\left(T_n \geq \vartheta + \frac{c}{n}\right). \tag{6.1}$$

In order to bound the right-hand side of this expression in terms of σ_ϑ, we first observe that for any $t \geq 0$ we have

$$\begin{aligned} Z_t(A) &= \sum_{\substack{\sigma\in A \\ \sigma \text{ simple}}} \int_0^\infty \cdots \int_0^\infty \mathbb{1}_{t_1+\cdots+t_{|\sigma|}\leq t} e^{-t_1-\cdots-t_{|\sigma|}}\, dt_1 \ldots dt_{|\sigma|} \\ &\leq \sum_{\sigma\in A} \int_0^\infty \cdots \int_0^\infty \mathbb{1}_{t_1+\cdots+t_{|\sigma|}\leq t}\, dt_1 \ldots dt_{|\sigma|} \\ &= \sum_{\sigma\in A} \frac{t^{|\sigma|}}{|\sigma|!} = X_t(A). \end{aligned}$$

Secondly, by the Cauchy-Schwarz inequality

$$\begin{aligned} X_{\vartheta+\frac{c}{n}}(A) &= \sum_{\sigma\in A} \frac{\vartheta^{|\sigma|}}{|\sigma|!} 1 \cdot \left(1+\frac{c}{\vartheta n}\right)^{|\sigma|} \\ &\leq \sqrt{\sum_{\sigma\in A} \frac{\vartheta^{|\sigma|}}{|\sigma|!}} \cdot \sqrt{\sum_{\sigma\in A} \frac{\vartheta^{|\sigma|}}{|\sigma|!} \left(1+\frac{c}{\vartheta n}\right)^{2|\sigma|}} \\ &= \sqrt{X_\vartheta(A)} \cdot \sqrt{X_{\vartheta\left(1+\frac{c}{\vartheta n}\right)^2}(A)} \\ &\leq \sqrt{X_\vartheta(A)} \cdot \sqrt{m\left(\hat{\mathbf{1}}, \vartheta\left(1+\frac{c}{\vartheta n}\right)^2\right)}. \end{aligned}$$

Note that $m\left(\hat{\mathbf{1}}, \vartheta\left(1+\frac{c}{\vartheta n}\right)^2\right)$ is bounded as $n \to \infty$. It follows from Lemma 5.4 that

$$\mathbb{P}\left(\Gamma_n \in A\right) \leq O\left(\sqrt{\mathbb{P}\left(\sigma_\vartheta \in A \middle| y(\vartheta) = \hat{\mathbf{1}}\right)}\right) + \mathbb{P}\left(T_n \geq \vartheta + \frac{c}{n}\right). \tag{6.2}$$

Now, consider any asymptotically almost sure property of σ_ϑ conditioned on $y(\vartheta) = \hat{\mathbf{1}}$. For each $n \geq 1$ let A_n denote the set of paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ in $\mathbb{Q}_n$ that do not have this property. Then, by taking lim sup of both sides in (6.2) we get

$$\limsup_{n\to\infty} \mathbb{P}\left(\Gamma_n \in A_n\right) \leq \limsup_{n\to\infty} \mathbb{P}\left(T_n \geq \vartheta + \frac{c}{n}\right). \tag{6.3}$$

The general case of Theorem 1.2 follows from Theorem 1.1 by letting $c \to \infty$. For the special case of the length of Γ_n, see the proof of Lemma 5.5. ■

Acknowledgements

I am very grateful to my supervisor Peter Hegarty for his constant support and for numerous insightful discussions and comments during this project. I would also like to thank Jeffrey Steif, Johan Tykesson and Johan Wästlund for valuable input on various aspects of this topic.

References

[1] Richard Durrett, *Lecture notes on particle systems and percolation*, The Wadsworth & Brooks/Cole Statistics/Probability Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1988. MR **940469 (89k:**60157)

[2] James Allen Fill and Robin Pemantle, *Percolation, first-passage percolation and covering times for Richardson's model on the n-cube*, Ann. Appl. Probab. **3** (1993), no. 2, 593–629. MR **1221168 (94h:**60145)

[3] Béla Bollobás and Yoshiharu Kohayakawa, *On Richardson's model on the hypercube*, Combinatorics, geometry and probability (Cambridge, 1993), Cambridge Univ. Press, Cambridge, 1997, pp. 129–137. MR **1476439 (98j:**05122)

[4] Peter Hegarty and Anders Martinsson, *On the existence of accessible paths in various models of fitness landscapes*, Ann. Appl. Probab. **24** (2014), no. 4, 1375–1395, DOI 10.1214/13-AAP949. MR 3210999

[5] David Aldous, *Probability approximations via the Poisson clumping heuristic*, Applied Mathematical Sciences, vol. 77, Springer-Verlag, New York, 1989. MR **969362 (90k:**60004)

[6] Rolf Schneider and Wolfgang Weil, *Stochastic and integral geometry*, Probability and its Applications (New York), Springer-Verlag, Berlin, 2008. MR **2455326 (2010g:**60002)

Paper III

Accessibility percolation and first-passage site percolation on the unoriented binary hypercube

Anders Martinsson

Abstract

Inspired by biological evolution, we consider the following so-called accessibility percolation problem: The vertices of the unoriented n-dimensional binary hypercube are assigned independent $U(0,1)$ weights, referred to as fitnesses. A path is considered accessible if fitnesses are strictly increasing along it. We prove that the probability that the global fitness maximum is accessible from the all zeroes vertex converges to $1 - \frac{1}{2}\ln\left(2+\sqrt{5}\right)$ as $n \to \infty$. Moreover, we prove that if one conditions on the location of the fitness maximum being $\hat{\boldsymbol{v}}$, then provided $\hat{\boldsymbol{v}}$ is not too close to the all zeroes vertex in Hamming distance, the probability that $\hat{\boldsymbol{v}}$ is accessible converges to a function of this distance divided by n as $n \to \infty$. This resolves a conjecture by Berestycki, Brunet and Shi in almost full generality.

As a second result we show that, for any graph, accessibility percolation can equivalently be formulated in terms of first-passage site percolation. This connection is of particular importance for the study of accessibility percolation on trees.

1 Introduction

A number of recent papers [4–10] have studied a percolation problem known as accessibility percolation, based on ideas of Kauffman and Levin for modeling biological evolution [1]. In its simplest form, accessibility percolation consists of a graph $G = (V, E)$, or more generally a digraph, together with a *fitness function* $\omega : V \to \mathbb{R}$ generated according to some random distribution. This is thought of as representing the landscape of possible evolutionary trajectories of a species. The vertices in G represent the possible genotypes for an organism whose fitness is a measure of how successful an individual of that genotype is, and the edges the possible ways the genome can change subject to a single mutation. Here it makes sense both to consider directed and undirected edges depending on whether or not a certain mutation is reversible. Of primary concern is the existence or distribution of so-called *accessible* paths.

Definition. Let $G = (V, E)$ and $\omega : V \to \mathbb{R}$ be a fitness landscape. We say that a path $v_0 \to v_1 \to \ldots \to v_l$ in G is *accessible* if

$$\omega(v_0) < \omega(v_1) < \ldots < \omega(v_l). \tag{1.1}$$

For $v, w \in V$ we say that w *is accessible from* v if there exists an accessible path from v to w.

For the distribution of ω we will in this paper consider two variations of Kingman's House-of-Cards model [3]. Both of which have previously been considered in accessibility percolation. In fact, all results in [6–10] consider some variation of the House-of-Cards model, whereas [4] and [5] also consider the so-called Rough Mount Fuji model. The first model we will consider here is the original formulation of the House-of-Cards model, in which the $\omega(v)$:s are independent and $U(0,1)$-distributed for all $v \in V$. Kauffman and Levin refers to this as an uncorrelated landscape. For the second distribution we modify the House-of-Cards model by introducing an a priori global fitness maximum $\hat{\boldsymbol{v}} \in V$ by changing $\omega(\hat{\boldsymbol{v}})$ to one. As accessibility percolation only depends on the relative order of fitnesses, this can be seen as equivalent to conditioning the House-of-Cards model on $\hat{\boldsymbol{v}}$ being the global fitness maximum. In particular, if $\hat{\boldsymbol{v}}$ is chosen uniformly at random among V, then this is equivalent to the House-of-Cards model with $\hat{\boldsymbol{v}}$ denoting the global fitness maximum.

Our first main result considers accessibility percolation on the unoriented n-dimensional binary hypercube. The question of primary concern is whether or not there exists an accessible path from the all zeroes vertex, $\hat{\mathbf{0}}$, to the fitness maximum $\hat{\boldsymbol{v}}$. We prove that, provided $\hat{\boldsymbol{v}}$ is not too close to $\hat{\mathbf{0}}$ in Hamming distance, the probability that such path exists converges to a non-trivial function of the Hamming distance between $\hat{\boldsymbol{v}}$ and $\hat{\mathbf{0}}$ divided by n, confirming a conjecture by Berestycki, Brunet and Shi [7] in almost full generality.

As a second result, we show that accessibility percolation for a general graph can be equivalently formulated in terms of first-passage site percolation. This lets us reformulate previous results in the literature in terms of first-passage site percolation. In particular, this relation has important implications for accessibility percolation on trees, as studied in [6, 8–10].

1.1 Notation

- Whenever talking about a general graph $G = (V, E)$, we allow both undirected and directed edges. For vertices $u, v \in V$, we write $u \sim v$ if there is either an undirected edge between u and v or a directed edge going from u to v.

- The unoriented n-dimensional binary hypercube, denoted by $\mathbb{Q}_n$, is the graph whose vertices are the binary n-tuplets $\{0, 1\}^n$ and where two vertices share an edge if their Hamming distance is one. The oriented n-dimensional binary hypercube, $\overrightarrow{\mathbb{Q}_n}$, is the directed graph obtained by directing each edge in $\mathbb{Q}_n$ towards the vertex with more ones.

- For a vertex v in the hypercube we let $|v|$ denote the number of coordinates of v that are one. Addition and subtraction of vertices in $\mathbb{Q}_n$ denotes coordinate-wise addition/subtraction modulo two. We let $\hat{\mathbf{0}}$ and $\hat{\mathbf{1}}$ denote the all zeroes and all ones vertices respectively, and let $e_1, \ldots, e_n$ denote the standard basis.

- Often when considering the House-of-Cards model, it is useful to condition on the fitness of $\hat{\mathbf{0}}$. Following the convention in [6, 7], for any $\alpha \in [0, 1]$ we let $\mathbb{P}^{\alpha}(\cdot)$ and $\mathbb{E}^{\alpha}[\cdot]$ denote conditional probability and expectation respectively, given $\omega(\hat{\mathbf{0}}) = \alpha$.

1.2 Recent work

Let us take a moment to summarize the results for accessibility percolation on the binary hypercube with House-of-Cards fitnesses in [5–7]. We start by consider the simplified version of the problem where we replace $\mathbb{Q}_n$ by $\overrightarrow{\mathbb{Q}_n}$. This is equivalent to only considering paths without backwards mutations. As any coordinate where $\hat{\boldsymbol{v}}$

is zero will be constantly zero along any such path, it suffices to consider the case where $\hat{\boldsymbol{v}} = \hat{\mathbf{1}}$.

Let X denote the number of oriented paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$ which are accessible. As there are $n!$ oriented paths from $\hat{\mathbf{0}}$ to $\hat{\mathbf{1}}$, and each path is accessible if and only if the n random fitnesses along the path are in ascending order, we see that $\mathbb{E}X = 1$. At first glance, this may seem to imply a positive probability of accessible paths existing. However, a much clearer picture of what occurs is obtained by conditioning on the fitness of the starting vertex. Indeed, conditioned on the fitness of $\hat{\mathbf{0}}$ being $\alpha \in [0,1]$, we have

$$\mathbb{E}^{\alpha} X = n(1-\alpha)^{n-1}. \tag{1.2}$$

We see that, for large n, this expression is 1 approximately at $\alpha = \frac{\ln n}{n}$, and rapidly decreasing as α increases. Informally, this means that unless the fitness of the starting vertex is below $\frac{\ln n}{n}$, accessible paths are highly unlikely. In fact, by considering (1.2) a bit more closely it follows that $\mathbb{P}\left(X \geq 1 \wedge \omega(\hat{\mathbf{0}}) > \frac{\ln n}{n}\right) \leq \frac{1}{n}$. On the other hand, the regime where α is smaller than $\frac{\ln n}{n}$ turns out to be more difficult to treat. In [5] it was shown by Hegarty and the author that the probability of accessible paths in this case tends to 1 as $n \to \infty$.

Theorem 1.1. *(Hegarty, Martinsson) For any sequence $\{\varepsilon_n\}_{n=1}^{\infty}$ such that $n\varepsilon_n \to \infty$, as $n \to \infty$ we have*

$$\mathbb{P}^{\frac{\ln n}{n}+\varepsilon_n}(X \geq 1) \to 0 \tag{1.3}$$

$$\mathbb{P}^{\frac{\ln n}{n}-\varepsilon_n}(X \geq 1) \to 1. \tag{1.4}$$

Furthermore,

$$\mathbb{P}(X \geq 1) \sim \frac{\ln n}{n}. \tag{1.5}$$

This theorem was later strengthened by Berestycki, Brunet and Shi in [6] who proved that, in the special case where $\omega(\hat{\mathbf{0}}) = O\left(\frac{1}{n}\right)$, X has a non-trivial limit distribution when scaled appropriately.

Let us now switch back to the unoriented hypercube and see how this analysis changes. Again, let X denote the number of accessible paths from $\hat{\mathbf{0}}$ to $\hat{\boldsymbol{v}}$. Here, paths are not as combinatorially well-behaved as for the oriented cube, and first moment estimates are not as easy to come by. Nevertheless, in a recent paper by Berestycki, Brunet and Shi [7] it was shown that $\mathbb{E}^{\alpha}X$ has the following asymptotic behavior:

Theorem 1.2. *(Berestycki, Brunet, Shi) Let $\alpha \in [0,1]$ be fixed, and let $\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}_n \in \mathbb{Q}_n$ be such that $x := \lim_{n\to\infty} |\hat{\boldsymbol{v}}_n| / n$ exists. We have that as $n \to \infty$*

$$(\mathbb{E}^\alpha X)^{1/n} \to \sinh(1-\alpha)^x \cosh(1-\alpha)^{1-x}. \tag{1.6}$$

As a consequence, for each x there is a critical value $\alpha^*(x) = 1 - \vartheta(x)$ for the fitness of $\hat{\mathbf{0}}$, given by the unique non-negative solution to

$$(\sinh \vartheta)^x (\cosh \vartheta)^{1-x} = 1, \tag{1.7}$$

such that

- *For $\alpha > 1 - \vartheta(x)$, $\mathbb{P}^\alpha (X \geq 1)$ goes to 0 exponentially fast as $n \to \infty$.*
- *For $\alpha < 1 - \vartheta(x)$, $\mathbb{E}^\alpha X$ diverges exponentially fast as $n \to \infty$.*

Hence, the unconditioned probability that $X \geq 1$ is at most $1 - \vartheta(x)$.

We see a similar behavior of $\mathbb{E}^\alpha X$ as for the oriented cube. One important difference though is that unlike the oriented cube the critical value has a nontrivial limit as $n \to \infty$. The function $\vartheta(x)$ is plotted in Figure 1. This function is continuous and increasing where $\vartheta(0) = 0$ and $\vartheta(1) = \ln\left(1 + \sqrt{2}\right) \approx 0.88$. In particular, it follows that if the a priori global fitness maximum is $\hat{\mathbf{1}}$, then the critical fitness is $1 - \ln\left(1 + \sqrt{2}\right) \approx 0.12$, and if chosen uniformly at random then $|\hat{\boldsymbol{v}}| / n$ will be tightly concentrated around $\frac{1}{2}$ and hence the critical fitness is $1 - \frac{1}{2} \ln\left(2 + \sqrt{5}\right) \approx 0.28$.

Berestycki et al. further gave two conjectures that (1.6) "tells the truth" in the sense that $\mathbb{P}^\alpha (X \geq 1)$ tends to 1 as $n \to \infty$ for $\alpha < 1 - \vartheta(x)$. Conjecture 1 of their paper proposes this in the special case where $\hat{\boldsymbol{v}} = \hat{\mathbf{1}}$, and Conjecture 2 in the more general setting of $\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}_n$ satisfying $|\hat{\boldsymbol{v}}_n| / n \to x \in [0,1]$.

1.3 Results

The first result of this paper fully resolves Conjecture 1 by Berestycki, Brunet and Shi [7], and Conjecture 2 under the additional condition that x is not too small.

Theorem 1.3. *Let $\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}_n \in \mathbb{Q}_n$ be a sequence of vertices such that $x := \lim_{n\to\infty} |\hat{\boldsymbol{v}}| / n$ exists. Let X denote the number of accessible*

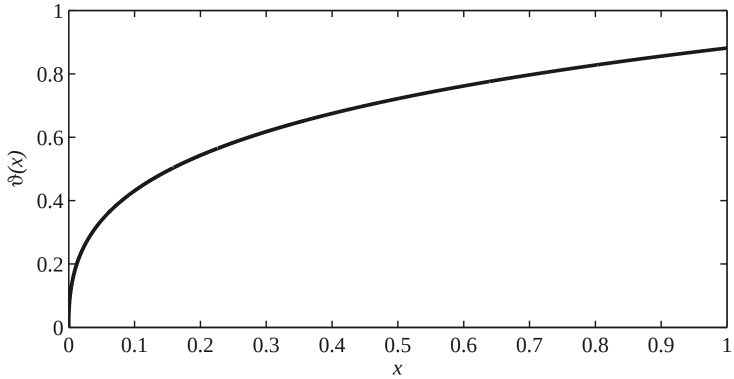


Figure 1: The function $\vartheta(x)$ as defined in (1.7).

paths from $\hat{\mathbf{0}}$ *to* $\hat{\boldsymbol{v}}$*. Let* $\vartheta(x)$ *be as defined in Theorem 1.2. Assuming* $x \geq 0.002$*, we have*

$$\lim_{n\to\infty} \mathbb{P}^{\alpha}\left(X \geq 1\right) = \begin{cases} 0 & \text{if } \alpha > 1 - \vartheta(x) \\ 1 & \text{if } \alpha < 1 - \vartheta(x). \end{cases} \tag{1.8}$$

In particular, if $\hat{\boldsymbol{v}} = \hat{\mathbf{1}}$*, then*

$$\mathbb{P}\left(X \geq 1\right) \to 1 - \ln\left(1 + \sqrt{2}\right) \text{ as } n \to \infty \tag{1.9}$$

and if $\hat{\boldsymbol{v}}$ *is chosen uniformly at random, then*

$$\mathbb{P}\left(X \geq 1\right) \to 1 - \frac{1}{2}\ln\left(2 + \sqrt{5}\right) \text{ as } n \to \infty. \tag{1.10}$$

The value 0.002 deserves some explanation. In the proof of Theorem 1.3, or more accurately the proof of Theorem 1.6 below which is shown to be equivalent to the former, we see that there is a value $x^* \approx 0.00167$ such that the proof goes through whenever $x > x^*$ and breaks down when $x < x^*$, see Remark 4.8. It seems likely however that this is simply an artifact of the technique used in the proof, and that the statement should hold true even for smaller x. Regardless of whether or not this is true, we can note that the two cases of most concern, $x = 1$ and $x = 0.5$, are far above x^*.

We now turn to the relation between accessibility percolation and first-passage site percolation for a general graph. Let $G =$

(V, E) be a graph with a distinguished vertex $\hat{\mathbf{0}}$. Note that each edge of G may either be directed or undirected. For each vertex $v \in G$ randomly assign a cost, denoted by $c(v)$, according to independent $U(0,1)$ random variables. For a path $u_0, u_1, \ldots, u_l$ in G we define the site passage time of the path by

$$\sum_{1 \le i \le l} c(u_i), \tag{1.11}$$

and similarly define its reduced site passage time by

$$\sum_{1 \le i < l} c(u_i). \tag{1.12}$$

Note that neither the passage time nor the reduced passage time of a path include the cost of the first vertex. For each $u, v \in G$ we define the site first-passage time from u to v, denoted by $\mathcal{T}_V(u, v)$, and the reduced first passage time from u to v, denoted by $\mathcal{T}'_V(u, v)$, as the minimum of the respective quantity over all paths from u to v.

Theorem 1.4. *Let G be a graph with two distinct vertices $\hat{\mathbf{0}}$ and $\hat{\boldsymbol{v}}$, and let $\alpha \in [0,1]$. Consider accessibility percolation on G. If fitnesses are assigned according to the House-of-Cards model with $\hat{\boldsymbol{v}}$ as the a priori global fitness maximum, then*

$$\mathbb{P}^\alpha \left(\hat{\boldsymbol{v}} \textit{ accessible from } \hat{\mathbf{0}} \right) = \mathbb{P}\left(\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}) \le 1 - \alpha \right). \tag{1.13}$$

If fitnesses are assigned according to the House-of-Cards model without an a priori global maximum, then for any vertex $v \in G$

$$\mathbb{P}^\alpha \left(v \textit{ accessible from } \hat{\mathbf{0}} \right) = \mathbb{P}\left(\mathcal{T}_V(\hat{\mathbf{0}}, v) \le 1 - \alpha \right). \tag{1.14}$$

Moreover, in the latter case this claim can be significantly strengthened. Conditioned on the fitness of $\hat{\mathbf{0}}$ being α, the set of vertices accessible from $\hat{\mathbf{0}}$ has the same distribution as the set of vertices v such that $\mathcal{T}_V(\hat{\mathbf{0}}, v) \le 1 - \alpha$.

Informally we can think of this theorem as saying that accessibility percolation is equivalent to first-passage site percolation with independent $U(0,1)$ vertex passage times. We need to be a bit careful there though; the theorem only deals with the question of whether or not a certain vertex is accessible from $\hat{\mathbf{0}}$ along any path,

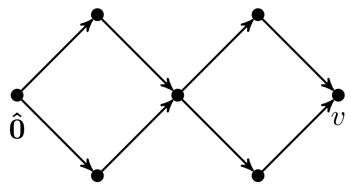


Figure 2: Example of a graph where accessible paths have a different distribution than paths with small passage time. We can for instance note that there can never be exactly three accessible paths from $\hat{\mathbf{0}}$ to v, whereas there can certainly be exactly three paths with reduced passage time at most $1-\alpha$.

and it does not for instance say anything about the number of accessible paths. Indeed, it is not true in general that the number of accessible paths from $\hat{\mathbf{0}}$ to v is distributed as the number of paths from $\hat{\mathbf{0}}$ to v with reduced passage time at most $1-\alpha$. For graphs containing non-simple paths this is clear as non-simple paths can have arbitrarily small passage time but cannot be accessible, but it can even be false for directed acyclic graphs, see for instance Figure 2. On the other hand, the connection is more general than just treating which vertices are accessible. For instance, using the proof ideas in Section 2 one can show that the minimal number of times you need to move to a less fit vertex to get from $\hat{\mathbf{0}}$ to v is distributed as the integer part of $\mathcal{T}_V(\hat{\mathbf{0}}, v) + \alpha$.

A problem with using Theorem 1.4 to relate known results from first-passage percolation to accessibility percolation is that the vast majority of the first-passage percolation literature assigns passage times to edges rather than vertices. However, a common property for percolation problems is that it is harder to percolate on vertices than edges [2]. The following proposition shows that something similar holds for first-passage percolation.

Proposition 1.5. *Suppose the edges of G are assigned independent $U(0,1)$ weights. Let $\mathcal{T}_E(u,v)$ denote the minimum total weight of any path from u to v in G. Then, it is possible to couple $\mathcal{T}_E(\hat{\mathbf{0}},\cdot)$ to $\mathcal{T}_V(\hat{\mathbf{0}},\cdot)$ such that $\mathcal{T}_E(\hat{\mathbf{0}},v) \le \mathcal{T}_V(\hat{\mathbf{0}},v)$ for all $v \in G$.*

For the special case when G is a rooted tree one can see that this coupling is exact; to go from site to bond percolation we can simply consider the passage time of each vertex to instead be assigned to the edge leading to it. Accessibility percolation on trees has been considered in [6, 8–10]. With the exception of [6], these articles have considered regular rooted trees with degree n and height h, and where fitnesses are assigned according to the House-of-Cards model conditioned on the fitness of the root being zero. Of principal concern is how the number of vertices in generation h that are accessible from the root varies as a function of n, and in particular whether this number is non-zero. Using Theorem 1.4 we can see that this is equivalent to assigning independent $U(0,1)$ passage times to the edges of the tree and considering the number of vertices v in generation h such that $\mathcal{T}_E(\hat{\mathbf{0}}, v) \leq 1$. In particular, the question of whether generation h is accessible from $\hat{\mathbf{0}}$ is equivalent to to asking if the first-passage time from the root to generation h is at most 1. It should be mentioned however that the usual setting in first-passage percolation on regular rooted trees keeps n fixed and considers the first-passage time from the root to generation h as $h \to \infty$. While the author is not aware of any results from this field that have appropriate error bounds to be directly applicable to accessibility percolation, there seems to be a significant overlap of ideas between [9, 10] and the literature on first-passage percolation on trees. See for instance [13].

Let us now consider the implications of Theorem 1.4 for the hypercube. Using this result, we can immediately translate the result from Theorem 1.1 to that, for the oriented hypercube, $\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\mathbf{1}})$ is concentrated around $1 - \frac{\ln n}{n}$ with fluctuations of order $\frac{1}{n}$. More importantly, we have that the following is equivalent to Theorem 1.3:

Theorem 1.6. *Let $G = \mathbb{Q}_n$ and let $\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}_n \in \mathbb{Q}_n$ be a sequence of vertices such that $x := \lim_{n\to\infty} |\hat{\boldsymbol{v}}|/n$ exists. Assuming $x \geq 0.002$, as $n \to \infty$ we have*

$$\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}) \to \vartheta(x) \tag{1.15}$$

in probability.

Note here that the fact that $\vartheta(x)$ is an asymptotic lower bound on the reduced passage time is already implied by Theorem 1.2.

It should be mentioned that basically the same results holds true for bond percolation. In [11] it was shown that for the oriented hypercube, we have $\mathcal{T}_E(\hat{\mathbf{0}}, \hat{\mathbf{1}}) \to 1$ in probability as $n \to \infty$. In a more recent result by the author [12], it was shown that for the unoriented hypercube $\mathcal{T}_E(\hat{\mathbf{0}}, \hat{\mathbf{1}}) \to \ln\left(1 + \sqrt{2}\right)$ as $n \to \infty$. Strictly speaking these results assume standard exponential edge weights, but it is not too hard to show that the limiting distribution of $\mathcal{T}_E(\hat{\mathbf{0}}, \hat{\mathbf{1}})$ only depends on the weight distribution as the righthand limit of its probability distribution function at 0, hence it will be the same for $U(0, 1)$ weights.

The remainder of the paper will be structured as follows: In Section 2 we prove Proposition 1.5 and Theorem 1.4. The remaining sections, Sections 3, 4 and 5, are dedicated to the proof of Theorem 1.6.

2 Proof of Proposition 1.5 and Theorem 1.4

We may, without loss of generality, assume that for any vertex v there exists a path from $\hat{\mathbf{0}}$ to v.

A key idea of the proofs of Proposition 1.5 and Theorem 1.4 is the following procedure for computing $\mathcal{T}_V(\hat{\mathbf{0}}, v)$. We initially consider $\mathcal{T}_V(\hat{\mathbf{0}}, v)$ to be unassigned for each v, except $\hat{\mathbf{0}}$ for which it is set to 0, and we let $U = \{\hat{\mathbf{0}}\}$ denote the set of vertices with assigned first-passage times. Until $\mathcal{T}_V(\hat{\mathbf{0}}, v)$ is assigned for all v, we do the following operation:

1. Find a pair of vertices u, v that minimizes $\mathcal{T}_V(\hat{\mathbf{0}}, u)$ subject to $(u, v) \in E$, $u \in U$ and $v \notin U$.
2. Let $\mathcal{T}_V(\hat{\mathbf{0}}, v) := \mathcal{T}_V(\hat{\mathbf{0}}, u) + c(v)$
3. Add v to U.

To see that this assigns first-passage times correctly, suppose that we are in the step where $\mathcal{T}_V(\hat{\mathbf{0}}, v)$ is assigned. As v is not in U, the passage time of any path from $\hat{\mathbf{0}}$ to v must include the passage time from $\hat{\mathbf{0}}$ to some vertex u' in U adjacent to some vertex outside U, as well as the cost v. Hence $\mathcal{T}_V(\hat{\mathbf{0}}, v) \geq \mathcal{T}_V(\hat{\mathbf{0}}, u') + c(v) \geq \mathcal{T}_V(\hat{\mathbf{0}}, u) + c(v)$. As there is a path from $\hat{\mathbf{0}}$ to v with passage time $\mathcal{T}_V(\hat{\mathbf{0}}, u) + c(v)$, this must be optimal. Hence, if all previous assignments are correct, $\mathcal{T}_V(\hat{\mathbf{0}}, v)$ will be assigned correctly as well.

Proof of Proposition 1.5. We can modify this algorithm to run on first-passage bond percolation by replacing $c(v)$ by the weight of the edge from u to v. In either case, as no vertex cost or edge weight respectively is accessed more than once, the accessed values form a sequence of independent and $U(0,1)$ random variables. Hence the distribution of $\mathcal{T}_V(\hat{\mathbf{0}}, v)$ is unaffected. On the other hand, for bond percolation we get that $\mathcal{T}_V(\hat{\mathbf{0}}, v)$ is the edge passage time of some path from $\hat{\mathbf{0}}$ to v (but not necessarily the shortest). ■

We now turn to the proof of Theorem 1.4. The coupling between first-passage site percolation and accessibility percolation we will consider is essentially to let $f(v) = \left\{\alpha + \mathcal{T}_V(\hat{\mathbf{0}}, v)\right\}$ be the fitness function, where $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part of x. We will however modify this slightly by putting $f(v) = 1$ whenever $\alpha + \mathcal{T}_V(\hat{\mathbf{0}}, v) = 1$. It is clear that the probability of such v other than $\hat{\mathbf{0}}$ existing is 0, so the only way this will change the distribution of f is that $f(\hat{\mathbf{0}}) = 1$ if $\alpha = 1$.

It is not too hard to see that, for any vertex v except $\hat{\mathbf{0}}$, $f(v)$ is $U(0,1)$-distributed. The following lemma shows that the $f(v)$:s are also independent, hence showing that f is distributed according to the House-of-Cards model without an a priori global fitness maximum, conditioned on $f(\hat{\mathbf{0}}) = \alpha$.

Lemma 2.1. *$f(v)$ are independent $U(0,1)$ random variables for $v \in V \setminus \{\hat{\mathbf{0}}\}$.*

Proof. Suppose that we generate vertex costs in the following way: Run the procedure above, but with the modification that whenever the algorithm tries to access $c(v)$, first generate a $U(0,1)$ random variable $\tilde{f}(v)$ and assign $c(v)$ the value $\left\{\tilde{f}(v) - \alpha - \mathcal{T}_V(\hat{\mathbf{0}}, u)\right\}$.

It is clear that the $c(v)$:s are independent and $U(0,1)$-distributed. The lemma follows by noting that, in the latter case, we have $f(v) = \tilde{f}(v)$ almost surely for all $v \in V \setminus \{\hat{\mathbf{0}}\}$. ■

Proof of Theorem 1.4. We begin by considering the case with no a priori global fitness maximum. In this case, we can consider $f : V \to \mathbb{R}$ to be the fitness function. For simplicity let us assume that no vertex cost is exactly 0.

Assume $\mathcal{T}_V(\hat{\mathbf{0}}, v) \leq 1 - \alpha$, and let $\hat{\mathbf{0}} = v_0, v_1, \ldots, v_l = v$ be the path with shortest passage time. Then, as $0 < \alpha + \mathcal{T}_V(\hat{\mathbf{0}}, v_i) \leq 1$

for $1 \leq i \leq l$ it follows that

$$f(v_i) = \alpha + \sum_{j=1}^{i} c(v_j) \tag{2.1}$$

for $0 \leq i \leq l$. Hence $v_0, v_1, \ldots, v_l$ is accessible. Conversely, suppose $\mathcal{T}_V(\hat{\mathbf{0}}, v) > 1 - \alpha$ and let $\hat{\mathbf{0}} = v_0, v_1, \ldots, v_l = v$ be any path between $\hat{\mathbf{0}}$ and v. Let i be the lowest index such that $\mathcal{T}_V(\hat{\mathbf{0}}, v_i) > 1 - \alpha$. Then $f(v_i) \leq \alpha + \mathcal{T}_V(\hat{\mathbf{0}}, v_i) - 1 \leq \alpha + \mathcal{T}_V(\hat{\mathbf{0}}, v_{i-1}) + c(v_i) - 1 < \alpha + \mathcal{T}_V(\hat{\mathbf{0}}, v_{i-1}) = f(v_{i-1})$. Hence the path is not accessible.

Now for the case where $\hat{\boldsymbol{v}} \in V \setminus \{\hat{\mathbf{0}}\}$ is the a priori global fitness maximum. We here keep the same coupling as before between $f(v)$ and $c(v)$ for $v \in V$, except that we fix $f(\hat{\boldsymbol{v}}) = 1$. Let U be the set of vertices $v \in V$ such that $(v, \hat{\boldsymbol{v}}) \in E$. Then $\hat{\boldsymbol{v}}$ being accessible from $\hat{\mathbf{0}}$ is almost surely equivalent to some vertex in U being accessible from $\hat{\mathbf{0}}$. Note that this last statement does not depend on the value of $f(\hat{\boldsymbol{v}})$. It follows that $\hat{\boldsymbol{v}}$ is accessible from $\hat{\mathbf{0}}$ is almost surely equivalent to that $\min_{v \in U} \mathcal{T}_V(\hat{\mathbf{0}}, v) \leq 1 - \alpha$. The theorem follows by noting that $\min_{v \in U} \mathcal{T}_V(\hat{\mathbf{0}}, v) = \mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}})$. ■

3 The Clustering Translation Process

Before proceeding, we will slightly modify $\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}})$ by replacing the $U(0,1)$ vertex costs by independent standard exponential such. Note that the standard exponential distribution stochastically dominates $U(0,1)$, and hence this modification will only increase $\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}})$. As the lower bound in Theorem 1.6 follows from Theorem 1.2, it suffices to show that, with this modification, asymptotically almost surely $\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}) \leq \vartheta(x) + o(1)$. To do this, we will mimic the argument in [12] for first-passage bond percolation on $\mathbb{Q}_n$.

Let us take a moment to describe some of the underlying machinery for first-passage bond percolation on $\mathbb{Q}_n$. We assume independent standard exponential edge weights. In [11], Durrett introduced the following process, which he called the the *branching translation process*, BTP: At time 0 we place one particle at $\hat{\mathbf{0}}$ in $\mathbb{Q}_n$. The system then evolves by each existing particle independently generating offspring at each vertex adjacent to its position at rate 1. One can show that for each vertex $v \in \mathbb{Q}_n$, the time at which the first particle at v is born is stochastically dominated by

$\mathcal{T}_E(\hat{\mathbf{0}}, v)$. This follows from the fact that the BTP dominates the so-called Richardson's model. The strategy in [12] is basically to show that, with a certain coupling, there is a probability bounded away from zero of these quantities being equal.

In order to translate this approach to first-passage site percolation, we need to find a corresponding process to the BTP for this case. We claim that the following is such a process: We initially have a finite number of particles, each located at a vertex in $\mathbb{Q}_n$. For each particle, we assign an independent Poisson clock with unit rate. When a particle's clock goes off, it simultaneously generates one new offspring at each vertex adjacent to its position. The new particles are then assigned new Poisson clocks and the process continues. We will refer to this process as the clustering translation process, CTP.

We see that in both the BTP and CTP each particle generates offspring at each neighboring vertex at rate 1. A big difference however is that in the BTP this is done independently for each neighboring vertex, whereas in the CTP a particle generates offspring all neighboring vertices simultaneously. Another difference is that the initial state of the CTP is not fixed.

The most important initial state of the CTP will be one particle at each neighbor of $\hat{\mathbf{0}}$. We will refer to a CTP initialized in this way as a *standard* CTP. Particles born due to the same Poisson clock tick will be referred to as *identical n-tuplets*. To simplify terminology we will also consider the initial n particles in a standard CTP as identical n-tuplets. Below we will use the terms *ancestor* and *descendant* of a particle to denote the natural partial order of particles generated by the CTP. For convenience, we say that a particle is both an ancestor and a descendant of itself. The terms *parent* and *child* are defined in the natural way. The *ancestral line* of a particle x is the ordered set of ancestors of x, and we say that the ancestral line of x *follows the path* $\hat{\mathbf{0}} = v_0, v_1, v_2, \ldots, v_l$ if the location of the ancestors of x in chronological order is given by $v_1, v_2, \ldots v_l$. Note that this path always starts at $\hat{\mathbf{0}}$ even though the first ancestor is located at a neighbor of $\hat{\mathbf{0}}$. We say that a particle *x originates from a particle y at a time t* if y is the last particle in the ancestral line of x that exists at time t.

We can immediately note some properties of this process. Firstly, it is Markovian. Secondly, let A be a set of vertices in $\mathbb{Q}_n$, and let

$M_A(v,t)$ denote the expected number of particles at vertex v at time $t \geq 0$ in the CTP initialized by placing one particle at each vertex in A. Then it is easy to see that $M_A(v,t)$ must solve the initial value problem

$$\frac{d}{dt} M_A(v,t) = \sum_{w \sim v} M_A(w,t) \text{ for } t > 0 \tag{3.1}$$

$$M_A(v,0) = \mathbb{1}_A(v). \tag{3.2}$$

In particular, if $A = \{\hat{\mathbf{0}}\}$, then the unique solution to this problem is

$$m(v,t) := (\sinh t)^{|v|} (\cosh t)^{n-|v|}, \tag{3.3}$$

and it follows by linearity that for any A, we have

$$M_A(v,t) = \sum_{w \in A} m(v-w,t). \tag{3.4}$$

Recall that addition/subtraction of vertices in $\mathbb{Q}_n$ are interpreted as coordinate-wise addition/subtraction modulo 2. It should be remarked that the exact same analysis holds for the BTP.

We now show that the standard CTP indeed has the desired relation to first-passage site percolation. To this end, we partition the particles in this process into two sets, the set of *alive* particles and the set of *ghosts*. Each initial particle is alive. Whenever a new particle is born, it is alive if its location does not already contain an alive particle and its parent is alive, and is a ghost otherwise. Note that at most one particle at each vertex can be alive. Furthermore, it is easy to show that each vertex will almost surely eventually contain an alive particle.

Proposition 3.1. *Consider first-passage site percolation on $\mathbb{Q}_n$ with exponentially distributed costs with unit mean. It is possible to couple this process to the standard CTP such that for each vertex v except $\hat{\mathbf{0}}$, $\mathcal{T}'_V(\hat{\mathbf{0}}, v)$ denotes the birth time of the alive particle at v.*

Proof. For each vertex v, we let $\tilde{\mathcal{T}}'(v)$ denote the first time $t \geq 0$ when v contains an alive particle, and we let $\tilde{c}(v)$ denote the time from the birth of this particle to the first arrival of its clock. Then $\tilde{c}(v)$ for $v \in \mathbb{Q}_n$ are independent exponentially distributed random variables with unit mean.

From the definitions of the CTP and alive particles, it follows that for any vertex v that is not a neighbor of $\hat{\mathbf{0}}$, the alive particle at v is born at the first arrival time of an alive particle at an adjacent vertex. Hence, for any v that is not a neighbor of $\hat{\mathbf{0}}$, we have

$$\tilde{\mathcal{T}}'(v) = \min_{w \sim v} \left(\tilde{\mathcal{T}}'(w) + \tilde{c}(w) \right), \tag{3.5}$$

and trivially $\tilde{\mathcal{T}}'(v) = 0$ when v is a neighbor of $\hat{\mathbf{0}}$. It is easy to see that this uniquely defines $\tilde{\mathcal{T}}'(v)$, and that for each vertex v except $\hat{\mathbf{0}}$, $\tilde{\mathcal{T}}'(v)$ denotes the reduced first-passage time from $\hat{\mathbf{0}}$ to v with respect to the vertex costs given by $\tilde{c}(v)$. ■

Given this proposition, we are able to proceed analogously to Sections 2 and 3 in [12]. In applying this coupling between the CTP and first-passage site percolation we will consider a stronger and more tractable property than aliveness. For any particle x in the CTP, we let $c(x)$ denote the number of pairs of particles y and z such that

- y and z occupy the same vertex
- y is an ancestor of x
- y was born after z.

We furthermore let $a(x)$ denote the number of such pairs where z is either an ancestor of x or an identical n-tuple of an ancestor of x, and define $b(x) = c(x) - a(x)$. We call a particle x uncontested if $c(x) = 0$.

It can be noted that $a(x)$ is defined differently for the BTP. This is because the strategy is loosely speaking to let $a(x)$ denote the number of pairs (y, z) that deterministically must exist given x. For the CTP we have additional such pairs, namely those corresponding to identical n-tuplets of ancestors of x.

Lemma 3.2. *If a particle is uncontested, then it is alive.*

Proof. If a particle x is a ghost, then it must have an earliest ancestor (possibly itself) which is a ghost, y. As y is a ghost but the parent of y is alive, it follows that the location of y must have already been occupied by some (alive) particle z. The pair (y, z) is then counted in $c(x)$. ■

It is not hard to see that $a(x)$ only depends on the path followed by the ancestral line of x. If we know this path, then we know the locations and order of births of all ancestors and identical n-tuplets of ancestors of x. Let σ be a path represented as a vertex sequence. We say that σ is vertex-minimal if there is no proper subsequence which is a path with the same end points.

Lemma 3.3. *Let x be a particle in the CTP. If the ancestral line of x is vertex-minimal, then $a(x) = 0$. The converse is true unless x is located at $\hat{\mathbf{0}}$.*

Proof. Denote the path followed by the ancestral line of x by $v_0, v_1, \ldots, v_l$ and the ancestors of x by $x_1, x_2, \ldots, x_l = x$. We have that $a(x) > 0$ if and only if there exist $1 \leq i < j \leq l$ such that x_j occupies the same vertex as either x_i or an identical n-tuplet of x_i, that is, v_{i-1} and v_j are adjacent. Hence, if $a(x) > 0$ the path is not vertex-minimal. Conversely, if $a(x) = 0$ it follows that the only pairs of adjacent vertices are consecutive in the path. It is straightforward to show that, unless the path starts and stops at the same vertex, this implies vertex-minimality. ■

What follows are two technical lemmas, corresponding to Lemmas 3.2 and 3.3 in [12]. Before presenting these, we need to specify how to formally describe the CTP. Firstly, by a (potential) particle we mean a word $\{v_1, z_1, v_2, z_2, \ldots, v_{l-1}, z_{l-1}, v_l\}$ where $v_1, \ldots, v_l$ denote vertices and $z_1, \ldots z_{l-1}$ positive real numbers. This is interpreted as the particle whose ancestors are located at $v_1, v_2, \ldots, v_l$ and born at times $0, z_1, z_1 + z_2$ and so on. The CTP is described by a random set $\mathbf{X}$ of potential particles, denoting the set of particles that will ever be born in the CTP. We will use $\oplus$ to denote concatenation of words. We remark that this representation means that the functions $c(x)$ and $b(x)$ are not functions only of x, and should more correctly be denoted by $c(\mathbf{X}, x)$ and $b(\mathbf{X}, x)$. On the other hand, $a(x)$ is really a function of x as it only depends on the location of the ancestors of x.

Lemma 3.4. *Let $\sigma = \{\hat{\mathbf{0}} = v_0, v_1, \ldots, v_{l-1}, v_l\}$ be a path. For $0 \leq i \leq l-1$ let $\mathbf{X}_i$ denote independent CTP:s where $\mathbf{X}_i$ is the CTP obtained by initially placing one particle at each neighbor of v_i. Let f be a function that maps pairs (X, x) to the non-negative real numbers where X is a realization of a CTP, and x is a particle* *in X. Similarly, let $V_\sigma(X)$ denote the set of particles in X whose ancestral lines follow σ. Then for a standard CTP, $\mathbf{X}$, we have*

$$\begin{aligned}&\mathbb{E}\sum_{x\in V_\sigma(\mathbf{X})} f(\mathbf{X},x)\\ &=\int_0^\infty\cdots\int_0^\infty \mathbb{E}f\left(\mathbf{X}^{z_1,\dots z_{l-1}},x^{z_1,\dots,z_{l-1}}\right)dz_1\dots dz_{l-1},\end{aligned}\tag{3.6}$$

where

$$\begin{aligned}&\mathbf{X}^{z_1,\dots z_{l-1}}\\ &=\mathbf{X}_0\cup(\{v_1z_1\}\oplus\mathbf{X}_1)\cup\dots\cup(\{v_1z_1v_2z_2\dots v_{l-1}z_{l-1}\}\oplus\mathbf{X}_{l-1})\end{aligned}\tag{3.7}$$

and $x^{z_1,\dots,z_{l-1}}=\{v_1z_1v_2z_2\dots v_l\}$*.*

For compactness, we will only sketch a proof. The reader unconvinced by this is referred to the proof of Lemma 3.2 in [12].

Proof sketch. Let us first consider the case when $f(X,x)$ only depends on x. In that case, we have

$$\mathbb{E}\sum_{x\in V_\sigma(\mathbf{X})} f(x)=\int_0^\infty\cdots\int_0^\infty f\left(x^{z_1,\dots,z_{l-1}}\right)dz_1\dots dz_{l-1}.\tag{3.8}$$

This is because the original particle at v_1 gives birth to particles at v_2 at rate one whereupon, after its birth, each child at v_2 of this original particle gives birth to particles at v_3 at rate one, and so on. When f also depends on the realization of the CTP, the idea is that we substitute $f(\mathbf{X},x)$ in the left-hand side of this sum by $\mathbb{E}\left[f(\mathbf{X},x)|x\in\mathbf{X}\right]$. Now, formally this conditioning does not really make sense, but its meaning is intuitively clear; it denotes the average value of $f(\mathbf{X},x)$ where the average is taken over all $\mathbf{X}$ that include x. We have

$$\begin{aligned}\mathbb{E}\sum_{x\in V_\sigma(\mathbf{X})} f(\mathbf{X},x)&=\mathbb{E}\sum_{x\in V_\sigma(\mathbf{X})}\mathbb{E}\left[f(\mathbf{X},x)|x\in\mathbf{X}\right]\\ &=\int_0^\infty\cdots\int_0^\infty \mathbb{E}\left[f(\mathbf{X},x^{z_1,\dots,z_{l-1}})|x^{z_1,\dots,z_{l-1}}\in\mathbf{X}\right]dz_1\dots dz_{l-1}.\end{aligned}$$

Now, $x^{z_1,\dots,z_{l-1}}$ exists in $\mathbf{X}$ if and only if certain Poisson clocks have arrivals at certain times. By the independent increment property, conditioning on these arrivals does not affect the Poisson clocks at any other times. Hence, the conditional distribution of $\mathbf{X}$ given the existence of $x^{z_1,\dots,z_{l-1}}$ is the same as that of a standard CTP, except with added arrivals, corresponding to the births of the ancestors of $x^{z_1,\dots,z_{l-1}}$. This is precisely the distribution of $\mathbf{X}^{z_1,\dots z_{l-1}}$. ■

Lemma 3.5. *Let $\mathbf{X}$ be a CTP, and let ϕ be an indicator function on the set of potential particles in $\mathbf{X}$. If $\phi(x) = 0$ for all original particles in the CTP, then*

$$\mathbb{P}\left(\sum_{x\in\mathbf{X}} \phi(x) = 0\right) \geq \exp\left(-\mathbb{E}\sum_{x\in\mathbf{X}} \phi(x)\right). \tag{3.9}$$

Proof. Let us refer to the set of original particles as generation one, their children as generation two and so on. Let $\mathbf{T}$ denote the set of birth times for particles in generation two in $\mathbf{X}$, and let $\mathbf{T}' \subseteq \mathbf{T}$ be the subset obtained by including $t \in \mathbf{T}$ if there exists a particle $x \in \mathbf{X}$ such that $\phi(x) = 1$ and x is an descendant of a particle in generation two born at time t. It is clear that $|\mathbf{T}'| \leq \sum_{x\in\mathbf{X}} \phi(x)$ and that $|\mathbf{T}'| = 0$ if and only if $\sum_{x\in\mathbf{X}} \phi(x) = 0$.

By definition of the CTP, it is clear that $\mathbf{T}$ is a Poisson point process. Furthermore, as the event that $t \in \mathbf{T}$ is included in $\mathbf{T}'$ only depends on descendants of particles in generation two born at time t, this occurs independently for each $t \in \mathbf{T}$. Hence, by the random selection property, $\mathbf{T}'$ is also a Poisson point process. This implies that

$$\begin{aligned}\mathbb{P}\left(\sum_{x\in\mathbf{X}} \phi(x) = 0\right) &= \mathbb{P}\left(\mathbf{T}' = \emptyset\right) = \exp\left(-\mathbb{E}\left|\mathbf{T}'\right|\right) \\ &\geq \exp\left(-\mathbb{E}\sum_{x\in\mathbf{X}} \phi(x)\right),\end{aligned} \tag{3.10}$$

as desired. ■

Theorem 3.6. *Consider a standard CTP. For any vertex v and any $t \geq 0$, let $B(v,t) = \mathbb{E}\sum_x b(x)$ where the sum goes over all particles at v at time t in the CTP, and let $S(v,t)$ denote the expected number of particles x at v at time t such that $a(x) = 0$. The probability that there is an uncontested particle at v at time t is at least $S(v,t)\exp\left(-\frac{B(v,t)}{S(v,t)}\right)$.*

Proof. Let $P(v,t)$ denote the probability that v contains an uncontested particle at time t. As at most one particle at each vertex can be uncontested, this is the same thing as the expected number of uncontested particles at v at time t. For each path σ from $\hat{\mathbf{0}}$ to v, let $P_\sigma(v,t)$, $B_\sigma(v,t)$ and $S_\sigma(v,t)$ denote the contribution to $P(v,t)$, $B(v,t)$ and $S(v,t)$ respectively from particles whose ancestral line follows σ.

The idea now is to bound $P_\sigma(v,t)$ in terms of $B_\sigma(v,t)$ and $S_\sigma(v,t)$ for each path σ from $\hat{\mathbf{0}}$ to v. Recall that $a(x)$ is constant over all particles x whose ancestral line follows a fixed σ. We will denote this constant by $a(\sigma)$.

Let σ be a path from $\hat{\mathbf{0}}$ to v such that $a(\sigma) = 0$. Applying Lemma 3.4 we see that

$$\begin{aligned} S_\sigma(v,t) &= \mathbb{E} \sum_{x \in V_\sigma(\mathbf{X})} \mathbb{1}_{T(x) \le t} \\ &= \int_0^\infty \dots \int_0^\infty \mathbb{1}_{z_1 + \dots + z_{l-1} \le t} \, dz_1 \dots dz_{l-1}, \end{aligned} \tag{3.11}$$

where $T(x)$ denotes the time of birth of x. Similarly, for any $x \in V_\sigma(\mathbf{X})$ we have

$$\begin{aligned} B_\sigma(v,t) = \mathbb{E} \sum_{x \in V_\sigma(\mathbf{X})} \mathbb{1}_{T(x) \le t} \, b(\mathbf{X}, x) = \int_0^\infty \dots \int_0^\infty \\ \mathbb{1}_{z_1 + \dots + z_{l-1} \le t} \, \mathbb{E} b\left(\mathbf{X}^{z_1,\dots,z_{l-1}}, x^{z_1,\dots,z_{l-1}}\right) dz_1 \dots dz_{l-1} \end{aligned} \tag{3.12}$$

and

$$\begin{aligned} P_\sigma(v,t) = \mathbb{E} \sum_{x \in V_\sigma(\mathbf{X})} \mathbb{1}_{T(x) \le t} \mathbb{1}_{b(\mathbf{X},x)=0} = \int_0^\infty \dots \int_0^\infty \\ \mathbb{1}_{z_1 + \dots + z_{l-1} \le t} \, \mathbb{P}\left(b\left(\mathbf{X}^{z_1,\dots,z_{l-1}}, x^{z_1,\dots,z_{l-1}}\right) = 0\right) dz_1 \dots dz_{l-1}. \end{aligned} \tag{3.13}$$

As $a(x^{z_1,\dots,z_{l-1}}) = 0$, no two ancestors of $x^{z_1,\dots,z_{l-1}}$ occupy the same vertex. It follows that any pair of particles y and z which is counted in $b\left(\mathbf{X}^{z_1,\dots,z_{l-1}}, x^{z_1,\dots,z_{l-1}}\right)$ is uniquely determined by z. Fixing σ and $z_1, \dots, z_{l-1}$, this means that we can define $\phi(x)$ as an indicator function such that

$$b\left(\mathbf{X}^{z_1,\dots,z_{l-1}}, x^{z_1,\dots,z_{l-1}}\right) = \sum_{x \in \mathbf{X}^{z_1,\dots,z_{l-1}}} \phi(x). \tag{3.14}$$

More precisely, $\phi(x)$ is the indicator for x occupying the same vertex as an ancestor of $x^{z_1,\dots,z_{l-1}}$ and being born before it. By the definition of $b(x^{z_1,\dots,z_{l-1}})$, we have that $\phi(x)$ is zero for any ancestor or identical n-tuplet of an ancestor of $x^{z_1,\dots,z_{l-1}}$. It follows by Lemma 3.5 that

$$\mathbb{P}\left(b\left(\mathbf{X}^{z_1,\dots,z_{l-1}}, x^{z_1,\dots,z_{l-1}}\right) = 0\right) \geq \exp\left(-\mathbb{E}b\left(\mathbf{X}^{z_1,\dots,z_{l-1}}, x^{z_1,\dots,z_{l-1}}\right)\right). \tag{3.15}$$

By convexity of the exponential function we have $e^{-r} \geq (1 + r_0 - r)e^{-r_0}$ for any $r, r_0 \in \mathbb{R}$. Hence

$$P_\sigma(v,t) \geq (1+r_0)e^{-r_0} S_\sigma(v,t) - e^{-r_0} B_\sigma(v,t), \tag{3.16}$$

for any path σ from $\hat{\mathbf{0}}$ to v such that $a(\sigma) = 0$. For any σ that satisfies $a(\sigma) \neq 0$ it is clear that $P_\sigma(v,t) = S_\sigma(v,t) = 0$ and $B_\sigma(v,t) \geq 0$, hence (3.16) holds in this case as well. Summing over all paths σ from $\hat{\mathbf{0}}$ to v and optimizing over r_0 yields $P(v,t) \geq S(v,t)e^{-\frac{B(v,t)}{S(v,t)}}$, as desired. ■

We will apply Theorem 3.6 as follows: Let $\{\hat{\boldsymbol{v}}_n\}_{n=1}^{\infty}$ be a sequence of vertices such that, for each n, $\hat{\boldsymbol{v}}_n \in \mathbb{Q}_n$ and $x = \lim_{n\to\infty} |\hat{\boldsymbol{v}}_n|/n$ exists and is non-zero. We may, without loss of generality, assume that $\hat{\boldsymbol{v}}_n$ is never equal to $\hat{\mathbf{0}}$. For each n, we let ϑ_n denote the unique non-negative solution to

$$m(\hat{\boldsymbol{v}}_n, \vartheta_n) = \frac{1}{n}. \tag{3.17}$$

Note that the expected number of particles at $\hat{\boldsymbol{v}}_n$ at time ϑ_n in a standard CTP on $\mathbb{Q}_n$ is $\Theta(1)$, and that $\vartheta_n \to \vartheta(x)$ as $n \to \infty$. By Theorem 3.6 we have that the probability that there is a uncontested particle at $\hat{\boldsymbol{v}}_n$ at time ϑ_n in the CTP on $\mathbb{Q}_n$ is at least $S(\hat{\boldsymbol{v}}_n, \vartheta_n)\exp\left(-\frac{B(\hat{\boldsymbol{v}}_n,\vartheta_n)}{S(\hat{\boldsymbol{v}}_n,\vartheta_n)}\right)$. Hence by Lemma 3.2 and Proposition 3.1 it follows that

$$\mathbb{P}\left(\mathcal{T}_V'(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}_n) \leq \vartheta_n\right) \geq S(\hat{\boldsymbol{v}}_n, \vartheta_n) \exp\left(-\frac{B(\hat{\boldsymbol{v}}_n, \vartheta_n)}{S(\hat{\boldsymbol{v}}_n, \vartheta_n)}\right). \tag{3.18}$$

This means that if we can show that $S(\hat{\boldsymbol{v}}_n, \vartheta_n) = \Theta(1)$ and $B(\hat{\boldsymbol{v}}_n, \vartheta_n) = O(1)$, then we know that $\mathcal{T}_V'(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}_n) \leq \vartheta_n$ with probability bounded away from 0 as $n \to \infty$.

Section 4 will be dedicated to estimating $S(\hat{\boldsymbol{v}}_n, \vartheta_n)$ and $B(\hat{\boldsymbol{v}}_n, \vartheta_n)$. The proof of Theorem 1.6 is then completed in Section 5 by showing that if $\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}_n) \leq \vartheta_n$ with probability bounded away from 0, then a slightly larger upper bound on $\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}_n)$ must hold asymptotically almost surely.

4 Calculus

4.1 Estimating S

We will prove that $S(\hat{\boldsymbol{v}}_n, \vartheta_n) = \Theta(1)$ in two steps. Firstly, we show that most particles at $\hat{\boldsymbol{v}}_n$ at time ϑ_n have ancestral lines which are close to vertex-minimal. Using this, we then give a combinatorial argument that shows that a positive proportion of these particles must have vertex-minimal ancestral lines.

Let us formalize the notion of paths being close to vertex-minimal. Let $v, w \in \mathbb{Q}_n$ be fixed distinct vertices and let $\sigma = \{v = v_0, v_1, \ldots, v_l = w\}$ be a path from v to w. Throughout this section, we will always think of a path as a finite sequence of vertices. In particular, by the length of a path we mean the number of vertices in the path. For any $0 < i \leq j < l$ we say that the subsequence $v_i, v_{i+1}, \ldots, v_j$ is a *detour* of σ if removing these elements from σ results in a valid path. Clearly, for $v \neq w$ a path is vertex-minimal if and only if it has no detours. Inspired by this, we say that a path is *almost vertex-minimal* if all detours have length at most 2. Note that as $\mathbb{Q}_n$ is bipartite, any detour must have even length. Hence, a path is almost vertex-minimal if it only has the shortest possible detours.

An important property of almost vertex-minimal paths is that any such path from v to w can be constructed by taking a vertex-minimal path with the same end-points and extending it as follows: Between each two adjacent elements in the sequence either do nothing or insert a detour of length 2.

Lemma 4.1. *Let $s, t \geq 0$ and $v \in \mathbb{Q}_n$. Then*

$$\sum_{w \in \mathbb{Q}_n} m(w, s) m(v + w, t) = m(v, s + t). \tag{4.1}$$

Proof. Fix s. Observe that equality holds when $t = 0$ and that both expressions solves (3.1) ∎

Proposition 4.2. *Let $\{\hat{\boldsymbol{v}}_n\}_{n=1}^{\infty}$ be a sequence of vertices, $\hat{\boldsymbol{v}}_n \in \mathbb{Q}_n$, such that $\alpha = \lim_{n\to\infty} |\hat{\boldsymbol{v}}_n| / n$ exists and is positive. Then, as $n \to \infty$, the expected number of particles in the standard CTP on $\mathbb{Q}_n$ which are at $\hat{\boldsymbol{v}}_n$ at time ϑ_n, but that do not have almost vertex-minimal ancestral lines tends to* 0.

Proof. Let X_n denote the number of triples of particles x, y, z in the CTP on $\mathbb{Q}_n$ such that

- x is at $\hat{\boldsymbol{v}}_n$ at time ϑ_n
- y and z are located at adjacent vertices
- z is an ancestor of y which is an ancestor of x.
- y and z are neither one nor three generations apart.

We note that if the ancestral line of a given particle x at $\hat{\boldsymbol{v}}_n$ at time ϑ_n can be constructed using some detour of length $d > 2$, then it is clear that x would have a pair of ancestors at adjacent vertices which are $d + 1$ generations apart. This means that any such x is counted at least once in X_n. Hence, it suffices to show that $\mathbb{E}X_n = o(1)$.

For each triple x, y, z as above there are uniquely defined particles c, the particle after z in the ancestral line of x, and p, the parent of y. Note that the requirement that y is neither the child, nor the grand-grandchild of z implies that p is a descendant of c, but not a child of c.

Let $T = \{0 = t_0 < t_1 < \cdots < t_k = \vartheta_n\}$ denote the end-points of a partition of $[0, \vartheta_n)$ into left-closed right-open subintervals, and let $X_{n,T}$ denote the number of triples as above where c and y are the only ancestors of x born during their respective time intervals. Pick a, b integers between 0 and $k - 1$. Consider the number of triples counted in $X_{n,T}$ where c is born during $[t_a, t_{a+1})$ and y is born during $[t_b, t_{b+1})$. Note that this is trivially 0 whenever $b \leq a$.

Let us count the expected number of corresponding triples for $a < b$. As z and y are located at adjacent vertices, for each such triple we may denote the locations of z, y, c and p by v, $v+e_i$, $v+e_j$ and $v + e_i - e_k$ respectively for some $v \in \mathbb{Q}_n$ and $1 \leq i, j, k \leq n$. A particle is a potential z if it is born before time t_a, hence there are on average $\sum_{l=1}^{n} m(v - e_l, t_a)$ potential z:s at v. For each z, a particle is a potential c if it is a child of z born during $[t_a, t_{a+1})$. Hence for each potential z at v, there are on average $t_{a+1} - t_a$ potential c:s at $v + e_j$. For each potential c, a particle is a potential p if it originates from c at time t_{a+1} and is born before t_b, but is not a child of c. Hence for each potential c at $v+e_j$ there are on average $m(e_i - e_k - e_j, t_b - t_{a+1})$ potential p:s at $v + e_i - e_k$ if $v + e_j$ and $v+e_i-e_k$ are not adjacent, and $m(e_i-e_k-e_j, t_b-t_{a+1})-(t_b-t_{a+1})$ if they are. Lastly, for each potential p, a particle is a potential y if it is a child of p born during $[t_b, t_{b+1})$, and for each potential y a particle is a potential x if it is located at $\hat{\boldsymbol{v}}_n$, originates from y at time t_{b+1}, and is born before time ϑ_n. Hence for each potential p at $v + e_i - e_k$ the expected number of potential y:s at $v + e_i$ is $t_{b+1} - t_b$, and for each potential y at $v + e_i$, the expected number of x:s is $m(\hat{\boldsymbol{v}}_n - v - e_i, \vartheta_n - t_{b+1})$. Combining all of these, we see that

$$
\begin{aligned}
\mathbb{E}X_{n,T} = \sum_{a<b} \sum_{v \in \mathbb{Q}_n} \sum_{i,j,k,l} & m(v - e_l, t_a)(t_{a+1} - t_a) \cdot \\
& \cdot \Big(m(e_i - e_k - e_j, t_b - t_{a+1}) - \mathbb{1}_{|e_i+e_j+e_k|=1}(t_b - t_{a+1}) \Big) \cdot \\
& \cdot (t_{b+1} - t_b) m(\hat{\boldsymbol{v}}_n - v - e_i, \vartheta_n - t_{b+1}).
\end{aligned}
\tag{4.2}
$$

where the sums over i, j, k and l all go from 1 to n. Letting $T_1, T_2, \dots$ be a sequence of increasingly finer partitions of $[0, \vartheta_n]$ such that the length of the longest interval in T_k tends to 0 as $k \to \infty$, it follows by monotone convergence that we have $\mathbb{E}X_n = \lim_{k\to\infty} \mathbb{E}X_{n,T}$. Combining this with equation (4.2), and recognizing the right-hand side as a Riemann sum, we get

$$
\begin{aligned}
\mathbb{E}X_n = \int_0^{\vartheta_n} \int_a^{\vartheta_n} \sum_{v \in \mathbb{Q}_n} \sum_{i,j,k,l} & m(v - e_l, a) \Big(m(e_i - e_k - e_j, b - a) \\
& - \mathbb{1}_{|e_i+e_j+e_k|=1}(b - a) \Big) m(\hat{\boldsymbol{v}}_n - v - e_i, \vartheta_n - b) \, da \, db.
\end{aligned}
\tag{4.3}
$$

Lemma 4.1 implies that we may replace the factor $\sum_{v\in\mathbb{Q}_n} m(v - e_l, a) m(\hat{\boldsymbol{v}}_n - v - e_i, \vartheta_n - b)$ in the integrand of equation (4.3) by $m(\hat{\boldsymbol{v}}_n + e_i + e_l, \vartheta_n - b + a)$. Hence, by the substitution $t = b - a$, the right-hand side of (4.3) simplifies to

$$\int_0^{\vartheta_n} (\vartheta_n - t) \sum_{i,j,k,l} m(\hat{\boldsymbol{v}}_n + e_i + e_l, \vartheta_n - t) \cdot \cdot \left(m(e_i + e_j + e_k, t) - \mathbb{1}_{|e_i+e_j+e_k|=1} t\right) dt. \tag{4.4}$$

Using the fact that $\sinh t \le \cosh t$ for all $t \in \mathbb{R}$, we have

$$\sum_{i,j,k,l} m(\hat{\boldsymbol{v}}_n + e_i + e_l, \vartheta_n - t) \left(m(e_i + e_j + e_k, t) - \mathbb{1}_{|e_i+e_j+e_k|=1} t\right)$$
$$\le n\, (\sinh(\vartheta_n - t))^{|\hat{v}_n|-2} (\cosh(\vartheta_n - t))^{n-|\hat{v}_n|+2} \cdot$$
$$\cdot \sum_{i,j,k} \left(m(e_i + e_j + e_k, t) - \mathbb{1}_{|e_i+e_j+e_k|=1} t\right).$$

It is straight-forward (but messy) to show that

$$\sum_{i,j,k} \left(m(e_i + e_j + e_k, t) - \mathbb{1}_{|e_i+e_j+e_k|=1} t\right) = (\cosh t)^n \, O\left(n^3 t^3\right).$$

As $\cosh(\vartheta_n - t)\cosh t \le \cosh \vartheta_n$ it follows that

$$\mathbb{E}X_n \le \int_0^{\vartheta_n} n\, (\sinh(\vartheta_n - t)\cosh t)^{|\hat{v}_n|-2} (\cosh \vartheta_n)^{n-|\hat{v}_n|+2} \, O\left(n^3 t^3\right) dt. \tag{4.5}$$

Recall that by the definition of ϑ_n we have

$$(\sinh \vartheta_n)^{|\hat{v}_n|} (\cosh \vartheta_n)^{n-|\hat{v}_n|} = \frac{1}{n}. \tag{4.6}$$

Define the function $f(t) = \ln \sinh(\vartheta_n - t) + \ln \cosh t$. Note that $f'(t) = -\coth(\vartheta_n - t) + \tanh t$, and $f''(t) = -\operatorname{csch}^2(\vartheta_n - t) + \operatorname{sech}^2 t$. As $0 \le \operatorname{sech} t \le 1$ and $\operatorname{csch} t \ge 1$ for all $0 < t \le \vartheta_n$ it follows that f is concave, and thus for any $0 \le t \le \vartheta_n$ we have $f(t) \le f(0) - t \coth \vartheta_n \le f(0) - t$. Hence

$$(\sinh(\vartheta_n - t)\cosh t)^{|\hat{v}_n|-2} \le (\sinh \vartheta_n)^{|\hat{v}_n|-2} e^{-(|\hat{v}_n|-2)t}. \tag{4.7}$$

Plugging this into equation (4.5), we get

$$\mathbb{E}X_n \le \int_0^{\vartheta_n} e^{-(|\hat{v}_n|-2)t} \, O\left(n^3 t^3\right) dt. \tag{4.8}$$

As $|\hat{\boldsymbol{v}}_n| \sim x \cdot n$ this implies that $\mathbb{E}X_n = O\left(\frac{1}{n}\right)$, as desired. ■

Proposition 4.3. *For any pair of sequences* $\{\hat{\boldsymbol{v}}_n\}_{n=1}^{\infty}$ *and* $\{\vartheta_n\}_{n=1}^{\infty}$ *as above, we have* $S(\hat{\boldsymbol{v}}_n, \vartheta_n) = \Theta(1)$.

Proof. Let Γ_n and $\tilde{\Gamma}_n$ denote the sets of vertex-minimal and almost vertex-minimal paths from $\hat{\mathbf{0}}$ to $\hat{\boldsymbol{v}}_n$ respectively. Using Lemma 3.4 with $f(X, x)$ as the indicator function of x being born at time ϑ_n and having ancestral line in $\tilde{\Gamma}_n$ and Γ_n respectively, we can write the expected number particles at $\hat{\boldsymbol{v}}_n$ at time ϑ_n in the CTP whose ancestral lines are almost vertex-minimal as

$$\sum_{\sigma \in \tilde{\Gamma}_n} \frac{\vartheta_n^{|\sigma|-2}}{(|\sigma|-2)!} \tag{4.9}$$

and the expected number that are vertex-minimal as

$$\sum_{\sigma \in \Gamma_n} \frac{\vartheta_n^{|\sigma|-2}}{(|\sigma|-2)!}. \tag{4.10}$$

As the total expected number of particles at $\hat{\boldsymbol{v}}_n$ at time ϑ_n in the CTP is $\Theta(1)$, Proposition 4.2 implies that the sum in (4.9) is also $\Theta(1)$.

The idea now is to group the terms of the sum in (4.9) according to which vertex-minimal path σ it is an extension of, that is we write

$$\sum_{\sigma \in \tilde{\Gamma}_n} \frac{\vartheta_n^{|\sigma|-2}}{(|\sigma|-2)!} \leq \sum_{\sigma \in \Gamma_n} \sum_{\substack{\tilde{\sigma} \in \tilde{\Gamma}_n \\ \tilde{\sigma} \supseteq \sigma}} \frac{\vartheta_n^{|\tilde{\sigma}|-2}}{(|\tilde{\sigma}|-2)!}. \tag{4.11}$$

Here $\tilde{\sigma} \supseteq \sigma$ denotes that $\tilde{\sigma}$ is an extension of σ. Note that the inequality comes from the fact that $\tilde{\sigma}$ may be an extension of more than one vertex-minimal path.

Let us fix a vertex-minimal path $\sigma \in \Gamma_n$ consisting of l vertices. It is straight-forward to show that the number of possible detours of length 2 that can be inserted between each adjacent pair of elements in σ is $3(n-1)$. Hence, there are at most $3^k(n-1)^k\binom{l-1}{k}$ ways to extend σ to an almost vertex-minimal path of length $l + 2k$. This means that

$$\sum_{\substack{\tilde\sigma\in\tilde\Gamma_n\\ \tilde\sigma\supseteq\sigma}} \frac{\vartheta_n^{|\tilde\sigma|-2}}{(|\tilde\sigma|-2)!} \le \sum_{k=0}^{l-1} 3^k (n-1)^k \binom{l-1}{k} \frac{\vartheta_n^{l-2+2k}}{(l-2+2k)!}$$

$$\le \frac{\vartheta_n^{l-2}}{(l-2)!} \sum_{k=0}^{l-1} 3^k (n-1)^k \binom{l-1}{k} \frac{\vartheta_n^{2k}}{(l-1)^{2k}}$$

$$= \frac{\vartheta_n^{l-2}}{(l-2)!} \left(1 + \frac{3\vartheta_n^2(n-1)}{(l-1)^2}\right)^{l-1}$$

$$\le \frac{\vartheta_n^{l-2}}{(l-2)!} \exp\left(\frac{3\vartheta_n^2(n-1)}{l-1}\right).$$

As any path from $\hat{\mathbf{0}}$ to $\hat{\boldsymbol{v}}_n$ must have length at least $|\hat{\boldsymbol{v}}_n|+1$, we conclude that

$$\sum_{\sigma\in\Gamma_n} \frac{\vartheta_n^{|\sigma|-2}}{(|\sigma|-2)!} \ge \exp\left(-3\vartheta_n^2 \frac{n-1}{|\hat{\boldsymbol{v}}_n|}\right) \sum_{\sigma\in\tilde\Gamma_n} \frac{\vartheta_n^{|\sigma|-2}}{(|\sigma|-2)!} = \Theta(1). \tag{4.12}$$

■

4.2 Estimating B

Proposition 4.4. *For any $\hat{\boldsymbol{v}} \in \mathbb{Q}_n$ and any $u > 0$ we have*

$$\begin{aligned}
B(\hat{\boldsymbol{v}}, u) \le &\int_0^u \sum_{\Delta\in\mathbb{Q}_n} \sum_{i,j,k} m(\Delta - e_k - e_i, t)\cdot \\
&\quad \cdot m(\Delta - e_j, t)\, m(\hat{\boldsymbol{v}} - \Delta, u - t)\, dt \\
&+ \int_0^u (u-t) \sum_{\Delta\in\mathbb{Q}_n} \sum_{i,j,k} m(\Delta - e_k - e_j, t)\cdot \\
&\quad \cdot m(\Delta, t)\, m(\hat{\boldsymbol{v}} - \Delta - e_i, u - t)\, dt \\
&+ \int_0^u (u-t) \sum_{\Delta\in\mathbb{Q}_n} \sum_{i,j,k} m(\Delta - e_k, t)\cdot \\
&\quad \cdot m(\Delta - e_j, t)\, m(\hat{\boldsymbol{v}} - \Delta - e_i, u - t)\, dt \\
&+ \int_0^t (u-t) \sum_{\Delta\in\mathbb{Q}_n} \sum_{i,j,k,l} m(\Delta - e_l - e_j, t)\cdot \\
&\quad \cdot m(\Delta - e_k, t)\, m(\hat{\boldsymbol{v}} - \Delta - e_i, u - t)\, dt,
\end{aligned} \tag{4.13}$$

where the sums over i, j, k and l go from 1 *to* n.

Proof. We observe that $B(\hat{\boldsymbol{v}}, u)$ is bounded by the expected number of triplets of particles x, y, z in the CTP such that

- x is at $\hat{\boldsymbol{v}}$ at time u
- y is an ancestor of x
- y and z occupy the same vertex
- z was born before y.

Note the similarity to the quantity X_n in Proposition 4.2. For the sake of compactness, we will be less rigorous here, and refer to the proof of that proposition to see how to formalize this argument.

Let us start by considering the number of such triples x, y and z where z has no ancestors in common with x and y, that is, for some $i \neq j$ we have that x and y originate from the original particle at e_i whereas z originates from the original particle at e_j. Denote the common location of y and z by v, and pick k such that the parent of y is located at $v - e_k$. Note that as z is strictly older than y, y cannot be an original particle and hence has a parent. The lineage of x, y, z is illustrated in Graph 1 of Figure 3.

Let us count the expected number of such triples corresponding to a fixed v and where y is born during the time interval $[t, t + dt)$. The potential z:s corresponding to a fixed j are simply the descendants of the original particle at e_j that are at v at time t. Hence the expected number of such particles is $m(v - e_j, t)$. Similarly, for a fixed i the expected number of potential y:s is given by $m(v - e_k - e_i, t)\, dt$, and for each potential y the expected number of potential x:s is $m(v, u - t)$. As the potential z:s are born independently of the pairs of potential x:s and y:s, we see that the expected number of triples x, y, z that do not have common ancestors, corresponding to a fixed vertex v and a fixed time interval $[t, t + dt)$ is given by

$$\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \sum_{k=1}^{n} m(v - e_k - e_i, t)\, m(v - e_j, t)\, m(\hat{\boldsymbol{v}} - v, u - t)\, dt. \quad (4.14)$$

The total expected number of triples x, y, z without common ancestors is hence given by summing this expression over all vertices

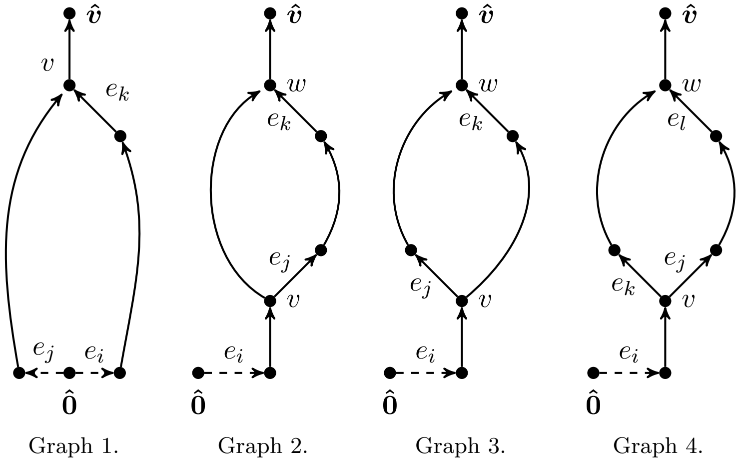


Figure 3: Illustration of the possible ways x, y and z can be related. The left-most arrows describe the ancestors of z and the right-most the ancestors of x and y. Graph 1 shows the case when z has no ancestor in common with x and y. Here v is the common location of y and z, and $v - e_k$ is the location of the parent of y. For Graphs 2-4, v denotes the location of the last common ancestor of x and z, and w the common location of y and z. Graph 2 shows the case where the ancestral lines of x and z split by the birth of a new ancestor of x, Graph 3 the case where this occurs by a new ancestor of z and Graph 4 the case where the first unique ancestors of x and of z are born simultaneously as part of the same group of identical n-tuplets.

$v \in \mathbb{Q}_n$ and integrating over t from 0 to u. This is clearly bounded from above by the first term in the right-hand side of equation (4.13).

We now consider the cases where the three particles x, y, z have common ancestors. Denote the last common ancestor of the particles by l and its location by v. As x and z have common ancestors but neither is a descendant of the other, there must be a time s when the ancestral lines of x and z split. There are three possible ways in which this can occur, as illustrated by Graphs 2-4 in Figure 3; either a new ancestor of x is born, a new ancestor of z is born, or new ancestors of x and z are identical n-tuplets and therefore born at the same time. Observe that, in all three cases, y must be born strictly after this time. We let w denote the common location of y and z.

We now count the expected number of such triples corresponding to fixed vertices v and w, where the ancestral lines split during the time interval $[s, s+dt)$ and such that y is born during $[s+t, s+t+dt)$. The potential l:s are the particles in the CTP at v at time s, hence the expected number of potential l:s is $\sum_{i=1}^{n} m(v-e_i, s)$. For each potential l, the probability that it gives birth during $[s, s+ds)$ is ds. Now, for each possibility for the ancestral lines of x and z to split, conditioned on the process at time $s + ds$, the pairs of potential x:s and y:s originate from a different particle than the potential z:s. Hence these are born independently. By following the ancestral lines as illustrated in Graphs 2-4 in a similar manner as above, we see that the expected number of triples with common ancestors corresponding to fixed v and w, fixed time intervals, and corresponding to each case for how the ancestral line splits are given by

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} m(v-e_i, s)\, m(w-e_k-v-e_j, t)\cdot \\ \cdot\, m(w-v, t)\, m(\hat{\boldsymbol{v}}-w, u-s-t)\, ds\, dt\end{aligned} \tag{4.15}$$

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} m(v-e_i, s)\, m(w-e_k-v, t)\cdot \\ \cdot\, m(w-v-e_j, t)\, m(\hat{\boldsymbol{v}}-w, u-s-t)\, ds\, dt\end{aligned} \tag{4.16}$$

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\substack{k=1\\k\neq j}}^{n}\sum_{l=1}^{n} m(v-e_i, s)\, m(w-e_l-v-e_j, t)\cdot \\ \cdot\, m(w-v-e_k, t)\, m(\hat{\boldsymbol{v}}-w, u-s-t)\, ds\, dt\end{aligned} \tag{4.17}$$

respectively. The total expected number of triples x, y, z with common ancestors is hence given by summing these three expressions over all pairs of vertices $v, w \in \mathbb{Q}_n$ and integrating over all s and t such that $s, t \geq 0$ and $s + t \leq u$.

It only remains to simplify these expressions. We observe that summing (4.15), (4.16) and (4.17) over all $v, w \in \mathbb{Q}_n$ removes all dependence on s. Consider in particular the sum of (4.15) over all $v, w \in \mathbb{Q}_n$. By substituting summing over w by summing over $\Delta = w - v$ and applying Lemma 4.1 we have

$$\begin{aligned}
&\sum_{v,\Delta\in\mathbb{Q}_n}\sum_{i,j,k} m(v-e_i,s)\,m(\Delta-e_k-e_j,t)\cdot \\
&\qquad \cdot\, m(\Delta,t)\,m(\hat{\boldsymbol{v}}-\Delta-v,u-s-t)\,ds\,dt \\
&= \sum_{\Delta\in\mathbb{Q}_n}\sum_{i,j,k} m(\Delta-e_k-e_j,t)\,m(\Delta,t)m(\hat{\boldsymbol{v}}-\Delta-e_i,u-t)\,ds\,dt.
\end{aligned} \tag{4.18}$$

Integrating this expression over all $s, t \geq 0$ such that $s + t \leq u$, we see that the expected number of triples of particles x, y, z as above corresponding to the case illustrated in Graph 2 in Figure 3 is given by

$$\int_0^u (u-t)\sum_{\Delta\in\mathbb{Q}_n}\sum_{i,j,k} m(\Delta-e_k-e_j,t)\,m(\Delta,t)m(\hat{\boldsymbol{v}}-\Delta-e_i,u-t)\,dt. \tag{4.19}$$

Proceeding analogously for (4.16) and (4.17) we see that the expected number of triples corresponding to Graphs 3 and 4 in Figure 3 are given respectively by

$$\int_0^u (u-t)\sum_{\Delta\in\mathbb{Q}_n}\sum_{i,j,k} m(\Delta-e_k,t)\,m(\Delta-e_j,t)m(\hat{\boldsymbol{v}}-\Delta-e_i,u-t)\,dt \tag{4.20}$$

and

$$\int_0^t (u-t)\sum_{\Delta\in\mathbb{Q}_n}\sum_{\substack{i,j,k,l\\ j\neq k}} m(\Delta-e_l-e_j,t)\,m(\Delta-e_k,t)m(\hat{\boldsymbol{v}}-\Delta-e_i,u-t)\,dt. \tag{4.21}$$

The expressions in (4.19)-(4.21) are clearly bounded from above by terms 2-4 respectively in the right-hand side of equation (4.13). ■

Consider the sum $\sum_{\Delta\in\mathbb{Q}_n} m(\Delta,a)^2\,m(\hat{\boldsymbol{v}}-\Delta,b)$. For any $v \in \mathbb{Q}_n$ we let v^i denote the i:th coordinate of v. Define the function $m_1 : \{0,1\} \times \mathbb{R} \to \mathbb{R}$ by $m_1(0,t) = \cosh t$ and $m_1(1,t) = \sinh t$. Using the fact that $m(v,t) = \prod_{i=1}^n m_1(v^i,t)$, we see that

$$\begin{aligned}
&\sum_{\Delta\in\mathbb{Q}_n} m(\Delta,a)^2\, m(\hat{\boldsymbol{v}}-\Delta,b)\\
&\qquad = \sum_{\Delta\in\mathbb{Q}_n}\prod_{i=1}^n m_1(\Delta^i,a)^2\, m_1(\hat{\boldsymbol{v}}^i+\Delta^i,b)\\
&\qquad = \prod_{i=1}^n\sum_{\delta=0}^1 m_1(\delta,a)^2\, m_1(\hat{\boldsymbol{v}}^i+\delta,b)\\
&\qquad = \big(\cosh(a)^2\,\sinh(b)+\sinh(a)^2\,\cosh(b)\big)^k\cdot\\
&\qquad\qquad \cdot\big(\sinh(a)^2\,\sinh(b)+\cosh(a)^2\,\cosh(b)\big)^{n-k}\\
&\qquad = e^{nb}\left(\frac{1}{2}\cosh 2a-\frac{1}{2}e^{-2b}\right)^k\left(\frac{1}{2}\cosh 2a+\frac{1}{2}e^{-2b}\right)^{n-k},
\end{aligned}$$

where $k = |\hat{\boldsymbol{v}}|$. Let

$$\begin{aligned}
G_x(a,b) &= x\ln\big(\cosh(a)^2\,\sinh(b)+\sinh(a)^2\,\cosh(b)\big)\\
&\quad + (1-x)\ln\big(\sinh(a)^2\,\sinh(b)+\cosh(a)^2\,\cosh(b)\big)\\
&= b + x\ln\left(\frac{1}{2}\cosh 2a-\frac{1}{2}e^{-2b}\right) \qquad (4.22)\\
&\quad + (1-x)\ln\left(\frac{1}{2}\cosh 2a+\frac{1}{2}e^{-2b}\right).
\end{aligned}$$

Then

$$\sum_{\Delta\in\mathbb{Q}_n} m(\Delta,a)^2\, m(\hat{\boldsymbol{v}}-\Delta,b) = \exp\left(nG_{\frac{k}{n}}(a,b)\right). \qquad (4.23)$$

Proposition 4.5. *For any $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ only depending on ε such that whenever $u \in [\varepsilon, 1]$ we have*

$$\begin{aligned}
B(\hat{\boldsymbol{v}},u) \le C_\varepsilon \int_0^u &\left[\left(n^4t^3+n^3t+n^2\right)(u-t)+\left(n^3t^3+n^2t+n\right)\right]\cdot\\
&\cdot\exp\left(nG_{\frac{k}{n}}(t,u-t)\right)\,dt,
\end{aligned} \qquad (4.24)$$

where $k = |\hat{\boldsymbol{v}}|$.

Proof. The idea of the proof is to use equation (4.23) to reformulate equation (4.13) in terms of partial derivatives of $G_x(a,b)$. Note that by the fact that $m(v,t)$ satisfies (3.1) we have

$$\begin{aligned}\sum_{\Delta\in\mathbb{Q}_n}\sum_{i,j,k} & m(\Delta-e_k-e_j,a)\,m(\Delta,a)m(\hat{\boldsymbol{v}}-\Delta-e_i,b)\\ &+\sum_{\Delta\in\mathbb{Q}_n}\sum_{i,j,k} m(\Delta-e_k,a)\,m(\Delta-e_j,a)m(\hat{\boldsymbol{v}}-\Delta-e_i,b)\\ &=\sum_{\Delta\in\mathbb{Q}_n} m''(\Delta,a)\,m(\Delta,a)m'(\hat{\boldsymbol{v}}-\Delta,b)\\ &\qquad+m'(\Delta,a)\,m'(\Delta,a)m'(\hat{\boldsymbol{v}}-\Delta,b)\\ &=\frac{1}{2}\frac{\partial^3}{\partial a^2\,\partial b}\sum_{\Delta\in\mathbb{Q}_n} m(\Delta,a)^2 m(\hat{\boldsymbol{v}}-\Delta,b)\\ &=\frac{1}{2}\frac{\partial^3}{\partial a^2\,\partial b}\exp\left(nG_{\frac{k}{n}}(a,b)\right).\end{aligned}$$

Similarly, using the fact that all derivatives of $m(v,t)$ are non-negative, we have

$$\begin{aligned}\sum_{\Delta\in\mathbb{Q}_n}\sum_{i,j,k,l} & m(\Delta-e_l-e_j,a)\,m(\Delta-e_k,a)m(\hat{\boldsymbol{v}}-\Delta-e_i,b)\\ &\leq\frac{1}{6}\frac{\partial^4}{\partial a^3\,\partial b}\exp\left(nG_{\frac{k}{n}}(a,b)\right),\end{aligned}$$

and

$$\begin{aligned}\sum_{\Delta\in\mathbb{Q}_n}\sum_{i,j,k} & m(\Delta-e_k-e_i,a)\,m(\Delta-e_j,a)m(\hat{\boldsymbol{v}}-\Delta,b)\\ &\leq\frac{1}{6}\frac{\partial^3}{\partial a^3}\exp\left(nG_{\frac{k}{n}}(a,b)\right).\end{aligned}$$

Let c denote the minimum of $\frac{1}{2}\cosh 2a-\frac{1}{2}e^{-2b}$ over all $a,b\geq 0$ such that $\varepsilon\leq a+b\leq 1$. It is clear that $c>0$. This means that for any a,b in this range and any $0\leq x\leq 1$ we have

$$\begin{aligned}&\left|\frac{\partial}{\partial a}G_x(a,b)\right|\\ &\qquad=\left|x\frac{\sinh 2a}{\frac{1}{2}\cosh 2a-\frac{1}{2}e^{-2b}}+(1-x)\frac{\sinh 2a}{\frac{1}{2}\cosh 2a+\frac{1}{2}e^{-2b}}\right| \qquad (4.25)\\ &\qquad\leq c^{-1}\sinh 2a.\end{aligned}$$

Hence, for sufficiently large $C > 0$ we have $\left|\frac{\partial}{\partial a} G_x(a,b)\right| \leq C\,a$ whenever $0 \leq x \leq 1$ and $a, b \geq 0$ such that $\varepsilon \leq a+b \leq 1$. Moreover, as $G_x(a,b)$ is smooth wherever it is defined, we know that for C sufficiently large all partial derivatives of order up to 4 of $G_x(a,b)$ are bounded in absolute value by C when the pair (a,b) is in this domain.

By explicitly writing out the partial derivatives of $\exp\left(nG_{\frac{k}{n}}(a,b)\right)$ above and combining this with Proposition 4.4 we see that (4.24) holds for sufficiently large C, as desired. ■

For a given sequence $\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}_n$ as above, we define

$$\begin{aligned} f_n(t) &= G_{\frac{k}{n}}(t, \vartheta_n - t) \\ &= \vartheta_n - t + \frac{k}{n} \ln\left(\frac{1}{2}\cosh 2t - \frac{1}{2}e^{-2\vartheta_n + 2t}\right) \\ &\qquad + \frac{n-k}{n} \ln\left(\frac{1}{2}\cosh 2t + \frac{1}{2}e^{-2\vartheta_n + 2t}\right) \end{aligned} \tag{4.26}$$

and

$$\begin{aligned} f(t) &= G_x(t, \vartheta(x) - t) \\ &= \vartheta(x) - t + x \ln\left(\frac{1}{2}\cosh 2t - \frac{1}{2}e^{-2\vartheta(x) + 2t}\right) \\ &\qquad + (1-x) \ln\left(\frac{1}{2}\cosh 2t + \frac{1}{2}e^{-2\vartheta(x) + 2t}\right). \end{aligned} \tag{4.27}$$

Note that f depends on x. From the definition of $G_x(a,b)$ we see that $f_n(0) = -\frac{\ln n}{n}$ and $f_n(\vartheta_n) = -2\frac{\ln n}{n}$, and that $f(0) = f(\vartheta(x)) = 0$, see (1.7).

Suppose that $f_n(t)$ is "asymptotically U-shaped" in the sense that exists a constant $\lambda > 0$ such that for sufficiently large n we have

$$f_n(t) \leq \max\left(f_n(0) - \lambda t, f_n(\vartheta_n) - \lambda(\vartheta_n - t)\right) \tag{4.28}$$

for any $0 \leq t \leq \vartheta_n$. If this holds, then by Proposition 4.5 we have

$$B(\hat{\boldsymbol{v}}_n, \vartheta_n) \leq \int_0^\infty O\left(n^3t^3 + n^2t + n\right) e^{-\lambda n t}\, dt, \tag{4.29}$$

which would imply that $B(\hat{\boldsymbol{v}}_n, \vartheta_n) = O(1)$ as desired. It remains to show for which sequences of vertices $\hat{\boldsymbol{v}} = \hat{\boldsymbol{v}}_n$, f_n is asymptotically U-shaped. We start by giving a simple sufficient condition for x.

Proposition 4.6. *Suppose that* $x > 1 - \frac{\ln(2\sqrt{2})}{\ln 3} \approx 0.054$. *Then* $f_n(t)$ *is asymptotically U-shaped.*

Proof. By some straight-forward but tedious calculations we see that

$$f_n''(t) = \frac{k}{n}\frac{4(1-2e^{-2\vartheta_n})}{(\cosh 2t - e^{-2\vartheta_n+2t})^2} + \frac{n-k}{n}\frac{4(1+2e^{-2\vartheta_n})}{(\cosh 2t + e^{-2\vartheta_n+2t})^2}. \tag{4.30}$$

Now, if we assume that $\vartheta_n \geq \ln\sqrt{2}$, then the first term in the right-hand side is non-negative, and so we have that $f_n''(t)$ is at least, say, $\frac{n-k}{100n}$ for all $0 \leq t \leq \vartheta_n$. It follows that if $\vartheta(x) > \ln\sqrt{2}$, then $f_n(t)$ is asymptotically U-shaped. The proposition follows by the easily verified fact that $\vartheta\left(1 - \frac{\ln(2\sqrt{2})}{\ln 3}\right) = \ln\sqrt{2}$. ■

It is clear from the proof of Proposition 4.6 that the limit $1 - \frac{\ln(2\sqrt{2})}{\ln 3}$ is not optimal, and can be lowered by considering $f_n''(t)$ more closely. It turns out however that there is a limit for x at which the convexity of f_n breaks down, and more importantly for sufficiently small x the asymptotic U-shape of f_n breaks down. In the remaining part of this section, we will investigate when this occurs.

By some more straight-forward but tedious calculations we see that

$$\begin{aligned} &f_n'(t)\cdot\left(\cosh 2t - e^{-2\vartheta_n+2t}\right)\left(\cosh 2t + e^{-2\vartheta_n+2t}\right) \\ &= \left(\frac{1}{4} - e^{-4\vartheta_n}\right)e^{4t} - \frac{3}{4}e^{-4t} - \frac{1}{2} + 2\frac{n-2k}{n}e^{-2\vartheta_n}. \end{aligned} \tag{4.31}$$

This expression has the same sign as $f_n'(t)$. We see that depending on the sign of $\frac{1}{4} - e^{-4\vartheta_n}$ it is either increasing or concave, hence f_n changes sign at most twice. Furthermore, if f_n changes sign twice it goes from negative to positive to negative. In the same way, since

$$\begin{aligned} &f'(t)\cdot\left(\cosh 2t - e^{-2\vartheta(x)+2t}\right)\left(\cosh 2t + e^{-2\vartheta(x)+2t}\right) \\ &= \left(\frac{1}{4} - e^{-4\vartheta(x)}\right)e^{4t} - \frac{3}{4}e^{-4t} - \frac{1}{2} + 2(1-2x)e^{-2\vartheta(x)}. \end{aligned} \tag{4.32}$$

the same must be true for $f(t)$.

Combining this observation with the fact that $f_n''(t)$ is bounded it follows that a necessary and sufficient condition for f_n being

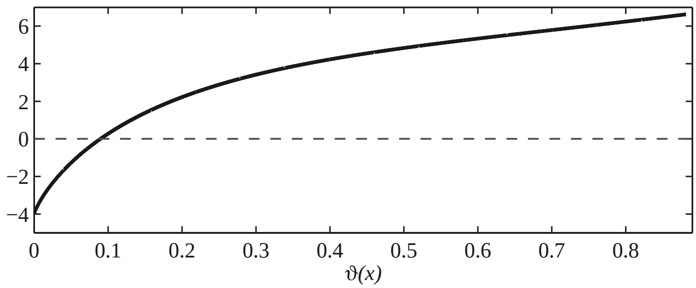


Figure 4: Plot of equation (4.34) divided by x as a function of $\vartheta(x)$. We see that as ϑ tends to 0, this converges to its limit of -4. The curve intersects the ϑ-axis at $\vartheta(x) \approx 0.0898$, that is at $x \approx 0.00167$.

asymptotically U-shaped is that $\lim_{n\to\infty} f_n'(0) = f'(0) < 0$ and $\lim_{n\to\infty} f_n'(\vartheta_n) = f'(\vartheta(x)) > 0$. In fact, the former condition is implied by the latter as then $f_n'(t)$ changes sign at most once, but $f(0) = f(\vartheta(x)) = 0$.

As $1 = (\cosh\vartheta)(\tanh\vartheta)^x$ we have

$$x(\vartheta) = \frac{\ln\cosh\vartheta}{-\ln\tanh\vartheta} = \frac{\vartheta^2}{-2\ln\vartheta} + O\left(\frac{\vartheta^4}{(\ln\vartheta)^2}\right). \tag{4.33}$$

Hence, we have an explicit expression for (4.32) as a function of $\vartheta(x)$. Plugging $t = \vartheta(x)$ into the right-hand side of this expression we get

$$\frac{1}{4}e^{4\vartheta} - \frac{3}{4}e^{-4\vartheta} + 2(1-2x)e^{-2\vartheta} - \frac{3}{2}. \tag{4.34}$$

Note that this has the same sign as $f'(\vartheta(x))$. By Taylor expanding this expression in x and ϑ we see that the dominating term for small ϑ is $-4x$. Hence f_n is not asymptotically U-shaped for sufficiently small x. To get a picture of what happens when x increases, we divide (4.34) by x and plot as a function of ϑ, see Figure 4. It is clear that there is a critical value x^* slightly less than 0.0017 such f_n is asymptotically U-shaped if and only if $x > x^*$. This proves the following proposition:

Proposition 4.7. *Let $\{\hat{\boldsymbol{v}}_n\}_{n=1}^{\infty}$ be a sequence of vertices, $\hat{\boldsymbol{v}}_n \in \mathbb{Q}_n$, such that $\lim_{n\to\infty} |\hat{\boldsymbol{v}}_n|/n$ exists and is strictly greater than x^*. Then for ϑ_n as defined in* (3.17) *we have $B(\hat{\boldsymbol{v}}_n, \vartheta_n) = O(1)$.*

Remark 4.8. Throughout this section we have only really been interested in deriving a tractable upper bound for $B(\hat{\boldsymbol{v}}_n, \vartheta_n)$ without discussing sharpness. Nevertheless, it is not too hard to convince oneself that the bound given in Proposition 4.5 is sharp up to, say, a polynomial factor in n. However, for $x < x^*$ we know that there exists an interval of positive length for t where $f_n(t)$ is positive, which would then imply that $B(\hat{\boldsymbol{v}}_n, \vartheta_n)$ diverges exponentially fast in n.

5 Completing the proof of Theorem 1.6

Let $\{\hat{\boldsymbol{v}}_n\}_{n=1}^{\infty}$ be a sequence of vertices, $\hat{\boldsymbol{v}}_n \in \mathbb{Q}_n$ for each n, such that $x = \lim_{n\to\infty} |\hat{\boldsymbol{v}}_n| / n$ exists and is at least 0.002 and let $\{\vartheta_n\}_{n=1}^{\infty}$ be as in (3.17). Applying the estimates of $S(\hat{\boldsymbol{v}}_n, \vartheta_n)$ and $B(\hat{\boldsymbol{v}}_n, \vartheta_n)$ from Propositions 4.3 and 4.7 to Theorem 3.6 it follows by Proposition 3.1 and Lemma 3.2 that there exists a constant $c_0 > 0$ such that

$$\liminf_{n\to\infty} \mathbb{P}\left(\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}_n) \le \vartheta_n\right) \ge c_0. \tag{5.1}$$

Since $\vartheta_n \to \vartheta(x)$ as $n \to \infty$, this means in particular that for any $\varepsilon > 0$ we have

$$\liminf_{n\to\infty} \mathbb{P}\left(\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}_n) \le \vartheta(x) + \varepsilon\right) \ge c_0. \tag{5.2}$$

Note that we can assume that c_0 is independent of the choice of sequence.

Proposition 5.1. *Let $\{\hat{\boldsymbol{v}}_n\}_{n=1}^{\infty}$ be a sequence as above, and let $x = \lim |\hat{\boldsymbol{v}}| / n$. Then, for any $\varepsilon > 0$ we have*

$$\mathbb{P}\left(\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}_n) \le \vartheta(x) + \varepsilon\right) \to 1 \tag{5.3}$$

as $n \to \infty$.

Proof. Let $\varepsilon > 0$ be arbitrary. Condition on the vertex passage times of all neighbors of $\hat{\mathbf{0}}$ and $\hat{\boldsymbol{v}}_n$. Assuming $|\hat{\boldsymbol{v}}_n| \ge 3$, it is easy to see that the number of coordinate places $1 \le i \le n$ with the property that the i:th coordinate of $\hat{\boldsymbol{v}}_n$ is 1, and the cost of both e_i and $\hat{\boldsymbol{v}}_n - e_i$ are at most $\varepsilon/3$, is distributed as $\text{Bin}\left(|\hat{\boldsymbol{v}}_n|, (1 - e^{-\varepsilon/3})^2\right)$. Hence as $n \to \infty$ it is clear that, with probability $1 - o(1)$, there are at least two such coordiantes. Pick a pair $i \ne j$.

Depending on the choice of i and j, we define Q_0 as the induced subgraph of $\mathbb{Q}_n$ with vertex set $\{v \in \mathbb{Q}_n : v_i = 1, v_j = 0\}$. We similarly define Q_1 as the induced subgraph of $\mathbb{Q}_n$ with vertex set $\{v \in \mathbb{Q}_n : v_i = 0, v_j = 1\}$. Note that Q_0 and Q_1 are vertex disjoint subgraphs of $\mathbb{Q}_n$, both isomorphic to $\mathbb{Q}_{n-2}$.

In light of Q_0 and Q_1, we have two natural upper bounds for $\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}_n)$, namely $c(e_i)+c(\hat{\boldsymbol{v}}_n-e_j)$ plus the smallest reduced vertex passage time for any path from e_i to $\hat{\boldsymbol{v}}_n-e_j$ in Q_0, and $c(e_j)+c(\hat{\boldsymbol{v}}_n-e_i)$ plus the smallest reduced vertex passage time for any path from e_j to $\hat{\boldsymbol{v}}_n - e_i$ in Q_1. As the only vertices of Q_0 and Q_1 which are neighbors of $\hat{\mathbf{0}}$ or $\hat{\boldsymbol{v}}_n$ are e_i, e_j, $\hat{\boldsymbol{v}}_n - e_i$ and $\hat{\boldsymbol{v}}_n - e_j$, the reduced first-passage times in Q_0 and Q_1 are independent of each other and each is distributed as the reduced first-passage time between two vertices at distance $|\hat{\boldsymbol{v}}_n| - 2$ in $\mathbb{Q}_{n-2}$. By applying (5.2) to the first-passage percolation problems in Q_0 and Q_1, we conclude that for any $\varepsilon > 0$ and for any sequence $\{\hat{\boldsymbol{v}}_n\}_{n=1}^{\infty}$ where $\hat{\boldsymbol{v}}_n \in \mathbb{Q}_n$ for each $n \geq 1$ such that $x = \lim_{n\to\infty} |\hat{\boldsymbol{v}}_n| / n$ exists and is at least 0.002, we have

$$\liminf_{n\to\infty} \mathbb{P}\left(\mathcal{T}'_V(\hat{\mathbf{0}}, \hat{\boldsymbol{v}}_n) \leq \vartheta(x) + \varepsilon\right) \geq 1 - (1 - c_0)^2. \tag{5.4}$$

Note that this is the same expression as (5.2), except that the right-hand side here is strictly larger. Hence, by iteratively applying this argument, we see that we can replace the right-hand side in (5.4) by $c_k = 1 - (1 - c_0)^{2^k}$ for any non-negative integer k. The Proposition follows by letting $k \to \infty$. ■

Acknowledgements

The author would like to thank his supervisor, Peter Hegarty, for helpful discussions.

References

[1] Stuart Kauffman and Simon Levin, *Towards a general theory of adaptive walks on rugged landscapes*, J. Theor. Biol. **128** (1987), no. 1, 11-45, DOI 10.1016/s0022-5193(87)80029-2.

[2] Geoffrey Grimmett, *Percolation*, 2nd ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 321, Springer-Verlag, Berlin, 1999. MR **1707339 (2001a:**60114)

[3] John Kingman, *A simple model for the balance between selection and mutation*, J. Appl. Probab. **15** (1978), no. 1, 1–12. MR 0465272 (57 #5177)

[4] Jasper Franke, Alexander Klözer, J. Arjan G. M. de Visser, and Joachim Krug, *Evolutionary accessibility of mutational pathways*, PLoS Comput. Biol. **7** (2011), no. 8, e1002134, 9, DOI 10.1371/journal.pcbi.1002134. MR 2845072

[5] Peter Hegarty and Anders Martinsson, *On the existence of accessible paths in various models of fitness landscapes*, Ann. Appl. Probab. **24** (2014), no. 4, 1375-1395, DOI 10.1214/13-AAP949.

[6] Julien Berestycki, Éric Brunet, and Zhan Shi, *The number of accessible paths in the hypercube*, available at `arXiv:1304.0246[math.PR]`.

[7] ———, *Accessibility percolation with backsteps*, available at `arXiv:1401.6894[math.PR]`.

[8] Stefan Nowak and Joachim Krug, *Accessibility percolation on n-trees*, Europhys. Lett. **101** (2013), no. 6.

[9] Matthew Roberts and Lee Zhuo Zhao, *Increasing paths in regular trees*, Electron. Commun. Probab. **18** (2013), no. 87, 1-10, DOI 10.1214/ECP.v18-2784.

[10] Xinxin Chen, *Increasing paths on N-ary trees*, available at `arXiv:1403.0843[math.PR]`.

[11] James Allen Fill and Robin Pemantle, *Percolation, first-passage percolation and covering times for Richardson's model on the n-cube*, Ann. Appl. Probab. **3** (1993), no. 2, 593–629. MR **1221168 (94h:**60145)

[12] Anders Martinsson, *First-passage percolation on the unoriented n-cube*, available at `arXiv:1402.2928[math.PR]`.

[13] Louigi Addario-Berry and Bruce Reed, *Minima in branching random walks*, Ann. Probab. **37** (2009), no. 3, 1044–1079, DOI 10.1214/08-AOP428. MR **2537549 (2011b:**60338)