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 $W_{vw}$ 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 $\gamma$ in $G$ we define the passage time of $\gamma$ as the sum of passage times of the edges along $\gamma$. 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

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

$$N(1)$$

**UNORIENTED FIRST-PASSAGE PERCOLATION ON THE $n$-CUBE**

ANDERS MARTINSSON

**Abstract.** The $n$-dimensional binary hypercube $Q_n$ 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 let $\hat{0}$ and $\hat{1}$ denote the all zeroes and all ones vertices respectively. For any vertex $v \in Q_n$, we let $|v|$ denote the number of coordinates of $v$ that are 1. A path $v_0, v_1, \ldots, v_k$ in $Q_n$ 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 $W_{vw}$ 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 $\gamma$ in $G$ we define the passage time of $\gamma$ as the sum of passage times of the edges along $\gamma$. 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

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

**Date:** February 13, 2014.

2010 Mathematics Subject Classification. 60K35, 60C05.

Key words and phrases. first-passage percolation, Richardson’s model, hypercube, branching translation process, lower bound on Richardson’s model.
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 $R(\cdot, 0) = \delta_{o_{\cdot}}$, and the transitions $\{R(\cdot) \to R(\cdot) + \delta_{o_{\cdot}}\}$ at rate equal to the number of infected neighbors of $v$ if $v$ is healthy, and 0 if $v$ is infected, see [1]. Here $\delta_{o_{\cdot}}$ 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 $\hat{0}$, though by symmetry of the hypercube it is clear that the analogous statements hold for any starting vertex. The quantities considered in these articles are the first-passage time from $\hat{0}$ to $\hat{1}$, which we denote by $T_n$, the oriented first-passage time from $\hat{0}$ to $\hat{1}$, and the covering time. Note that, in terms of Richardson’s model, $T_n$ is the time until the vertex antipodal to 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 $\hat{0}$ to $\hat{1}$. The covering time is the random amount of time in Richardson’s model on $Q_n$ until all vertices are infected or, equivalently, $\max_{v \in Q_n} d_W(\hat{0}, v)$, the maximum first-passage time from $\hat{0}$ to any other vertex in $Q_n$.

For various random processes on the hypercube it turns out to be simpler to consider the oriented hypercube, where each edge is oriented towards $\hat{1}$, rather than the full hypercube. In the oriented hypercube we have $n!$ paths from $\hat{0}$ to $\hat{1}$, all of length $n$ and all equivalent up to permutation of coordinates. There is also a nice parametrization of oriented paths as permutations of $\{1, 2, \ldots, n\}$ which makes it possible to treat how pairs of paths intersect in a combinatorial manner in terms of permutations. These properties of the oriented paths allows us to approach many questions regarding random processes on the oriented hypercube by a moment analysis. Examples of where arguments of this type have been used are oriented first-passage percolation, oriented Bernoulli percolation, and, more recently, oriented accessibility percolation, see [2] [4] [5]. In contrast, general paths from $\hat{0}$ to $\hat{1}$ do not seem to have a similar parametrization, and in any case there are a lot more paths with a lot more variation between them. Hence, path counting arguments appear not to be feasible for the unoriented hypercube.

In case of first-passage percolation, it was shown by Fill and Pemantle that the oriented first-passage time converges to 1 in probability as $n \to \infty$. Their proof is essentially a second moment analysis on the number of oriented paths from $\hat{0}$ to $\hat{1}$ with passage time at most $t$, 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, the authors consider a “variance reduction trick”, which effectively means that they consider a slightly different random variable.

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

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

with probability $1 - o(1)$. The upper bound follows directly from the oriented first-passage time. The authors 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 $Q_n$, which Durrett calls a branching translation process (BTP). This is a kind of branching process in which each individual is placed at one of the vertices in $Q_n$. 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 occupied vertices in the BTP. He proves that the time when the first particle at $\hat{1}$ is born 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 $1 - o(1)$. The advantage of considering this process is that the expected number of
particles at each vertex at a given time must satisfy a certain initial value problem which can be solved explicitly, hence circumventing the problem of counting paths. However, beyond the fact that BTP stochastically dominates Richardson’s model the relation between the two models is fairly subtle, and there is therefore no clear canonical approach to finding a corresponding upper bound.

Bollobás and Kohayakawa [3] showed that many global first-passage percolation properties on $Q_n$, 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

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

The authors referred to this quantity as simply the first-passage percolation time between two antipodal vertices in $Q_n$. From the results by Fill and Pemantle it follows that $\ln (1 + \sqrt{2}) \leq T_\infty \leq 1$. Furthermore, it is easy to see that if $T_n$ converges in probability as $n \to \infty$, then it must converge to $T_\infty$. In fact, it was conjectured by Bollobás and Kohayakawa that this is the case. Their main result is that with probability $1-o(1)$, the covering time is at most $T_\infty + \ln 2 + o(1)$ and the graph diameter is at most $T_\infty + 2 \ln 2 + o(1)$.

In this paper, we return to the problem of determining $T_n$. We prove that, as $n \to \infty$, we have $T_n \to \ln (1 + \sqrt{2})$ not only in probability but in $L^p$-norm for all $1 \leq p < \infty$, confirming the conjecture by Bollobás and Kohayakawa. A direct consequence of this is that $T_\infty = \ln (1 + \sqrt{2})$, which in particular improves the best known upper bound on the covering time to $\ln (1 + \sqrt{2}) + \ln 2 + o(1) = 1.574 \cdots + o(1)$. One can compare this with the best known lower bound $\frac{1}{2} \ln (2 + \sqrt{5}) + \ln 2 - o(1) = 1.414 \cdots - o(1)$, as shown by Fill and Pemantle. The central idea of our proof is to consider a certain subset of particles in the BTP, which we call the uncontested particles. These particles form a subprocess of the BTP which is stochastically dominated by Richardson’s model. Hence Richardson’s model is stochastically sandwiched between the full BTP and this subprocess. We derive an explicit lower bound on the probability that the first particle to arrive at a vertex is uncontested, which in turn gives a lower bound on the probability that the time of the first arrival at a vertex is equal to the first-passage time to that vertex. This is used to show that $T_n \leq \ln (1 + \sqrt{2})$ with probability bounded away from 0 as $n \to \infty$. After this we employ a bootstrapping argument similar to a technique used in [4] to show that a slightly larger upper bound holds with probability $1-o(1)$.

2. Richardson’s model and uncontested particles

Before presenting our results we first give an overview of the technique used by Durrett to obtain the lower bound on $T_n$ in [2]. To accommodate our first main result, 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 a finite connected simple graph.

The branching translation process (BTP), as introduced by Durrett, is a branching process where each particle is considered to be located at one of the vertices of some given graph $G$, defined in the following way: At time 0 we place a particle at $v_0$ in $G$. 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 $m(v,t) = E Z(v,t)$. 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 $\{Z(\cdot) \to Z(\cdot) + \delta_{v,w}\}$ 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 parametrization 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 parametrizing 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 formal definition 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 \( \epsilon \)-child of \( x \) to denote a child of \( x \) which at time of its birth was displaced along an edge \( \epsilon \). We define the ancestral line of a particle \( x \) as the ordered set of all ancestors of \( x \) (including \( x \) itself). If \( \sigma \) 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 \( \sigma \), 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 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 vertices at \( v \) at time \( t \).

The proof of the lower bound on \( T_n \) in [2] can now be summarized as follows. Consider a BTP on \( Q_n \) originating at \( \hat{0} \). Since BTP dominates Richardson’s model it suffices to show that with probability \( 1 - o(1) \), no particle occupies \( \hat{1} \) at time \( \ln \left( 1 + \sqrt{2} \right) - \epsilon \) for all \( \epsilon > 0 \) 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

\[
\frac{d}{dt}m(v,t) = \sum_{w \in N(v)} m(w,t), \quad t > 0
\]

\[
m(v,0) = \delta_{v,v_0}.
\]

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

\[
m(v,t) = (\sinh t)^{|v|} (\cosh t)^{n-|v|}.
\]

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

4
Our first main result, Theorem 2.2, is a means to complete this analysis. 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 \( x \) and \( y \) to be the same particle. 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 \( x \) 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

\[
Z_0 \leq \frac{d}{d} \text{Richardson’s model} \leq Z,
\]

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.

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 go 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 \( a(x) = 0 \).

**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

\[
P(Z_0(v, t) > 0) \geq S(v, t)e^{-\frac{B(v)}{S(v, t)}}.
\]

Loosely speaking, 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 the right-hand side of (2.4) is not an increasing function of $t$, and it seems that this expression generally attains its maximum for $t$ such that $m(v, t) \approx 1$.

We now apply this result to the hypercube. We set $G = Q_n$, $v_0 = \hat{0}$ and $t = u = \ln (1 + \sqrt{2})$. In this case, the quantities $A(\hat{1}, u)$ and $B(\hat{1}, u)$ can be expressed analytically in a similar manner as the variance calculations for the BTP in [2]. This will be done in Section 3. We have the following estimates:

**Proposition 2.3.** Let $u = \ln (1 + \sqrt{2})$. Then

$$A(\hat{1}, u) = \frac{n}{\sqrt{2}} + o(1) = 0.623 \cdots + o(1)$$

$$B(\hat{1}, u) = u + \frac{1}{3 - 2\sqrt{2}} + o(1) = 6.709 \cdots + o(1).$$

In order to bound $S(\hat{1}, u)$, we observe that $A(v, t)$ is an upper bound on the expected number of particles at $v$ at time $t$ whose ancestral line is 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

$$m(v, t) - A(v, t) \leq S(v, t) \leq m(v, t),$$

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

Plugging these values into Theorem 2.2 we conclude the following:

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

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

In order to complete proof that $T_n$ converges to $\ln (1 + \sqrt{2})$ we will employ a bootstrapping argument similar to one presented in [4]. Using a slightly stronger version of Corollary 2.4 this proves an explicit upper bound on $T_n$ in form of a recursion inequality. Combining this with the lower bound shown by Durrett allows us to obtain a good estimate on the $L^p$-norm of $T_n - u$. This argument will be given in Section 3. We have the following result:

**Theorem 2.5.** Let $T_n$ denote the first-passage time from $\hat{0}$ to $\hat{1}$ in $Q_n$ and let $u = \ln (1 + \sqrt{2})$. For any $1 \leq p < \infty$ we have $\|T_n - u\|_p = \Theta \left( \frac{1}{n} \right)$. In particular, we have $\mathbb{E}T_n = u + O \left( \frac{1}{n} \right)$ and $\text{Var} \left( T_n \right) = \Theta \left( \frac{1}{n} \right)$.

**Remark 2.6.** Bollobás and Kohayakawa have in principle already shown that Corollary 2.4 implies convergence in probability in [3]. However, their article does not point this out explicitly and it would require some albeit small changes to a fairly technical argument.

The remainder of the article will be structured as follows: In Section 3 we discuss how to formally parametrize the BTP and use this to prove of Theorem 2.2. In Section 4 we show Proposition 2.3. Lastly, in Section 5 we derive a recursion inequality on $T_n$ and use this to show convergence in $L^p$-norm, as stated in Theorem 2.5.

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 $\emptyset$, 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 \( x = \{e_1t_1e_2t_2\ldots\ e_{k-1}t_{k-1}\} \) 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 \( x \) is a child of \( v \) is interpreted as the event that the original particle has an \( e \)-child at time \( t_1 \), that this child has an \( e \)-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 \( \exp(-\frac{B_v(x,t)}{S(v,t)}) \). 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 [8].

**Theorem 3.1.** (Slivnyak-Mecke formula) Let \( T \) 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

\[
\mathbb{E} \sum_{t \in T} G(T,t) = \int_0^\infty \mathbb{E} G(T \cup \{t\},t) \, dt.
\]

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

\[
\mathbb{E} \sum_{t \in T} G(T,t) = \sum_t \mathbb{P}(t \in T) \mathbb{E} [G(T,t)|t \in T].
\]

By the standard way to translate this statement, if \( T \) 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 \( \mathbb{P}(t \in T) \) to \( dt \), the Lebesgue measure on \( \mathbb{R}_+ \). Hence, the theorem states that if \( T \) is a Poisson process as above, then we should translate \( \mathbb{E} G(T,t)|t \in T) \) to \( \mathbb{E} G(T \cup \{t\},t) \), and so we may interpret \( T \cup \{t\} \) as the conditional distribution of \( T \) given \( t \in T \).

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 \( x \) exists in the BTP \( X_0 \), the conditional distribution of the process is given by \( X^\{z_2 \ldots z_i\} \), where \( x \) and \( X^\{z_2 \ldots z_i\} \) are as defined below. This result may be well-known from the properties of more general processes.

**Lemma 3.2.** Let \( \sigma \) 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 \( X_0, \ldots, X_l \) be independent branching translation processes where \( X_i \) 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 $v_1$ (no matter when they are born) whose ancestral line follows $\sigma$. Then, we have

\begin{equation}
\mathbb{E} \sum_{x \in V_\sigma(X_0)} f(X_0, x) = \int_0^\infty \cdots \int_0^\infty \mathbb{E} f(X^{z_1, \ldots, z_l}, x^{z_1, \ldots, z_l}) \, dz_1 \cdots dz_l
\end{equation}

where

\begin{equation}
X^{z_1, \ldots, z_l} = X_0 \cup (\{\sigma_1 z_1 \oplus X_1\} \cup (\{\sigma_2 z_2 \oplus X_2\} \cup \cdots \cup (\{\sigma_l z_l \oplus X_l\}
\end{equation}

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

\textbf{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 $T_e$ denote the set of birth times of the $e$-children of the original particle in $X_0$. Clearly, $T_e$ 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 independent of any other particle. Hence we can express $X_0$ recursively by

\begin{equation}
X_0 = \bigcup_{e \ni v_0} \bigcup_{t \in T_e} \{e t_1\} \oplus Y_{e,i}
\end{equation}

where for each $e \ni v_0$ and each $i = 1, 2, \ldots$, we have $Y_{e,i}$ 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 $X_0(T)$ denote the random variable obtained by replacing $T_\sigma$ by $T$ in (3.5). Then $X_0(T)$ is a random function independent of $T_\sigma$, and we have $X_0 = X_0(T_\sigma)$. Note that, by independence, $X_0(T)$ is a version of the conditional distribution of $X_0$ given $T_\sigma = T$.

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

\begin{equation}
F(T) = \mathbb{E} \sum_{x \in V_\sigma(X_0(T))} f(X_0(T), x)
\end{equation}

and

\begin{equation}
F(T, t) = \mathbb{E} \sum_{x \in V_\sigma(X_0(T))} f(X_0(T), x).
\end{equation}

It is clear from the definition that, for any fixed $T$, we have $F(T) = \sum_{t \in T} F(T, t)$. Furthermore, as $T_\sigma$ and $X_0(\cdot)$ are independent we have $\mathbb{E} F(T_\sigma) = \mathbb{E} \sum_{x \in V_\sigma(X_0)} f(X_0, x)$. Hence by the Slivnyak-Mecke formula we have

\begin{equation}
\mathbb{E} \sum_{x \in V_\sigma(X_0)} f(X_0, x) = \mathbb{E} \sum_{t \in T_\sigma} F(T_\sigma, t) = \int_0^\infty \mathbb{E} F(T_\sigma \cup \{z_1\}, z_1) \, dz_1.
\end{equation}

By independence of $X_0(\cdot)$ and $T_\sigma \cup \{z_1\}$ we can conclude that

\begin{equation}
\mathbb{E} \sum_{x \in V_\sigma(X_0)} f(X_0, x) = \int_0^\infty \mathbb{E} \sum_{x \in V_\sigma(X_0(T_\sigma \cup \{z_1\}))} f(X_0(T_\sigma \cup \{z_1\}), x) \, dz_1.
\end{equation}

Let us now consider the random process $X_0(T_\sigma \cup \{z_1\})$. We can interpret the expression for $X_0$ in (3.3) and the subsequent definition of $X_0(T)$ 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 $X_0$ and $X_0(T_\sigma \cup \{z_1\})$ is that the latter has an additional particle in generation 1. Hence, $X_0(T_\sigma \cup \{z_1\})$ has the same distribution as $X_0 \cup \{\sigma_1 z_1 \oplus X_1\}$, and so we can replace $X_0(T_\sigma \cup \{z_1\})$ in (3.9) by this other random process.
Letting $\sigma = \{\sigma_2, \sigma_3, \ldots, \sigma_l\}$, we note that the subset of elements in $V_\sigma (X_0 \cup \{\{\sigma_1 z_1\} \oplus X_1\})$ that are descendants of $\{\sigma_1 z_1\} \oplus X_1$ is precisely the set $\{\sigma_1 z_1\} \oplus V_\sigma (X_1)$. Hence (3.10) simplifies to
\[
(3.10) \quad \mathbb{E} \sum_{x \in V_\sigma (X_0)} f(x, x) = \int_0^\infty \mathbb{E} \sum_{x \in V_\sigma (X_1)} f(x, x) \, dz_1.
\]

The lemma follows by induction. If $l = 1$, then the only particle in $V_\sigma (X_1)$ is $\{\}$, the original particle in $X_1$, and so equation (3.10) simplifies to
\[
(3.11) \quad \mathbb{E} \sum_{x \in V_\sigma (X_0)} f(x, x) = \int_0^\infty \mathbb{E} f(X_1, \{\sigma_1 z_1\}) \, dz_1
\]
as desired.

Now assume $l > 1$. By the induction hypothesis we have for any non-negative function $\tilde{f}$
\[
(3.12) \quad \mathbb{E} \sum_{x \in V_\sigma (X_1)} \tilde{f}(X_1, x) = \int_0^\infty \int_0^\infty \mathbb{E} \tilde{f}(\tilde{X}^{z_2 \ldots z_l}, \tilde{x}^{z_2 \ldots z_l}) \, dz_2 \ldots dz_l,
\]
where
\[
(3.13) \quad \tilde{X}^{z_2 \ldots z_l} = X_1 \cup (\{\sigma_2 z_2\} \oplus X_2) \cup (\{\sigma_2 \sigma_3 z_3\} \oplus X_3) \cup \cdots \cup (\{\sigma_2 \sigma_3 \ldots \sigma_l z_l\} \oplus X_l)
\]
and $\tilde{x}^{z_2 \ldots z_l} = \{\sigma_2 \sigma_3 \ldots \sigma_l z_l\} \oplus \{\}$.

Let us consider the expression $\mathbb{E} \sum_{x \in V_\sigma (X_1)} f(x, x) \cup (\{\sigma_1 z_1\} \oplus X_1), \{\sigma_1 z_1\} \oplus x$, the integrand on the right-hand side of equation (3.10). If we fix $z_1 > 0$ and condition on $X_0 = X_0$, then $f(X_0 \cup \{\sigma_1 z_1\} \oplus X_1), \{\sigma_1 z_1\} \oplus x$ is a function of $X_1$ and $x$ only. By the induction hypothesis,
\[
\mathbb{E} \sum_{x \in V_\sigma (X_1)} f(x, x) = \int_0^\infty \int_0^\infty \mathbb{E} f(X_1, \{\sigma_1 z_1\} \oplus \tilde{X}^{z_2 \ldots z_l}, \{\sigma_1 z_1\} \oplus \tilde{x}^{z_2 \ldots z_l}) \, dz_2 \ldots dz_l.
\]
Hence, by integrating this expression over $z_1$ and $X_0$ we conclude that
\[
\mathbb{E} \sum_{x \in V_\sigma (X_0)} f(x, x) = \int_0^\infty \int_0^\infty \mathbb{E} f(X_0 \cup \{\sigma_1 z_1\} \oplus \tilde{X}^{z_2 \ldots z_l}, \{\sigma_1 z_1\} \oplus \tilde{x}^{z_2 \ldots z_l}) \, dz_2 \ldots dz_l,
\]
where clearly $X^{z_1 \ldots z_l} = X_0 \cup \{\sigma_1 z_1\} \tilde{X}^{z_2 \ldots z_l}$ and $x^{z_1 \ldots z_l} = \{\sigma_1 z_1\} \oplus \tilde{x}^{z_2 \ldots z_l}$.

**Lemma 3.3.** Let $X$ be a BTP originating at a vertex $v$. Let $\varphi$ be an indicator function defined over the set of potential particles. If $\varphi(\{\}) = 0$, then
\[
(3.14) \quad \mathbb{P}(\varphi(x) = 0 \forall x \in X) \geq \exp \left( -\mathbb{E} \sum_{x \in X} \varphi(x) \right).
\]

**Proof.** For any particle $x \in X$, 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 $\sum_{x \text{ in gen 1}} \psi(x) \leq \sum_{x \in X} \varphi(x)$. Furthermore, $\sum_{x \text{ in gen 1}} \psi(x) = 0$ if and only if $\sum_{x \in X} \varphi(x) = 0$.

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 $x \in X$ is $e^{-\mathbb{E} \sum_{x \text{ in gen 1}} \psi(x)}$, which is at least $e^{-\mathbb{E} \sum_{x \in X} \varphi(x)}$. 

9
Prologue of Theorem 3.2. For any path \( \sigma \) 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 \( \sigma \). Similarly, we define \( P(v, t) = E Z_0(v, t) \) and \( P_\sigma(v, t) \) the contribution to \( P(v, t) \) from particles whose ancestral lines follow \( \sigma \). 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 \( \sigma \) 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 \( \sigma \) we have \( S_\sigma(v, t) = P_\sigma(v, t) = 0 \), and trivially \( B_\sigma(v, t) \geq 0 \).

Let us now fix \( \sigma \), a simple path from \( v_0 \) to \( v \). We denote the length of \( \sigma \) by \( l \). For any realization \( X \) of \( X_0 \) and \( x \in X \), let \( T(X, x) \) denote the birth time of \( x \). Then it follows from Lemma 3.2 that

\[
E \sum_{x \in V_\sigma(X_0)} 1_{T(X_0, x) \leq t} = \int_0^\infty \cdots \int_0^\infty 1_{z_1 + \cdots + z_l \leq t} dz_1 \cdots dz_l.
\]

In order to express \( B_\sigma \) and \( P_\sigma \) 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 \( b(X_0, x) \). Using this notation we have

\[
B_\sigma(v, t) = E \sum_{x \in V_\sigma(X_0)} 1_{T(X_0, x) \leq t} b(X_0, x)
\]

and

\[
P_\sigma(v, t) = E \sum_{x \in V_\sigma(X_0)} 1_{T(X_0, x) \leq t} 1_{b(X_0, x) = 0}.
\]

Hence, again by Lemma 3.2

\[
B_\sigma(v, t) = \int_0^\infty \cdots \int_0^\infty 1_{z_1 + \cdots + z_l \leq t} E [b(X^{z_1, \cdots, z_l}, x^{z_1, \cdots, z_l})] dz_1 \cdots dz_l
\]

and

\[
P_\sigma(v, t) = \int_0^\infty \cdots \int_0^\infty 1_{z_1 + \cdots + z_l \leq t} \mathbb{P} (b(X^{z_1, \cdots, z_l}, x^{z_1, \cdots, z_l}) = 0) dz_1 \cdots dz_l.
\]

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

\[
b(X^{z_1, \cdots, z_l}, x^{z_1, \cdots, z_l}) = \sum_{i=0}^l \sum_{x \in X_i} \varphi_i(x).
\]

As the original particles in \( X_0, \ldots, X_l \) correspond to ancestors of \( x^{z_1, \cdots, z_l} \), these are never counted in \( b \) and hence the corresponding indicator functions are always zero. Furthermore, as \( X_0, \ldots, X_l \) are independent processes, we have by Lemma 3.3

\[
\mathbb{P} (b(X^{z_1, \cdots, z_l}, x^{z_1, \cdots, z_l}) = 0) = \prod_{i=0}^l \mathbb{P} (\varphi_i(x) = 0 \forall x \in X_i)
\]

\[
\geq \prod_{i=0}^l \exp \left( -E \sum_{x \in X_i} \varphi_i(x) \right)
\]

\[
= \exp \left( -Eb(X^{z_1, \cdots, z_l}, x^{z_1, \cdots, z_l}) \right)
\]
Hence, by (3.19),
\begin{equation}
(3.22) \quad P_\sigma(v, t) \geq \int_0^\infty \cdots \int_0^\infty 1_{z_1, \ldots, z_t \leq t} \exp \left( -\mathcal{E}(X_{z_1, \ldots, z_t}, x_{z_1, \ldots, z_t}) \right) \, dz_1 \cdots dz_t.
\end{equation}

Let \( r_0 \in \mathbb{R} \) be fixed. By convexity we have \( e^{-r} \geq e^{-r_0} (1 + r_0) - e^{-r_0} r \). Applying this inequality to the integrand in (3.22) and comparing to (3.15) and (3.18) we get, for any simple path \( \sigma \),
\begin{equation}
(3.23) \quad P_\sigma(v, t) \geq e^{-r_0} (1 + r_0) S_\sigma(v, t) - e^{-r_0} B_\sigma(v, t).
\end{equation}

As remarked, for non-simple paths \( \sigma \) we have \( P_\sigma = S_\sigma = 0 \) and \( B_\sigma \geq 0 \), so clearly (3.23) holds for all paths \( \sigma \) from \( v_0 \) to \( v \). Summing this inequality over all such paths \( \sigma \), we get
\begin{equation}
(3.24) \quad P(v, t) \geq e^{-r_0} (1 + r_0) S(v, t) - e^{-r_0} B(v, t).
\end{equation}

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 \( P(v, t) \geq S(v, t) e^{-\frac{B(v, t)}{S(v, t)}} \) as desired.

\section*{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 \( \hat{0} \). We will accordingly let \( m(v, t) \) denote the expected number of particles at \( v \) at time \( t \) for a BTP originating at \( \hat{0} \), 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 \( \mathbb{Z}_n^d \), the \( n \)-fold group product of \( \mathbb{Z}_2 \), and we let \( e_1, e_2, \ldots, e_n \in \mathbb{Z}_n^d \) 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 \( \hat{0} \). 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 \( \mathbb{Z}_n^d \), we will sometimes make a formal distinction between addition and subtraction in order to indicate direction.

\begin{lemma}
For any \( t > 0 \) and \( v \in \mathbb{Q}_n \) we have
\begin{equation}
(4.1) \quad \frac{d}{dt} \frac{d}{dt} m(v, t) = \sum_{i=1}^n \sum_{j=1}^n m(v + e_i + e_j, t)
\end{equation}
and
\begin{equation}
(4.2) \quad \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).
\end{equation}
\end{lemma}

\begin{proof}
Recall that \( m(v, t) \) satisfies
\begin{equation}
(4.3) \quad \frac{d}{dt} m(v, t) = \sum_{i=1}^n m(v + e_i, t).
\end{equation}
This directly implies that
\begin{align*}
\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{align*}

The second equation now follows from \( \frac{1}{2} \frac{d^2}{dt^2} m(v, t)^2 = m''(v, t)m(v, t) + m'(v, t)m'(v, t) \).
\end{proof}
Lemma 4.2. Let \( s, t \geq 0 \) and \( v \in \mathbb{Q}_n \). Then
\[
(4.4) \quad \sum_{w \in \mathbb{Q}_n} m(w, s)m(v + w, t) = m(v, s + t).
\]

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
\[
(4.5) \quad \mathbb{E}[Z(v, s + t)|Z(s)] = \sum_{w \in \mathbb{Q}_n} Z(w, s)m(v + w, t).
\]
The lemma follows by taking the expected value of this expression.

We now turn to the problem of expressing \( A(\hat{1}, u) \) and \( B(\hat{1}, u) \) in terms of \( m(v, t) \). Fix \( u > 0 \) and let \( X \) be a BTP on \( \mathbb{Q}_n \) originating at \( \hat{0} \). Let \( T \) denote the random set of triples of particles \((x, y, z)\) in \( X \) such that
- \( x \) is located at \( \hat{1} \) 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 \( T_a \), the set of all such triples where \( y \) is a descendant of \( z \), and \( T_b \), the set of all such triples where \( y \) is not a descendant of \( z \). For any \( x \) at \( \hat{1} \) at time \( u \) in \( X \), it is clear that \( c(x) \) gives the number of triples in \( T \) where the first element is \( x \). Hence by summing \( c(x) \) over all particles at \( \hat{1} \) at time \( u \) we obtain the size of \( T \). Similarly we see that \( \sum_x a(x) \) and \( \sum_x b(x) \) where \( x \) goes over all particles \( x \) at \( \hat{1} \) at time \( u \) gives the size of \( T_a \) and \( T_b \) respectively. Hence \( A(\hat{1}, u) = \mathbb{E}[|T_a|], B(\hat{1}, u) = \mathbb{E}[|T_b|] \) and \( A(\hat{1}, u) + B(\hat{1}, u) = \mathbb{E}[|T|] \).

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

Proposition 4.3. For any \( u > 0 \) we have
\[
(4.6) \quad A(\hat{1}, u) = \sum_{v \in \mathbb{Q}_n} \sum_{i=1}^n \sum_{j=1}^n \int_0^\infty \int_0^\infty 1_{s+t\leq u} m(v, s)m(\hat{1} - v, u - s - t)m(e_j + e_i, t) \, ds \, dt
\]
\[
(4.7) \quad A(\hat{1}, u) + B(\hat{1}, u) = \sum_{v \in \mathbb{Q}_n} \sum_{w \in \mathbb{Q}_n} \sum_{i=1}^n \sum_{j=1}^n \int_0^\infty \int_0^\infty 1_{s+t\leq u} m(v, s)m(\hat{1} - w, u - s - t) \cdot \left( m(w - v, t)m(w - v - e_i + e_j, t) + m(w - v - e_i, t)m(w - v + e_j, t) \right) \, ds \, dt.
\]

Proof. Let us start by considering \( A(\hat{1}, u) \). For any \( (x, y, z) \in T_a \) 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 \( v \in \mathbb{Q}_n, 1 \leq i, j \leq n \), and infinitesimal time intervals \((s, s + ds]\) and \((s + t, s + t + dt]\) where \( 0 \leq s < s + t < u \). 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 \(\hat{1}\) at time \(u\) which is a descendant of \(y\) forms a triple in \(T_a\). By computing the expected number of potential particles in each step, we see that

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 \(v\). For each potential \(z\), \(c\), \(p\), \(y\) we note that we cannot have \(x\) first particle to be an ancestor of precisely one of \(x\) or \(z\). Here \(z\) is a child of \(z + e\) and \(c\) is born during \((s, s + ds]\) and \((s, s + t + dt]\). Note that \(c\) must be a child of \(l\). Similar to the case of \(T_a\), we note that we cannot have \(c = y\).

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 \(z\). 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 3.

We now fix \(v, w \in \mathbb{Q}_n\), \(1 \leq i, j \leq n\) and time intervals \((s, s + ds]\) and \((s, s + t + dt]\) 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

![Figure 1](image-url)

**Figure 1.** Illustration of the possible configurations of ancestral lines of elements in \(T_a\) and \(\mathcal{T}\) respectively. Graph 1 shows the configuration of elements in \(T_a\). Here \(z\) is located at \(v\), \(c\) is a child of \(z\) at \(v + e\), \(p\) is a descendant of \(c\) at \(v - e_j\), \(y\) is a child of \(p\) at \(v\), and \(x\) a descendant of \(y\) at \(\hat{1}\). 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 \(x\) or \(z\). As \(c = y\) is the first particle to be an ancestor of precisely one of \(x\) and \(z\), 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.
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 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 $m(w - e_j - v - e_i, t) dt m(\hat{1} - w, u - s - t)$, and the expected number of potential $z$'s is $m(w - v, t)$. Combining this, we see that the expected number of elements in $T$ corresponding to fixed $v, w, i, j$, fixed time intervals as above and where $c$ is an ancestor of $x$ is

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

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 $T$ is

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

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 $T_u$ and $T$). Hence this can directly be generalized to any loop-free graph.

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

**Proof.** By reordering the sums and integrals in (4.10) we have

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

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

$$\int_0^\infty \int_0^\infty \mathbb{1}_{s+t \leq u} m(\hat{1}, u - t) \frac{d^2}{dt^2} m(\hat{0}, t) ds dt = \int_0^u (u - t) m(\hat{1}, u - t) \frac{d^2}{dt^2} m(\hat{0}, t) dt,$$

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

$$A(\hat{1}, u) = \int_0^u (u - t) (\sinh(u - t))^n \frac{d^2}{dt^2} (\cosh t)^n dt$$

$$= \int_0^u (u - t) (\sinh(u - t))^n \left( n + n(n - 1) (\tanh t)^2 \right) (\cosh t)^n dt$$

$$= \int_0^u (u - t) \left( n + n(n - 1) (\tanh t)^2 \right) e^{nf(t)} dt,$$

where $f(t) := \ln (\sinh(u - t) \cosh t)$.

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

$$f'(t) = - \coth(u - t) + \tanh t$$

$$f''(t) = - \csch(u - t)^2 + \sech(t)^2.$$

As $\sech t \leq 1$ for all $t \in \mathbb{R}$ and $\csch t \geq 1$ for $0 < t < u$, it follows that $f''(t) < 0$ for $0 < t < u$. 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

$$A(\hat{1}, u) = \int_0^\infty \mathbb{1}_{z \leq nu} \left( u - \frac{z}{n} \right) \left( 1 + (n - 1) \tanh \left( \frac{z}{n} \right)^2 \right) e^{nf\left( \frac{z}{n} \right)} dz.$$
It is clear that the integrand is bounded for all \( n \) by \( u \left( 1 + C z^2 \right) e^{-\sqrt{2} s} \), which is integrable over \([0, \infty)\). Hence, by dominated convergence, it follows that

\[
A(\hat{1}, u) \to \int_0^\infty u e^{-\sqrt{2} z} \, dz = \frac{u}{\sqrt{2}} \quad \text{as} \quad n \to \infty.
\]

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

\[
A(\hat{1}, u) + B(\hat{1}, u) \to \frac{ue^u}{\sqrt{2}} + \frac{1}{3 - 2\sqrt{2}} \quad \text{as} \quad n \to \infty.
\]

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

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

\[
\frac{1}{2} \sum_{v \in \mathbb{Q}_n} \sum_{w \in \mathbb{Q}_n} \int_0^\infty \int_0^\infty 1_{s+t\leq u} m(v, s)m(\hat{1} - w, u - s - t) \frac{d^2}{dt^2} m(w-v, t)^2 \, ds \, dt.
\]

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

\[
\frac{1}{2} \sum_{v \in \mathbb{Q}_n} \sum_{\Delta \in \mathbb{Q}_n} \int_0^\infty \int_0^\infty 1_{s+t\leq u} m(v, s)m(\hat{1} - \Delta + v, u - s - t) \frac{d^2}{dt^2} m(\Delta, t)^2 \, ds \, dt,
\]

which by Lemma 4.2 simplifies to

\[
\frac{1}{2} \int_0^u (u - t) \sum_{\Delta \in \mathbb{Q}_n} m(\hat{1} - \Delta, u - t) \frac{d^2}{dt^2} m(\Delta, t)^2 \, dt.
\]

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

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

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

\[
\sum_{\Delta \in \mathbb{Q}_n} m(\hat{1} - \Delta, z) m(\Delta, t)^2 = \sum_{k=0}^n \binom{n}{k} (\sinh z)^k (\cosh z)^{n-k} (\sinh t)^{2n-2k} (\cosh t)^{2k}
\]

\[
= \sum_{k=0}^n \binom{n}{k} (\sinh z (\cosh t)^2)^k (\cosh z (\sinh t)^2)^{n-k}
\]

\[
= \left( \sinh z (\cosh t)^2 + \cosh z (\sinh t)^2 \right)^n
\]

\[
= \left( \frac{1}{2} e^z \cosh 2t - \frac{1}{2} e^{-z} \right)^n.
\]

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

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

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

we can write
\[ \text{(4.25)} \quad A(\hat{1}, u) + B(\hat{1}, u) = \frac{1}{2} \int_0^u (u - t) (n g(t) + n(n - 1) h(t)) e^{nf(t)} \, dt. \]

One can check that \( f(0) = f(u) = 0 \), \( f(\frac{1}{2}) < -\frac{1}{2} \), and that \( f \) has derivatives
\[ \text{(4.26)} \quad f'(t) = -1 + 2 \frac{\sinh 2t - e^{-2u+2t}}{\cosh 2t - e^{-2u+2t}} \]
and
\[ \text{(4.27)} \quad f''(t) = 4 \frac{1 - 2e^{-2u}}{(\cosh 2t - e^{-2u+2t})^2}. \]

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

To evaluate the integral in equation (4.25), we divide it into two integrals, one over the interval \([0, \frac{1}{2}]\), and one over \([\frac{1}{2}, u]\), that is into the two integrals
\[ \text{(4.28)} \quad \int_0^{\frac{1}{2}} (u - t) (n g(t) + n(n - 1) h(t)) e^{nf(t)} \, dt \]
\[ = \int_0^\infty 1_{z \leq \frac{1}{2}} \left( u - \frac{z}{n} \right) \left( g \left( \frac{z}{n} \right) + (n - 1) h \left( \frac{z}{n} \right) \right) e^{nf\left( \frac{z}{n} \right)} \, dz \]
and
\[ \text{(4.29)} \quad \int_\frac{1}{2}^u (u - t) (n g(t) + n(n - 1) h(t)) e^{nf(t)} \, dt \]
\[ = \int_0^\infty 1_{z \geq \frac{u-1}{n}} \left( \frac{1}{n} g \left( u - \frac{z}{n} \right) + \frac{n-1}{n} h \left( u - \frac{z}{n} \right) \right) e^{nf\left( u - \frac{z}{n} \right)} \, dz. \]

Now, using the convexity of \( f(t) \) it is a standard calculation to show that the integrands of these expressions are uniformly dominated by \( C(1 + t^2) e^{-\lambda t} \) and \( Cte^{-\lambda t} \) respectively, for appropriate positive constants \( \lambda \) and \( C \). Hence, by the Lebesgue dominated convergence theorem, these integrals converge to
\[ \text{(4.30)} \quad \int_0^\infty 2ue^{u+f'(0)z} \, dz = \frac{2ue^u}{-f'(0)} = \sqrt{2}ue^u \]
and
\[ \text{(4.31)} \quad \int_0^\infty (\sinh 2u)^2 e^{-f'(u)z} \, dz = \frac{8}{f'(u)^2} = \frac{2}{3 - 2\sqrt{2}} \]
respectively, as \( n \to \infty \). We conclude that
\[ \text{(4.32)} \quad A(\hat{1}, u) + B(\hat{1}, u) \to \frac{1}{2} \left( \sqrt{2}ue^u + \frac{2}{3 - 2\sqrt{2}} \right) \quad \text{as } n \to \infty. \]
5. Proof of Theorem 2.5

Throughout this section, we will use $u$ to denote $\ln(1 + \sqrt{2})$. In order to bound $\|T_n - u\|_p$, it is natural to treat the problems of bounding $T_n - u$ from above and below separately. To this end, we let $T_n^+$ and $T_n^-$ denote the positive and negative part of $T_n - u$ respectively, that is, $T_n^+$ is the maximum of $T_n - u$ and 0 and $T_n^-$ is the maximum of $u - T_n$ and 0. Hence, we can bound $\|T_n - u\|_p$ by $\|T_n^+\|_p + \|T_n^-\|_p$. We will begin by proving two simple propositions. The first shows that the fluctuations of $T_n$ are indeed of order $\Omega \left( \frac{1}{\sqrt{n}} \right)$, hence the variance of $T_n$ and the $L^p$-norm of $T_n - u$ for any $1 \leq p < \infty$ are $\Omega \left( \frac{1}{n^{1/p}} \right)$. 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 bound $\|T_n^+\|_p$.

**Proposition 5.1.** For any constant $c \in \mathbb{R}$, there exists an $\varepsilon_c > 0$ such that

\[(5.1) \quad \mathbb{P}\left( T_n \geq u + \frac{c}{n} \right) \geq \varepsilon_c - o(1).\]

**Proof.** Let $X$ be a BTP originating at $\hat{0}$. For any potential particle $x$, we define $\varphi(x)$ as the indicator function for the event that $x$ is located at $\hat{1}$ and born at time at most $u + \frac{c}{n}$. Note that $\sum_{x \in X} \varphi(x) = Z \left( \hat{1}, u + \frac{c}{n} \right)$, and hence

\[(5.2) \quad \mathbb{E} \sum_{x \in X} \varphi(x) = m \left( \hat{1}, u + \frac{c}{n} \right) = \left( 1 + \frac{\sqrt{2}c}{n} + O \left( \frac{1}{n^2} \right) \right)^n = e^{\sqrt{2}c} + o(1).\]

Applying Lemma 5.3, we conclude that the probability that $Z \left( \hat{1}, u + \frac{c}{n} \right) = 0$ is at least $e^{-e\sqrt{2}c} - o(1)$. The Proposition follows from the relation between the BTP and Richardson’s model. ■

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

**Proof.** We have

\[(5.3) \quad \mathbb{E} \left( (T_n^-)^p \right) = \mathbb{E} \int_0^\infty \mathbb{1}_{t \leq T_n} p t^{p-1} dt = \int_0^\infty p t^{p-1} \mathbb{P}(T_n \leq u - t) dt.\]

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

\[(5.4) \quad \mathbb{E} \left( (T_n^-)^p \right) \leq \int_0^u p t^{p-1} e^{-\sqrt{2}nt} dt = O \left( \frac{1}{n^p} \right).\]

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

Condition on the weights of all edges connected to either $\hat{0}$ or $\hat{1}$. We pick vertices $a_1$ and $a_2$ adjacent to $\hat{0}$ such that $W_{a_1}$ and $W_{a_2}$ have the smallest and second smallest edge weights respectively among all edges adjacent to $\hat{0}$. Among all $n - 2$ neighboring vertices of $\hat{1}$ 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 $W_{a_1} - W_{a_2}, W_{b_1}$, 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 $\hat{0}$ and $b_1$ and $b_2$ are adjacent to $\hat{1}$, 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 $Q_n$ consisting of all vertices $v \in Q_n$ 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 $Q_n$ consisting of all vertices $v \in Q_n$ 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 $Q_n$ 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 $\{W_e\}_{e \in E(Q_n)}$. As $H_1$ and $H_2$ are both isomorphic to $Q_{n-2}$, where $a_1$ and $b_1$ are antipodal in $H_1$ and $a_2$ and $b_2$ are antipodal in $H_2$, Corollary 2.1 implies that the corresponding first-passage times in each of $H_1$ and $H_2$ are at most $u$ 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 2.1. 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 $\varepsilon_2$ the first-passage time in $H'_2$ from $a_2$ to $b_2$ is at most $u$.

Now, let $\xi$ denote the indicator function for the event that the first-passage time from $a_2$ to $b_2$ in $H'_2$ is at most $u$. As $H_1$ is isomorphic to $Q_{n-2}$ 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 $\{W_e\}_{e \in E(Q_n)}$ such that $T_{n-2}$ denotes this quantity. Note that this means that $\xi$ 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 $\hat{0}$ 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 $\hat{1}$. Furthermore, if $\xi = 1$ we similarly see that $T_n \leq W_{a_2} + W_{b_2} + u$. Combining these bounds we see that for any $n \geq 4$ we have

$$T_n \leq \xi (W_{a_2} + W_{b_2} + u) + (1 - \xi) (W_{a_1} + W_{b_1} + T_{n-2}).$$

We may interpret this inequality as follows. We flip a coin $\xi$. If the coin turns up heads then $T_n$ is bounded by $u$ 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 $\xi$. 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 (5.5) to bound the $L^p$-norm of $T_n$. By subtracting $u$ and taking the positive part of both sides we get

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

As $W_{a_2} \geq W_{a_1}$ and $W_{b_2} \geq W_{b_1}$ we can replace $\xi (W_{a_2} + W_{b_2}) + (1 - \xi) (W_{a_1} + W_{b_1})$ in the right-hand side of (5.6) by $W_{a_2} + W_{b_2}$. Taking the $L^p$-norm of both sides we obtain the inequality

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

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 $\xi$ and $T_{n-2}$ are independent we have $\| (1 - \xi) T_{n-2}^+ \|_p \leq (1 - \varepsilon_2)^{\frac{1}{p}} \| T_{n-2}^+ \|_p$. Hence, for any fixed $p$ we have the inequality

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

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 $\|T_n - u\|_p = O \left( \frac{1}{n} \right)$, as desired.

It only remains to prove Proposition 5.3.

In the following argument, we will identify $H_2$ with $Q_{n-2}$ 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 $Q_{n-2}$ 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 $u$, 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 $X$ denote a BTP on $\mathbb{Q}_n$ originating at $\hat{0}$. 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 $\hat{0}$ with rate $n$, and for each $t \geq 0$ let $\sigma_t$ denote the path that the random walk has followed up to time $t$.

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

$$X_t(S') = \mathbb{P}(\sigma_t \in S' | y(t) = \hat{1}).$$

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

$$X_t(\{\sigma\}) = \mathbb{E} \sum_{x \in V_\sigma(X)} 1_{T(X,x) \leq t} = \int_0^\infty \cdots \int_0^\infty \mathbb{1}_{z_1 + \cdots + z_l \leq t} dz_1 \cdots dz_l = \frac{t^l}{l!},$$

where $T(X,x)$ denotes the birth time of $x$. In comparison, it is straightforward to see that

$$\mathbb{P}(\sigma_t = \sigma) = e^{-nt \frac{l}{n}}.$$  

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

$$X_t(S') = \frac{\mathbb{P}(\sigma_t \in S')}{\mathbb{P}(\sigma_t \in S)} = \mathbb{P}(\sigma_t \in S' | y(t) = \hat{1}),$$

as desired. \hfill \blacksquare

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

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

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 $y(u) = \hat{1}$, each coordinate $y_i(t)$ is an independent simple random walk on $\{0,1\}$ conditioned on the event that $y_i(u) = 1$. It is easy to see that the expected number of steps taken by such a process up to time $u$ is

$$e^{-u} u + \frac{u^3}{3!} + \frac{u^5}{5!} + \cdots = u \coth u = \sqrt{2} u.$$  

The lemma follows by the law of large numbers. \hfill \blacksquare

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

19
As $H_2$ is graph isomorphic to $Q_{n-2}$, we know from Corollary 2.4 that, with probability bounded away from zero, there exists an uncontested particle in $X$ at $b_2$ at time $u$. Furthermore, by Lemma 5.3 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 $X$ at $\hat{1}$ at time $u$ whose ancestral line is of length at most $1.25(n - 2)$. As a path from $\hat{0}$ to $\hat{1}$ must traverse edges in each of the $n - 2$ directions of $Q_{n-2}$ 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 $\hat{1}$ at time $u$ in $X'$. The proposition follows from the fact that Richardson’s model stochastically dominates the set of uncontested particles in a BTP.

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 and Johan Tykesson 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.
[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(2):593–629, 1993.
[3] Béla Bollobás and Yoshiharu Kohayakawa. *On Richardson’s model on the hypercube*. In Combinatorics, geometry and probability (Cambridge, 1993), pages 129–137. Cambridge Univ. Press, Cambridge, 1997.
[4] Peter Hegarty and Anders Martinsson. *On the existence of accessible paths in various models of fitness landscapes*. To appear in Ann. Appl. Probab., arXiv:1210.4798 [math.PR]
[5] David Aldous. *Probability approximations via the Poisson clumping heuristic*, volume 77 of Applied Mathematical Sciences. Springer-Verlag, New York, 1989.
[6] Rolf Schneider and Wolfgang Weil. *Stochastic and integral geometry*. Probability and its Applications (New York). Springer-Verlag, Berlin, 2008.