RANDOM WALK ON BARELY SUPERCRITICAL BRANCHING RANDOM WALK

REMCO VAN DER HOFSTAD, TIM HULSHOF, AND JAN NAGEL

Abstract. Let $T$ be a supercritical Galton-Watson tree with a bounded offspring distribution that has mean $\mu > 1$, conditioned to survive. Let $\varphi_T$ be a random embedding of $T$ into $\mathbb{Z}^d$ according to a simple random walk step distribution. Let $T_p$ be percolation on $T$ with parameter $p$, and let $p_c = \mu - 1$ be the critical percolation parameter. We consider a random walk $(X_n)_{n \geq 1}$ on $T_p$ and investigate the behavior of the embedded process $\varphi_{T_p}(X_n)$ as $n \to \infty$ and simultaneously, $T_p$ becomes critical, that is, $p = p_n \searrow p_c$. We show that when we scale time by $n/(p_n - p_c)^3$ and space by $\sqrt{(p_n - p_c)/n}$, the process $(\varphi_{T_p}(X_n))_{n \geq 1}$ converges to a $d$-dimensional Brownian motion. We argue that this scaling can be seen as an interpolation between the scaling of random walk on a static random tree and the anomalous scaling of processes in critical random environments.

MSC 2010. Primary: 60K37, 82C41. Secondary: 60F17, 60K40.
Key words and phrases. Branching random walk, random walk indexed by a tree, percolation, scaling limit, supercriticality.

1. Introduction

Percolation has long stood as a simple and tractable model for random media in physics. Starting with the infinite integer lattice $\mathbb{Z}^d$ with nearest-neighbor edges, each edge is kept independently with probability $p \in (0, 1)$. The resulting random subgraph already has many properties of real random media, see e.g. [Sta85, Sah94]. Moreover, the model has a rich mathematical structure. Its main feature is the existence of a phase transition: there exists a $p_c(\mathbb{Z}^d) \in (0, 1)$ such that when $p > p_c$ the model has an infinite connected component or cluster, while it does not when $p < p_c$. The behavior for $p$ close to and at the critical point $p_c$, moreover, has many remarkable features.

Random walk on percolation clusters of the lattice $\mathbb{Z}^d$ at or near the critical point has been a central model in both physics and modern probability ever since de Gennes proposed it more than forty years ago [dG76] (see [BAF16] for an overview). Computer simulations and non-rigorous studies suggest that the model, which de Gennes dubbed "the ant in the labyrinth", has many intriguing features, such as the observation that random walk on large critical clusters exhibits anomalous diffusion. This fact has since been rigorously verified [Kes86, KN09]. Yet many questions still remain unanswered. For instance, almost all rigorous results we have are either for the two-dimensional lattice or when the dimension is "high enough" (depending on the problem this is

Date: April 13, 2018.
typically above either six or ten dimensions [HS90, FvdH17]). Dimensions three to
six are terra incognita. And even in high dimensions, where the picture is perhaps
most complete, there are still many big open problems. For instance, it is currently
unknown what the scaling limit of random walk on large critical clusters is (although
there are good conjectures [HS00a, Sla02], on which much progress has been made
recently [BACF16a, BACF16b]).

It is, however, broadly believed that percolation is just one model in a much larger
universality class. In high dimensions, this class contains many more tractable models,
such as lattice trees and lattice animals, oriented percolation, the contact process, and,
perhaps most tractable of all, branching random walk (see [Sla06] for an overview).
There is a wealth of evidence that around the critical point these models all behave in
largely the same way on large scales.

One subject that is now, after a long line of research, rather well understood for
percolation in any dimension, is that random walk on the infinite supercritical percolation
cluster (i.e., when \( p > p_c \)) has a Brownian scaling limit when the temporal and
spatial scaling factors are \( n \) and \( 1/\sqrt{n} \), respectively, just as for random walk on
\( \mathbb{Z}^d \) [KV86, DMFGW85, DMFGW89, SS04, BB07, MP07]. That is, writing
\( (\tilde{X}_n)_{n \geq 0} \) for a random walk on an infinite percolation cluster \( C_\infty \) at supercriticality, one obtains a scaling limit of the form
\[
(n^{-1/2} \tilde{X}_{[tn]})_{t \geq 0} \xrightarrow{d} (\bar{\sigma}(p) B_t)_{t \geq 0}.
\]

The diffusion constant \( \bar{\sigma}(p) \), which describes the typical fluctuations of this Brownian
motion, is not known but believed to depend on \( p \). Remarkably, it is at this time not
even known whether \( \bar{\sigma}(p) \) is monotonically increasing for \( p > p_c \). There is, however,
clear evidence that the above scaling limit cannot hold for (an infinite version of) critical
clusters, because there it is known that \( (n^{-1/6} |\tilde{X}_n|)_{n \geq 1} \) is a tight sequence of random
variables [HvdHH14]. This suggests that either \( \bar{\sigma}(p) \to 0 \) as \( p \searrow p_c \), or that there exists
a discontinuity for \( \bar{\sigma}(p) \).

In the current paper we investigate this barely supercritical regime for a different
model that we believe to be in the same universality class, namely for random walk on
a branching random walk (defined below). Here, we show that \( \sigma(p) = O(p - p_c) \)
for \( p \) slightly above \( p_c \). The tree structure of the branching random walk even allows
for a precise analysis of the process as \( p \searrow p_c \), which we believe to be representative
of a random walk on a barely supercritical percolation cluster. We establish a scaling
limit for the first \( n \) steps of a random walk, when, as \( n \to \infty \), the underlying branching
random walk becomes more and more critical, and \( p = p_n \searrow p_c \). We show that the
Brownian scaling remains observable in that limit, if we apply the temporal and spatial
scaling factors \( n/(p - p_c)^3 \) and \( \sqrt{(p - p_c)/n} \), respectively.

1.1. Random walk on a randomly embedded random tree. Before we proceed
with our main results, let us define the model.

Let \( T \) be a Galton-Watson tree rooted at \( \varrho \) with law \( \mathbf{P} \), and let \( \xi \) be the random
number of offspring of the root. Suppose that the tree is supercritical, i.e., \( \mu := \mathbf{E}[\xi] > 1. \)
We consider percolation on the edges of the tree and let $T_p$ denote the connected component of the root $\varrho$ in $T$, when each edge is deleted with probability $p \in (0,1)$, independently of every other edge and of $T$. Then $T_p$ is again a Galton-Watson tree, whose distribution we denote by $P_p$. The root of $T_p$ now has $\xi_p$ offspring, such that, conditioned on $\{\xi = n\}$, $\xi_p$ is a binomial random variable with parameters $n$ and $p$.

The percolated tree $T_p$ has offspring mean $p\mu$ and is supercritical if and only if $p > p_c = 1/\mu$. In this setting, denote the distribution of $T_p$, conditioned on non-extinction by $P_p$. Given a tree $T$ rooted at $\varrho$, let $(X_n)_{n \geq 0}$ be a simple random walk on $T$ started at $\varrho$. Given $v \in T$, we write $P_T^v$ for the law of $(X_n)_{n \geq 0}$ with $X_0 = v$, and write $P_T$ if $v = \varrho$. Given a realization of the percolated Galton-Watson tree $T_p$, we call $P_{\varrho}$ the quenched law of the random walk, and we call $P_p := P_p \times P_{\varrho}$ the annealed law of the random walk on the tree (writing $P_p^v := P_p \times P_{\varrho}^v$, with the convention that $P_p^v = \delta_{\{X_0 = v\}_{n \geq 0}}$ if $v \notin T$).

We embed the tree $T_p$ into $\mathbb{Z}^d$ by means of a branching random walk, which we will define now: Let $D$ denote a non-degenerate probability distribution on $\mathbb{Z}^d$. Given a tree $T$ rooted at $\varrho$, set $Z(\varrho) = 0$ and assign to each vertex $v \neq \varrho$ of $T_p$ an independent random variable $Z(v)$ with law $D$. For any $v \in T$ there exists a unique path $\varrho = v_0, v_1, \ldots, v_m = v$. The branching random walk embedding $\varphi = \varphi_T$ is defined such that $\varphi(v) := Z(v_0) + \cdots + Z(v_m)$. If we write $(Y_n)_{n \geq 1}$ for a random walk on $\mathbb{Z}^d$ with step distribution $D$ started at 0, then two vertices $v, w \in T$ with the unique path $v = v_0, v_1, \ldots, v_k = w$ between them are mapped such that $(\varphi(v_0), \ldots, \varphi(v_k))$ has the same law as the translation of $(Y_1, \ldots, Y_k)$ by $\varphi(v_0)$, i.e., the marginal of the branching random walk embedding of any simple path in $T$ is a random walk path. This type of branching random walk is also sometimes referred to in the literature as “random walk indexed by a tree”.

The process that we consider in this paper is that of $(\varphi_T(X_i))_{1 \leq i \leq n}$ under $P_p$ as $p_n \searrow p_c$, i.e., the first $n$ steps of an (annealed) random walk on a barely supercritical infinite tree, embedded by a branching random walk $\varphi$ into $\mathbb{Z}^d$.

In our main results, Theorems 1.1 and 1.2 below, we show how the speed of the random walk changes as $p_n \searrow p_c$, as well as the correct scaling to obtain a Brownian motion.

The main challenges in this paper are due to the behavior of random walk on a barely supercritical tree. Taking the embedding into account is more straightforward, and is mainly done so that we may give a more geometric interpretation to the process, allowing for a comparison to high-dimensional percolation. (In a certain sense the embedding even simplifies some computations.)

To understand the problem of random walk on a barely supercritical BRW, and to give a clearer context to our main results, let us first discuss the behavior of random walk on a BRW with $p > p_c$ fixed. For simplicity, let us also consider for the moment

\footnote{Note that this is a different process than if we were to consider a simple random walk on the subgraph of $\mathbb{Z}^d$ traced out by the branching random walk, as is for instance the topic of [BACF16a] (for a different kind of trees): random walk on the trace is, in our setting, a vacuous complication, because $T_p$ is supercritical, and thus grows at an exponential rate, while $\mathbb{Z}^d$ has polynomial growth, so that $\varphi_T(T_p) = \mathbb{Z}^d P_p$-almost surely.}
that $D$ is the nearest-neighbor step distribution, i.e., it is the uniform distribution on $\pm e_1, \ldots, \pm e_d$, where $e_i$ is the $i$-th unit vector of $\mathbb{Z}^d$.

Given $v \in T$, let $|v|$ denote the distance of a vertex $v$ to the root. It is a well-known result [LPP95] that the effective speed

$$v(p) := \lim_{n \to \infty} \frac{|X_n|}{n}$$

exists $\mathbb{P}_p$-almost surely and is non-zero when $p > p_c$. Moreover, it is shown in [LPP95] that there exists an infinite sequence of regeneration times $(\tau_k)_{k \geq 1}$ such that $|X_n| < |X_{\tau_k}|$ for any $n < \tau_k$ and $|X_n| > |X_{\tau_k}|$ for any $n > \tau_k$, $\mathbb{P}_p$-almost surely. These regeneration times decompose the trajectory of the random walk into independent increments with good moment bounds. Since $\varphi(X_{\tau_k+1}) - \varphi(X_{\tau_k})$ are then independent as well, and this increment is given by the displacement of the embedding random walk after $\tau_k + 1 - \tau_k$ steps, standard arguments allow us to conclude the convergence

$$(n^{-1/2} \varphi(X_{\lfloor tn \rfloor}))_{t \geq 0} \overset{d}{\underset{n \to \infty}{\to}} (\sigma(p) B_t)_{t \geq 0}$$

in distribution under $\mathbb{P}_p$, with $(B_t)_{t \geq 0}$ a standard $d$-dimensional Brownian motion. (The proof of Theorem 1.2 in Section 5 below also establishes (1.2)). The variance of the limit is determined by

$$\sigma^2(p) := \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_p[\|\varphi(X_n)\|^2] = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_p[|X_n|] = v(p),$$

where $\| \cdot \|$ denotes the 2-norm on $\mathbb{Z}^d$. So we know that $v(p) > 0$ for any fixed $p > p_c$. It is not known what happens to $v(p_n)$ when $p_n \searrow p_c$. The first result of this paper is that we establish that $v(p_n) \to 0$, and at what rate it does so.

To this end, we define a few more pieces of notation. To simplify what is to come, we frequently omit $p$ or $p_n$ from subscripts when they are clear from context, e.g., we write $\mathbb{P}$ instead of $\mathbb{P}_p$. For $k \in \mathbb{N}$, define the $k$-th factorial moment of the (unpercolated) offspring distribution as

$$m_k := \mathbb{E} \left[ \prod_{i=0}^{k-1} (\xi - i) \right].$$

Our first main result is a limit for the speed of the random walk on $T_p$ as $p \searrow p_c$:

**Theorem 1.1** (The speed of random walk on a barely supercritical tree). Consider a random walk $(X_n)_{n \geq 1}$ on a Galton-Watson tree with $m_1 > 1$ and $m_3 < \infty$. For any sequence $p_n \searrow p_c$,

$$\frac{v(p_n)}{(p_n - p_c)^2} \overset{n \to \infty}{\longrightarrow} \frac{m_1^2}{3m_2} =: \kappa.$$

We prove this theorem in Section 11 by analyzing the Taylor expansion of the extinction probability in terms of the generating function.

Our second main result is a scaling limit for the random walk on a branching random walk as $p \searrow p_c$. For this we regard $(\varphi(X_{\lfloor tn \rfloor}))_{t \geq 0}$ as a random element in the space $\mathcal{D}_{\mathbb{R}^d}[0, \infty)$, endowed with the Skorokhod topology and the Borel $\sigma$-algebra.
We write $\Delta$ for the maximal number of children that a single individual in the (unpercolated) tree can have, i.e.,
$$\Delta := \sup\{ n : \Pr(\xi = n) > 0 \}.$$

**Theorem 1.2.** Consider random walk $(X_n)_{n \geq 0}$ on a Galton-Watson tree with $m_1 > 1$ and $\Delta < \infty$, embedded into $\mathbb{Z}^d$ with $d \geq 1$ by $\varphi$ whose one-step distribution $D$ satisfies $\sum_{x \in \mathbb{Z}^d} x D(x) = 0$ and $\sum_{x \in \mathbb{Z}^d} e^{\|x\|} D(x) < \infty$ for some $c > 0$. For any sequence $p_n \searrow p_c$ satisfying $e^{\delta \sqrt{n}} (p_n - p_c) \to \infty$ for any $\delta > 0$ as $n \to \infty$,
$$\left( \sqrt{\frac{p_n - p_c}{n}} \varphi(X_{\lfloor tn(p_n-p_c)^{-1} \rfloor}) \right)_{t \geq 0} \xrightarrow{n \to \infty} ((\kappa \Sigma)^{1/2} B_t)_{t \geq 0}$$
under $\mathbb{P}_{p_n}$, with $(B_t)_{t \geq 0}$ a standard Brownian motion in $\mathbb{R}^d$, $\Sigma$ the covariance matrix of $D$, and $\kappa$ as in Theorem 1.1 above.

The rescaling of time and space in Theorem 1.2 can be seen as an interpolation between the scaling for random walk on a static supercritical tree and, on the other hand, in a critical random environment. Indeed, for $p = p_n$ fixed, the scaling is the same as in (1.2), while in the barely supercritical regime $p_n \searrow p_c$ the additional factors $(p_n - p_c)$ bring the scaling close to that on a critical tree. We discuss the latter point in more detail in Section 1.2.

Heuristically, the fact that rescaling space by $\sqrt{(p - p_c)/n}$ and time by $n/(p - p_c)^3$ yields a Brownian limit can be understood as follows: The duality principle tells us that any infinite supercritical GW-tree with $\Pr(\xi = 0) > 0$ can be decomposed into a supercritical GW-tree with no leaves (which we call the “backbone tree”) to which, at each vertex, are attached a random number of i.i.d. subcritical GW-trees (which we call “traps”), see Section 2 below for details. Percolating the tree to $p$ close to $p_c$ (but conditioning on survival) makes the backbone tree become thinner, while the traps become bigger (relative to the full tree) as $p$ comes closer to the critical point. Because the growth rate of the backbone tree is equal to that of the full (percolated) tree (see Remark 2.2 below), and because the growth rate of the full tree is proportional to $1 + O((p - p_c)^{-1})$, the furcations in the backbone tree must be separated by path-like segments with lengths of order $(p - p_c)^{-1}$. Simple random walk on a GW-tree with no leaves is transient because each time the walk reaches a furcation, it is more likely to take a step away from the root than towards it. If the walk in $n$ steps is likely to see a number of furcations that grows with $n$, then a limit of the form (1.2) can be expected. For this, however, we would need to rescale the tree by $(p - p_c)/n$ to accommodate for the low growth rate of the tree. For the BRW embedding this means we need to rescale space by $\sqrt{(p - p_c)/n}$.

If the walk were restricted to the backbone tree, then it would take the walk order $(p - p_c)^{-2}$ steps to visit the next furcation, because that is how long it takes for a random walk to travel distance $O((p - p_c)^{-1})$ on a line. We can view the walk as moving on the backbone tree and at each step having the option of walking into a trap. Between two furcations, we can expect a tight number of large traps of size $O((p - p_c)^{-2})$, each visited $O((p - p_c)^{-1})$ times. An application of electrical network theory for Markov
chains tells us that the time to exit such a large trap is of order \((p - p_c)^{-2}\). The time spent in traps between furcations thus accumulates to \(O((p - p_c)^{-3})\). Hence, to see a scaling limit like (1.2), we need to rescale time by a factor \(n/(p - p_c)^3\).

**Remark 1.3** (About the assumptions in Theorem 1.2). (1) We only assume that there exists a finite maximal degree \(\Delta\) to simplify the proofs below. We believe that our proof can be modified (albeit with some lengthy computations) to admit any offspring distribution whose generating function \(f(s) = E[s^\xi]\) has derivatives that satisfy \(f^{(n+1)}(s) \leq Cnf^{(n)}(s)\) for all \(n \geq 1\) and \(0 \leq s \leq 1\) (we do not know if this condition is necessary). Examples of distributions that satisfy this condition are Poisson and Geometric.

(2) The assumption on \(D\) that \(\sum_{x \in \mathbb{Z}^d} xD(x) = 0\) is necessary, otherwise the BRW would have a drift, which would require a different analysis and yield a different limit.

(3) The assumption that \(\sum_{x \in \mathbb{Z}^d} e^{c\|x\|}D(x) < \infty\) prevents the BRW from making very large jumps. The strength of the assumption is for technical reasons. We believe that our result should remain valid when \(\sum_{x \in \mathbb{Z}^d} \|x\|^a D(x) < \infty\). It is known that the behavior of BRW (and other statistical mechanical models) alters dramatically when this restriction is relaxed further, see e.g. [JM05, HvdHS08, CS15]. We believe that when there instead exists an \(\alpha < 2\) such that \(\alpha = \sup\{a : \sum_{x \in \mathbb{Z}^d} \|x\|^a D(x) < \infty\}\), then a limit can still be achieved, but the limiting process would be an \(\alpha\)-stable motion instead.

(4) We require that \((p_n)_{n \geq 1}\) satisfies that for any \(\delta > 0\) we have \(e^{\delta \sqrt{n}}(p_n - p_c) \to \infty\), so that we may apply a result of Neuman and Zheng regarding the maximal displacement of subcritical BRW [NZ17], which we crucially use to estimate the size of the traps after embedding into \(\mathbb{Z}^d\).

### 1.2. Towards anomalous scaling

As mentioned, one of the main motivations for studying random walk on a slightly supercritical branching random walk is that we believe it to be in the same universality class as random walk on high-dimensional percolation (and various other important models from statistical mechanics). Supporting evidence for this belief is that it appears to be true at criticality. In particular, Aldous [Ald93] proved that if we consider \((\mathcal{T}_n, \varphi_n)_{n \geq 1}\) to be branching random walks indexed by critical Galton-Watson trees conditioned to have \(n\) vertices, then there exists a scaling limit for the trace of this BRW, the Integrated Super Brownian Excursion, or ISE. The ISE is now understood to be a canonical random object. It can be viewed as a Brownian embedding of a continuum random tree of a fixed size (cf. [Ald91]), as the scaling limit of high-dimensional lattice trees, cf. [DS98] (a model understood to be in the universality class of percolation in high dimensions), and it has many other deep connections in spatial probability and statistical mechanics, cf. [LG05]. Moreover, its two-point function coincides with that of the scaling limit of large critical percolation clusters in high dimensions [HS00a, HS00b]. There are several other connections between statistical mechanics models at criticality and critical BRWs, depending on the precise conditioning (see e.g. [Hol08, HHP17] or [HvdH17, Section 15.1]).

Not only is there evidence that critical BRW has a scaling limit that is in the right universality class, there is also evidence that random walk on a critical BRW also has a universal scaling limit. Croydon [Cro09] showed that the critical version of our model, namely a random walk \((Y_m)_{m \geq 1}\) on a sequence of critical trees \(\mathcal{T}_n\) of size \(n\) randomly
embedded by \( \varphi_n \) has the scaling limit

\[
\left( n^{-1/4} \varphi_n(Y_{tn^{3/2}}) \right)_{t \geq 0} \xrightarrow{d} \left( \sqrt{\kappa} B^{ISE}_t \right)_{t \geq 0},
\]

where \( B^{ISE} \) is a Brownian motion on the ISE (an object constructed in [Cro09]), and \( \kappa \) and \( \nu \) are constants. Note the “anomalous” temporal and spatial scaling factors \( n^{3/2} \) and \( n^{-1/4} \). Recently, Ben-Arous, Cabezas, and Fribergh [BACF16b] have shown that the same scaling limit holds in high dimensions (with different \( \kappa \) and \( \nu \)) for the model where the random walk is on the trace of the critical BRW (rather than on the tree). This latter model is thus a true random walk in a random environment on \( \mathbb{Z}^d \).

Moreover, in [BACF16a] they identify four conditions that give rise to a scaling limit as in (1.4), and they conjecture that these four conditions are valid for all models in the universality class of high-dimensional percolation.

Roughly speaking, since our model considers random walk on “increasingly critical” infinite trees, whereas the above result considers random walk on “increasingly infinite” critical trees, we think it is a natural question to ask whether the two scaling limits somehow connect. In relation to this, let us observe the following: If we rescale time in (1.4) by setting \( b_n = n^{3/2} \), we obtain instead

\[
\left( b_n^{-1/6} \varphi_n(Y_{tb_n}) \right)_{t \geq 0} \xrightarrow{d} \left( \sqrt{\kappa} B^{ISE}_t \right)_{t \geq 0},
\]

so the scaling factors are \( b_n \) and \( b_n^{-1/6} \). We can also change the rescaling in Theorem 1.2 by replacing both instances of \( n \) by \( a_n \) and then set \( b_n = a_n(p_n - p_c)^{-3} \). The time rescaling is then \( b_n \) and the spacial factor becomes \( a_n^{-2/3} b_n^{-1/6} = o(1) b_n^{-1/6} \). Note that the assumption in Theorem 1.2 on \( p_n \) then require \( b_n \) to not grow too fast (compared to \( a_n \)). An equivalent representation of the convergence of Theorem 1.2 then reads

\[
\left( a_n^{-2/3} b_n^{-1/6} \varphi(X_{[tb_n]}) \right)_{t \geq 0} \xrightarrow{d} \left( \kappa \Sigma \right)^{1/2} B_t \]

So by letting \( a_n \) tend to infinity arbitrarily slowly, we “almost” obtain the anomalous scaling factors of (1.4) from Theorem 1.2 (but the limit remains a Brownian motion, of course). We require that \( a_n \) tends to infinity because then the RW moves on such a large time scale that it can visit a growing number of furcation points, i.e., it can “feel” the drift that arises from the fact that the barely supercritical tree is in fact supercritical. It would be interesting to study the behavior of the random walk on time scales where \( a_n \) does not tend to infinity, but this would require a very different approach than the current paper takes.

1.3. The structure of the proof of Theorem 1.2. We start by stating some preliminary lemmas on the structure of slightly supercritical trees in Section 2. We conclude that the typical length scale of the trees is \( (p - p_c)^{-1} \), and introduce a rescaling by this length that we use throughout the paper. In Section 3, we state some further lemmas with preliminary bounds on escape probabilities of random walk on such trees. The proofs of these lemmas can be found in Section 6.

In Section 4 we define a new regeneration structure. This regeneration structure is based on that of [LPP96], but is specifically designed to account for the length rescaling of the tree. In particular, our regeneration structure is inspired by the regenerations
of [GMP12], and allows for a limited amount of backtracking that is related to how close \( p_n \) is to \( p_c \). This backtracking makes it harder to recover the independence structure that we need. We state moment bounds on the regeneration distances in Lemma 4.4 (which we prove in Section 7), and moment bounds on regeneration distances distances in Lemma 4.5 (which we prove in Section 8).

In Section 5 we give the main steps of the proof of Theorem 1.2. Utilizing the decoupling effect of the regenerations, we first prove that the limit holds if we consider the increments of the walk between regeneration times. For this we need the moment bounds on the regeneration distances of Lemma 4.4, and also bounds on the inter-regeneration times of Lemma 4.5. Because this process does not take into account where the random walk is between regeneration times, and the regeneration times by definition never occur in the traps of the tree, the final step of the proof is to show that the random walk is not able to walk great distances in traps. For this we use a recent result by Neuman and Zheng [NZ17] on the maximal displacement of subcritical BRW, which yields a bound on the maximal displacement of the traps in Lemma 5.1. We also need bounds on the size and shape of the trace of the backbone of the BRW, in Lemma 5.2 (which we prove in Section 10).

1.4. New and interesting aspects of the proofs. The proof of Theorem 1.1 in Section 11 is a straightforward Taylor expansion of a generating function for the effective speed of simple random walk on a supercritical GW-tree derived in [LPP95], applied to our setting. It is, however, interesting to note that in two seemingly far removed parts of the proof we see a term involving the third moment of \( \xi \) arise, but these terms cancel perfectly, leaving us with an expression of the asymptotic speed that only involves the first two moments. It is unclear to us why the third moment should drop out like this.

The proof of Theorem 1.2 is, at its core, close to the classical proofs of Brownian limits for random walk in random environment. In particular, we construct a sequence of regeneration times to find (almost) independent increments, and follow the standard approach from there.

What is new about our proof is that we require all the necessary moment bounds to hold uniformly for \( p \in (p_c, p_d) \) for some \( p_d > p_c \). To obtain such uniform bounds involves a careful analysis of the structure of the backbone tree. In particular, we need to take into account that furcations occur only on a length scale of order \( (p - p_c)^{-1} \), which means that in the neighborhood of a given furcation, the random walk typically behaves diffusively as \( p \searrow p_c \). In this setting, we were not able to construct regeneration times \( (\tilde{\tau}_n) \) such that, uniformly in \( p \), we have (1) that the paths of the past and of the future of the walk at \( \tilde{\tau}_n \) do not intersect, and (2) that \( (\tilde{\tau}_{n+1} - \tilde{\tau}_n) \) is an i.i.d. sequence. In fact, we believe that such regeneration times do not occur on the desired time scale of order \( n/(p - p_c)^3 \). Instead we construct a sequence of regeneration times \( (\tau_n) \) that allows the random walk to backtrack a short distance shortly after a regeneration time, to increase the density of the regeneration times to the right scale. This relaxation comes at the cost of (1) having a small and localized intersection between the past and future of the walk at \( \tau_n \), and (2) making \( (\tau_{n+1} - \tau_n) \) a stationary and 1-dependent sequence, rather than i.i.d. (but these are not serious complications). From this new
construction of regeneration times we are then able to derive all the moment bounds on
the regeneration times and distances uniformly for \( p \in (p_c, p_d) \).

Besides controlling the regeneration structure, we also need exponentially tight control
over the displacement of the random walk in the traps (Lemma 5.1), and a bound on all
moments of the size of the trace of the random walk on the backbone tree (Lemma 5.2).
Uniformity for \( p \) is again an important requirement here.

2. Preliminaries: the shape of the tree

In this section we establish some facts about the percolated trees, such as their growth
rate. We also discuss a useful decomposition of the tree.

To simplify notation we frequently drop the subscript \( p \). We write \( c \) and \( C \) for generic
constants whose value may change from line to line. Central bounds appearing in the
a-priori estimates in this section are denoted by \( a \) and \( c \). The constants \( a, c, C, \) and \( c \)
may depend on the original offspring distribution and on \( d \), but we stress that they are
independent of \( p \). Since we are interested in \( p \) close to \( p_c \), we will state many results
only for \( p \in (p_c, p_d) \) for some unspecified \( p_d > p_c \).

Given a tree \( T \) rooted at \( \varrho \) and \( v, w \in T \), we say \( w \) is a descendant of \( v \) if the unique
path from \( \varrho \) to \( w \) passes through \( v \), and we call \( w \) an ancestor of \( v \) if the path from
\( \varrho \) to \( v \) passes through \( w \). We say a vertex \( v \) has an infinite line of descent if \( T_v \) is an
infinite tree. We can decompose any infinite tree \( T \) into its backbone tree \( T^{Bb} \), the tree
induced by all vertices with an infinite line of descent, and the forest \( T \setminus T^{Bb} \), which
consists only of finite trees. Given \( v \in T^{Bb} \), we call the connected component of \( v \) in
\( T \setminus (T^{Bb} \setminus \{v\}) \) the trap at \( v \) and write \( T_v^{trap} \).

Lyons [Lyo92, Proposition 4.10] gives an explicit description of this decomposition
for Galton-Watson trees using the “duality principle” (cf. [AN72, Chapter 12]). For \( f \) a
generating function of an offspring distribution of a Galton-Watson tree with extinction
probability \( q \in (0, 1) \), set

\[
\hat{f}(s) := \frac{f((1-q)s + q) - q}{1-q} \quad \text{and} \quad f^*(s) := \frac{f(qs)}{q}.
\]

A supercritical Galton-Watson tree \( T \) with generating function \( f \) conditioned on survival
can be generated by sampling a tree \( T^{Bb} \) with generating function \( \hat{f} \), and then adding
every vertex \( v \in T^{Bb} \) a random number \( U_v \) of edges, and to the other ends of those edges
independent Galton-Watson trees \( T_v^*, i \in \{1, \ldots, U_v\} \) with generating function \( f^* \).
While \( T^{Bb} \) has no leaves, the trees generated by \( f^* \) go extinct with probability 1. The
distribution of \( U_v \) can be characterized as follows: for \( v \in T^{Bb} \) write \( \delta_v = \deg_{T^{Bb}}(v) - 1 \)
and write \( f^{(n)}(s) \) for the \( n \)-th derivative of \( f(s) \). Then,

\[
E[s^{U_v}] = \frac{f^{(\delta_v)}(qs)}{f^{(\delta_v)}(q)}.
\]

Then \( T_v^{trap} \) is the random subtree consisting of \( v \), the \( U_v \) edges attached to \( v \), and the
finite trees generated by \( f^* \) attached to the edges.
The distribution of the traps, however, is such that $P$ (or non-existent), while a tight number of them are macroscopically large, having a size edge has been replaced by a path whose length is distributed as an independent geometric constants of this heuristic directly in the proofs that follow, we do use them often as guiding

$$f_p'(0) \geq c_0, \quad \mu_p = \mu_p' = f_p'(1) = p/p_c,$$  \hspace{1cm} (2.3)

and

$$q_p = 1 - c_1(p - p_c)(1 + o(1)), \quad \mu_p'' = f_p''(q_p) = 1 - c_2(p - p_c)(1 + o(1)), \quad \hat{f}_p''(0) = c_3(p - p_c)(1 + o(1)).$$  \hspace{1cm} (2.4)

The proof of this lemma is standard and can be found in Section 11.

Remark 2.2 (The topology of $T_{p_n}$, its backbone tree and its traps as $p_n \searrow p_c$). It is shown in [AN72, Lemma 10.1] that the ratio of backbone tree vertices to all vertices in generation $n$ of $T$ tends to $1 - q$, the survival probability, as $n \to \infty$. Writing $\mu, \hat{\mu},$ and $\mu^*$ for the mean number of offspring of trees with generating functions $f, \hat{f},$ and $f^*$, respectively, we can easily compute that $\mu = f'(1) = \hat{\mu}$, and $\mu^* = f'(q)$. Writing $(\hat{p}_k)_{k \geq 0}$ for the offspring distribution of the backbone tree, a simple computation moreover shows that $\hat{p}_0 = 0$, $\hat{p}_1 = f'(q)$, and $\hat{p}_2 = (1 - q)\hat{f}''(q)/2$.

From the above lemma we can infer that, heuristically speaking, $T_p$ (conditioned on survival) and its decomposition look as follows for $p$ close to $p_c$: Both $T_p$ and $T_p^{Bb}$ grow at rate $f'_p(1) = p/p_c$, but only a fraction $p - p_c$ of the vertices in $T_p$ is contained in $T_p^{Bb}$. By a similar computation to the bounds above, we can show that $\sum_{k \geq 3} f_p^{(k)}(0) = o(p - p_c)$, so we are unlikely to see any vertices in $T_p^{Bb}$ with out-degree three or more up to a height of order $(p - p_c)^{-1}$. Up to this height, the tree looks like a binary tree where each edge has been replaced by a path whose length is distributed as an independent geometric random variable with a parameter of order $p - p_c$. The trap at any vertex $v$ consists of a random number of i.i.d. subcritical Galton-Watson trees with mean offspring distribution $\mu^*_v \approx c_1(p - p_c)$, whose expected size and depth are both known to be of order $(p - p_c)^{-1}$. The distribution of the traps, however, is such that $\mathbb{P}(|T_v^{trap}| \geq A(p - p_c)^{-2}) \approx \frac{p - p_c}{A} e^{-A/2}$, which implies that among $O((p - p_c)^{-1})$ typical traps, almost all traps will be very small (or non-existent), while a tight number of them are macroscopically large, having a size and depth of the order of $(p - p_c)^{-2}$. Although we do not use any of the computations of this heuristic directly in the proofs that follow, we do use them often as guiding principles.

For a tree $T$ with backbone tree $T^{Bb}$, let $G'_m(T) = \{ v \in T^{Bb} : |v| = m \}$ be the backbone-tree vertices in generation $m \geq 0$. As discussed above, the relevant spatial scaling factor of the random tree will be of order $p - p_c$, so throughout this paper we will often consider the backbone tree only at generations that are multiples of $L/(p - p_c)$, for some $L \geq 1$ to be determined later. We call these generations the $L$-levels of $T$, and for $m \in \mathbb{N}$ write

$$G_{[m]} := G_{[m]}(T) := G'_{[Lm/(p - p_c)]}(T)$$  \hspace{1cm} (2.5)
for the $m$-th $L$-level, and write $G_{[m]} := \{\emptyset\}$. We tacitly ignore the dependency on $L$ in the definition of $G_{[m]}$, the reason being that we soon fix the value of $L$. Roughly speaking, the choice of $L$ will be such that with good probability, the random walk starting from a vertex $v \in G_{[m]}$ does never hit $G_{[m-1]}$. For $m < n$, let

$$T^{Bb}_{[m,n]} := \{v \in T^{Bb} : |Lm/(p - p_c)| \leq |v| \leq |Ln/(p - p_c)|\} \quad (2.6)$$

denote the forest segment of the backbone tree between the $m$-th and $n$-th $L$-level. Let

$$T_{[m,n]} := \bigcup_{v \in T^{Bb}_{[m,n]}} T^{\text{trap}}_v \quad (2.7)$$

denote the same forest segment with all the traps attached.

The next lemma shows that on the scale of a single $L$-level, the backbone tree $T^{Bb}$ looks like a non-degenerate tree:

**Lemma 2.3** (The size of the first $L$-level). For any $L > 0$ there exists $p_d > 0$ and constants $0 < a_1 \leq a_2 < 1$ and $a_3 < \infty$, such that for all $p \in (p_c, p_d)$

$$a_1 \leq \P(|G_{[1]}(T)| = 1) \leq a_2 \quad \text{and} \quad \E[|G_{[1]}(T)|] \leq a_3.$$

**Proof.** By (2.1) we know that the backbone tree $T^{Bb}$ is a Galton-Watson tree with generating function $\hat{f}(s)$, so the probability that a vertex in the backbone tree has exactly one child in $T^{Bb}$ is $\hat{p}_1 = f'(q_p)$. By Lemma 2.1, $f'(q_p) = 1 - c_2(p - p_c)(1 + o(1))$ as $p \searrow p_c$. Therefore,

$$\lim_{p \searrow p_c} \P(|G_{[1]}(T)| = 1) = \lim_{p \searrow p_c} (1 - c_2(p - p_c)(1 + o(1)))^{L/(p - p_c)} = e^{-c_2L}. \quad (2.8)$$

This proves the first part of the lemma.

For the second part, recall that $T_p$ has offspring mean $\mu_p = \mu p$ and that $p_c = 1/\mu$, so that

$$\E[|G_{[1]}(T)|] = \mu_p^{L/(p - p_c)} \leq (1 + \mu(p - p_c))^{L/(p - p_c)},$$

which is uniformly bounded as $p \searrow p_c$. \qed

---

**3. Preliminaries: escape probabilities**

In this section we establish three useful bounds on the probability that the random walk escapes to infinity before returning to the previous $L$-level.

For $A$ a set of vertices of $T$, let $\eta(A) = \inf\{n \geq 0 : X_n \in A\}$ be the hitting time of the set $A$. For $m \geq 0$, we denote by $\eta_m := \eta(G_{[m]}(T))$ the hitting time of the $m$th $L$-level. A crucial step in defining the regeneration times is the following uniform bound for the probability that the random walk backtracks more than $L/(p - p_c)$ in tree distance:

**Lemma 3.1** (Annealed escape probability). There exists an $L_0 \geq 1$ depending only on the original offspring distribution, such that for any $L \geq L_0$, for any tree $T$ with $G_{[1]}(T) \neq \emptyset$ and any $v \in G_{[1]}(T)$,

$$\P^v(\eta_0 = \infty | T_{[0,1]} = T_{[0,1]}) \geq \frac{1}{3}.$$
We prove this lemma in Section 6.1.

In what follows, fix an $L$ such that the estimate in Lemma 3.1 holds.

It would simplify our argument substantially if we could get a quenched version of Lemma 3.1. Unfortunately, such a quenched version does not hold. We do, however, have the following statement, which shows that the backtracking probabilities are small with a high probability, even when additionally delete all outgoing edges at the starting point $v \in G_{[1]}(T)$, so that the walker has to escape to infinity via a different vertex in $G_{[1]}(T)$:

**Lemma 3.2** (Quenched indirect escape probability). There exists a function $h$ independent of $p$ with $h(\alpha) \to 0$ as $\alpha \to 0$, such that

$$\mathbb{E}\left[ \sum_{v \in G_{[1]}(T)} 1\{p_{v \setminus \gamma_c}(\eta_0 = \infty) < \alpha\} \right] \leq a_2 + h(\alpha),$$

with $a_2$ as in Lemma 2.3.

We prove this lemma in Section 6.2.

For a vertex $v \in G_{[m]}(T)$ with $m \geq 1$, let $\text{anc}(v) \in G_{[m-1]}(T)$ denote the $L$-ancestor of $v$, its ancestor in $L$-level $m - 1$, and define the number of $L$-siblings of $v$ as

$$\text{sib}(v) := |G_{[1]}(T_{\text{anc}(v)})| - 1. \quad (3.1)$$

Note that $\text{anc}(v)$ and $\text{sib}(v)$ are only defined for vertices on the backbone tree, and that $\text{sib}(v)$ only counts the siblings on the backbone tree. The following bound shows that we have a uniformly bounded probability of hitting vertices that have no $L$-siblings:

**Lemma 3.3** (Escape probability on thin parts of the backbone tree). There exists a constant $a_4 > 0$ such that for any tree $T$ with $G_{[1]}(T) \neq \emptyset$ and any $v \in G_{[1]}(T)$,

$$\mathbb{P}^v(\eta_2 < \eta_0, \text{sib}(X_{\eta_2}) = 0 \mid T_{[0,1]} = T_{[0,1]}) \geq a_4.$$

We prove this lemma in Section 6.3.

4. The regeneration structure

Regeneration times are a classical tool to decouple the increments of random walks in random environments. For random walks on GW-trees, they were utilized already in the papers [LPP95, LPP96] or [PZ08]. We follow in our definition the formulation of [SZ99]. There are two main changes from classical regeneration time structure of the above mentioned papers: First, we have to allow the random walk to backtrack a distance of order $(p - p_c)^{-1}$, similar to the construction in [GMP12, GGN17]. Second, to obtain a stationary sequence even with backtracking, we have to control the environment where the walker regenerates. For the first point, we can rely on Lemma 3.1 to see that we have a good probability of not backtracking too far. For the second point, we want to ensure that a regeneration point has no siblings, such that the random walk can backtrack only on a branch where the backbone has no furcations.

Intuitively, the regeneration times are constructed as follows: We wait until the random walker for the first time reaches a vertex with the potential for regeneration,
namely a vertex in an $L$-level with no siblings. We call this time $S_1$ and the associated $L$-level $M_1$. If the previous $L$-level $M_1 - 1$ is never visited again, then we call $S_1$ the first regeneration time $\tau_1$. But suppose that the walker does revisit $L$-level $M_1 - 1$ again at time $R_1$. Then $S_1$ is not a regeneration time. Instead, we denote by $N_1$ the highest $L$-level that the walker visited between $S_1$ and $R_1$ on the backbone tree, and we wait until the walker first reaches an $L$-level with no siblings with generation greater than $N_1$. We call this potential regeneration time $S_2$. If $S_2$ is never visited again, then we set $\tau_1 = S_2$. If it is revisited again, then we repeat the above procedure. Because the walk on the tree is transient, the first regeneration time $\tau_1$ is finite almost surely. We repeat the entire procedure to construct the sequence $(\tau_k)_{k \geq 1}$ of regeneration times. See Figure 1 for a sketch of this construction.

To define the above construction formally, we start with some more notation. Let $(z_n)_{n \geq 0}$ be an infinite path, and let $\theta_n$ be the time shift on the path such that $(\theta_m z_n)_{n \geq 0} = z_{m+n}$ and such that for any function $f$ that takes an infinite path as its argument, $f \circ \theta_m$ denotes the same function applied to the time-shifted path. Define the backtracking time $\eta' := \inf\{ n \geq 0 : |X_n| = |X_0| - (L/(p-p_c)) \}$. We need $\eta' = \infty$ to regenerate.

**Definition 4.1.** Given a tree $T$ and a random walk $(X_n)_{n \geq 0}$ on $T$, we define a sequence of stopping times $S_1 \leq R_1 \leq S_2 \leq R_2 \leq \ldots$ and distances $M_k, N_k$, beginning with

\[
M_1 := \min\{ m \geq 1 : \text{sib}(X_{\eta_m}) = 0 \}, \quad S_1 := \eta_{M_1},
\]

\[
R_1 := S_1 + \eta' \circ \theta_{S_1}, \quad N_1 := \max\{ m \geq 1 : \eta_m < R_1 \},
\]

and recursively, for $k \geq 1$,

\[
M_{k+1} := \min\{ m \geq N_k + 2 : \text{sib}(X_{\eta_m}) = 0 \}, \quad S_{k+1} := \eta_{M_{k+1}},
\]

\[
R_{k+1} := M_{k+1} + \eta' \circ \theta_{S_{k+1}}, \quad N_{k+1} := \max\{ m : \eta_m < R_{k+1} \}.
\]

These definitions make sense until for some $k$ we have $R_k = \infty$ and we set

\[
K := \inf\{ k : R_k = \infty \}, \quad \tau_1 := S_K.
\]

We call $\tau_1$ the first regeneration time. By Lemma 3.1, $K < \infty$ almost surely so that $\tau_1$ is well-defined. We set $\tau_0 := 0$ and for $k \geq 2$ we define the subsequent regeneration times as

\[
\tau_k := \tau_{k-1} + \tau_1 \circ \theta_{\tau_{k-1}}.
\]

Finally, denote by $\pi_k := \eta(\text{anc}(X_{\tau_k}))$ the times when the $L$-ancestors of the regeneration points are visited for the first time, and by $\Lambda_k$ the $L$-generation at which the $k$-th regeneration time occurs, i.e., $\Lambda_k := |X_{\tau_k}|/(L/(p-p_c))$.

We can make the following observations.

- $(X_k)_{n \leq \pi_k}$ and $(X_k)_{n \geq \tau_k}$ visit disjoint parts of the tree.
- The only part of the environment that is visited by both $(X_k)_{n \leq \tau_k}$ and $(X_k)_{n \geq \tau_k}$ is the tree segment $T_{X_{\tau_k}} \setminus T_{X_{\pi_k}}$, consisting of the tree rooted at $X_{\pi_k}$ with $X_{\tau_k}$ and all descendants of the latter vertex removed.
Figure 1. An idealized sketch of the sample path of the random walk (in red) on a portion of the backbone tree (in black). The labels correspond to hitting times of vertices, $L$-levels are indicated by dashed lines. The hitting time $\eta_1$ of the first $L$-level is not the first potential regeneration time $S_1$, since $\text{sib}(X_{\eta_1}) = 1$. Observe that $\eta_2$ is the first potential regeneration time $S_1$ because $X_{\eta_1}$ has no $L$-siblings, but that the random walk then backtracks more than one $L$-level. Finally, $\eta_4 = S_2$ satisfies the condition for a regeneration time and $\pi_1 = \eta_3$.

The definition above allows for the random walk to move distance at most $\lfloor L/(p - p_c) \rfloor$ towards the root after a regeneration time. This is allowed because by construction we have $\text{sib}(X_{\tau_k}) = 0$, so we know that $\mathcal{T}_{X_{\tau_k}}^{B_b} \setminus \mathcal{T}_{X_{\tau_k}}^{B_b}$ is isomorphic to a line graph of length $\lfloor L/(p - p_c) \rfloor$, and because the inter-regeneration distances $|X_{\tau_k} - X_{\tau_{k-1}}|$ are independent of the traps, we may conclude that the inter-regeneration distances are independent and, with the exception of $k = 1$, they are identically distributed. (These first three observations are formalized in Lemma 4.2 below.)

The inter-regeneration times $\tau_k - \tau_{k-1}$ are not independent, since two subsequent time intervals both depend on the traps in the tree segment $\mathcal{T}_{X_{\sigma_{k-1}}} \setminus \mathcal{T}_{X_{\tau_k}}$. The inter-regeneration times are, however, stationary and 1-dependent. (This is formalized in Lemma 4.3 below.)

To state the above observations generally and unambiguously, we introduce the $\sigma$-fields

$$G_k := \sigma(\mathcal{T} \setminus \mathcal{T}_{X_{\pi_k}}, (X_n)_{0 \leq n < \pi_k}, \pi_k)$$

and the time shift $\tilde{\theta}_m$ defined for a set

$$B = \{ \mathcal{T} \in B_1, (X_n)_{n \geq 0} \in B_2 \}$$
by
\[ B \circ \bar{\theta}_k = \{ T_{X_{\bar{\theta}_k}} \in B_1, (X_n - X_{\tau_k})_{n \geq \tau_k} \in B_2 \}. \]

The following two lemmas are standard, so we omit their proofs (see e.g. [Guo16, Lemma 17 and Proposition 18]).

**Lemma 4.2** (Stationarity of the tree and walk). For any measurable set \( B = \{ T \in B_1, (X_n)_{n \geq 0} \in B_2 \} \) and \( k \geq 1 \),
\[ P(B \circ \bar{\theta}_k | G_k) = P^v(B | G_1(T) = \{ v \}, \eta_0 = \infty). \]

**Lemma 4.3** (Stationarity and 1-dependence of the regeneration times). Under \( P \), the sequence
\[ (T_{X_{\bar{\theta}_k}} \backslash T_{X_{\bar{\theta}_{k+1}}}, (X_n - X_{\tau_k})_{\tau_k \leq n < \tau_{k+1}}, \tau_{k+1} - \tau_k)_{k \geq 1} \]
is stationary and 1-dependent. Furthermore, the marginal distribution of this sequence is given by
\[ P(T_{X_{\bar{\theta}_k}} \backslash T_{X_{\bar{\theta}_{k+1}}} \in B_1, (X_n - X_{\tau_k})_{\tau_k \leq n < \tau_{k+1}} \in B_2, \tau_{k+1} - \tau_k \in B_3 | G_1(T) = \{ v \}, \eta_0 = \infty). \]

The moment bounds for the regeneration distances and the regeneration times in the two following lemmas are crucial ingredients for the proof of our main result:

**Lemma 4.4** (Moment bounds on regeneration distances). For any \( a \geq 1 \), there exists a finite constant \( C_a \) such that
\[ (p - p_c)^a E[|X_{\tau_1}|^a] \leq C_a, \quad (p - p_c)^a E[(|X_{\tau_2}| - |X_{\tau_1}|)^a] \leq C_a, \]
for any \( p > p_c \).

We prove this lemma in Section 7 below. Note that we also have the trivial lower bound \( |X_{\tau_1}| \geq c(p - p_c)^{-1} \).

**Lemma 4.5** (Moment bounds on regeneration times). There exists a constant \( C \) such that
\[ (p - p_c)^6 E[(\tau_1)^2] \leq C, \quad (p - p_c)^6 E[(\tau_2 - \tau_1)^2] \leq C, \]
for any \( p > p_c \). Furthermore, there exists a constant \( c_0 > 0 \), such that
\[ (p - p_c)^3 E[(\tau_2 - \tau_1)] \geq c_0. \]

We prove this lemma in Section 8 below.

**Remark 4.6** (Robust bounds on the effective speed near criticality). Given the moment bounds of Lemmas 4.4 and 4.5, we may apply the Law of Large Numbers to obtain an expression for the effective speed in terms of regeneration times, as
\[ v(p) = \lim_{n \to \infty} \frac{|X_n|}{n} = \lim_{n \to \infty} \frac{|X_{\tau_n}|}{\tau_n} = \frac{E[|X_{\tau_2}| - |X_{\tau_1}|] }{E[\tau_2 - \tau_1]}. \]
In particular, the $p$-independence of the moment bounds implies

$$0 < \liminf_{p_n \searrow p_c} (p_n - p_c)^2 v(p_n) \leq \limsup_{p_n \searrow p_c} (p_n - p_c)^2 v(p_n) < \infty.$$  (4.5)

This gives the order of the effective speed $v(p)$ close to criticality. The result of Theorem 1.1 is stronger and shows that limit inferior and limit superior in (4.5) agree, but the proof of Theorem 1.1 in Section 11 below relies on an explicit formula for the effective speed of SRW on GW-trees obtained in [LPP95], while the above arguments are quite robust against changes to tree or the behavior of the random walk.

5. The scaling limit: proof of Theorem 1.2

In this section we prove Theorem 1.2, subject to the proofs of Theorem 1.1, the Lemmas in Sections 2, 3, and 4, and subject to the proofs of two further lemmas, Lemmas 5.1 and 5.2, which are stated below as we need them.

The proof of Theorem 1.2 will go in four steps. In the first two, we consider the scaling limit of a simple random walk whose jumps have the same size distribution as the increments of the random walk on the BRW at regeneration times, but with jumps at a fixed rate. We show that under a rescaling equivalent to the one of Theorem 1.2, this process converges to a Brownian motion with diffusion $(\kappa \Sigma)^{1/2}$ as desired. Then, we will apply a time change to have the jumps occur at random times with the same distribution as the regeneration times and show that the difference with the process of the first step vanishes in the limit. In the final step, we show that the trajectories of the random walk during excursions between regeneration times also vanishes in the limit, thus yielding the scaling limit for the process that we are after.

We assume in this section some familiarity with the theory of convergence for Markov processes. We refer the reader who is insufficiently knowledgeable about this topic to [Bil99, EK86], where all convergence-related topics of this section that are not explicitly cited are defined and discussed.

To start, let us recall that in the setting of Theorem 1.2 we consider $p = p_n$ with $p_n \searrow p_c$ as $n \to \infty$. Therefore, the distribution of the random walk changes with $n$. We will thus indicate this dependency on $n$ clearly in this section.

For the first step of the proof we define the process

$$W_{t,n}^{(1)} := \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} W_k$$

as a random element of the Skorokhod space $D_{\mathbb{R}^d}[0,T]$. Without loss of generality, we take $T \in \mathbb{N}$ and then, rescaling $n$ linearly and using the scale invariance of the limiting Brownian motion, may restrict our proof further to the case $T = 1$. The increments of $W_{t,n}$ are given by

$$W_1 := \frac{\varphi(X_{\tau_1})}{\sqrt{(p_n - p_c)^2 \mathbb{E}_{p_n}[\tau_1]}}, \quad W_k := \frac{\varphi(X_{\tau_k}) - \varphi(X_{\tau_{k-1}})}{\sqrt{(p_n - p_c)^2 \mathbb{E}_{p_n}[\tau_2 - \tau_1]}} \quad \text{for } k \geq 2.$$

By Definition 4.1 and Lemma 4.2, the $W_k$ are independent and $W_2, W_3, \ldots$ are identically distributed. Using that conditioned on $(X_n)_{n \geq 0}$, $\varphi(X_{\tau_k}) - \varphi(X_{\tau_{k-1}})$ is the
This process jumps \( \nu \) times in \( [0, 1] \), the step distribution \( D \) after \( |X_{\tau_k} - X_{\tau_{k-1}}| \) steps, so by the assumption in Theorem 1.2 that \( \sum_{x \in \mathbb{Z}^d} xD(x) = 0 \), we get \( \mathbb{E}[W_k] = 0 \) and by Lemma 4.4 and Lemma 4.5 we get

\[
\mathbb{E}_{p_n} [||W_k||^2] = \frac{\mathbb{E}_{p_n} [|X_{\tau_k} - X_{\tau_{k-1}}|] \mathbb{E}_{p_n} [|Y_1|^2]}{(p_n - p_c)^2 \mathbb{E}_{p_n} [\tau_2 - \tau_1]} \leq C. \tag{5.1}
\]

Writing \( v^T \) for the transpose of \( v \), we also have the convergence

\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E}_{p_n} \left[ \left( \sum_{k=1}^n W_k \right) \left( \sum_{k=1}^n W_k \right)^T \right] = \lim_{n \to \infty} \frac{\mathbb{E}_{p_n} [|X_{\tau_k}|] \mathbb{E}_{p_n} [Y_1 Y_1^T]}{n(p_n - p_c)^2 \mathbb{E}_{p_n} [\tau_2 - \tau_1]} \mathbb{E}_{p_n} [Y_1 Y_1^T] \quad \text{by Lemma 4.5}, \tag{5.2}
\]

where in the last step we used Theorem 1.1 and the assumption in Theorem 1.2 that \( \sum_{x \in \mathbb{Z}^d} ||x||^2 D(x) < \infty \), so that \( \Sigma \) exists. The Invariance Principle [Bil56] implies that because \( \mathbb{E}[W_k] = 0 \) and (5.1) and (5.2) hold, we may conclude the convergence

\[
(W_{t,n}^{(1)})_{t \in [0,1]} \xrightarrow{n \to \infty} ((\kappa \Sigma)^{1/2} B_1)_{t \in [0,1]}, \tag{5.3}
\]

under \( \mathbb{P}_{p_n} \), with \((B_t)_{t \geq 0}\) a standard Brownian motion on \( \mathbb{R}^d \). This concludes the first step.

The process \( W_{t,n}^{(1)} := (W_{t,n}^{(1)})_{t \in [0,1]} \) is a piecewise constant function in \( t \), jumping exactly \( n \) times in \( [0, 1] \). The second step is to construct a process that will also jump at regular intervals, but has a random number of jumps, determined by the random number of regeneration times we see in \( n \) steps. Given \( n \geq 0 \), let \( k_n \) be the integer satisfying

\[
\tau_{k_n} \leq n < \tau_{k_n+1}, \tag{5.4}
\]

where we recall \( \tau_0 = 0 \). Set \( \nu_n := (p_n - p_c)^{-3} k_n \) and define

\[
W_{t,n}^{(2)} := W_{t,\nu_n}^{(1)}. \tag{5.5}
\]

This process jumps \( \nu_n \) times and we next show that this number of jumps is asymptotically equal to

\[
a_n := n(p_n - p_c)^{-3} \mathbb{E}_{p_n} [\tau_2 - \tau_1]^{-1}. \tag{5.6}
\]

Note that by Lemma 4.5, \( a_n \to \infty \). Now, by definition of \( k_n \),

\[
a_n = \frac{n}{\nu_n} \leq \frac{\tau_{k_n+1}}{k_n \mathbb{E}_{p_n} [\tau_2 - \tau_1]} \leq \frac{1}{k_n \mathbb{E}_{p_n} [\tau_2 - \tau_1]} \sum_{m=1}^{k_n+1} (\tau_m - \tau_{m-1}). \tag{5.7}
\]
The right hand side of (5.7) converges to 1 by the Weak Law of Large Numbers, showing that \( \limsup_{n \to \infty} a_n / \nu_n \leq 1 \) in probability under \( \mathbb{P}_{p_n} \). Bounding \( n \geq \tau_{k,n} \), the same argument shows \( \liminf_{n \to \infty} a_n / \nu_n \geq 1 \), which implies \[
abla \frac{a_n}{\nu_n} \mathbb{P}_{p_n} \to 1. \tag{5.8}
\]

We may then apply [Bil99, Theorem 14.4] to conclude that \( \mathcal{W}_{t,n}^{(2)} \) has the same limit as \( \mathcal{W}_{t,n}^{(1)} \), that is, (5.3) with \( \mathcal{W}_{t,n}^{(1)} \) replaced by \( \mathcal{W}_{t,n}^{(2)} \). This concludes the second step.

The process \( \mathcal{W}_{t,n}^{(2)} := (\mathcal{W}_{t,n}^{(2)})_{t \in [0,1]} \) still jumps with equal intervals, at times \( t = \frac{1}{\nu_n}, \frac{2}{\nu_n}, \ldots, 1 \). The third step of the proof is to consider instead the process \( \mathcal{W}_{n}^{(3)} := (\mathcal{W}_{t,n}^{(3)})_{t \in [0,1]} \), which has the same increments as \( \mathcal{W}_{n}^{(2)} \), but jumps at times \( t = \tau_{1}/\nu_n, \tau_{2}/\nu_n, \ldots, 1 \). That is,

\[
\mathcal{W}_{t,n}^{(3)} := \mathcal{W}_{k/\nu_n,n}^{(3)} = \frac{\varphi(X_{\tau_k})}{\sqrt{\nu_n(p_n-p_c)^2 \mathbb{E}_{p_n}[\tau_2 - \tau_1]}} \quad \text{for} \quad \tau_k \leq \lfloor t \nu_n \rfloor < \tau_{k+1},
\]

so \( \mathcal{W}_{n}^{(3)} \) is a random time change of \( \mathcal{W}_{n}^{(2)} \). We can bound their distance in the Skorokhod metric \( d_S \) as

\[
d_S(\mathcal{W}_{n}^{(3)}, \mathcal{W}_{n}^{(2)}) \leq \max_{k=1,\ldots,\nu_n} \frac{\left| \frac{\tau_k}{\nu_n} - \frac{k}{\nu_n} \right|}{\nu_n} \leq \max_{k=1,\ldots,\nu_n} \frac{\left| \tau_k - \mathbb{E}_{p_n}[\tau_k - \tau_1] \right|}{\nu_n} + \max_{k=1,\ldots,\nu_n} \frac{\left| \mathbb{E}_{p_n}[\tau_k - \tau_1] - \frac{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1]}{\nu_n} \right|}{\nu_n}. \tag{5.9}
\]

By Lemma 4.3 we have \( \mathbb{E}_{p_n}[\tau_k - \tau_1] = k \mathbb{E}_{p_n}[\tau_2 - \tau_1] \), so we may write the second term as

\[
\max_{k=1,\ldots,\nu_n} \left| \frac{\mathbb{E}_{p_n}[\tau_k - \tau_1]}{\nu_n} - \frac{k}{\nu_n} \right| = \left| \frac{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1]}{\nu_n} - 1 \right|.
\]

Using again the Weak Law of Large Numbers and the moment bounds of Lemma 4.5,

\[
\frac{\tau_{\nu_n}}{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1]} = \frac{\tau_1}{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1]} + \frac{1}{\nu_n} \sum_{k=2}^{\nu_n} \frac{\tau_k - \tau_{k-1}}{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1]} \mathbb{P}_{p_n} \to 1. \tag{5.10}
\]

This shows that the second term in (5.9) vanishes in probability.

We write the first term in (5.9) as

\[
\max_{k=1,\ldots,\nu_n} \frac{\tau_1}{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1]} + \frac{(\tau_k - \tau_1) - \mathbb{E}_{p_n}[\tau_k - \tau_1]}{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1]} \left| \frac{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1]}{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1]} \right|^{-1}.
\]

The second factor converges in probability to 1 by (5.10). That the first factor also converges to 1 in probability follows from Lemma 4.3 combined with an Invariance Principle for triangular arrays of 1-dependent random variables due to Chen and Romano [CR99, Theorem 2.1]. This implies that (5.9) vanishes in probability and thus (5.3) also holds with \( \mathcal{W}_{n}^{(1)} \) replaced by \( \mathcal{W}_{n}^{(3)} \). This concludes the third step.
The fourth and last main step of the proof is now to transfer the convergence to

\[
W_{n}^{(4)} := \frac{\varphi(X_{\lfloor t\tau_{n} \rfloor})}{\sqrt{\nu_{n}(p_{n} - p_{c})^{2}E_{p_{n}}[\tau_{2} - \tau_{1}]}}
\]

i.e., to the full (rescaled) process of random walk on a BRW, where we now also incorporate the fluctuations of \((X_{n})_{n \geq 0}\) between regeneration times. Note that by definition of \(a_{n}\) and \(\nu_{n}\) and by (5.8), we have

\[
\frac{1}{\sqrt{\nu_{n}(p_{n} - p_{c})^{2}E_{p_{n}}[\tau_{2} - \tau_{1}]}} = \sqrt{\frac{a_{n}(p_{n} - p_{c})}{\nu_{n}n}} \leq C\sqrt{\frac{p_{n} - p_{c}}{n}},
\]

so bounding

\[
d_{S}(W_{n}^{(4)}, W_{n}^{(3)}) \leq C\sqrt{\frac{p_{n} - p_{c}}{n}} \max_{k = 1, \ldots, \nu_{n}} \max_{1 \leq i < \tau_{k}} \|\varphi(w) - \varphi(X_{\tau_{k-1}})\|,
\]

it suffices to show that the right-hand side vanishes in probability. Again by the moment bounds of Lemma 4.5, \(\nu_{n} \leq Cn\) with probability converging to 1, so that we may replace the remaining \(\nu_{n}\) in (5.11) by \(Cn\). We need a lemma that controls the maximal displacement of the embedded traps. As we show in Section 9, this is a direct consequence of a precise estimate by Neuman and Zheng [NZ17] for the maximal displacement of subcritical branching random walks.

Recall that in the decomposition of \(T\) into the backbone tree and the traps given in Section 2, for any \(v \in T_{\text{Bb}}\) we have that the law of \(T_{\text{trap}}^{v}\) only depends on \(v\) through \(\deg_{T}(v)\), and that we assumed in Theorem 1.2 that \(\deg_{T}(v) \leq \Delta + 1\) almost surely.

**Lemma 5.1 (Maximal displacement inside traps).** Under the assumptions of Theorem 1.2, there exists a \(\gamma > 0\) and \(C < \infty\), both independent of \(p\), such that

\[
\max_{1 \leq \delta \leq \Delta - 1} E\left[\exp\left(\gamma\sqrt{p - p_{c}} \sup_{w \in T_{\text{trap}}^{v}} \|\varphi(w) - \varphi(v)\|\right) \right] \deg_{T_{\text{trap}}^{v}}(v) = \delta \leq C. \quad (5.12)
\]

We prove this lemma in Section 9 below.

We also need to bound the number of visited backbone-tree vertices. For this we write

\[
\text{BBT}_{k} := \{X_{\tau_{k-1}}, X_{\tau_{k-1} + 1}, \ldots, X_{\tau_{k}}\} \cap T_{\text{Bb}},
\]

i.e., \(\text{BBT}_{k}\) is the trace of \((X_{n})_{\tau_{k-1} \leq n \leq \tau_{k}}\) restricted to the backbone tree.

**Lemma 5.2 (Size of and maximal distance from \(\text{BBT}_{k}\)).** Under the assumptions of Theorem 1.2, for any \(a \geq 1\) there exists a \(C_{a}\) such that for any \(k \geq 1\)

\[
E[|\text{BBT}_{k}|^{a}] \leq C_{a}(p - p_{c})^{-a}, \quad (5.13)
\]

and

\[
E[\max_{v \in \text{BBT}_{k}} \|\varphi(v) - \varphi(X_{\tau_{k-1}})\|^{a}] \leq C_{a}(p - p_{c})^{-a/2}. \quad (5.14)
\]
We prove this lemma in Section 10 below.

Using the estimates of Lemmas 5.1 and 5.2, we can show that the right-hand side of (5.11) vanishes. Indeed,

\[
\sqrt{\frac{p_n - p_c}{n}} \max_{\tau_{k-1} \leq i < \tau_k} \| \varphi(X_i) - \varphi(X_{\tau_{k-1}}) \|
\leq \sqrt{\frac{p_n - p_c}{n}} \left( \max_{v \in \text{BBT}_k} \| \varphi(v) - \varphi(X_{\tau_{k-1}}) \| + \max_{v \in \text{BBT}_k} \sup_{w \in \mathcal{T}_{\text{trap}}} \| \varphi(w) - \varphi(v) \| \right). \tag{5.15}
\]

By (5.14), we have for any \( \varepsilon > 0 \)

\[
P_{p_n} \left( \sqrt{\frac{p_n - p_c}{n}} \max_{v \in \text{BBT}_k} \| \varphi(v) - \varphi(X_{\tau_{k-1}}) \| > \varepsilon \right) \leq \frac{1}{\varepsilon^2 n^{3/2}} C. \tag{5.16}
\]

Using further that by the decomposition (2.1), the GW-trees of the traps are independent of the backbone tree,

\[
P_{p_n} \left( \sqrt{\frac{p_n - p_c}{n}} \max_{v \in \text{BBT}_k} \sup_{w \in \mathcal{T}_{\text{trap}}} \| \varphi(w) - \varphi(v) \| > \varepsilon \right)
\leq \mathbb{E}_{p_n}[|\text{BBT}_k|] \max_{1 \leq \delta \leq \Delta_-1} P_{p_n} \left( \sqrt{\frac{p_n - p_c}{n}} \sup_{w \in \mathcal{T}_{\text{trap}}} \| \varphi(w) - \varphi(v) \| > \varepsilon \mid \deg_{\mathcal{T}_{\text{trap}}}(v) = \delta \right)
\leq e^{-\gamma \sqrt{n}} \mathbb{E}_{p_n}[|\text{BBT}_k|] \times \max_{1 \leq \delta \leq \Delta_-1} \mathbb{E}_{p_n} \left[ \exp \left( \gamma \sqrt{\frac{p_n - p_c}{n}} \sup_{w \in \mathcal{T}_{\text{trap}}} \| \varphi(w) - \varphi(v) \| \right) \mid \deg_{\mathcal{T}_{\text{trap}}}(v) = \delta \right]
\leq C \left( (p_n - p_c)e^{\gamma \sqrt{n}} \right)^{-1}, \tag{5.17}
\]

where the last inequality follows from (5.12) and (5.13). From (5.15), (5.16), and (5.17) we may conclude

\[
P_{p_n} \left( \sqrt{\frac{p_n - p_c}{n}} \max_{k=1, \ldots, C_n} \max_{\tau_{k-1} \leq i < \tau_k} \| \varphi(X_i) - \varphi(X_{k-1}) \| > \varepsilon \right)
\leq \frac{C}{\varepsilon^2 n^{1/2}} + C \left( (p_n - p_c)e^{\gamma \sqrt{n}} \right)^{-1}. \tag{5.18}
\]

Recall that in Theorem 1.2 we assumed that for any \( \delta > 0 \), \( e^{\delta \sqrt{n}} (p_n - p_c) \rightarrow \infty \). Under this assumption the right-hand side converges to zero, which implies that the right-hand side of (5.11) vanishes in probability. Therefore, the convergence (5.3) also holds with \( W_n^{(1)} \) replaced by \( W_n^{(4)} \), that is,

\[
\left( \frac{\varphi(X_{t_{\tau_{n}}})}{\sqrt{\nu_n \mathbb{E}_{p_n}[\tau_2 - \tau_1](p_n - p_c)^2}} \right)_{t \in [0, 1]} \overset{d}{\underset{n \rightarrow \infty}{\rightarrow}} \left( (\kappa \Sigma)^{1/2} B_t \right)_{t \in [0, 1]}. \tag{5.19}
\]

This completes the fourth step.
All that remains to finish the proof of Theorem 1.2 is to apply the time change
\( t \mapsto tn(p_n - p_c)^{-3}\nu_n^{-1} \), which, by the scale invariance of Brownian motion, yields
\[
\left( \nu_n t \right)^{3\nu_n} \varphi(X_{\lfloor tn(p_n - p_c)^{-3} \rfloor}) \xrightarrow{d} \left( (\nu \Sigma)^{1/2} B_t \right)_{t \in [0,1]} .
\] (5.20)
The Law of Large Numbers for \( \tau_n \) then allows us to replace the prefactor by \( \sqrt{(p - p_c)/n} \), which completes the proof. \( \square \)

6. Proofs of the escape time estimates

In this section we prove the lemmas of Section 3. We assume in this section some
familiarity with the theory of reversible Markov chains and electrical networks. We
refer the reader who is insufficiently knowledgable about this topic to [AF02, LP16],
where all reversibility-related topics of this section that are not explicitly cited are
defined and discussed.

6.1. Proof of Lemma 3.1. Let \( T \) be an infinite tree and \( v \in \mathcal{G}[1](T) \). We have
\[
P^v_T(\eta_0 < \infty) = \lim_{n \to \infty} P^v_T(\eta_0 < \eta_n).
\]
Denote by \( C_G(x, A) \) the effective conductance in a graph \( G = (V, E) \) with unit edge
weights between a vertex \( x \in V \) and a subset \( A \subset V \), and write \( \mathcal{R}_G(x, A) := C_G(x, A)^{-1} \)
for the associated effective resistance. For two disjoint sets \( A, B \) and \( x \notin A, B \) we have the bound
\[
P^x_G(\eta(A) < \eta(B)) \leq \frac{C_G(x, A)}{C_G(x, A \cup B)} \leq \frac{C_G(x, A)}{C_G(x, B)},
\] (6.1)
see [LP16, Exercise 2.34] or [BGP03, Fact 2]. In fact, the arguments in the latter
reference even show that if it is impossible that the random walk hits both \( A \) and \( B \)
during the same excursion starting from \( x \), then
\[
P^x_G(\eta(A) < \eta(B)) = \frac{C_G(x, A)}{C_G(x, A \cup B)} = \frac{C_G(x, A)}{C_G(x, A) + C_G(x, B)}.
\] (6.2)
In our setting, (6.1) gives
\[
P^v_T(\eta_0 < \eta_n) \leq \frac{C_T(v, \emptyset)}{C_T(v, \mathcal{G}[n](T))}.
\] (6.3)
The Series Law for conductances implies that for any \( v \in \mathcal{G}[1](T) \) we have \( C_T(v, \emptyset) = [L/(p - p_c)]^{-1} \). Rayleigh’s Monotonicity Principle, moreover, implies that for any
\( v \in \mathcal{G}[1](T) \) we have
\[
C_T(v, \mathcal{G}[n](T)) \geq C_{T_v}(v, \mathcal{G}[n-1](T_v)).
\]
Observe that \( C_{T_v}(v, \mathcal{G}[n-1](T_v)) \) (and any bound we formulate from here on) depends
only on the subtree rooted at \( v \), and is thus independent of the tree segment \( T_{[0,1]} \) under
\( \mathbb{P} \). This implies
\[
P^v(\eta_0 < \infty \mid T_{[0,1]} = T_{[0,1]}) \leq \lim_{n \to \infty} \frac{p - p_c}{L} \mathbb{E} \left[ C_{T_v}(v, \mathcal{G}[n-1](T_v)) \right].
\] (6.4)
Now given an infinite tree $T$ and $v \in G_{[1]}(T)$, we prune $T_v$ as follows: first, remove any vertex $w \in T_v$ that does not have an infinite line of descent to obtain $T_{v,\text{BB}}$, and second, at every vertex $w \in T_{v,\text{BB}}$ with more than two children in $T_{v,\text{BB}}$, keep the first two in Ulam-Harris ordering and delete all other children (and their subtrees). Call the resulting tree $\tilde{T}_v$. Since we only removed edges, Rayleigh’s Monotonicity Principle once more implies
\[ R_{\tilde{T}_v}(v, G_{[n]}(T)) \leq R_{\tilde{T}_v}(v, G_{[n]}(\tilde{T}_v)). \]  
(6.5)

Now let $\tilde{G}_{[n]}(\tilde{T}_v)$ denote the set of vertices $w \in \tilde{T}_v$ such that $\deg_{\tilde{T}_v}(w) = 3$ and such that the unique path $v = v_0, v_1, \ldots, v_m = w$ in $\tilde{T}_v$ satisfies $|\{v_i : \deg_{\tilde{T}_v}(v_i) = 3\}| = n$ (i.e., there are exactly $n$ degree-three vertices between $v$ and $w$ in $\tilde{T}_v$). For $\bar{P}$-almost all trees, $\tilde{G}_{[n]}(\tilde{T}_v)$ is nonempty for all $n$. Moreover, for any $T$,
\[ \lim_{n \to \infty} R_{\tilde{T}_v}(v, G_{[n]}(\tilde{T}_v)) = \lim_{n \to \infty} R_{\tilde{T}_v}(v, \tilde{G}_{[n]}(\tilde{T}_v)), \]  
(6.6)

because this limit does not depend on the choice of the sequence of exhausting subgraphs of $\tilde{T}_v$ (see [LP16, Exercise 2.4]).

Under $\bar{P}$, the subtree of $\tilde{T}_v$ of all paths connecting $v$ to $\tilde{G}_{[n]}(\tilde{T}_v)$ now has a particularly easy structure: vertices with two children are connected by line segments (a sequence of vertices with one child) up to the $n$-th “generation” of branching points. By the Series Law, the effective resistances of these line segments are just their respective lengths, so they are independent $\text{Geo}(1 - \hat{p}_1)$ random variables (which take values on $\{1, 2, \ldots \}$), where $\hat{p}_1$ is the probability that the root of a Galton-Watson tree with generating function $\hat{f}(s)$ as described in (2.1) has exactly one child. By Lemma 2.1, $\hat{p}_1 = \hat{f}'(0) = f'(q_p) = 1 - c_1(p - p_c)(1 + o(1))$, so the mean length of a path segment between two branching points in $\tilde{T}_v$ is at most $\lambda(p - p_c)^{-1}$ for some $\lambda > 0$.

Suppose $G_{[1]}(\tilde{T}_v) = \{w\}$ and call $\tilde{T}_w^{(1)}, \tilde{T}_w^{(2)}$ the two subtrees of $\tilde{T}_v$ rooted at $w$. An application of the Series Law and the Parallel Law gives rise to the recursion
\[ R_{\tilde{T}_v}(v, \tilde{G}_{[n]}(\tilde{T}_v)) = d_{\tilde{T}_v}(v, w) + \left( \frac{1}{R_{\tilde{T}_w^{(1)}}(v, \tilde{G}_{[n-1]}(\tilde{T}_w^{(1)})))} \right)^{-1} + \frac{1}{R_{\tilde{T}_w^{(2)}}(v, \tilde{G}_{[n-1]}(\tilde{T}_w^{(2)}))} \]  
(6.7)

Now write $\tilde{R}_{[n]} := R_{\tilde{T}_v}(v, \tilde{G}_{[n]}(\tilde{T}_v))$ for the effective resistance between $v$ and $\tilde{G}_{[n]}(\tilde{T}_v)$, and write $\tilde{R}_{[n-1]}^{(1)}$ and $\tilde{R}_{[n-1]}^{(2)}$ for independent copies of $\tilde{R}_{[n-1]}$. Then, under $\bar{P}$, we have the following equality in distribution:
\[ \tilde{R}_{[n]} \overset{d}{=} S + \left( \frac{1}{\tilde{R}_{[n-1]}^{(1)}} + \frac{1}{\tilde{R}_{[n-1]}^{(2)}} \right)^{-1}, \]  
(6.8)

where $S \sim \text{Geo}(1 - \hat{p}_1)$. Bounding the harmonic mean by the arithmetic mean, i.e., using that $(\frac{1}{x} + \frac{1}{y})^{-1} \leq \frac{1}{2}(x + y)$ for $x, y > 0$, this yields the stochastic domination
\[ \tilde{R}_{[n]} \leq S + \frac{1}{4}(\tilde{R}_{[n-1]}^{(1)} + \tilde{R}_{[n-1]}^{(2)}), \]  
(6.9)
Iterating (6.9), we obtain
\[ \tilde{R}_{[n]} \leq \sum_{i=0}^{n-1} 4^{-i} \left( \sum_{j=1}^{2^i} S_{i,j} \right), \tag{6.10} \]
where \((S_{i,j})_{i,j \geq 1}\) is an array of i.i.d. copies of \(S\). Taking the expectation,
\[
\lim_{n \to \infty} \tilde{\mathbb{E}}[\tilde{R}_{[n]}] \leq \lim_{n \to \infty} \sum_{i=0}^{n-1} 2^{-i} \mathbb{E}[S] = 2\mathbb{E}[S] \leq \frac{2\lambda}{p - p_c}. \tag{6.11}
\]
Combining (6.4), (6.5), (6.6) and (6.11), we obtain
\[
\mathbb{P}^\nu(\eta_0 < \infty \mid T_{[0,1]} = T_{[0,1]}) \leq \frac{p - p_c}{L} \frac{2\lambda}{p - p_c} = \frac{2\lambda}{L},
\]
and this upper bound is smaller than \(\frac{3}{2}\) when we choose \(L \geq L_0 := 3\lambda\). \(\Box\)

6.2. Proof of Lemma 3.2. Define the event \(G := \{|G_{[1]}(T)| > 1\}\) and, for \(\alpha < \frac{1}{2}\), let \(\mathcal{F}(\alpha) := \{|G'_{[\sqrt{\alpha}L/(p-p_c)]}(T)| = 1\}\) denote the event that the backbone tree does not branch before the \([\sqrt{\alpha}L/(p-p_c)]\)-th generation. Splitting the expectation along \(G\) and \(\mathcal{F}(\alpha)\), we bound
\[
\tilde{\mathbb{E}}\left[ \sum_{v \in G_{[1]}(T)} 1_{\{P^v_{T \setminus T_v}(\eta_0 = \infty) < \alpha\}} \right] \leq \tilde{\mathbb{P}}(G^c) + \tilde{\mathbb{E}}\left[ 1_{\mathcal{F}(\alpha)^c} |G_{[1]}| \right] \tag{6.12}
\]
\[ \quad + \tilde{\mathbb{E}}\left[ 1_{G \cap \mathcal{F}(\alpha)} \sum_{v \in G_{[1]}(T)} \tilde{\mathbb{P}}(P^v_{T \setminus T_v}(\eta_0 = \infty) < \alpha \mid T_{[0,1]}) \right]. \]
By Lemma 2.3 the first term on the right-hand side is bounded from above by \(a_2\), as desired. It remains to show that the second and third term on the right-hand side above both vanish as \(\alpha \to 0\).

We start with the second term. Let \(B_k\) be the event that the first branching of the backbone tree occurs in generation \(k\), that is, \(|G'_{1}| = \cdots = |G'_{k-1}| = 1\), but \(|G'_{k}| > 1\). Then, writing \(G'_{[i]}\), \(i \geq 0\) for i.i.d. copies of \(G_{[1]}\),
\[
\tilde{\mathbb{E}}\left[ 1_{\mathcal{F}(\alpha)^c} |G_{[1]}| \right] = \sum_{k=1}^{[\sqrt{\alpha}L/(p-p_c)]} \tilde{\mathbb{E}}\left[ 1_{B_k} |G_{[1]}| \right]
\]
\[ \leq \sum_{k=1}^{[\sqrt{\alpha}L/(p-p_c)]} \tilde{\mathbb{E}}\left[ 1_{B_k} \sum_{i=1}^{[G'_{[1]}]} |G'_{[i]}| \right] = \tilde{\mathbb{P}}(\mathcal{F}(\alpha)^c) \tilde{\mathbb{E}}\left[ |G_{[1]}|^2 \right], \]
where we used Wald’s identity for the last equality. By Lemma 2.3, \(\tilde{\mathbb{E}}[|G_{[1]}(T)|] \leq a_3\) and, by an argument similar to (2.8), there exists a constant \(c > 0\) such that \(\tilde{\mathbb{P}}(\mathcal{F}(\alpha)^c) \leq C \sqrt{\alpha}\) when \(\alpha\) is sufficiently close to 0, so we conclude
\[
\tilde{\mathbb{E}}\left[ 1_{\mathcal{F}(\alpha)^c} |G_{[1]}| \right] \leq C \sqrt{\alpha}a_3^2. \tag{6.13}
\]
Now we show that the third term on the right-hand side of (6.12) also tends to 0 as \( \alpha \to 0 \). Let \( T \in \mathcal{G} \cap \mathcal{F}(\alpha) \) and denote by \( v_0 \) the unique vertex of \( G'_{\sqrt{\alpha}L/(p-p_c)}(T) \). Then, for any \( v \in G_{[1]} \),

\[
P_{T \setminus T_v}(\eta_0 = \infty) \geq P_{T \setminus T_v}^{v_0}(\eta_0 = \infty) = \frac{C_{T \setminus T_v}(v_0, \infty)}{C_{T \setminus T_v}(v_0, v) + C_{T \setminus T_v}(v_0, \infty)}
\]

by (6.2). By the Series Law and the fact that \( T \in \mathcal{G} \cap \mathcal{F}(\alpha) \) we have \( C_{T \setminus T_v}(v_0, v) = \sqrt{\alpha}L/(p-p_c) \). It follows that \( P_{T \setminus T_v}^{v_0}(\eta_0 = \infty) < \alpha \) can only hold if

\[
R_{T \setminus T_v}(v_0, \infty) > \frac{1 - \alpha}{\sqrt{\alpha}} \frac{L}{p - p_c}.
\]

This implies that for any vertex \( w \in G_{[1]} \setminus \{v\} \),

\[
\mathbb{P} \left( P_{T \setminus T_v}^{w}(\eta_0 = \infty) < \alpha \mid T_{[0,1]} \right) \leq \mathbb{P} \left( R_{T \setminus T_v}(v_0, \infty) > \frac{1 - \alpha}{\sqrt{\alpha}} \frac{L}{p - p_c} \mid T_{[0,1]} \right)
\]

\[
\leq \mathbb{P} \left( \frac{L}{p - p_c} + R_{T_w}(w, \infty) > \frac{1 - \alpha}{\sqrt{\alpha}} \frac{L}{p - p_c} \mid T_{[0,1]} \right)
\]

\[
\leq 4\sqrt{\alpha} \frac{p - p_c}{L} \mathbb{E} \left[ R_{T_w}(w, \infty) \right],
\]

where we have used the independence of \( T_w \) and \( T_{[0,1]} \) and the fact that \( \alpha < 1/4 \). We established above in (6.11) that \( \mathbb{E} \left[ R_{T_w}(w, \infty) \right] \leq C(p - p_c)^{-1} \). Inserting this and (6.13) into (6.12) to obtain

\[
\mathbb{E} \left[ \sum_{v \in G_{[1]}(T)} \mathbb{1}_{\{P_{T \setminus T_v}^{\eta_0}(\eta_0 = \infty) < \alpha\}} \right] \leq a_2 + C \sqrt{\alpha},
\]

which completes the proof. \( \square \)

6.3. **Proof of Lemma 3.3.** The main idea in this proof is to decompose the random walk on \( T_{[0,2]} \) into random walks on \( T_{[0,1]} \) and walks on the disjoint trees of \( T_{[1,2]} \). Let \( (Y_n)_{n \geq 1} \) be the simple random walk on \( T_{[0,1]} \) started at \( v \), and write \( H = \sigma(Y_1, Y_2, \ldots, Y_{\eta_0}) \) for the history of this process until it hits the root \( v \). For \( i \geq 1 \), let \( h_i \) be the \( i \)-th time that \( (Y_n)_{n \geq 1} \) visits a vertex in \( G_{[1]} \cup \{v\} \) and let \( V_i = Y_{h_i} \). Furthermore, let \( H \) denote the number of visits of \( (Y_n)_{n \geq 1} \) to \( G_{[1]} \cup \{v\} \) until its first visit to the root, i.e., \( H = \inf\{i : V_i = v\} \). We can decompose the walk on \( T_{[0,2]} \) into the walk \( (Y_n)_{n \geq 1} \) and the walks on the branches of \( T_{[0,2]} \setminus T_{[0,1]} \), and write

\[
\mathbb{P}^v(\eta_2 < \eta_0, \text{sib}(X_{\eta_2}) = 0 \mid T_{[0,1]} = T_{[0,1]} \times \mathbb{E} \left[ \sum_{m=1}^{H-1} P_{T}^{Y_{m}}(\eta_2 < \eta_0^H) \mathbb{1}_{\{|G_{[1]}(T_{\eta_0^H})| = 1\}} \prod_{i=1}^{m-1} P_{T}^{Y_{i}}(\eta_2^i < \eta_2) \mid H \right] \mid T_{[0,1]} = T_{[0,1]} \right],
\]

where we write \( \eta_0^H \) for the hitting time of a parent vertex of \( G_{[1]} \) (i.e., the hitting time of level \( \lfloor L(p - p_c)^{-1} \rfloor - 1 \)). Observe that conditioned on \( V_i, P_{T}^{Y_{i}}(\eta_2 < \eta_0^H) \) and \( \mathbb{1}_{\{|G_{[1]}(T_{\eta_0^H})| = 1\}} \)
are independent of $\mathcal{T}_{[0,1]}$. By (6.2) we have

$$P_{T}^{V_{m}}(\eta_2 < \eta_1^p) \mathbb{1}_{\{|G_{[1]}(T_{V_m})| = 1\}} = \frac{C_{T_{V_m}}(V_{m}, G_{[1]}(T_{V_i}))}{1 + C_{T_{V_m}}(V_{m}, G_{[1]}(T_{V_i}))} \mathbb{1}_{\{|G_{[1]}(T_{V_m})| = 1\}} \geq c(p - p_c) \mathbb{1}_{\{|G_{[1]}(T_{V_m})| = 1\}},$$

(6.16)

and similarly

$$P_{T}^{V_{i}}(\eta_1^p < \eta_2) = \frac{1}{1 + C_{T_{V_m}}(V_{m}, G_{[1]}(T_{V_i}))} \geq 1 - c|G_{[1]}(T_{V_i})|(p - p_c).$$

(6.17)

Define for $m < H$ and $v \in G_{[1]}$

$$H_m(v) := |\{k < m : Y_{h_k} = v\}|,$$

(6.18)

which counts the number of visits of $(Y_n)_{n \geq 1}$ to vertex $v$ before the $m$-th visit to a vertex in $G_{[1]}$. Reordering the product in (6.15) according to the vertices and applying the bound in (6.16) gives the following lower bound,

$$\mathbb{E} \left[ \sum_{m=1}^{H-1} P_{T}^{V_{m}}(\eta_2 < \eta_1^p) \mathbb{1}_{\{|G_{[1]}(T_{V_m})| = 1\}} \prod_{i=1}^{m-1} P_{T}^{V_{i}}(\eta_1^p < \eta_2) \left| \mathcal{H} \right. \right] \geq c(p - p_c) \mathbb{E} \left[ \sum_{m=1}^{H-1} \prod_{i=1}^{m-1} P_{T}^{V_{i}}(\eta_1^p < \eta_2) \mathbb{1}_{\{|G_{[1]}(T_{V_m})| = 1\}} \prod_{v \in G_{[1]}: v \neq V_{m}} P_{T}^{V_{i}}(\eta_1^p < \eta_2) \mathbb{1}_{|H_{m}(v)|} \left| \mathcal{H} \right. \right].$$

(6.19)

Now note that $H$, $H_{m}(v)$ and $V_{m}$ are measurable with respect to $\mathcal{H}$, and the probabilities $P_{T}^{V_{i}}(\eta_1^p < \eta_2)$ are independent of $\mathcal{H}$ and independent under $\tilde{P}(\cdot | \mathcal{T}_{[0,1]})$ for different $v$, so

$$\mathbb{E} \left[ \sum_{m=1}^{H-1} P_{T}^{V_{m}}(\eta_2 < \eta_1^p) \mathbb{1}_{\{|G_{[1]}(T_{V_m})| = 1\}} \prod_{i=1}^{m-1} P_{T}^{V_{i}}(\eta_1^p < \eta_2) \left| \mathcal{H} \right. \right] \geq c(p - p_c) \sum_{m=1}^{H-1} \mathbb{E}_{\mathcal{T}_{[0,1]}} \left[ \mathbb{1}_{\{|G_{[1]}(T_{V_m})| = 1\}} P_{T}^{V_{m}}(\eta_1^p < \eta_2) \mathbb{1}_{|H_{m}(V_{m})|} \left| \mathcal{H} \right. \right] \times \prod_{v \in G_{[1]}: v \neq V_{m}} \mathbb{E}_{\mathcal{T}_{[0,1]}} \left[ P_{T}^{V_{i}}(\eta_1^p < \eta_2) \mathbb{1}_{|H_{m}(v)|} \left| \mathcal{H} \right. \right].$$

(6.20)
Applying now the bound (6.17), we see that for $p$ sufficiently close to $p_c$, (6.20) is bounded from below by

$$c(p - p_c) \sum_{m=1}^{H-1} \tilde{P}(|G_{[1]}| = 1)(1 - c(p - p_c))^{H_m(V_m)} \prod_{v \in G_{[1]}: v \neq V_m} (1 - c\mathbb{E}[|G_{[1]}|(p - p_c)])^{H_m(v)}$$

$$\geq c(p - p_c)\tilde{P}(|G_{[1]}| = 1) \sum_{m=1}^{H-1} \prod_{v \in G_{[1]}} (1 - c\mathbb{E}[|G_{[1]}|(p - p_c)])^{H_m(v)}$$

$$= c(p - p_c)\tilde{P}(|G_{[1]}| = 1) \sum_{m=1}^{H-1} (1 - c\mathbb{E}[|G_{[1]}|(p - p_c)])^m$$

$$= c\mathbb{E}[|G_{[1]}|]^{-1} \tilde{P}(|G_{[1]}| = 1) \left( (1 - c\mathbb{E}[|G_{[1]}|(p - p_c)]) - (1 - c\mathbb{E}[|G_{[1]}|(p - p_c)])^H \right).$$

(6.21)

By Lemma 2.3, $\tilde{P}(|G_{[1]}(T)| = 1) \geq a_1$ and $\mathbb{E}[|G_{[1]}(T)|] \leq a_3$. So to show a uniform lower bound for (6.15), it remains to show that

$$\mathbb{E} \left[ (1 - c\mathbb{E}[|G_{[1]}|(p - p_c)])^H \bigg| T_{[0,1]} = T_{[0,1]} \right] < 1 - a$$

(6.22)

for an $a > 0$ independent of $p$. While the exact distribution of $H$ depends on $T_{[0,1]}$, we may give an easy lower bound. Recall that $H - 1$ is the number of visits of $(Y_n)_{n \geq 1}$ to $G_{[1]}$ until it hits the root, starting at some $v \in G_{[1]}$. For any tree $T$ with $v \in G_{[1]}(T)$ we have

$$P_T^n(\eta_0 < \eta_1^+) \leq C_T(v, g) \leq c(p - p_c),$$

(6.23)

where $\eta_1^+$ denotes the first hitting time of $G_{[1]}$ after time 0. This bound implies that $H$ stochastically dominates a Geometric random variable $H'$ with success probability $c(p - p_c)$ and generating function

$$\mathbb{E}[\theta^{H'}] = \frac{c(p - p_c)\theta}{1 - \theta + c\theta(p - p_c)}$$

for $0 \leq \theta \leq 1$, which in turn implies that

$$\mathbb{E} \left[ (1 - c\mathbb{E}[|G_{[1]}|(p - p_c)])^H \bigg| T_{[0,1]} = T_{[0,1]} \right] \leq \mathbb{E} \left[ (1 - c\mathbb{E}[|G_{[1]}|(p - p_c)])^H \right]$$

$$= 1 - \frac{c\mathbb{E}[|G_{[1]}|]}{c\mathbb{E}[|G_{[1]}|] + c(1 - c\mathbb{E}[|G_{[1]}|(p - p_c)]).}$$

(6.24)

The last term is bounded away from 0 uniformly in $p$, which shows (6.22) and completes the proof.
7. Moment bounds on regeneration distances: proof of Lemma 4.4

This proof is inspired by a similar moment estimate of [DGPZ02] and a regularity estimate for trees from [GK01]. By Lemmas 2.3, 3.1, and 4.3,

\[
\mathbb{E}[(|X_{\tau_1}| - |X_{\tau_1}|)^q] = \mathbb{E}^v[|X_{\tau_1}|^q | G_{[1]}(T) = \{v\}, \eta_0 = \infty] \\
\leq C\mathbb{E}[|X_{\tau_1}|^q],
\]

so it suffices to find a moment estimate for $|X_{\tau_1}|$. Recall that by Definition 4.1, the first regeneration occurs at one of the potential regeneration times $S_k$. More precisely, the regeneration time is set to equal the last $S_k$ that is finite, such that

\[
\mathbb{E}[|X_{\tau_1}|^q] \leq \sum_{k \geq 1} \mathbb{E}[|X_{S_k}|^q \mathbbm{1}_{\{S_k < \infty\}}] \\
\leq \sum_{k \geq 1} \mathbb{E}[|X_{S_k}|^{2q} \mathbbm{1}_{\{S_k < \infty\}}]^{1/2} \mathbb{P}(S_k < \infty)^{1/2}.
\]

For $S_k$ to be finite, it has to occur at least $k$ times that the walker sees a new part of the tree but then backtracks a generation, so by Lemma 3.1,

\[
\mathbb{P}(S_k < \infty) \leq \left( \sup_{T: v \in G_{[1]}(T)} \mathbb{P}^v(\eta_0 < \infty | T_{[0,1]} = T_{[0,1]}) \right)^k \leq \left( \frac{3}{4} \right)^k.
\]

By Definition 4.1, we can also write

\[
\frac{|X_{S_k}|}{[L/(p - p_c)]} = M_k =: N_{k-1} + H_k, \quad M_1 =: H_1 \quad (7.1)
\]

so that $H_k = M_k - N_{k-1}$. That is, $H_k$ counts the number of $L$-levels that are visited by the walk after it surpasses the previous maximum generation $N_{k-1}$, until it arrives at the next possible regeneration point, i.e., at a vertex $v$ in an $L$-level with $\text{sib}(v) = 0$. Then $H_k$ can be stochastically dominated by $2 + \tilde{H}_k$, where $\tilde{H}_k$ is a Geometric random variable with success probability

\[
\inf_{T: v \in G_{[1]}(T)} \mathbb{P}^v(\eta_2 < \eta_0, \text{sib}(X_{\eta_2}) = 0 | T_{[0,1]} = T_{[0,1]}) \geq a_3,
\]

where the lower bound holds by Lemma 3.3. Since Geometric random variables have all moments bounded, it also holds for any $q > 0$ that $\mathbb{E}[H_k^{2q}] < \infty$.

Write $\tilde{N}_k := N_k - M_k$. Iterating the recursion (7.1), and noting that $\{S_k < \infty\} \subset \{R_{k-1} < \infty\}$, we arrive at

\[
(p - p_c)^q \mathbb{E}[|X_{\tau_1}|^q] \leq L^q \sum_{k \geq 1} \mathbb{E} \left[ \left( H_1 + \sum_{i=1}^{k-1} (\tilde{N}_i + H_{i+1}) \right)^{2q} \mathbbm{1}_{\{R_{k-1} < \infty\}} \right]^{1/2} \left( \frac{2}{3} \right)^{k/2} \\
\leq L^q \left( \sum_{k \geq 1} (2k)^{2q-1} \left( \sum_{i=1}^k \mathbb{E}[H_i^{2q}] + \sum_{i=1}^{k-1} \mathbb{E}[\tilde{N}_i^{2q} \mathbbm{1}_{\{R_i < \infty\}}] \right) \right)^{1/2} \left( \frac{2}{3} \right)^{k/2}.
\]
Since the \((2q)\)-th moment of \(H_i\) is uniformly bounded, the proof will be completed once we show that
\[
\mathbb{E}\left[e^{s\tilde{N}_i 1_{\{R_1<\infty\}}}\right] \leq C < \infty, \tag{7.2}
\]
for some \(s > 0\). Observe that \(\tilde{N}_i\) counts how many new \(L\)-levels are reached by the walker at \(X_{\eta_{M_i}}\) before backtracking to its \(L\)-ancestor. Since for different \(i\) these excursions happen in disjoint parts of the tree, the \(\tilde{N}_i\) are in fact i.i.d. under \(\mathbb{P}\). It thus suffices to bound
\[
\mathbb{E}\left[e^{s\tilde{N}_i 1_{\{R_1<\infty\}}}\right] = \sum_{n \geq 1} e^{sn}\mathbb{P}(\tilde{N}_1 = n, R_1 < \infty).
\]

On the event \(\{R_1 < \infty\}\), a large value for \(\tilde{N}_1\) implies that the random walk backtracks a long distance towards the root. We will bound the probability by decomposing this trajectory into level-sized chunks. Since the segment of the backbone tree between \(X_{\eta_{M_i}}\) and its \(L\)-ancestor is by definition of \(|M_i|\) isomorphic to a line graph of length \([L/(p_c-p)]\), the first step in this decomposition is to bound
\[
\mathbb{P}(\tilde{N}_1 = n, R_1 < \infty) \leq \mathbb{E}\left[\sum_{v \in G_{[n]}(T)} P_{\pi_T}^v(X_{\eta_n} = v)P_{\pi_T}^v(\eta_0 < \infty) \left| G_{[1]}(T) \right| = 1\right],
\]
where \(v_1\) is the unique vertex of \(G_{[1]}(T)\). Note that Lemma 3.1 is not sufficient to bound \(P_{\pi_T}^v(\eta_0 < \infty)\), because the lemma gives an annealed bound, whereas we need a quenched bound. We instead use Lemmas 2.3 and 3.2 to show that the event
\[
B_n(\alpha, \beta) := \left\{\sum_{i=1}^n 1_{\{P_{\pi_{\tau_{\eta_i}}}^{v_i}(\eta_{i-1}=\infty) \geq \alpha\}} \geq \beta n \text{ for all } v \in G_{[n]}(T)\right\},
\]
has a high probability provided \(\alpha\) and \(\beta\) are small enough, where, for \(v = v_n \in G_{[n]}(T)\), we denote by \(v_i\) its ancestor in \(G_{[i]}(T)\) for \(0 \leq i < n\). If we show this we are done, because, for \(T \in B_n(\alpha, \beta)\) and \(v \in G_{[n]}(T)\),
\[
P_{\pi_T}^v(\eta_0 < \infty) \leq \prod_{i=1}^n P_{\pi_{\tau_{\eta_i}}}^{v_i}(\eta_{i-1} < \infty) \leq (1-\alpha)^{\beta n}, \tag{7.3}
\]
so that
\[
\mathbb{P}(\tilde{N}_1 = n, R_1 < \infty) \leq (1-\alpha)^{\beta n} \mathbb{E}\left[\sum_{v \in G_{[n]}(T)} P_{\pi_T}^v(X_{\eta_n} = v) \left| G_{[1]}(T) \right| = 1\right] + \mathbb{P}(B_n(\alpha, \beta)^c | \left| G_{[1]}(T) \right| = 1) \tag{7.4}
\]
\[
= (1-\alpha)^{\beta n} + \mathbb{P}(B_n(\alpha, \beta)^c | \left| G_{[1]}(T) \right| = 1),
\]
where for the equality we have used that conditionally on \(\text{sib}(v_1) = 0\), the probabilities add to 1 because \(G_{[n]}(T)\) is a cutset on \(T\) separating the root from infinity. To show
that $B_n(\alpha, \beta)^c$ also has exponentially small probability, we introduce

$$A_n(\alpha, v) := \sum_{i=1}^{n} \mathbb{1}_{\{P^*_{T \setminus T_v}(\eta_{i-1}=\infty) \geq \alpha\}},$$

and

$$Z_n(\alpha, \theta) := \sum_{v \in G_n(T)} e^{-\theta A_n(\alpha, v)}.$$

We want to show that $Z_n(\alpha, \theta)$ decays exponentially for $\theta$ large enough. Note that by asking for the event \{\eta_{i-1} = \infty\} only on the cropped tree $T \setminus T_v$ we have independence of $P^*_{T \setminus T_v}(\eta_{i-1} = \infty)$ for different $i$ under the environment law, which allows us to calculate recursively

$$\mathbb{E}[Z_n(\alpha, \theta) | \text{sib}(v_1) = 0]$$

$$= \mathbb{E} \left[ \sum_{v \in G_{n-1}(T)} e^{-\theta A_{n-1}(\alpha, v)} \mathbb{E} \left[ \sum_{w \in G_{[1]}(T_v)} \mathbb{1}_{\{P^*_{T \setminus T_v}(\eta_{i-1}=\infty) < \alpha\}} \right] + e^{-\theta} \mathbb{1}_{\{P^*_{T \setminus T_v}(\eta_{i-1}=\infty) \geq \alpha\}} \mathbb{E} \left[ |G_{[1]}(T)| \right] |G_{[1]}(T)| = 1 \right]$$

$$= \mathbb{E}[Z_{n-1}(\alpha, \theta) | |G_{[1]}(T)| = 1] \zeta(\alpha, \theta) = \zeta(\alpha, \theta)^{n-1},$$

where

$$\zeta(\alpha, \theta) := \mathbb{E} \left[ \sum_{v \in G_{[1]}(T)} \mathbb{1}_{\{P^*_{T \setminus T_v}(\eta_0=\infty) < \alpha\}} + e^{-\theta} \mathbb{1}_{\{P^*_{T \setminus T_v}(\eta_0=\infty) \geq \alpha\}} \right].$$

In the last step in the recursion we have used that $\mathbb{E}[Z_1(\alpha, \theta) | |G_{[1]}(T)| = 1] = 1$, since $|G_{[1]}(T)| = 1$ implies that $P^*_{T \setminus T_v}(\eta_0 = \infty) = 0$.

We proceed by bounding $\zeta(\alpha, \theta)$ as

$$\zeta(\alpha, \theta) \leq \mathbb{E} \left[ \sum_{v \in G_{[1]}(T)} \mathbb{1}_{\{P^*_{T \setminus T_v}(\eta_0=\infty) < \alpha\}} + e^{-\theta} \mathbb{E}[|G_{[1]}(T)|] \right]. \quad (7.5)$$

By Lemma 2.3 we have $\mathbb{E}[|G_{[1]}(T)|] \leq a_3$. With the bound of Lemma 3.2, this means that we can bound

$$\zeta(\alpha, \theta) \leq a_2 + h(\alpha) + e^{-\theta} a_3 =: \gamma(\alpha, \theta).$$
with $\gamma(\alpha, \theta) < 1$ if $\alpha$ is small enough and $\theta$ is large enough. Fix such a sufficient choice of $\alpha$ and $\theta$, then, by Markov’s inequality,

$$\tilde{P}(B_n(\alpha, \beta)^c | |G_1(T)| = 1) = \tilde{P}\left( \min_{v \in G_1(T)} A_n(\alpha, v) < \beta n | |G_1(T)| = 1 \right)$$

$$= \tilde{P}\left( e^{-\theta \min_v A_n(\alpha, v)} > e^{-\theta \beta n} | |G_1(T)| = 1 \right)$$

$$\leq e^{n^2 \beta} E \left[ e^{-\theta \min_v A_n(\alpha, v)} | |G_1(T)| = 1 \right]$$

$$\leq e^{n^2 \beta} E[Z_n(\alpha, \theta) | |G_1(T)| = 1]$$

$$\leq e^{n^2 \gamma(\alpha, \theta) n - 1},$$

and this bound decays exponentially in $n$ if $\beta$ is small enough. Inserting this bound into (7.4) we see that (7.2) holds for some $s > 0$ sufficiently small, which completes the proof. \hfill \Box

8. Moment bounds on the regeneration times: proof of Lemma 4.5

We start by establishing moment bounds for the time spent in the finite trees $T_v^{\text{trap}}$, which as the name suggests, act as traps for the random walk. Let $H_v := \inf\{ n \geq 0 : X_n = v \}$ be the hitting time of vertex $v$ and $H_v^+ = \inf\{ n > 0 : X_n = v \}$. When $v = \emptyset$ we will suppress the subscript. If $\pi_T$ is the invariant distribution for random walk on a tree $T$, then the well-known formula

$$E_T^v[H_v^+] = \pi_T(v)^{-1}$$

holds. For the second moment we have the following identity, see [AF02, (2.21)],

$$E_T^v[(H_v^+)^2] = \pi_T(v)^{-1} \left( 2 \sum_{w \in T} \pi_T(w) E_T^w[H_v] + 1 \right).$$

(8.2)

Furthermore, we may bound $E_T^v[H_v]$ by the commute time $E_T^v[H_v] + E_T^v[H_w]$. The Commute Time Identity of [CRR+97] applied to simple random walk on a finite tree $T$ gives

$$E_T^v[H_v] + E_T^v[H_w] = 2(|T| - 1) d_T(v, w),$$

(8.3)

with $|T|$ the vertex cardinality of the tree and $d_T(v, w)$ the graph distance between $v$ and $w$ on $T$. This implies

$$E_T^v[(H_v^+)^2] \leq 4 \pi_T(v)^{-1} (|T| - 1) \sum_{w \in T} \pi_T(w) d_T(v, w).$$

(8.4)

Moreover, for simple random walk on a finite connected tree $T$,

$$\pi_T(v) = \frac{\deg_T(v)}{\sum_{w \in T} \deg_T(w)} = \frac{\deg_T(v)}{2(|T| - 1)}.$$

So we arrive at

$$E_T^v[(H_v^+)^2] \leq 4 \deg_T(v)^{-1} (|T| - 1) \sum_{w \in T} \deg_T(w) d_T(v, w).$$

(8.5)
Applying Lemma 2.1 we obtain

\[ U \]

where we used \( n \) (given by (2.2)) and where to the \( i \)th edge, \( i \in \{1, \ldots, U_v\} \), coming out of the root, there is an independent tree \( T_i^* \) with generating function \( f(s) \) as defined in (2.1) and with root \( g_i \) attached to the other end of the \( i \)th edge. We know that \( T_i^* \) is almost surely finite. Write \( H_v^\text{trap} \) for \( H_v^+ \) of random walk started at \( v \), restricted to \( T_v^\text{trap} \), with the convention that \( H_v^\text{trap} = 0 \) if \( T_v^\text{trap} = \{v\} \), that is, when \( U_v = 0 \). Then, by (8.1),

\[
\mathbb{E}[H_v^\text{trap}|U_v] = \mathbb{E}[\pi_T^\text{trap}(v)^{-1}] \mathbb{1}_{\{U_v \neq 0\}}
= 2\mathbb{E} \left[ U_v^{-1} \sum_{i=1}^{U_v} |T_i^*| \right] \mathbb{1}_{\{U_v \neq 0\}}
= 2\mathbb{E}[|T^*|] \mathbb{1}_{\{U_v \neq 0\}}.
\]

Applying Lemma 2.1 we obtain

\[
\mathbb{E}[H_v^\text{trap}|U_v] = 2(1 - \mu^*)^{-1} \mathbb{1}_{\{U_v \neq 0\}} = C(p - p_c)^{-1} \mathbb{1}_{\{U_v \neq 0\}} + o((p - p_c)^{-1}).
\]

From (8.5) we get for the second moment

\[
\mathbb{E}[(H_v^\text{trap})^2|U_v] \leq 4\mathbb{E} \left[ U_v^{-1} \left( \sum_{i=1}^{U_v} |T_i^*| \right) \left( \sum_{i=1}^{U_v} \sum_{w \in T_i^*} \deg_{T_v^\text{trap}}(w) d_{T_v^\text{trap}}(w, v) \mathbb{1}_{\{U_v \neq 0\}} \right) \right] U_v
\]

\[
\leq 4 \sum_{i=1}^{\Delta} \mathbb{E} \left[ |T_i^*| \sum_{w \in T_i^*} (\deg_{T_i^*}(w) + 1)(d_{T_i^*}(w, g_i) + 1) \right]
+ 4 \sum_{i \neq j} \mathbb{E} \left[ |T_i^*| \sum_{w \in T_j^*} (\deg_{T_j^*}(w) + 1)(d_{T_j^*}(w, g_j) + 1) \right]
= 4 \Delta \mathbb{E} \left[ |T^*| \sum_{w \in T^*} (\deg_{T^*}(w) + 1)(d_{T^*}(w, g) + 1) \right]
+ 4 \Delta(\Delta - 1) \mathbb{E} \left[ |T^*| \right] \mathbb{E} \left[ \sum_{w \in T^*} (\deg_{T^*}(w) + 1)(d_{T^*}(w, g) + 1) \right]
=: (I) + (II)
\]

where we used \( U_v \leq \Delta \), the independence of the different GW-trees in a trap and in the second step we added one to the degree because \( \deg_{T_i^\text{trap}}(g_i) = \deg_{T_i^*}(g_i) + 1 \). Now write \( Z_n^* \) for the size of the \( n \)th generation of \( T^* \) and observe that \( |T^*| = \sum_{w \in T^*} 1 = \sum_{n \geq 0} Z_n^* \), and that \( \sum_{w \in T^*} \deg_{T^*}(w) = \sum_{n \geq 0} (Z_n^* + Z_{n+1}^*) \) (each vertex has degree equal to the number of its offspring plus one, the latter accounting for its ancestor), so that we may
bound

\[(I) \leq 4\Delta \mathbb{E}\left[\left(\sum_{n \geq 0} Z_n^*\right) \left(\sum_{n \geq 0} (2Z_n^* + Z_{n+1}^*)(n + 1)\right)\right] \]

\[\leq C \mathbb{E}\left[\left(\sum_{n \geq 0} Z_n^*\right) \left(\sum_{n \geq 0} nZ_n^*\right)\right],\]

where we used in the last step our assumption that \(\Delta < \infty\). By conditioning on the earlier generations we can write the right-hand side as

\[(I) \leq C \sum_{0 \leq m < n} \mathbb{E}[Z_m^* \mathbb{E}[nZ_n^* | Z_m^*]] + C \sum_{n \geq 0} \mathbb{E}[n(Z_n^*)^2] \]

\[+ C \sum_{m > n \geq 0} \mathbb{E}[Z_m^* | Z_n^*] nZ_n^* \]

\[= C \sum_{0 \leq m < n} n\mathbb{E}[(Z_m^*)^2]\mathbb{E}[Z_{n-m}^*] + C \sum_{n \geq 0} n\mathbb{E}[(Z_n^*)^2] \]

\[+ C \sum_{m > n \geq 0} n\mathbb{E}[Z_{n-m}^*] \mathbb{E}[(Z_n^*)^2].\]

With the inequality \(\mathbb{E}[(Z_n^*)^2] \leq Cn(\mu^*)^{2n}\), this simplifies to

\[(I) \leq C \sum_{0 \leq m < n} nm(\mu^*)^{n+m} + C \sum_{n \geq 0} n^2(\mu^*)^{2n} + C \sum_{m > n \geq 0} n^2(\mu^*)^{n+m} \]

\[\leq C \left(\sum_{n \geq 0} n(\mu^*)^n\right)^2 + C \sum_{m > n \geq 0} n^2(\mu^*)^{n+m} \]

\[= C(\mu^*)^2(2 + \mu^*)^2 \leq \frac{C}{(p - p_c)^4}, \quad (8.9)\]

where the final inequality is due to Lemma 2.1.

An easier computation shows that

\[(II) \leq C\mathbb{E}[T^*] \mathbb{E}\left[\sum_{n \geq 0} nZ_n^*\right] \leq \frac{C}{(p - p_c)^3}. \quad (8.10)\]

Combining the bounds (8.9) and (8.10), we may thus conclude that

\[\mathbb{E}[(H_{\text{trap}}^v)^2 | U_v] \leq \frac{C}{(p - p_c)^4}. \quad (8.11)\]

We now need to combine the time spent in the traps with the time the random walk spends on the backbone tree between regeneration times. For this, let \((X_{m}^{\text{BB}})_{m \geq 0}\) be the random walk restricted to the backbone tree without holding times, i.e., writing \(t(m)\) for the time that the random walk \((X_n)_{n \geq 0}\) visits a vertex of \(T^{\text{BB}}\) distinct from the previously visited vertex for the \(m\)th time, we let \(X_{m}^{\text{BB}} := X_{t(m)}\). If \(v \in T^{\text{BB}}\), then we
write \( \ell_n(v) \) for the local time of \((X^B_m)_{m \geq 0}\) at \(v\) until time \(n\). During a visit of \((X^B_m)_{m \geq 0}\) to \(v \in T^B\), the total time \((X_n)_{n \geq 0}\) spends in \(T_v^{\text{trap}}\) is given by

\[
L_v = 1 + \sum_{i=1}^{Y_v} H_{v,i}^{\text{trap}},
\]

(8.12)

where for each \(v \in T^B\), \(H_{v,1}^{\text{trap}}, H_{v,2}^{\text{trap}}, \ldots\) are independent random variables that, conditioned on \(T_v^{\text{trap}}\), have the same distribution as \(H_v^{\text{trap}}\) under \(P^{T_v^{\text{trap}}}\). That is, they count the number of steps spent in the trap until returning to \(v\). Furthermore, conditioned on \(T\), we let \(Y_v\) be a \(\text{Geo}(1 - \sigma(1 - U_v/\deg_T(v)))\) random variable (of the type that takes values in \(\{0, 1, 2, \ldots\}\)), independent of \(H_{v,i}^{\text{trap}}\) for all \(i\). So \(Y_v\) counts the number of times that the random walk \((X_{n})_{n \geq 0}\) at a visit to \(v\) enters and exits the trap before moving on to a different vertex on the backbone tree.

If \(\tau\) is a random time measurable with respect to \(\sigma(X^B_m, m \geq 0)\), then these definitions allow us to write

\[
E[\tau] = E\left[\sum_{v \in T^B} \sum_{i=1}^{\ell_\tau(v)} L_v^{(i)}\right],
\]

(8.13)

where \(L_v^{(i)}\) are i.i.d. copies of \(L_v\), with the convention that if \(\ell_\tau(v) = 0\), then the sum equals 0.

We first apply this general formula to determine a lower bound for the inter-regeneration times. Recall from Definition 4.1 the regeneration structure and the various definitions. We bound

\[
E[\tau_2 - \tau_1] \geq E[(\tau_2 - \tau_1) \mathbb{1}_{\{|G_{|A_{1+1}|}| = 1\}}]
\]

\[
\geq E[\eta_{A_{1+1}} - \tau_1] \mathbb{1}_{\{|G_{|A_{1+1}|}| = 1\}}
\]

\[
\geq a_1 E[\eta_{A_{1+1}} - \tau_1 | |G_{|A_{1+1}|}| = 1]
\]

\[
= a_1 E[\eta_{A_{1+1}} - \tau_1 | |G_{|A_{1+1}|}| = 1, \eta_{A_{1+1}} < \eta_{A_{1-1}}]
\]

\[
= a_1 E^w [\eta_2 - \eta_1 | G_{[1]} = \{w\}, |G_{[2]}| = 1, \eta_2 < \eta_0]
\]

(8.14)

where in the third step we have used Lemma 2.3, in the fourth step we have conditioned on further events that are guaranteed to occur by the definition of the regeneration times, and in the fifth step we have used that \(\tau_1\) is a regeneration time to justify the time shift.

Conditionally on the event \(|G_{[1]}| = |G_{[2]}| = 1\) the graph \(T^B_{[0,2]}\) is isomorphic to a line graph of length \(2|L/(p - p_v)|\), so every vertex, except for the root, and possibly the vertex at the second \(L\)-level, has degree 2 in \(T^B_{[0,2]}\). Conditionally also on \(U_v\), the \(Y_{v,i}\) are
distributed as independent \( \text{Geo}(1 - U_v / (U_v + 2)) \) random variables. From (8.12) we get
\[
E^v[L_v \mid U_v, \deg(T_{bb}(v)) = 2] = 1 + E^v[Y_v \mid U_v, \deg(T_{bb}(v)) = 2]E^v[H_{v,1}^{\text{trap}} \mid U_v, \deg(T_{bb}(v)) = 2] \\
= 1 + \frac{1}{2} U_v E^v[H_{v,1}^{\text{trap}} \mid U_v] \geq c(p - p_c)^{-1} U_v(2)
\]
where \( U_v(2) \) is distributed as \( U_v \) conditionally on \( \deg(T_{bb}(v)) = 2 \), and we used (8.7) for the last step.

Thus, we have for any \( v \in T_{bb}^{[0,2]} \setminus \{ \emptyset, X_{\eta_2} \} \) that
\[
E^w \left[ \sum_{i=1}^{\ell_{\eta_2 - \eta_1}(v)} L_v \mid U_v, \ell_{\eta_2 - \eta_1}(v), G[1] = \{ w \}, |G[2]| = 1, \eta_2 < \eta_0 \right] \\
\geq \ell_{\eta_2 - \eta_1}(v)c(p - p_c)^{-1} U_v(2).
\]

Using (2.2) and Lemma 2.1, we can bound
\[
E[U_v(2)] \geq \mathbb{P}(U_v(2) > 0) = 1 - f'(0)/f'(q_p) \geq c.
\]

Combining this bound with (8.16), inserting that into (8.14), and writing \( \eta_{Bb}^j \) for the hitting time of \( G[j] \) for \( (X_{Bb}^m)_{m \geq 0} \), we obtain
\[
E[\tau_2 - \tau_1] \geq \frac{c}{p - p_c} E^w \left[ \eta_{Bb}^2 - \eta_{Bb}^1 - 2 \mid |G[1]| = \{ w \}, |G[2]| = 1, \eta_{Bb}^2 < \eta_{Bb}^1 \right] \geq \frac{c'}{(p - p_c)^3},
\]
where, in the last step, we have used that \( (X_{Bb}^m)_{m \geq 0} \) restricted to \( T_{bb}^{[0,2]} \), conditionally on \( \{|G[1]| = |G[2]| = 1\} \), is equivalent to a simple random walk on a line graph of length \( 2[L/(p - p_c)] \). This proves the lower bound in Lemma 4.5.

For the upper bounds, we first notice that by the same reasoning as in Lemma 4.4 it suffices to show the bound for \( \tau_1 \). Let us again first consider \( L_v \) as in (8.12), for which we can bound
\[
E^v[L_v^2 \mid T_{bb}] \leq 1 + E^v[Y_v^2 \mid T_{bb}]E^v[(H_{v,1}^{\text{trap}})^2 \mid T_{bb}]
\]
For the Geometric random variable \( Y_v \), the parameter satisfies \( 1 - U_v / \deg(T(v)) \geq 1/\Delta \), which implies
\[
E^v[Y_v^2 \mid T_{bb}] \leq \Delta^2,
\]
while for the second expectation in (8.17) we may apply (8.11) to conclude
\[
E^v[L_v^2 \mid T_{bb}] \leq C(p - p_c)^{-4}.
\]
We can use this upper bound to state a bound on the second moment of a generic random time \( \tau \) that is measurable with respect to \( \sigma(X_n^{Bb}, n \geq 0) \),

\[
E[\tau^2] = E \left[ \left( \sum_{v \in T^{Bb}} \sum_{i=1}^{\ell_v} L_v^{(i)} \right)^2 \right] 
\]

\[
= E \left[ \sum_{v \in T^{Bb}} \left( \sum_{i=1}^{\ell_v} L_v^{(i)} \right)^2 \right] + E \left[ \sum_{v \neq w \in T^{Bb}} \left( \sum_{i=1}^{\ell_v} L_v^{(i)} \right) \left( \sum_{i=1}^{\ell_w} L_w^{(i)} \right) \right] 
\]

\[
= E \left[ \sum_{v \in T^{Bb}} \ell_v E^v [L_v^2 | T^{Bb}] \right] + \ell_v (v)(\ell_v (v) - 1) E^v [L_v | T^{Bb}]^2 
\]

\[
+ E \left[ \sum_{v \neq w \in T^{Bb}} \ell_v (v) \ell_w (w) E^v [L_v | T^{Bb}] E^w [L_w | T^{Bb}] \right] 
\]

\[
\leq C(p - p_c)^{-4} E[\tau^{Bb}] + C(p - p_c)^{-2} E[(\tau^{Bb})^2], \quad (8.20) 
\]

where, for the third equality, we have used that traps at distinct vertices of the backbone tree are independent, and for the final inequality we used (8.17). Applying this upper bound to \( \tau_1^{Bb} \), which denotes the number of steps of \( (X_n^{Bb})_n \) until \( \tau_1 \), the upper bounds in Lemma 4.5 will therefore follow once we show

\[
E[\tau_1^{Bb}]^2 \leq C(p - p_c)^{-4}. \quad (8.21) 
\]

Since \( |X_n^{Bb}| \) stochastically dominates a simple random walk on \( \mathbb{N}_0 \) reflected at the origin, \( \eta_m^{Bb} \) is stochastically dominated by the time the simple random walk hits the vertex \( m[L/(p - p_c)] \). We have \( E[(\eta_m^{Bb})^k] \leq C m^{2k} (p - p_c)^{-2k} \) for any \( k \geq 1 \), so

\[
E[\tau_1^{Bb}]^2 = \sum_{m=1}^{\infty} E[(\eta_m^{Bb})^2 1\{X_{\tau_1} \in G_m\}] 
\]

\[
\leq \sum_{m=1}^{\infty} E[(\eta_m^{Bb})^2]^{1/2} P(X_{\tau_1} \in G_m)^{1/2} 
\]

\[
\leq C(p - p_c)^{-2q} \sum_{m=1}^{\infty} m^{2q} P(X_{\tau_1} \in G_m)^{1/2}. \quad (8.22) 
\]

By the moment bound of Lemma 4.4, \( P((p - p_c)|X_{\tau_1}| \geq m) \leq C m^{-4q-4} \). This implies (8.21), which concludes the proof. \( \square \)

9. The maximal displacement of a trap: proof of Lemma 5.1

The claimed moment bound follows from the following bound for the projected processes,

\[
\max_{1 \leq \delta \leq \Delta - 1} E \left[ \sup_{w \in T^{\text{trap}}} \varphi(w) \cdot e \left| \text{deg}_{T^{\text{trap}}}(\varphi) = \delta \right. \right] \leq C. \quad (9.1) 
\]
for constants $\gamma, C$, arbitrary unit vectors $e$ with $\|e\| = 1$ and where, without loss of generality, we consider a trap at the root, such that $\varphi(\varrho) = 0$.

Recall from the decomposition in (2.1) that $T_{\varrho}^{\text{trap}}$ consists of at most $\Delta - 1$ GW-trees attached by a single edge to $\varrho$, each tree having generating function $f^*$ and mean number $\mu^*_p$ of offspring. Thus, it suffices to show (9.1) with $T_{\varrho}^{\text{trap}}$ replaced by a single tree $T^*$ with generating function $f^*$. Let $W$ denote a $\mathbb{Z}^d$-valued random variable with distribution $D$, and write $K(s) = \mathbb{E}[s^{W \cdot e}]$ for the probability generating function of $W \cdot e$. For $x > 1$, let $\theta(x)$ be the unique solution in $(1, \infty)$ to

$$K(\theta(x)) = x. \quad (9.2)$$

Neuman and Zheng [NZ17, Theorem 1.2] show that

$$\limsup_{n \to \infty} \frac{\theta((\mu^*_p)^{-1} n)}{\mathbb{P}(\sup_{w \in T^*} \varphi(w) \cdot e \geq n)} \leq 1 \quad (9.3)$$

(it is here that we need the assumption $\sum_{x \in \mathbb{Z}^d} e^{\|x\|} D(x) < \infty$). The assertion of Lemma 5.1 thus easily follows if it holds that

$$\theta((\mu^*_p)^{-1} n^{-1/2}) \leq e^{-cn}, \quad (9.4)$$

for $n$ sufficiently large, uniformly in $p$ close enough to $p_c$. We proceed by showing that this is the case.

From Lemma 2.1 we obtain

$$(\mu^*_p)^{-1} = 1 + c(p - p_c)(1 + o(1)). \quad (9.5)$$

Since $\mathbb{E}[W] = \sum_{x \in \mathbb{Z}^d} x D(x) = 0$ by our assumption in Theorem 1.2 we have $\mathbb{E}[W \cdot e] = 0$, which implies $K'(1) = 0$, so that expansion around $s = 1$ gives

$$K(s) \leq 1 + c(s - 1)^2, \quad (9.6)$$

for $s$ close enough to $1$. This in turn implies that for $x$ close enough to $1$,

$$\theta(x) \geq 1 + c(x - 1)^{1/2}. \quad (9.7)$$

Combining (9.5) and (9.7), as well as $\mu^{-1} \geq 1 + c(p - p_c)$, we arrive at

$$\theta((\mu^{-1}_p)^{-1} n^{-1/2}) \leq (1 + c(p - p_c)^{1/2})^{-n(p - p_c)^{-1/2}} \leq e^{-cn}, \quad (9.8)$$

for $p$ close enough to $p_c$. This implies (9.4), which finished the proof.

10. The Trace of the Walk on the Backbone: Proof of Lemma 5.2

We will show a general moment bound for the trace until the second regeneration, that is, a moment bound for the cardinality of

$$\text{BBT}_1 \cup \text{BBT}_2 = \{X_0, X_1, \ldots, X_\tau_2\} \cap \mathcal{T}^{\text{BB}} = \{X_0^{\text{BB}}, X_1^{\text{BB}}, \ldots, X_\tau_2^{\text{BB}}\}. \quad (10.1)$$

The bound on $\mathbb{E}[\| \text{BBT}_k \|]$ for any $k \geq 1$ then follows from Lemma 4.3. We denote by $\mathcal{B} = \mathcal{B}(\mathcal{T})$ the set of branch points of the backbone tree, that is,

$$\mathcal{B}(\mathcal{T}) := \{\varrho\} \cup \{v \in \mathcal{T}^{\text{BB}} : \deg_{\mathcal{T}^{\text{BB}}}(v) \geq 3\}. $$
The bound for the trace on the backbone tree will follow from a bound on the number of visited branch points of the backbone. Between those branchpoints, the random walk has to travel across a section of the tree isomorphic to a line segment. For \( v \in T^{\text{Bb}} \), let \( \text{anc}_{B(T)}(v) \) denote the first vertex on the shortest path between \( v \) and \( q \) that is in \( B(T) \). Given \( v \in B(T) \), we call the path in \( T \) between \( v \) and \( \text{anc}_{B(T)}(v) \) the backbone branch to \( v \).

Fix a small \( \delta > 0 \) (to be determined below) and let \( S(T) \) be the set of all vertices \( v \in T^{\text{Bb}} \) for which there exists a path connecting \( v \) to the root \( q \), such that for any vertex \( w \) on this path,

\[
d_T(w, \text{anc}_{B(T)}(w)) \leq \lfloor \delta/(p - p_c) \rfloor,
\]

i.e., \( v \in S(T) \) if the path from \( v \) to \( \text{anc}_{B(T)}(v) \) and all backbone branches along the path to \( q \) have length at most \( \lfloor \delta/(p - p_c) \rfloor \).

The motivation for these definitions is as follows: Until the random walk exits \( S(T) \), the trace of the random walk on the backbone is a subset of \( S(T) \), and in order to exit \( S(T) \) the random walk has to cross a backbone branch of length at least \( \lfloor \delta/(p - p_c) \rfloor \). We can find a lower bound for the time it takes to traverse such a backbone branch, which implies an upper bound for the size of the trace on the backbone until that time. When the random walk exits \( S(T) \) and crosses a long backbone branch, it enters a new, independent subtree, and in this new tree we can iterate this upper bound.

So set \( S_1 := S(T) \) and let \( E_1 := H^{\text{Bb}}_{S_1} \) be the exit time of the set \( S_1 \) for the random walk \( (X_n^{\text{Bb}})_{n \geq 1} \) restricted to the backbone tree. Define recursively for \( k > 1 \),

\[
S_k := S_{k-1} \cup S(T_{X_{E_{k-1}}}) \quad \text{and} \quad E_k := H^{\text{Bb}}_{S_k}.
\]

(10.2)

Then \( E_k \) stochastically dominates the sum \( \tilde{E}_k = \tilde{E}_1^{(1)} + \cdots + \tilde{E}_1^{(k)} \) of \( k \) i.i.d. copies of \( \tilde{E}_1 = \inf\{n \geq 0 : Y_n = \lfloor \delta/(p - p_c) \rfloor \} \),

(10.3)

where \( (Y_n)_{n \geq 1} \) is a simple random walk on \( \{0, \ldots, \lfloor \delta/(p - p_c) \rfloor \} \) starting in and reflected at 0. Moreover, \( |S_k| \) stochastically dominates \( \{|X_0, \ldots, X_{E_k}| \cap T^{\text{Bb}}\} \). Since at each time \( E_i \) the set \( S(T_{X_{E_i}}) \) is independent of the tree explored so far, we can bound

\[
\mathbb{E}[|\{X_0, \ldots, X_{E_k}\} \cap T^{\text{Bb}}|^q] \leq \mathbb{E}[|S_k|^q] \leq \mathbb{E}\left[\left(\sum_{i=1}^k \tilde{S}^{(i)}\right)^q\right] \leq k^q \mathbb{E}[|S_1|^q],
\]

(10.4)

where we wrote \( \tilde{S}^{(i)} \) for i.i.d. copies of \( S_1 \).

The next step is to show that, for any \( q \geq 1 \),

\[
\mathbb{E}[|S_1|^q] \leq C_q(p - p_c)^{-q}.
\]

(10.5)

Recall that \( \Delta \) is the maximal number of children in the original Galton-Watson tree. By definition of \( S_1 \),

\[
|S_1| \leq \Delta \cdot |S_1 \cap B(T)| \cdot \lfloor \delta/(p - p_c) \rfloor,
\]

(10.6)

where we have used that \( \Delta \), the maximal number of offspring in the unpercolated tree, is finite. By construction, the set \( |S_1 \cap B(T)| \) is stochastically dominated by the total
progeny of a Galton-Watson process, denoted by \( \tilde{Z} \), with offspring distribution given by the law of

\[
\tilde{\xi} := \sum_{i=1}^{\Delta} \mathbb{1}_{\{\ell_i \leq \lfloor \delta/(p - p_c) \rfloor\}}
\]  

(10.7)

where \( \ell_i \) are independent random variables corresponding to the length of a backbone branch, so that each \( \ell_i \) is an independent Geometric random variable, having mean \( c(p - p_c)^{-1} \) (see Remark 2.2). Then \( \tilde{\xi} \) is a Binomial random variable with \( \Delta \) trials and probability of success

\[
P(\ell_i \leq \lfloor \delta/(p - p_c) \rfloor) \leq 1 - (1 - c(p - p_c))^{\lfloor \delta/(p - p_c) \rfloor} \leq 1 - e^{-c\delta},
\]  

(10.8)

where we have used Lemma 2.1. This means that we can choose \( \delta > 0 \) small enough so that

\[
E[\tilde{\xi}^2] \leq c < 1  
\]  

(10.9)

uniformly in \( p \), i.e., \( \tilde{Z} \) is a subcritical Galton-Watson process. Writing \( \tilde{Z}_n \) for the size of the \( n \)-th generation of \( \tilde{Z} \) and using

\[
E[|\tilde{Z}|^q] \leq \sum_{N=0}^{\infty} \sum_{n=0}^{N} N^{q-1} E\left[|\tilde{Z}_n|^{q/2}\mathbb{1}_{\{\tilde{Z}_N \neq 0, \tilde{Z}_{N+1} = 0\}}\right]  
\]  

(10.10)

Applying the Cauchy-Schwarz inequality and using \( P(\tilde{Z}_N \neq 0) \leq E[\tilde{Z}_N] = E[\tilde{\xi}]^N \),

\[
E[|\tilde{Z}|^q] \leq \sum_{N=0}^{\infty} \sum_{n=0}^{N} N^{q-1} E[\tilde{Z}_n^{2q}]^{1/2} P(\tilde{Z}_N \neq 0)^{1/2}  
\]  

(10.11)

Since \( E[\tilde{\xi}] \leq E[\tilde{\xi}^{2q}] < 1 \) uniformly in \( p \), (10.10) is bounded uniformly in \( p \) also. This implies that by the Cauchy-Schwarz inequality, (10.4) and (10.5),

\[
E[|S_1 \cap B(T)|^q] \leq E[|\tilde{Z}|^q] \leq C_q
\]  

(10.12)

and by (10.6), (10.12) shows that (10.5) holds.
So far, we have shown that (10.4) is bounded by $C_q k^q (p - p_c)^{-q}$. In order to turn this into a bound for (10.1), we note that

\[
\mathbb{E}[\{X_{0}^{\mathbb{B}b}, X_{1}^{\mathbb{B}b}, \ldots, X_{\tau_{2}^{\mathbb{B}b}}^{\mathbb{B}b}\}]^{\gamma/2} \leq \sum_{k=1}^{\infty} \mathbb{E}[\{X_{0}^{\mathbb{B}b}, X_{1}^{\mathbb{B}b}, \ldots, X_{E_{k}}^{\mathbb{B}b}\}]^{\gamma/2} \mathbb{P}(E_{k-1} \leq \tau_{2}^{\mathbb{B}b} < E_{k})^{1/2}
\]

\[
\leq C_q (p - p_c)^{-q} \sum_{k=1}^{\infty} k^{q} \mathbb{P}(E_{k-1} \leq \tau_{2}^{\mathbb{B}b})^{1/2}
\]

where we set $E_0 = 0$. It remains to show that the sum is bounded.

Consider first the random variable $\bar{E}_k$, which is stochastically dominated by $E_k$ and is a sum of $k$ i.i.d. random variables each distributed as $\bar{E}_1$ in (10.3). Letting the random walk $(Y_n)_{n \geq 1}$ in the definition (10.3) run for $(p - p_c)^{-2}$ steps and using the scaling limit of a simple random walk, it is easy to see that $(p - p_c)^2 \bar{E}_1$ dominates a Bernoulli distributed random variable with a success probability independent of $p$. This implies that by Cramér’s Theorem, there is thus a $\gamma > 0$ such that

\[
\mathbb{P}(\bar{E}_k \leq k(p - p_c)^{-2}\gamma) \leq Ce^{-ck}.
\]  

(10.14)

Since $E_k$ dominates $\bar{E}_k$, we also get

\[
\mathbb{P}(E_k \leq k(p - p_c)^{-2}\gamma) \leq Ce^{-ck}.
\]  

(10.15)

For this choice of $\gamma$, we use that

\[
\mathbb{P}(E_k \leq \tau_{2}^{\mathbb{B}b}) \leq \mathbb{P}(E_k \leq k(p - p_c)^{-2}\gamma) + \mathbb{P}(\tau_{2}^{\mathbb{B}b} \geq k(p - p_c)^{-2}\gamma).
\]  

(10.16)

On the other hand, the arguments in (8.22) with $\tau_1^{\mathbb{B}b}$ replaced by $\tau_{2}^{\mathbb{B}b}$ show that

\[
(p - p_c)^2 \mathbb{E}[\tau_{2}^{\mathbb{B}b}]^{2q+4} \leq C \cdot \infty,
\]

so Markov’s inequality implies

\[
\mathbb{P}(\tau_{2}^{\mathbb{B}b} \geq k(p - p_c)^{-2}\gamma) \leq Ck^{-2q-4}.
\]  

(10.17)

The bounds (10.15) and (10.17) thus together with (10.16) imply

\[
\mathbb{P}(E_k \leq \tau_{2}^{\mathbb{B}b}) \leq Ck^{-2q-4},
\]  

(10.18)

which, when inserted into (10.13), proves the first assertion of Lemma 5.2.

To prove (5.14), the second assertion of Lemma 5.2, we first observe that the number of visited branchpoints until time $E_k$ satisfies

\[
\mathbb{E}[\{X_0, \ldots, X_{E_k}\} \cap \mathcal{T}^{\mathbb{B}b} \cap \mathcal{B}(\mathcal{T})]^q] \leq k^q \mathbb{E}[^{\mathcal{S}_1} \cap \mathcal{B}(\mathcal{T})]^q],
\]  

(10.19)

and similarly to (10.2), and by (10.12), this moment is bounded by $C_q k^q$. If $\mathcal{N}_{E_k}$ denotes the number of distinct backbone branches visited by $(X_n^{\mathbb{B}b})_{n \geq 0}$ before time $E_k$, then $\mathcal{N}_{E_k}$ is bounded by the number of visited branchpoints multiplied by the bound $\Delta$ for the number of offspring, which implies

\[
\mathbb{E}[(\mathcal{N}_{E_k})^q] \leq C_q \Delta k^q.
\]  

(10.20)
By the same arguments as in the preceding paragraph, we may then go on to conclude that

$$\mathbb{E}[(\mathcal{N}_{gb})^q] \leq Cq\Delta. \quad (10.21)$$

Recall that the embedding \( \varphi \) of the tree \( T \) into \( \mathbb{R}^d \) is generated by the random walk one-step distribution \( D \). Following the branching random walk interpretation, we can think of the process as a random walker that splits into \( \xi \) random particles at each step. Therefore, we arrive at an upper bound for \( \|\varphi(v) - \varphi(X_{\tau_1})\| \) by considering the maximal displacement among \( \mathcal{N}_{gb} \) independent \(|BBT_2|\)-step random walks started at 0 with step distribution \( D \). Let \((W_n^{(i)})_{n \geq 0}\) for \( 1 \leq i \leq \mathcal{N}_{gb} \) denote these independent random walks. Then

$$\mathbb{E} \left[ \max_{v \in BBT_2} \|\varphi(v) - \varphi(X_{\tau_1})\|^q \right] \leq \mathbb{E} \left[ \max_{1 \leq i \leq \mathcal{N}_{gb}} \max_{0 \leq k \leq m \leq |BBT_2|} \|W^{(i)}_k - W^{(i)}_m\|^q \right]
$$

$$\leq \mathbb{E} \left[ \mathcal{N}_{gb} \right] \mathbb{E} \left[ \max_{0 \leq k \leq m \leq |BBT_2|} \|W^{(1)}_k - W^{(1)}_m\|^q \right]
$$

$$\leq C \mathbb{E} \left[ \mathcal{N}_{gb} \right] \mathbb{E} \left[ |BBT_2|^{q/2} \right] \quad (10.22)$$

where we bounded the maximum over \( i \) by the sum and used Wald’s identity. For the last inequality we have used the independence between the increments of the embedding and \( |BBT_2| \), and that \( \mathbb{E} \left[ \max_{0 \leq k \leq m \leq N} \|W^{(1)}_k - W^{(1)}_m\|^q \right] \leq CN^{q/2} \) by our assumption that \( D \) has finite exponential moments. By (10.21), the first expectation in (10.22) is bounded by a constant, and by (5.13) the second expectation is bounded by \( C(p - p_c)^{q/2} \). This completes the proof of Lemma 5.2. \( \square \)

11. The speed of the walk: proof of Theorem 1.1 and Lemma 2.1

In this section we prove Theorem 1.1 and Lemma 2.1. We start with Lemma 2.1, and also derive a result that will be crucially used in the proof of Theorem 1.1. In the following, let

$$m_{p,k} := f_p^{(k)}(1) = \mathbb{E}[\xi_p(\xi_p - 1) \cdots (\xi_p - k + 1)]$$

denote the \( k \)-th factorial moment of \( \xi_p \). We abbreviate \( m_k = m_{1,k} \), noting that \( m_{p,k} = p^k m_k \). Recall that \( q_p \) is the extinction probability, i.e., the unique fixed point of \( f_p \) in \((0, 1)\).

**Proof of Lemma 2.1.** We start by deriving the generating function \( f_p \), for which we use that, conditioned on \( \{\xi = n\} \), the number of offspring \( \xi \) in the pruned tree is distributed as a Binomial with \( n \) trials and success probability \( p \). Therefore,

$$f_p(s) = \mathbb{E}[\xi_p^s] = \sum_{k=0}^{\infty} \mathbb{E}[\xi_p^s \mid \xi = n] p_k$$

$$= \sum_{k=0}^{\infty} (ps + (1 - p))^k p_k = f(ps + (1 - p)), \quad (11.1)$$
where $f$ is the generating function of the offspring distribution $(p_k)_{k \geq 0}$. Thus, we obtain that

$$f_p'(0) = pf'(1 - p) \geq c_0 > 0,$$  \hspace{1cm} (11.2)

uniformly for $p \geq p_c$, as required. Furthermore, setting $s = 1$ yields $f'_p(1) = pf'(1) = pm_1 = p/p_c$. This proves the two equalities in (2.3).

We proceed with the three asymptotic relations in (2.4), where we start with the first.

Since $q_p$ is the fixed point of $f_p$, we have

$$1 - q_p = 1 - f_p(q_p) = 1 - f(pq_p + 1 - p) = 1 - f(1 - p(1 - q_p)).$$  \hspace{1cm} (11.3)

Taking the Taylor expansion of the right-hand side in $1 - q_p$ and using that $q_p \to 1$ as $p \to p_c$ yields

$$1 - q_p = pf'(1)(1 - q_p) - \frac{1}{2}p^2 f''(1)(1 - q_p)^2 + \frac{1}{6}p^3 f'''(1)(1 - q_p)^3 + o((1 - q_p)^3).$$  \hspace{1cm} (11.4)

Here we used our assumption that $E[\xi^3] < \infty$. Using that $f^{(k)}(1) = m_k$ and dividing through by $1 - q_p > 0$ we obtain

$$\frac{1}{2}p^2m_2(1 - q_p) = (pm_1 - 1) + \frac{1}{6}p^3m_3(1 - q_p)^2 + o((1 - q_p)^2).$$  \hspace{1cm} (11.5)

We will use these asymptotics in the proof of Theorem 1.1 below. Replacing the last two terms by $O((1 - q_p)^2)$, the first asymptotics in (2.4) follows, as

$$1 - q_p = \frac{2}{m_2p^2}(pm_1 - 1) + O((1 - q_p)^2) = \frac{2m_3}{m_2}(p - p_c) + O((1 - q_p)^2),$$  \hspace{1cm} (11.6)

with $c_2 = 2m_3/m_2$, since $p = p_c + O(p - p_c) = 1/m_1 + O(p - p_c)$.

To prove the second asymptotics in (2.4), we note that

$$f'_p(q_p) = pf'(q_p) = pf'(1 - (1 - q_p)) = pf'(1) - pf''(1)(1 - q_p) + O((1 - q_p)^2),$$

so that, by (11.6),

$$f'_p(q_p) = pm_1 - (1 - q_p)pm_2 + O((1 - q_p)^2)$$

$$= 1 + m_1(p - p_c) - 2m_1(p - p_c) + O((p - p_c)^2)$$

$$= 1 - (p - p_c)(2m_1^2 - m_1) + O((p - p_c)^2).$$  \hspace{1cm} (11.8)

Note that $c_3 = 2m_1^2 - m_1 > 0$, since $m_1 > 1$. The third asymptotics in (2.4) similarly follows, since

$$f''_p(0) = (1 - q_p)f''_p(q_p) = (1 - q_p)p^2f''_p(q_p)$$

$$= (1 - q_p)p^2c f''_p(1)(1 + o(1)) = m_2/m_1^2(p - p_c)(1 + o(1)).$$

We proceed with the proof of Theorem 1.1, for which we rely on an explicit formula for the effective speed due to Lyons, Pemantle and Peres [LPP95, Page 601],

$$v(p) = \sum_{k=0}^{\infty} \frac{k - 1}{k + 1} p_k(p) \frac{1 - q_p^{k+1}}{1 - q_p^2},$$

with $p_k(p) = P_p(\xi_p = k)$. Using Lemma 2.1, we expand this expression for $p$ close to $p_c$. 

\end{proof}
Proof of Theorem 1.1. We again expand in terms of $1 - q_p$. We rewrite

$$ (1 + q_p)v(p) = \sum_{k=0}^{\infty} \frac{k - 1}{k + 1} p_k(p) \sum_{i=0}^{k} q^i_p = \sum_{k=0}^{\infty} \frac{k - 1}{k + 1} p_k(p) \sum_{i=0}^{k} (1 - (1 - q_p))^i. \quad (11.9) $$

Substituting the expansion

$$ (1 - (1 - q_p))^i = 1 - i(1 - q_p) + \frac{1}{2}i(i - 1)(1 - q_p)^2 + o(i^2(1 - q_p)^2). \quad (11.10) $$
gives three terms and an error term. Using $p_c = 1/m_1 = 1/E[\xi]$, the first term equals

$$ \sum_{k=0}^{\infty} \frac{k - 1}{k + 1} p_k(p)(k + 1) = E_p[\xi_p - 1] = pE[\xi] - 1 = m_1(p - p_c) \quad (11.11) $$

Using $\sum_{i=0}^{k} i = \frac{1}{2}k(k + 1)$, the second equals

$$ - \frac{1}{2}(1 - q_p) \sum_{k=0}^{\infty} k(k - 1)p_k(p) = -\frac{1}{2}(1 - q_p)m_{p,2} = -\frac{1}{2}(1 - q_p)p^2m_2. \quad (11.12) $$

Using $\sum_{i=0}^{k} i(i - 1) = \frac{1}{3}(k + 1)k(k - 1)$, the third equals

$$ \frac{1}{2}(1 - q_p)^2 \sum_{k=0}^{\infty} \frac{k - 1}{k + 1} p_k(p) \sum_{i=2}^{k} i(i - 1) \quad (11.13) $$

$$ = \frac{1}{2}(1 - q_p)^2 \sum_{k=0}^{\infty} \frac{k - 1}{k + 1} p_k(p) \frac{1}{3}(k + 1)k(k - 1) $$

$$ = \frac{1}{6}(1 - q_p)^2 \left( \sum_{k=0}^{\infty} k(k - 1)(k - 2)p_k(p) + \sum_{k=0}^{\infty} k(k - 1)p_k(p) \right) $$

$$ = \frac{1}{6}(1 - q_p)^2 (m_{p,3} + m_{p,2}). $$

Using $E[\xi^3] < \infty$ and $|(k - 1)/(k + 1)| \leq 1$, the error term can be estimated as

$$ o((1 - q_p)^2) \sum_{k=0}^{\infty} p_k(p) \sum_{i=0}^{k} i^2 = o((1 - q_p)^2). \quad (11.14) $$

We conclude that

$$ (1 + q_p)v(p) = m_1(p - p_c) - \frac{1}{2}(1 - q_p)p^2m_2 $$

$$ + \frac{1}{6}(1 - q_p)^2 (m_{p,3} + m_{p,2}) + o((1 - q_p)^2). \quad (11.15) $$

Substituting (11.5) for the second term, we obtain

$$ m_1(p - p_c) - \frac{1}{2}(1 - q_p)p^2m_2 = -\frac{1}{6}p^3m_3(1 - q_p)^2 + o((p - p_c)^2), \quad (11.16) $$

which cancels up to leading order with the term $\frac{1}{6}(1 - q_p)^2 m_{p,3}$. (We have no intuitive explanation for why the third moment drops out.)
By (11.6), we arrive at

\[
(1 + q_p)v(p) = \frac{1}{6}(1 - q_p)^2 m_{p,2} + o((p - p_c)^2) \tag{11.17}
\]

\[
= \frac{2}{3}(p - p_c)^2 m_{10}^4 m_{20}^4 + o((p - p_c)^2).
\]

Using that \(1 + q_p = 2 + O(p - p_c)\) completes the proof. \(\square\)

**Acknowledgements.** The work of JN was supported by the Deutsche Forschungsgemeinschaft (DFG) through grant NA 1372/1. RvdH was supported by NWO through VICI-grant 639.033.806. The work of RvdH and TH in this paper is also supported by the Netherlands Organisation for Scientific Research (NWO) through Gravitation-grant NETWORKS-024.002.003.

**References**

[AF02] David Aldous and Jim Fill. Reversible markov chains and random walks on graphs. available at https://www.stat.berkeley.edu/~aldous/RWG/book.pdf, 2002.

[Ald91] David Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.

[Ald93] David Aldous. Tree-based models for random distribution of mass. *J. Statist. Phys.*, 73(3-4):625–641, 1993.

[AN72] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[BACF16a] Gérard Ben Arous, Manuel Cabezas, and Alexander Fribergh. Scaling limit for the ant in a simple labyrinth. arXiv:1609.03980, 2016.

[BACF16b] Gérard Ben Arous, Manuel Cabezas, and Alexander Fribergh. Scaling limit for the ant in high-dimensional labyrinths. arXiv:1609.03977, 2016.

[BAF16] Gérard Ben Arous and Alexander Fribergh. Biased random walks on random graphs. In *Probability and statistical physics in St. Petersburg*, volume 91 of *Proc. Sympos. Pure Math.*, pages 99–153. Amer. Math. Soc., Providence, RI, 2016.

[BB07] Noam Berger and Marek Biskup. Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Related Fields*, 137(1-2):83–120, 2007.

[BGP03] Noam Berger, Nina Gantert, and Yuval Peres. The speed of biased random walk on percolation clusters. *Probab. Theory Related Fields*, 126(2):221–242, 2003.

[Bil56] Patrick Billingsley. The invariance principle for dependent random variables. *Trans. Amer. Math. Soc.*, 83:250–268, 1956.

[Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[CR99] Hui Chen and Joseph P. Romano. An invariance principle for triangular arrays of dependent variables with application to autocovariance estimation. *Canad. J. Statist.*, 27(2):329–343, 1999.

[Cro09] David A. Croydon. Hausdorff measure of arcs and Brownian motion on Brownian spatial trees. *Ann. Probab.*, 37(3):946–978, 2009.

[CRR+97] Ashok K. Chandra, Prabhakar Raghavan, Walter L. Ruzzo, Roman Smolensky, and Prasoon Tiwari. The electrical resistance of a graph captures its commute and cover times. *Comput. Complexity*, 6(4):312–340, 1996/97.

[CS15] Lung-Chi Chen and Akira Sakai. Critical two-point functions for long-range statistical-mechanical models in high dimensions. *Ann. Probab.*, 43(2):639–681, (2015).

[dG76] Pierre Gilles de Gennes. Percolation: un concept unificateur. *La Recherche*, 7:919–927, 1976.
Random walk on barely supercritical branching random walk

[KV86] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. Comm. Math. Phys., 104(1):1–19, 1986.

[LG05] Jean-François Le Gall. Random trees and applications. Probab. Surv., 2:245–311, 2005.

[LP16] Russell Lyons and Yuval Peres. Probability on trees and networks, volume 42 of Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, New York, 2016.

[LPP95] Russell Lyons, Robin Pemantle, and Yuval Peres. Ergodic theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure. Ergodic Theory Dynam. Systems, 15(3):593–619, 1995.

[LPP96] Russell Lyons, Robin Pemantle, and Yuval Peres. Biased random walks on Galton-Watson trees. Probab. Theory Related Fields, 106(2):249–264, 1996.

[Lyo92] Russell Lyons. Random walks, capacity and percolation on trees. Ann. Probab., 20(4):2043–2088, 1992.

[MP07] P. Mathieu and A. Piatnitski. Quenched invariance principles for random walks on percolation clusters. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 463(2085):2287–2307, 2007.

[NZ17] Eyal Neuman and Xinghua Zheng. On the maximal displacement of subcritical branching random walks. Probab. Theory Related Fields, 167(3-4):1137–1164, 2017.

[PS08] Yuval Peres and Ofer Zeitouni. A central limit theorem for biased random walks on Galton–Watson trees. Probability Theory and Related Fields, 140(3-4):595–629, 2008.

[Sah94] M. Sahimi. Applications of percolation theory. Taylor & Francis, (1994).

[Slad02] G. Slade. Scaling limits and super-Brownian motion. Notices Amer. Math. Soc., 49(9):1056–1067, (2002).

[Slad06] G. Slade. The lace expansion and its applications, volume 1879 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, (2006). Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, Edited and with a foreword by Jean Picard.

[SS04] Vladas Sidoravicius and Alain-Sol Sznitman. Quenched invariance principles for walks on clusters of percolation or among random conductances. Probab. Theory Related Fields, 129(2):219–244, 2004.

[Sta85] Dietrich Stauffer. Introduction to percolation theory. Taylor & Francis, Ltd., London, 1985.

[SZ99] Alain-Sol Sznitman and Martin Zerner. A law of large numbers for random walks in random environment. Annals of probability, pages 1851–1869, 1999.

Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, the Netherlands.

E-mail address: r.w.v.d.hofstad@tue.nl, w.j.t.hulshof@tue.nl, j.h.nagel@tue.nl