Bernoulli **14**(1), 2008, 277–299
DOI: 10.3150/07-BEJ107

Small value probabilities via the branching tree heuristic

PETER MÖRTERS* and MARCEL ORTGIESE**

Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, England.
E-mail: *maspm@bath.ac.uk; **ma2mo@bath.ac.uk

In the first part of this paper, we give easy and intuitive proofs for the small value probabilities of the martingale limit of a supercritical Galton–Watson process in both the Schröder and the Böttcher cases. These results are well known, but the most cited proofs rely on generating function arguments which are hard to transfer to other settings. In the second part, we show that the strategy underlying our proofs can be used in the quite different context of self-intersections of stochastic processes. Solving a problem posed by Wenbo Li, we find the small value probabilities for intersection local times of several Brownian motions, as well as for self-intersection local times of a single Brownian motion.

Keywords: branching process; Brownian motion; embedded random walk; embedded tree; intersection local time; intersection of Brownian motions; local time; lower tail; martingale limit; random tree; self-intersection local time; small ball problem; small deviation; supercritical Galton–Watson process

1. Introduction

The *small value problem* is to find, for a non-negative random variable X, the speed of decay of the left tail $\mathbb{P}\{X < \varepsilon\}$ as $\varepsilon \downarrow 0$. Important examples are the *small ball problem*, where X is the norm of a random variable with values in a Banach space, the *lower level problem*, where X is the maximum of a continuous random process $(X(t): t \in [0, 1])$ and *boundary crossing problems*, where X is the first exit time of a stochastic process from a general space–time domain.

Small value problems arise in a great variety of contexts in probability and analysis. Examples include approximation and quantization problems (Li and Linde (1999), Dereich *et al.* (2003) and Graf *et al.* (2003)), Brownian pursuit problems (Li and Shao (2001a)), polymer measures (Hofstad *et al.* (1997)) and convex geometry (Klartag and Vershynin (2007)). A systematic theory of small value problems, however, is only available when X is the norm of a Gaussian random variable. For other cases, some isolated techniques are known, but a bigger picture has not yet emerged. A survey of Gaussian methods in this field is Li and Shao (2001b) and an updated bibliography on small value problems is maintained at Lifshits (2006).

In this paper, we contribute to the theory of small value problems by systematically presenting an approach which we found successful in a variety of cases. We illustrate our technique by three main examples. The first example is the most natural one for our approach – the martingale limit of a supercritical Galton–Watson process. In this case, the small value problem has been solved – by Dubuc (1971a, 1971b) in the Schröder case and, up to a Tauberian theorem of Bingham (1988), also in the Böttcher case. These proofs use an integral transformation approach, together with some non-trivial complex analysis, a powerful method, but inflexible and not very intuitive.

1350-7265

Our method, by contrast, is very simple and based on easy intuition. From this example, we derive the term *branching tree heuristic* for the general approach.

The second example is our main result and is treated here for the first time: we solve a problem posed by Wenbo Li at the mini-workshop "Small deviation probabilities and related topics" at Oberwolfach in October 2003. The problem is to identify the small value probability of the random variable

$$X = \int_{-\infty}^{\infty} \prod_{i=1}^{m} L_i^{q_i}(x, 1)\, \mathrm{d}x,$$

where $L_1(x,t), \ldots, L_m(x,t)$ are the local times of $m \geq 2$ independent Brownian motions. We explain very carefully how a heuristic embedding of a tree in the Brownian motion framework leads to a proof based on the same principles as in the Schröder case of the first example.

Our third example also appears to be new, although it is really quite elementary. We consider the L^q-norm of the local time of a single Brownian motion stopped when it exits a bounded interval for the first time, which, for q an integer, may be interpreted as the q-fold self-intersection local time of the motion. We again find a relation to a Galton–Watson tree, this time of Böttcher type, and exploit this relation to find a strikingly simple proof of the small value probability.

We believe that our method can be used in a number of further cases, when the optimal strategy for a random variable to obtain small values is inhomogeneous. We conclude the paper with a discussion of possible future research.

2. Small value probabilities for the martingale limit of a Galton–Watson tree

Consider a Galton–Watson branching process $(Z_n : n \geq 0)$ with offspring distribution $(p_k : k \geq 0)$ starting with a single founding ancestor, called ρ, in generation 0. We suppose that the offspring variable N is non-degenerate and satisfies $\mu := \mathbb{E}N > 1$ and $\mathbb{E}[N \log N] < \infty$. By the famous Kesten–Stigum theorem, these conditions ensure that the martingale limit

$$W := \lim_{n\to\infty} \frac{Z_n}{\mu^n}$$

exists and is non-trivial almost surely on survival. Except in the case when N is geometric, the distribution of W is not known explicitly and one relies on asymptotic results to describe its behaviour.

For the formulation of our results, we further assume that $p_0 = 0$, without loss of generality: removing all finite subtrees from a Galton–Watson tree does not change its martingale limit, but the resulting tree is still a Galton–Watson tree (with a modified offspring variable); see Athreya and Ney (1972), Chapter 1, Section 12.

As usual, we distinguish between the *Schröder* case and the *Böttcher* case, depending on whether $p_1 > 0$ or $p_1 = 0$. These two cases yield very different lower tail behaviours for W. In the following, $a(\varepsilon) \asymp b(\varepsilon)$ means that there exist constants $0 < c < C < \infty$ such that

$$ca(\varepsilon) \leq b(\varepsilon) \leq Ca(\varepsilon) \qquad \text{for all } 0 < \varepsilon < 1.$$

Theorem 1 (Dubuc (1971b)).

(a) *In the Schröder case, define* $\tau := -\log p_1/\log\mu > 0$. *Then,*

$$\mathbb{P}\{W<\varepsilon\} \asymp \varepsilon^{\tau}.$$

(b) *In the Böttcher case, define* $\nu := \min\{i \geq 0 : p_i \neq 0\} \geq 2$ *and* $\beta := \frac{\log\nu}{\log\mu} < 1$. *Then,*

$$-\log\mathbb{P}\{W<\varepsilon\} \asymp \varepsilon^{-\beta/(1-\beta)}.$$

In this paper, we offer simple proofs of both parts of Theorem 1 and show how the idea behind these proofs can be adapted to obtain small value probabilities for situations which might look quite different at first glance.

The main idea of the proofs is to understand the optimal strategy by which the tree keeps the generation size small. It turns out that the best strategy consists of producing as few offspring as possible at the beginning and then, once the necessary reduction in size is achieved, letting the tree grow normally. If the tree produces a larger number of children at the beginning, it will be more expensive to control the growth later on since every additional child is also likely to produce more than one child. This effect is illustrated in Figure 1.

By $(Z_n(v): n \geq 0)$, we denote the generation sizes of the subtree consisting of all of the descendants of the individual v. Note that for each fixed v, the process $(Z_n(v): n \geq 0)$ is again a Galton–Watson process and we can hence define the martingale limit

$$W(v) := \lim_{n\to\infty} \frac{Z_n(v)}{\mu^n}.$$

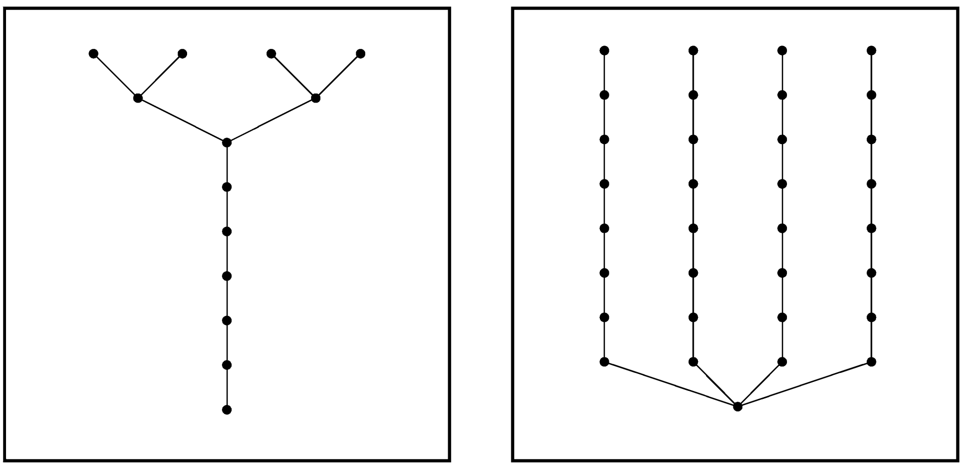

Figure 1. The picture on the left illustrates the optimal strategy to keep the final generation size small. By comparison, in the picture on the right, the offspring of more individuals must be kept under control to produce the same effect.

Let $v_k(1), \ldots, v_k(Z_k)$ be the individuals in the kth generation. By decomposing the individuals in the nth generation according to their ancestors in the kth generation, we obtain, for all $n \geq k$,

$$Z_n = \sum_{i=1}^{Z_k} Z_{n-k}(v_k(i)).$$

Hence, we obtain

$$W = \lim_{n\to\infty} \frac{Z_n}{\mu^n} = \lim_{n\to\infty} \mu^{-k} \sum_{i=1}^{Z_k} \frac{Z_{n-k}(v_k(i))}{\mu^{n-k}} = \mu^{-k} \sum_{i=1}^{Z_k} W(v_k(i)), \tag{2.1}$$

where all of the random variables $W(v_k(i))$ are i.i.d. with the same distribution as W.

This section is organized as follows. We first investigate the Schröder case. We start by showing that the suggested strategy is successful, which proves the lower bound. We then give a rough argument which produces the precise logarithmic asymptotics. This argument is then refined, exploiting the self-similarity of the tree, to complete the proof of Theorem 1(a). The arguments leading to the result in the Böttcher case, Theorem 1(b), are easier and are given in the final two subsections.

2.1. The Schröder case: the lower bound

For the lower bound, suppose that $0 < \varepsilon < 1$ and choose n such that $\mu^{-n} \leq \varepsilon < \mu^{-n+1}$. Using (2.1), we obtain

$$\begin{aligned}\mathbb{P}\{W < \varepsilon\} &\geq \mathbb{P}\{W < \mu^{-n} | Z_n = 1\}\mathbb{P}\{Z_n = 1\} \\ &= \mathbb{P}\{\mu^{-n} W(v_n(1)) < \mu^{-n}\} p_1^n = c p_1^n \geq (c p_1)\varepsilon^{\tau},\end{aligned}$$

where $c := \mathbb{P}\{W < 1\} > 0$.

2.2. The Schröder case: the logarithmic upper bound

As the first step in the proof of the upper bound, we show that

$$\limsup_{\varepsilon \downarrow 0} \frac{\log \mathbb{P}\{W < \varepsilon\}}{-\log \varepsilon} \leq -\tau. \tag{2.2}$$

Remark. In the second step of the argument, we only use the fact that $\mathbb{P}\{W < \varepsilon\}$ decreases like *some* positive power of ε. Other instances of our method, however, make use of lower bounds on this power, so it is instructive to show the 'best possible' argument here.

Fix a large m for the moment and let $n \geq m$. By decomposing the set of individuals in the nth generation of the branching process according to their last common ancestor with the 'spine'

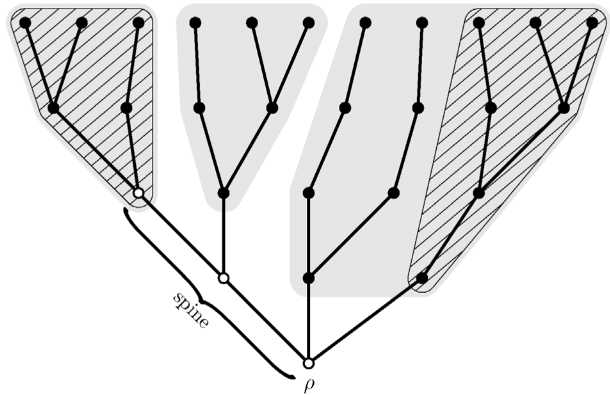

Figure 2. Decomposition of the tree according to the ancestry from a spine with length $m = 2$. The shaded parts of the tree are discarded in our calculation.

$\rho = v_0(1), v_1(1), v_2(1), \ldots, v_m(1)$ consisting of the leftmost individual in each of the first $m+1$ generations (see Figure 2 for illustration), we obtain a decomposition

$$Z_n = \sum_{k=1}^{m} \sum_{j=2}^{Z_1(v_{k-1}(1))} Z_{n-k}(v_k(j)) + Z_{n-m}(v_m(1)).$$

Discarding the contributions for $j \geq 3$, if they exist, and also the last summand, dividing by μ^n and letting $n \uparrow \infty$ gives

$$W \geq \sum_{k=1}^{m} \mu^{-k} W_k, \tag{2.3}$$

where $W_k = 0$ if $v_{k-1}(1)$ has only one offspring and $W_k = W(v_k(2))$ otherwise. Note that $W_1, \ldots, W_k$ are independent and identically distributed with distribution given by $\mathbb{P}\{W_k = 0\} = p_1$ and

$$\mathbb{P}\{W_k < x | W_k \neq 0\} = \mathbb{P}\{W < x\} \qquad \text{for all } x > 0.$$

Now suppose that $\delta > 0$ is given. As $W > 0$ almost surely, there exists $\theta > 0$ such that $\mathbb{P}\{W < \theta\} \leq \delta p_1$. We fix the integer ℓ such that $\mu^\ell \leq \theta < \mu^{\ell+1}$. Let $\varepsilon > 0$ be arbitrary and define n by $\mu^{-n-1} < \varepsilon \leq \mu^{-n}$. Then, using (2.3) for $m = n + \ell$,

$$\begin{aligned}\mathbb{P}\{W < \varepsilon\} &\leq \mathbb{P}\{W < \mu^{-n}\} \leq \mathbb{P}\left\{\sum_{k=1}^{m} \mu^{-k} W_k < \mu^{-n}\right\} \leq \prod_{k=1}^{m} \mathbb{P}\{W_k < \mu^{-n+m}\} \\ &\leq (p_1 + \mathbb{P}\{W < \theta\})^m \leq \big(p_1^\ell (1+\delta)^\ell\big) p_1^n (1+\delta)^n \leq C \varepsilon^\tau \mathrm{e}^{\delta n},\end{aligned}$$

for $C := p_1^\ell (1+\delta)^\ell \mu^\tau$, from which (2.2) follows, as $\delta > 0$ was arbitrary.

2.3. The Schröder case: up-to-constants asymptotics

We are now in a position to refine the upper bound and prove Theorem 1(a). Define a sequence $(a(n) : n \geq 0)$ by setting

$$a(n) := \mathbb{P}\{W < \mu^{-n}\} p_1^{-n}.$$

For arbitrary $0 < \varepsilon < 1$, we pick the integer $n \geq 0$ such that $\mu^{-n-1} \leq \varepsilon < \mu^{-n}$. Then,

$$\mathbb{P}\{W < \varepsilon\} \leq \mathbb{P}\{W < \mu^{-n}\} = a(n) p_1^n \leq a(n)(1/p_1)\varepsilon^\tau$$

and hence to complete the proof, it suffices to show that $(a(n) : n \geq 0)$ is bounded.

Denote by N_n the number of offspring of the leftmost individual in generation n and let

$$T := \min\{n \geq 0 : N_n \neq 1\}.$$

Obviously, $\mathbb{P}\{T = j\} = p_1^j(1 - p_1)$. Let $j < n$ be non-negative integers. Applying (2.1), we obtain

$$\begin{aligned} \mathbb{P}\{W < \mu^{-n}, T = j\} &\leq \mathbb{P}\big\{\mu^{-(j+1)}\big(W(v_{j+1}(1)) + W(v_{j+1}(2))\big) < \mu^{-n}, T = j\big\} \\ &\leq p_1^{j+1}\mathbb{P}\big\{W < \mu^{-(n-j-1)}\big\}\beta(n - j - 1), \end{aligned} \tag{2.4}$$

where $\beta(i) := p_1^{-1}\mathbb{P}\{W < \mu^{-i}\}$. By the a priori estimate (2.2), we have $\sum \beta(i) < \infty$.

Using (2.4), we obtain, for any positive integer n,

$$\begin{aligned} \mathbb{P}\{W < \mu^{-n}\} &\leq \sum_{j=0}^{n-1} \mathbb{P}\{W < \mu^{-n}, T = j\} + \mathbb{P}\{T \geq n\} \\ &\leq \sum_{j=0}^{n-1} p_1^{j+1}\mathbb{P}\{W < \mu^{-(n-j-1)}\}\beta(n - j - 1) + p_1^n. \end{aligned} \tag{2.5}$$

We deduce from (2.5) that $a(n) \leq \sum_{j=0}^{n-1} a(n - j - 1)\beta(n - j - 1) + 1$. Define $\tilde{a}(-1) := 1$, $\beta(-1) := 1$ and, inductively, for non-negative n,

$$\tilde{a}(n) := \sum_{j=0}^{n-1} \tilde{a}(n - j - 1)\beta(n - j - 1) + 1 = \sum_{j=-1}^{n-1} \tilde{a}(j)\beta(j).$$

Then, since $a(n) \leq \tilde{a}(n)$ for all $n \geq 0$, it suffices to show that $(\tilde{a}(n) : n \geq 0)$ is bounded. From the definition, it easily follows that $\tilde{a}(n) = \tilde{a}(n - 1)(1 + \beta(n - 1))$, hence $\tilde{a}(n) = \prod_{i=0}^{n-1}(1 + \beta(i))$, which converges as $\sum_{i=0}^{\infty} \beta(i)$ converges. Hence, $(\tilde{a}(n) : n \geq 0)$ is bounded and the proof is complete.

2.4. The Böttcher case: the lower bound

We now consider the case when $p_1 = 0$. Recall that $\nu := \min\{j \geq 0 : p_j \neq 0\} \geq 2$ and $\nu < \mu$. For every n, there are at least ν^n individuals in generation n, hence

$$\mathbb{P}\{Z_n = \nu^n\} = \mathbb{P}\{Z_n = \nu^n | Z_{n-1} = \nu^{n-1}\}\mathbb{P}\{Z_{n-1} = \nu^{n-1}\} = p_\nu^{\nu^{n-1}}\mathbb{P}\{Z_{n-1} = \nu^{n-1}\}.$$

Also, $\mathbb{P}\{Z_1 = \nu\} = p_\nu$ and therefore

$$\mathbb{P}\{Z_n = \nu^n\} = p_\nu^{1+\nu+\cdots+\nu^{n-1}} = p_\nu^{(\nu^n-1)/(\nu-1)}. \tag{2.6}$$

Given $\varepsilon > 0$, we look at the lower bound of the probability $\mathbb{P}\{W < \varepsilon\}$. Choose the integer n such that $(\frac{\nu}{\mu})^n \leq \varepsilon < (\frac{\nu}{\mu})^{n-1}$. Invoking (2.1) and (2.6), we obtain

$$\begin{aligned}\mathbb{P}\{W < \varepsilon\} &\geq \mathbb{P}\left\{W < \left(\frac{\nu}{\mu}\right)^n \;\middle|\; Z_{n+1} = \nu^{n+1}\right\}\mathbb{P}\{Z_{n+1} = \nu^{n+1}\}\\ &= \mathbb{P}\left\{W(v_{n+1}(1)) + \cdots + W(v_{n+1}(\nu^{n+1})) < \left(\frac{\mu}{\nu}\right)\nu^{n+1}\right\}p_\nu^{(\nu^{n+1}-1)/(\nu-1)}\\ &\geq \mathbb{P}\left\{\left|\sum_{j=1}^{\nu^{n+1}} W(v_{n+1}(j)) - \nu^{n+1}\right| < \delta\nu^{n+1}\right\}p_\nu^{(\nu^{n+1}-1)/(\nu-1)},\end{aligned}$$

where $\delta := \frac{\mu}{\nu} - 1 > 0$. By the weak law of large numbers, we may choose $N \in \mathbb{N}$ such that

$$\mathbb{P}\left\{\left|\sum_{j=1}^{\nu^{m+1}} W(v_{m+1}(j)) - \nu^{m+1}\right| < \delta\nu^{m+1}\right\} \geq p_\nu^{1/(\nu-1)} \qquad \text{for all } m \geq N.$$

Then, for all $n \geq N$, we have

$$-\log\mathbb{P}\{W < \varepsilon\} \leq (-\log p_\nu)\frac{\nu^{n+1}}{\nu-1} \leq C\varepsilon^{-\beta/(1-\beta)},$$

where $C := (-\log p_\nu)\frac{\nu^2}{\nu-1}$, using the fact that $(\frac{\nu}{\mu})^{-\beta/(1-\beta)} = \nu$, by the definition of β.

2.5. The Böttcher case: the upper bound

Given $\varepsilon > 0$, we continue with an upper bound for the probability $\mathbb{P}\{W < \varepsilon\}$. Choose the integer n such that $(\frac{\nu}{\mu})^{n+1} \leq \varepsilon < (\frac{\nu}{\mu})^n$. Once again using (2.1), we obtain

$$\mathbb{P}\{W < \varepsilon\} \leq \mathbb{P}\left\{\mu^{1-n}\sum_{j=1}^{\nu^{n-1}} W(v_{n-1}(j)) < \left(\frac{\nu}{\mu}\right)^n\right\} = \mathbb{P}\{S(\nu^{n-1}) > 0\}, \tag{2.7}$$

where $X_j := \frac{\nu}{\mu} - W(v_{n-1}(j))$ and $S(k) := \sum_{j=1}^{k} X_j$.

We now estimate the right-hand side by a simple large deviation bound, which only uses the fact that X_j is bounded from above and has negative mean. By the exponential Chebyshev inequality,

$$\mathbb{P}\{S(k) \geq 0\} \leq \mathbb{P}\{\mathrm{e}^{\tau S(k)} \geq 1\} \leq \mathbb{E}\mathrm{e}^{\tau S(k)} = (\mathbb{E}\mathrm{e}^{\tau X_1})^k. \tag{2.8}$$

We claim there exists $\tau > 0$ such that $\mathbb{E}\mathrm{e}^{\tau X_1} < 1$. Indeed, denoting $\varphi(\tau) := \mathbb{E}\mathrm{e}^{\tau X_1}$ and using Lebesgue's dominated convergence theorem, we have

$$\lim_{\tau \downarrow 0} \frac{\varphi(\tau) - \varphi(0)}{\tau} = \lim_{\tau \downarrow 0} \mathbb{E}\left[\frac{\mathrm{e}^{\tau X_1} - 1}{\tau}\right] = \mathbb{E} \lim_{\tau \downarrow 0}\left(\frac{\mathrm{e}^{\tau X_1} - 1}{\tau}\right) = \mathbb{E}X_1 = \frac{\nu}{\mu} - 1 < 0.$$

Since $\varphi(0) = 1$, we can thus choose $\tau > 0$ such that $\varphi(\tau) < 1$. Combining this with (2.7) and (2.8), we obtain $-\log \mathbb{P}\{W < \varepsilon\} \geq (-\log \varphi(\tau))\nu^{n-1} \geq c\varepsilon^{-\beta/(1-\beta)}$, where $c := -\nu^{-2} \times \log \varphi(\tau) > 0$.

3. Small value probabilities for mutual intersection local times

In this section, we identify the small value probability of the random variables

$$X(t_1, \ldots, t_m) := \int_{-\infty}^{\infty} \prod_{i=1}^{m} L_i^{q_i}(x, t_i)\, \mathrm{d}x,$$

where $(L_1(x,t) : x \in \mathbb{R}, t \geq 0), \ldots, (L_m(x,t) : x \in \mathbb{R}, t \geq 0)$ are the local time fields of m independent Brownian motions started at the origin. For $q_1 = \cdots = q_m = 1$, the random variable $X(t_1, \ldots, t_m)$ measures the amount of intersection between the motions up to times $t_1, \ldots, t_m$ and it is therefore called *(mutual) intersection local time*.

Our solution to the small value problem for intersection local times is based on an analogy between the martingale limit W of a Galton–Watson tree in the Schröder case and the random variables $X(\sigma^{(1)}, \ldots, \sigma^{(m)})$, where $\sigma^{(1)}, \ldots, \sigma^{(m)}$ are the first exit times of the Brownian motions from the interval $(-1, 1)$. This analogy allows us to carry over the crucial steps in the proof of Theorem 1(a) to the new situation and hence to prove the following theorem.

Theorem 2. *Suppose that $L_1, \ldots, L_m$ are the local times of $m \geq 2$ independent Brownian motions and that $q_j \geq 1$ for all $1 \leq j \leq m$. Then, for $q := \sum_{j=1}^{m} q_j$,*

$$\text{(a)} \quad \mathbb{P}\left\{\int_{-\infty}^{\infty} \prod_{i=1}^{m} L_i^{q_i}(x, \sigma^{(i)})\, \mathrm{d}x < \varepsilon\right\} \asymp \varepsilon^{2/(1+q)},$$

$$\text{(b)} \quad \mathbb{P}\left\{\int_{-\infty}^{\infty} \prod_{i=1}^{m} L_i^{q_i}(x, 1)\, \mathrm{d}x < \varepsilon\right\} \asymp \varepsilon^{2/(1+q)}.$$

Remark. The excluded case $m = 1$ is entirely different, as the small value probabilities decay exponentially. This will be discussed in Section 4 using the technique of the Böttcher case.

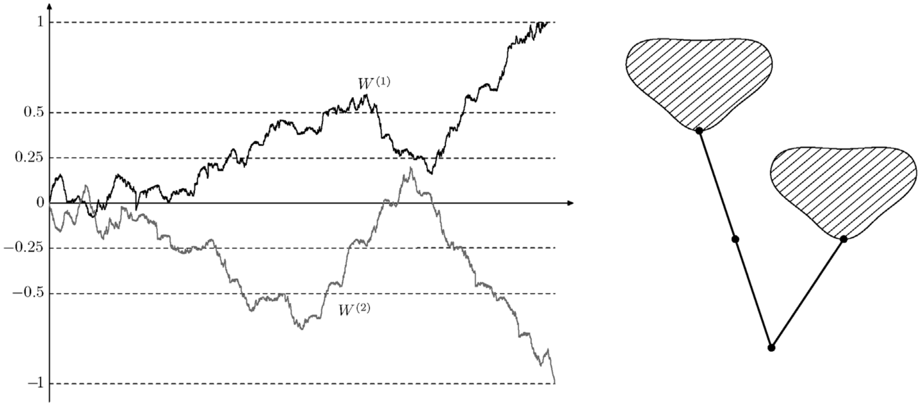

Figure 3. The tree associated with two Brownian paths for $\eta = 2$, up to 2nd generation. The intervals $W^{(1)}[\tau_1^{(1)}, \tau_0^{(1)}]$ and $W^{(2)}[\tau_1^{(2)}, \tau_0^{(2)}]$ have a non-empty intersection and therefore the root has more than one offspring; in contrast, the intervals $W^{(1)}[\tau_2^{(1)}, \tau_1^{(1)}]$ and $W^{(2)}[\tau_2^{(2)}, \tau_1^{(2)}]$ are disjoint and therefore the second vertex on the spine has just one offspring.

Before giving the detailed proof, we show how the analogy to the martingale limit of a Galton–Watson tree arises. From the Brownian paths, we need to recognize the particular elements of the tree featuring in the proof of the Schröder case: for each vertex of the spine, we first need to decide whether a subtree splits off from the vertex (this happens independently with probability $1 - p_1$) and, supposing this happens at the vertex in the kth generation, we need to see that this subtree gives rise to a summand of the intersection local time, which, in distribution, equals μ^{-k} times the intersection local time. Once an inequality analogous to (2.3) is established, we obtain lower tail asymptotics featuring the parameters μ and p_1 used in the construction of the tree.

To sketch the actual construction, focusing on $m = 2$ for the moment, we let $W^{(1)}, W^{(2)}$ be two independent Brownian motions started at the origin and assume that $W^{(1)}$ exits $(-1, 1)$ at the upper end, and $W^{(2)}$ exits $(-1, 1)$ at the lower end of the interval. Fix $\eta > 1$ and divide the Brownian paths according to the stopping times

$$\tau_k^{(1)} := \inf\{t \geq 0 : W^{(1)}(t) = \eta^{-k}\} \quad \text{and} \quad \tau_k^{(2)} := \inf\{t \geq 0 : W^{(2)}(t) = -\eta^{-k}\}.$$

To build the tree from its spine $v_0(1), \dots, v_n(1)$ of leftmost particles in the first n generations, we let the kth individual $v_0(k)$ on this spine have more than one offspring if

$$W^{(1)}[\tau_{k+1}^{(1)}, \tau_k^{(1)}] \cap W^{(2)}[\tau_{k+1}^{(2)}, \tau_k^{(2)}] \neq \varnothing.$$

If the intervals intersect, then the intersection local time of the two Brownian motions $W^{(j)}$, started at time $\tau_{k+1}^{(j)}$ and stopped at time $\tau_k^{(j)}$, for $j \in \{1, 2\}$, gives rise to a summand of the total intersection local time which is distributed approximately like a scaled copy of the total intersection local time (see Figure 3 for illustration).

3.1. Intersection local times: the parameters μ and p_1

We start with a basic scaling property of intersection local times. For any points $x_1, \dots, x_m \in \mathbb{R}$, we suppose that under $\mathbb{P}_{(x_j)}$, the Brownian motion $W^{(j)}$ is started in x_j and for $\eta > 0$, we denote by

$$\tau^{(j)}(\eta) = \inf\{t > 0 \colon W^{(j)}(t) = \eta\}$$

the first hitting time of η by the Brownian motion $W^{(j)}$.

Lemma 3.1. *For every $\varepsilon > 0$ and for $q := \sum_{j=1}^m q_j$, we have*

$$\mathbb{P}_{(x_j/\eta)}\left\{\int_{-\infty}^{\infty} \prod_{j=1}^m L_j^{q_j}\big(x, \tau^{(j)}(1)\big)\,\mathrm{d}x < \varepsilon\right\} = \mathbb{P}_{(x_j)}\left\{\int_{-\infty}^{\infty} \prod_{j=1}^m L_j^{q_j}\big(x, \tau^{(j)}(\eta)\big)\,\mathrm{d}x < \varepsilon\eta^{1+q}\right\}.$$

Proof. By Brownian scaling, we have

$$\mathbb{P}_{x_j/\eta}\big\{L\big(x, \tau^{(j)}(1)\big) < \varepsilon\big\} = \mathbb{P}_{x_j}\big\{\eta^{-1}L\big(\eta x, \tau^{(j)}(\eta)\big) < \varepsilon\big\}.$$

Hence,

$$\begin{aligned}
&\mathbb{P}_{(x_j/\eta)}\left\{\int_{-\infty}^{\infty} \prod_{j=1}^m L_j^{q_j}\big(x, \tau^{(j)}(1)\big)\,\mathrm{d}x < \varepsilon\right\} \\
&\quad = \mathbb{P}_{(x_j)}\left\{\eta^{-\sum_{j=1}^m q_j} \int_{-\infty}^{\infty} \prod_{j=1}^m L_j^{q_j}\big(\eta x, \tau^{(j)}(\eta)\big)\,\mathrm{d}x < \varepsilon\right\} \\
&\quad = \mathbb{P}_{(x_j)}\left\{\eta^{-(1+q)} \int_{-\infty}^{\infty} \prod_{j=1}^m L_j^{q_j}\big(x, \tau^{(j)}(\eta)\big)\,\mathrm{d}x < \varepsilon\right\}
\end{aligned}$$

and this proves the lemma. □

Fix $\eta > 1$ and let $W^{(1)}, \dots, W^{(m)}$ be Brownian motions started in the origin. Fix a set $M \subset \{1, \dots, m\}$ and define stopping times

$$\tau_k^{(j)} := \tau_k^{(j)}(M) := \begin{cases} \inf\{t \geq 0 \colon W^{(j)}(t) = \eta^{-k}\}, & \text{if } j \in M, \\ \inf\{t \geq 0 \colon W^{(j)}(t) = -\eta^{-k}\}, & \text{if } j \notin M \end{cases}$$

and abbreviate $\tau^{(j)} := \tau_0^{(j)}(M)$. Suppose that under $\mathbb{P}_{(\pm\varepsilon)}$, the Brownian motion $W^{(j)}$ is started in the point $+\varepsilon$ if $j \in M$ and in the point $-\varepsilon$ otherwise.

For $0 < s < t$, define local times $L_j(x, s, t) := L_j(x, t) - L_j(x, s)$ over the time interval $[s, t]$ and

$$L_k := \int_{-\infty}^{\infty} \prod_{j=1}^m L_j^{q_j}\big(x, \tau_{k+1}^{(j)}, \tau_k^{(j)}\big)\,\mathrm{d}x.$$

By the previous lemma, for every k, we have

$$\eta^{k(1+q)} L_k \stackrel{d}{=} L_0. \tag{3.1}$$

This identifies the parameter μ as η^{1+q}. Recall that in the tree model, this parameter corresponds to the mean offspring number.

Lemma 3.2. *If M is a proper, nonempty subset of* $\{1, \dots, m\}$, *we have*

$$\mathbb{P}_{(\pm\varepsilon)}\big\{W^{(1)}\big[0, \tau^{(1)}\big] \cap \cdots \cap W^{(m)}\big[0, \tau^{(m)}\big] = \varnothing\big\} \asymp \varepsilon^2.$$

Proof. On the one hand, if $\{W^{(1)}[0, \tau^{(1)}] \cap \cdots \cap W^{(m)}[0, \tau^{(m)}] = \varnothing\}$, then at least one of the motions $W^{(j)}$, $j \in M$, does not reach level $-\varepsilon$ before level 1, the probability of this being $2\varepsilon/(1+\varepsilon)$ per motion by the gambler's ruin probability. Analogously, one of the motions $W^{(j)}$, $j \notin M$, does not reach level ε before level -1, which has the same probability. This gives the upper bound

$$\mathbb{P}_{(\pm\varepsilon)}\big\{W^{(1)}\big[0, \tau^{(1)}\big] \cap \cdots \cap W^{(m)}\big[0, \tau^{(m)}\big] = \varnothing\big\} \le \frac{\varepsilon^2}{(1+\varepsilon)^2} 4\ell(m-\ell),$$

where ℓ is the cardinality of M. For the lower bound, note that if one of the motions in each of the two groups does not reach level 0 before level 1 (resp., -1), this implies that $W^{(1)}[0, \tau^{(1)}] \cap \cdots \cap W^{(m)}[0, \tau^{(m)}] = \varnothing$. As, for each motion, this event has probability ε, we obtain

$$\mathbb{P}_{(\pm\varepsilon)}\big\{W^{(1)}\big[0, \tau^{(1)}\big] \cap \cdots \cap W^{(m)}\big[0, \tau^{(m)}\big] = \varnothing\big\} \ge \varepsilon^2. \qquad \square$$

Remark. A refined calculation along the same lines shows that, as $\varepsilon \downarrow 0$,

$$\mathbb{P}_{(\pm\varepsilon)}\big\{W^{(1)}\big[0, \tau^{(1)}\big] \cap \cdots \cap W^{(m)}\big[0, \tau^{(m)}\big] = \varnothing\big\} \sim \varepsilon^2 2\ell(m-\ell),$$

where ℓ is the cardinality of M, but we do not need this here.

By Brownian scaling, we infer from Lemma 3.2 that there are constants $0 < c < C$ such that, if $M \subset \{1, \dots, m\}$ is proper and nonempty, then for any non-negative integer k and $\eta > 1$,

$$c\eta^{-2} \le \mathbb{P}\big\{W^{(1)}\big[\tau_{k+1}^{(1)}, \tau_k^{(1)}\big] \cap \cdots \cap W^{(m)}\big[\tau_{k+1}^{(1)}, \tau_k^{(m)}\big] = \varnothing\big\} \le C\eta^{-2}$$

and thus the parameter p_1 is identified (with sufficient accuracy) as η^{-2}. Recall that in the tree model, p_1 corresponds to the probability that a vertex has only one offspring.

3.2. Intersection local times: the lower bound

Let $W^{(1)}, \dots, W^{(m)}$ be Brownian motions started at the origin and fix $M \subset \{1, \dots, m\}$ such that $1 \in M$ and $2 \notin M$. We propose a sufficient strategy to realize the event $\{X(\sigma^{(1)}(1), \dots,$

$\sigma^{(m)}(1)) < \varepsilon\}$, which is time-inhomogeneous and consists of two phases. Given $\varepsilon > 0$, the phases are separated by the stopping times

$$\omega^{(j)} := \inf\{t \geq 0 : W^{(j)} \notin (-\varepsilon^{1/(1+q)}, \varepsilon^{1/(1+q)})\} \qquad \text{for } j \in \{1, \dots, m\}.$$

The first phase is described by the event

$$E_1 := \Big\{W^{(j)}\big(\omega^{(j)}\big) = \pm\varepsilon^{1/(1+q)}, \inf\big\{\pm W^{(j)}(s) : 0 \leq s \leq \omega^{(j)}\big\} > -\tfrac{1}{2}\varepsilon^{1/(1+q)} \\ \text{for all } j \text{ and } X\big(\omega^{(1)}, \dots, \omega^{(m)}\big) < \varepsilon\Big\},$$

where $\pm$ indicates $+$ if $j \in M$ and $-$ otherwise. By the scaling verified in Lemma 3.1, the probability $\delta := \mathbb{P}(E_1) > 0$ does not depend on ε. The second phase is described by the event

$$E_2 := \Big\{W^{(j)}\big(\tau^{(j)}\big) = \pm 1 \text{ for all } j \text{ and } \inf\big\{W^{(1)}(s) : \omega^{(1)} \leq s \leq \tau^{(1)}\big\} \geq \tfrac{1}{2}\varepsilon^{1/(1+q)} \\ \text{and } \sup\big\{W^{(2)}(s) : \omega^{(2)} \leq s \leq \tau^{(2)}\big\} \leq -\tfrac{1}{2}\varepsilon^{1/(1+q)}\Big\}.$$

Observe that if E_1 and E_2 hold, then we have

$$X\big(\sigma^{(1)}, \dots, \sigma^{(m)}\big) = X\big(\tau^{(1)}, \dots, \tau^{(m)}\big) = X\big(\omega^{(1)}, \dots, \omega^{(m)}\big) < \varepsilon,$$

as required. Moreover, using the strong Markov property and the gambler's ruin estimate,

$$\begin{aligned}\mathbb{P}(E_1 \cap E_2) &= \mathbb{E}\big[1_{E_1}\mathbb{P}_{(W^{(j)}(\omega^{(j)}))}(E_2)\big] \\ &= \mathbb{P}(E_1)\left(\frac{1+\varepsilon^{1/(1+q)}}{2}\right)^{m-2}\left(\frac{(1/2)\varepsilon^{1/(1+q)}}{1-(1/2)\varepsilon^{1/(1+q)}}\right)^2,\end{aligned}$$

so the lower bound holds with $c := \delta(1/2)^m$.

3.3. Intersection local times: the logarithmic upper bound

We now give an upper bound for the small value probability of $X(\sigma^{(1)}, \dots, \sigma^{(m)})$ along the lines of the argument leading to (2.2). Fix an arbitrarily small $\delta > 0$. Let $C \geq 1$ be the constant in the implied upper bound of Lemma 3.2. Choose and fix an integer $\eta > (2C)^{1/\delta}$.

For any subset $M \subset \{1, \dots, m\}$, define the event

$$E(M) := \big\{W^{(j)}\big(\sigma^{(j)}\big) = 1 \text{ for all } j \in M, W^{(j)}\big(\sigma^{(j)}\big) = -1 \text{ for all } j \notin M\big\}.$$

Recall the definition of the stopping times $\tau_k^{(j)} := \tau_k^{(j)}(M)$. Then,

$$\mathbb{P}\big\{X\big(\sigma^{(1)}, \dots, \sigma^{(m)}\big) < \varepsilon\big\} = \sum_{M \subset \{1,\dots,m\}} \mathbb{P}\big(\big\{X\big(\tau_0^{(1)}, \dots, \tau_0^{(m)}\big) < \varepsilon\big\} \cap E(M)\big). \tag{3.2}$$

It therefore suffices to fix $M \subset \{1, \dots, m\}$ and give upper bounds for $\mathbb{P}\{X(\tau_0^{(1)}, \dots, \tau_0^{(m)}) < \varepsilon\}$. Define, for $0 < s < t$, local times $L_j(x, s, t) := L_j(x, t) - L_j(x, s)$ over the time interval $[s, t]$. Let

$$L_k := \int_{-\infty}^{\infty} \prod_{j=1}^{m} L_j^{q_j}\big(x, \tau_{k+1}^{(j)}, \tau_k^{(j)}\big)\, dx.$$

The random variables $X_k = \eta^{k(1+q)} L_k$ are then independent, by the Markov property, and identically distributed, by (3.1). By Lemma 3.2, we have $\mathbb{P}\{X_0 = 0\} \le C\eta^{-2}$ if M is a proper, nonempty subset of $\{1, \dots, m\}$, and otherwise, obviously, $\mathbb{P}\{X_0 = 0\} = 0$. This implies that there exists a $\theta > 0$ such that

$$\mathbb{P}\{X_0 < \theta\} \le 2C\eta^{-2}.$$

Now, given $\varepsilon > 0$, choose the integer n such that

$$\theta\eta^{-(n+1)(1+q)} < \varepsilon \le \theta\eta^{-n(1+q)}.$$

Note that for $q_i \ge 1$, by superadditivity of $x \mapsto x^{q_i}$, $x \ge 0$, we obtain

$$L_j^{q_j}\big(x, \tau_0^{(j)}\big) \ge \left(\sum_{k=0}^{n-1} L_j\big(x, \tau_{k+1}^{(j)}, \tau_k^{(j)}\big)\right)^{q_j} \ge \sum_{k=0}^{n-1} L_j^{q_j}\big(x, \tau_{k+1}^{(j)}, \tau_k^{(j)}\big).$$

Applying this to the intersection local times, it follows that

$$X\big(\tau_0^{(1)}, \dots, \tau_0^{(m)}\big) = \int_{-\infty}^{\infty} \prod_{j=1}^{m} L_j^{q_j}\big(x, \tau_0^{(j)}\big)\, dx \ge \int_{-\infty}^{\infty} \prod_{j=1}^{m} \left(\sum_{k=0}^{n-1} L_j^{q_j}\big(x, \tau_{k+1}^{(j)}, \tau_k^{(j)}\big)\right) dx \ge \sum_{k=0}^{n-1} L_k.$$

Hence, we can estimate

$$\begin{aligned} \mathbb{P}\big\{X\big(\tau_0^{(1)}, \dots, \tau_0^{(m)}\big) < \varepsilon\big\} &\le \mathbb{P}\left\{\sum_{k=0}^{n-1} L_k < \varepsilon\right\} \le \mathbb{P}\left\{\sum_{k=0}^{n-1} \eta^{-k(1+q)} X_k < \theta\eta^{-n(1+q)}\right\} \\ &\le \mathbb{P}\left\{\sum_{k=0}^{n-1} X_k < \theta\right\} \le (\mathbb{P}\{X_0 < \theta\})^n \le (2C)^n \eta^{-2n} \le K\varepsilon^{(2-\delta)/(1+q)} \end{aligned}$$

for the constant $K := \eta^{2-\delta}\theta^{(-2+\delta)/(1+q)}$. As $\delta > 0$ can be chosen arbitrarily small, this shows that

$$\limsup_{\varepsilon \downarrow 0} \frac{\log \mathbb{P}\{X(\sigma^{(1)}, \dots, \sigma^{(m)}) < \varepsilon\}}{-\log \varepsilon} \le \frac{-2}{1+q}. \tag{3.3}$$

Note (for use in Lemma 3.3) that the proof also shows that (3.3) holds if $W^{(1)}, \dots, W^{(m)}$ are started in arbitrary points of the interval $[-\eta^{-n}, \eta^n]$ instead of the origin.

3.4. Intersection local times: up-to-constant asymptotics

Fix the set $M \subset \{1, \dots, m\}$, the integer $\eta > 1$ and recall the notation from the previous section. Define a sequence $(a(n) : n \geq 0)$ by

$$a(n) := \mathbb{P}\big\{X\big(\tau_0^{(1)}, \dots, \tau_0^{(m)}\big) < \theta\eta^{-n(1+q)}\big\}\eta^{2n}.$$

Given $0 < \varepsilon < 1$, we again choose the integer n such that $\theta\eta^{-(n+1)(1+q)} \leq \varepsilon < \theta\eta^{-n(1+q)}$. Then,

$$\begin{aligned}\mathbb{P}\big\{X\big(\tau_0^{(1)}, \dots, \tau_0^{(m)}\big) < \varepsilon\big\} &\leq \mathbb{P}\big\{X\big(\tau_0^{(1)}, \dots, \tau_0^{(m)}\big) < \theta\eta^{-n(1+q)}\big\} \\ &= a(n)\eta^{-2n} \leq a(n)\eta^2\theta^{-2/(1+q)}\varepsilon^{2/(1+q)}\end{aligned}$$

and hence, to complete the proof, it suffices to show that $(a(n) : n \geq 0)$ is bounded. Define

$$T := \min\big\{k \geq 0 : W^{(1)}\big[\tau_{k+1}^{(1)}, \tau_0^{(1)}\big] \cap \dots \cap W^{(m)}\big[\tau_{k+1}^{(m)}, \tau_0^{(m)}\big] \neq \varnothing\big\}.$$

In our tree heuristic, T is the first generation in which a tree is branching off the spine. The next lemma controls the behaviour of this tree and plays a similar role to (2.4).

Lemma 3.3. *There exists a sequence $(\beta(i) : i \in \mathbb{N})$ of non-negative numbers with $\sum \beta(i) < \infty$ such that, for $0 \leq j \leq n-1$,*

$$\mathbb{P}_{(y_i)}\big\{X\big(\tau_0^{(1)}, \dots, \tau_0^{(m)}\big) < \theta\eta^{-n(1+q)}, T = j\big\} \leq \eta^{-2j-2}\beta(n-j-1),$$

where $y_i = \pm\eta^{-j-1}$ with the sign chosen according to whether or not $i \in M$.

Proof. For $i \in \{1, \dots, m\}$ and $k \in \{-\eta^{n-j-1}, \dots, \eta^{n-j-1} - 1\}$, we introduce stopping times

$$\varrho_k^{(i)} := \inf\big\{t \geq 0 : W^{(i)}(t) \in [k\eta^{-n}, (k+1)\eta^{-n}]\big\}.$$

The assumption $T = j$ implies that there exists $k \in \{-\eta^{n-j-1}, \dots, \eta^{n-j-1} - 1\}$ such that $\varrho_k^{(i)} < \tau_0^{(i)}$ for all $i \in \{1, \dots, m\}$. If this holds, then let $\sigma_j^{(i)} := \inf\{t \geq \varrho_k^{(i)} : W^{(i)}(t) = \pm\eta^{-j}\}$ (with the usual convention on $\pm$). Hence, for any $0 < \delta < 1$ and sufficiently large $n - j$, first using Lemma 3.2 with $\varepsilon = \eta^{-j}$, then (3.3) and the subsequent remark in combination with Lemma 3.1 and, of course, the strong Markov property, we have

$$\begin{aligned}&\mathbb{P}_{(y_i)}\big\{X\big(\tau_0^{(1)}, \dots, \tau_0^{(m)}\big) < \theta\eta^{-n(1+q)}, T = j\big\} \\ &\quad\leq \sum_{k=-\eta^{n-j-1}}^{\eta^{n-j-1}-1} \mathbb{E}_{(y_i)}\Big[\mathbb{1}\big\{X\big(\sigma_j^{(1)}, \dots, \sigma_j^{(m)}\big) < \theta\eta^{-n(1+q)}\big\} \\ &\qquad\qquad\qquad \times \mathbb{P}_{(W^{(i)}(\sigma_j^{(i)}))}\big\{W^{(1)}\big[0, \tau_0^{(1)}\big] \cap \dots \cap W^{(m)}\big[0, \tau_0^{(m)}\big] = \varnothing\big\}\Big]\end{aligned}$$

$$\leq \sum_{k=-\eta^{n-j-1}}^{\eta^{n-j-1}-1} \mathbb{P}_{(W^{(i)}(\varrho_k^{(i)}))}\big\{X\big(\sigma_j^{(1)},\dots,\sigma_j^{(m)}\big) < \theta\eta^{-n(1+q)}\big\} C\eta^{-2j}$$
$$\leq 2\eta^{n-j-1}\eta^{(-2+\delta)(n-j)}C\eta^{-2j},$$

which gives the result with $\beta(i) := 2C\eta^{\delta}\eta^{(-1+\delta)i}$. □

We now argue, as in (2.5) of the Schröder case, using the upper bound of Lemma 3.2 in the second step and denoting the implied constant there by $C > 0$, that

$$\begin{aligned} &\mathbb{P}\big\{X\big(\tau_0^{(1)},\dots,\tau_0^{(m)}\big) < \theta\eta^{-n(1+q)}\big\} \\ &\quad \leq \mathbb{P}\{T \geq n\} + \sum_{j=0}^{n-1} \mathbb{P}\big\{X\big(\tau_0^{(1)},\dots,\tau_0^{(m)}\big) < \theta\eta^{-n(1+q)}, T = j\big\} \\ &\quad \leq C\eta^{-2n} + \sum_{j=0}^{n-1} \mathbb{P}\big\{X\big(\tau_0^{(1)},\dots,\tau_0^{(m)}\big) < \theta\eta^{-n(1+q)}, T = j\big\}. \end{aligned} \tag{3.4}$$

To estimate the remaining probability, we first use the strong Markov property, then Lemma 3.3 to estimate the inner probability and finally the definition of $(a(n) : n \geq 0)$ in combination with Lemma 3.1, to obtain

$$\begin{aligned} &\mathbb{P}\big\{X\big(\tau_0^{(1)},\dots,\tau_0^{(m)}\big) < \theta\eta^{-n(1+q)}, T = j\big\} \\ &\quad \leq \mathbb{E}\big\{\mathbb{1}\big\{X\big(\tau_{j+1}^{(1)},\dots,\tau_{j+1}^{(m)}\big) < \theta\eta^{-n(1+q)}\big\} \\ &\qquad\quad \times \mathbb{P}_{(W^{(i)}(\tau_{j+1}^{(i)}))}\big\{X\big(\tau_0^{(1)},\dots,\tau_0^{(m)}\big) < \theta\eta^{-n(1+q)}, T = j\big\}\big\} \\ &\quad \leq \eta^{-2j-2}\beta(n-j-1)\mathbb{P}\big\{X\big(\tau_{j+1}^{(1)},\dots,\tau_{j+1}^{(m)}\big) < \theta\eta^{-n(1+q)}\big\} \\ &\quad \leq \eta^{-2n}\beta(n-j-1)a(n-j-1). \end{aligned}$$

Substituting this into (3.4), we obtain a recursion formula for $a(n)$, namely

$$a(n) \leq \sum_{j=0}^{n-1} \beta(n-j-1)a(n-j-1) + C \qquad \text{for } n \geq 0.$$

As before, boundedness of $(a(n) : n \geq 0)$ follows from the recursion and the fact that $\sum \beta(j) < \infty$.

3.5. Intersection local times at fixed times

In this section, we use a technique adapted from Lawler (1996) to transfer our results from hitting times to fixed times, thus proving Theorem 2(b). Recall the following simple tail estimates for the first exit times $\sigma^{(j)}(x)$ from the interval $(-x, x)$ by a Brownian motion $W^{(j)}$ started in x_j.

Lemma 3.4. *There exist constants $\beta > 0$ and $\kappa > 0$ such that, for all $x > 0$, $|x_j| \le x/2$ and $a > 0$,*

(a) $\mathbb{P}_{(x_j)}\left\{\min_{j=1}^{m} \sigma^{(j)}(x) \le ax^2\right\} \le \kappa \mathrm{e}^{-\beta/a}$;

(b) $\mathbb{P}_{(x_j)}\left\{\max_{j=1}^{m} \sigma^{(j)}(x) \ge ax^2\right\} \le \kappa \mathrm{e}^{-\beta a}$.

Proof. By scaling, we may assume that $x = 1$. On the one hand, using the reflection principle, we obtain

$$\mathbb{P}_{x_j}\big\{\sigma^{(j)}(1) \le a\big\} \le \mathbb{P}_0\left\{\sup_{t\le a}\big|W^{(j)}(t)\big| \ge \frac{1}{2}\right\} \le 2\mathbb{P}_0\left\{\big|W^{(j)}(a)\big| \ge \frac{1}{2}\right\} = 2\mathbb{P}_0\left\{\big|W^{(j)}(1)\big| \ge \frac{1}{2\sqrt{a}}\right\}$$

and hence (a) follows from a standard estimate for the tail of a normal distribution. On the other hand, (b) follows from $\mathbb{P}_{x_j}\{\sigma^{(j)}(1) \ge k | \sigma^{(j)}(1) \ge k-1\} \le \mathbb{P}_0\{|W^{(j)}(1)| \le 2\} < 1$, by iteration. □

For the lower bound we obtain, for any $a > 0$, using Lemma 3.1 in the second step,

$$\begin{aligned}
&\mathbb{P}\{X(1,\dots,1) < \varepsilon\}\\
&\quad \ge \mathbb{P}\big\{X\big(\sigma^{(1)}(a),\dots,\sigma^{(m)}(a)\big) < \varepsilon\big\} - \mathbb{P}\left\{X\big(\sigma^{(1)}(a),\dots,\sigma^{(m)}(a)\big) < \varepsilon, \min_{j=1}^{m}\sigma^{(j)}(a) \le 1\right\}\\
&\quad = \mathbb{P}\big\{X\big(\sigma^{(1)}(1),\dots,\sigma^{(m)}(1)\big) < a^{-(1+q)}\varepsilon\big\}\\
&\qquad - \mathbb{P}\left\{X\big(\sigma^{(1)}(a),\dots,\sigma^{(m)}(a)\big) < \varepsilon, \min_{j=1}^{m}\sigma^{(j)}(a) \le 1\right\}.
\end{aligned}$$

First using Theorem 2(a) in combination with Lemma 3.1 and then Lemma 3.4(a), we have

$$\begin{aligned}
&\mathbb{P}\left\{X\big(\sigma^{(1)}(a),\dots,\sigma^{(m)}(a)\big) < \varepsilon, \min_{j=1}^{m}\sigma^{(j)}(a) \le 1\right\}\\
&\quad \le \mathbb{E}\left[1\big\{X\big(\sigma^{(1)}(a/2),\dots,\sigma^{(m)}(a/2)\big) < \varepsilon\big\}\mathbb{P}_{(W^{(j)}(\sigma^{(j)}(a/2)))}\left\{\min_{j=1}^{m}\sigma^{(j)}(a) \le 1\right\}\right]\\
&\quad \le 4Ca^{-2}\varepsilon^{2/(1+q)} \sup_{|x_j|=a/2} \mathbb{P}_{(x_j)}\left\{\min_{j=1}^{m}\sigma^{(j)}(a) \le 1\right\} \le 4Ca^{-2}\varepsilon^{2/(1+q)}\kappa\mathrm{e}^{-\beta a^2},
\end{aligned}$$

where $C > 0$ is the implied constant in the upper bound of Theorem 2(a). Substituting this into the previous equation and applying the lower bound of Theorem 2(a) with the implied constant denoted by $c > 0$, we obtain

$$\begin{aligned}\mathbb{P}\{X(1,\dots,1) < \varepsilon\} &\ge \mathbb{P}\big\{X\big(\sigma^{(1)}(1),\dots,\sigma^{(m)}(1)\big) < a^{-(1+q)}\varepsilon\big\} - 4Ca^{-2}\varepsilon^{2/(1+q)}\kappa \mathrm{e}^{-\beta a^2} \\ &\ge (ca^{-2} - 4Ca^{-2}\kappa \mathrm{e}^{-\beta a^2})\varepsilon^{2/(1+q)}\end{aligned}$$

and the result follows if we choose a large enough to ensure that the bracket is positive.

For the upper bound, given $\varepsilon > 0$, we choose the integer n such that

$$\mathrm{e}^{-\beta 2^n} \le \varepsilon^{2/(1+q)} < \mathrm{e}^{-\beta 2^{n-1}}. \tag{3.5}$$

We base the argument on the decomposition

$$\begin{aligned}&\mathbb{P}\{X(1,\dots,1) < \varepsilon\} \\ &\quad \le \mathbb{P}\big\{X\big(\sigma^{(1)}(1),\dots,\sigma^{(m)}(1)\big) < \varepsilon\big\} \\ &\qquad + \sum_{i=0}^{n-1}\sum_{j=1}^{m} \mathbb{P}\big\{X\big(\sigma^{(1)}(2^{-i-1}),\dots,\sigma^{(m)}(2^{-i-1})\big) < \varepsilon, \sigma^{(j)}(2^{-i}) \ge 1\big\} \\ &\qquad + \mathbb{P}\Big\{\max_{j=1}^{m} \sigma^{(j)}(2^{-n}) \ge 1\Big\}.\end{aligned} \tag{3.6}$$

We bound the first term on the right-hand side using Theorem 2(a) and the last term using Lemma 3.4(b) and (3.5). It remains to bound the sum in the middle. To this end, we write

$$\sigma^{(j)}(2^{-i}) = \sum_{k=i}^{n}\big(\sigma^{(j)}(2^{-k}) - \sigma^{(j)}\big(2^{-(k+1)}\big)\big) + \sigma^{(j)}\big(2^{-(n+1)}\big)$$

and note that, as $2^{-2n-2}2^{n+i} + \sum_{k=i}^{n} 2^{i-k-1} \le 1$, we obtain

$$\begin{aligned}&\mathbb{P}\big\{X\big(\sigma^{(1)}(2^{-i-1}),\dots,\sigma^{(m)}(2^{-i-1})\big) < \varepsilon, \sigma^{(j)}(2^{-i}) \ge 1\big\} \\ &\quad \le \sum_{k=i}^{n} \mathbb{P}\big\{X\big(\sigma^{(1)}(2^{-i-1}),\dots,\sigma^{(m)}(2^{-i-1})\big) < \varepsilon, \sigma^{(j)}(2^{-k}) - \sigma^{(j)}\big(2^{-(k+1)}\big) \ge 2^{i-k-1}\big\} \\ &\qquad + \mathbb{P}\big\{\sigma^{(j)}\big(2^{-(n+1)}\big) \ge 2^{-2n-2}2^{n+i}\big\}.\end{aligned}$$

Again, the contribution from the last summand can be bounded using Lemma 3.4(b). For the remaining term, we use the strong Markov property to obtain, if $n \ge k \ge i+1$,

$$\begin{aligned}&\mathbb{P}\big\{X\big(\sigma^{(1)}(2^{-i-1}),\dots,\sigma^{(m)}(2^{-i-1})\big) < \varepsilon, \sigma^{(j)}(2^{-k}) - \sigma^{(j)}(2^{-k-1}) \ge 2^{i-k-1}\big\} \\ &\quad \le \mathbb{P}\big\{X\big(\sigma^{(1)}(2^{-k-1}),\dots,\sigma^{(m)}(2^{-k-1})\big) < \varepsilon\big\} \sup_{|x_j|=2^{-k-1}} \mathbb{P}_{x_j}\big\{\sigma^{(j)}(2^{-k}) \ge 2^{i-k-1}\big\}\end{aligned} \tag{3.7}$$

$$\times \sup_{|x_j|=2^{-k}} \mathbb{P}_{(x_j)}\big\{X\big(\sigma^{(1)}(2^{-i-1}),\dots,\sigma^{(m)}(2^{-i-1})\big) < \varepsilon\big\}.$$

If n is large enough (or, equivalently, if $\varepsilon > 0$ small enough) to satisfy $\mathrm{e}^{-\beta 2^{n-2}} \le 2^{-n}$, then we get that

$$\begin{aligned}
&\sup_{|x_j|=2^{-k}} \mathbb{P}_{(x_j)}\big\{X\big(\sigma^{(1)}(2^{-i-1}),\dots,\sigma^{(m)}(2^{-i-1})\big) < \varepsilon\big\}\\
&\quad= \sup_{|x_j|=2^{i-k+1}} \mathbb{P}_{(x_j)}\big\{X\big(\sigma^{(1)}(1),\dots,\sigma^{(m)}(1)\big) < \varepsilon 2^{(i+1)(1+q)}\big\}\\
&\quad\le \sup_{|x_j|=2^{i-k+1}} \mathbb{P}_{(x_j)}\big\{X\big(\sigma^{(1)}(1),\dots,\sigma^{(m)}(1)\big) < 2^{(i-k+1)(1+q)}\big\}.
\end{aligned}$$

Recall that $\tau^{(j)}(x) = \inf\{t \ge 0 : W^{(j)}(t) = x\}$ and note that, for $|x_j| = 2^{-k}$,

$$\begin{aligned}
&\mathbb{P}_{(x_j)}\big\{X\big(\sigma^{(1)}(1),\dots,\sigma^{(m)}(1)\big) < 2^{(i-k+1)(1+q)}\big\}\\
&\quad\le \mathbb{P}_{(2^{i-k+1})}\big\{X\big(\sigma^{(1)}(1),\dots,\sigma^{(m)}(1)\big) < 2^{(i-k+1)(1+q)}\big\}\\
&\qquad+ \mathbb{P}_{(-2^{i-k+1})}\big\{X\big(\sigma^{(1)}(1),\dots,\sigma^{(m)}(1)\big) < 2^{(i-k+1)(1+q)}\big\}\\
&\qquad+ \sum_{j=1}^{m}\sum_{\ell=1}^{m} \mathbb{P}_{x_j}\big\{\tau^{(j)}(2^{i-k+1}) > \sigma^{(j)}(1)\big\}\mathbb{P}_{x_\ell}\big\{\tau^{(\ell)}(-2^{i-k+1}) > \sigma^{(\ell)}(1)\big\}.
\end{aligned}$$

While the first two probabilities are bounded by constant multiples of $2^{2(i-k+1)}$, by Theorem 2(a), the double sum is bounded by $m^2 2^{2(i-k+2)}$, by the gambler's ruin probability. Hence, for a suitable constant $C_0 > 1$ and all $n \ge k \ge i+1$,

$$\sup_{|x_j|=2^{-k}} \mathbb{P}_{(x_j)}\big\{X\big(\sigma^{(1)}(2^{-i-1}),\dots,\sigma^{(m)}(2^{-i-1})\big) < \varepsilon\big\} \le C_0 2^{2(i-k)}.$$

Combining this with Lemma 3.4(b) and substituting into (3.7), we obtain for all $n \ge k \ge i$,

$$\begin{aligned}
&\mathbb{P}\big\{X\big(\sigma^{(1)}(2^{-i-1}),\dots,\sigma^{(m)}(2^{-i-1})\big) < \varepsilon, \sigma^{(j)}(2^{-k}) - \sigma^{(j)}(2^{-k-1}) \ge 2^{i-k-1}\big\}\\
&\quad\le C_1 \varepsilon^{2/(1+q)}\big[2^{2k+2}\mathrm{e}^{-\beta 2^{k+i-1}} 2^{2(i-k)}\big]
\end{aligned}$$

for $C_1 := C_0 C\kappa$. After summing over $k \ge i$, $0 \le i \le n-1$ and $1 \le j \le m$, the square bracket on the right remains bounded and this completes the proof of Theorem 2(b).

4. Small value probabilities for self-intersection local times

In this section, we look at a single Brownian motion and its q-fold self-intersection local time

$$X(t) := \int_{-\infty}^{\infty} L^q(x,t)\,\mathrm{d}x.$$

This corresponds to the case $m = 1$ of the scenario described in Section 3 and, as mentioned there, this is quite different from the case $m > 1$. The argument used to study the Böttcher case of the Galton–Watson limit can be used to give an extremely simple proof of the following result.

Theorem 3. *Suppose that* $(L(x,t) : x \in \mathbb{R}, t \geq 0)$ *is the local time field and* $\sigma := \inf\{t \geq 0 : |B(t)| = 1\}$ *the first hitting time of level one of a Brownian motion. Then, for every* $q \geq 1$, *we have*

$$-\log \mathbb{P}\left\{\int_{-\infty}^{\infty} L^q(x,\sigma)\,\mathrm{d}x < \varepsilon\right\} \asymp \varepsilon^{-1/q}.$$

Remark. The behaviour is radically different when the Brownian motion is stopped at a fixed time instead of a fixed level. Indeed, in the proof of Theorem 3, we will see that the optimal strategy to make $X(\sigma)$ small is simply to make σ small, an option which cannot be used to make $X(1)$ small. It was shown, for $q = 2$ in Hofstad *et al.* (1997), Proposition 1, and extended to general $q > 1$ by Xia Chen and Wenbo Li (unpublished), that there is a constant $c(q) > 0$ such that

$$-\log \mathbb{P}\left\{\int_{-\infty}^{\infty} L^q(x,1)\,\mathrm{d}x < \varepsilon\right\} \sim c(q)\varepsilon^{-2/(q-1)}.$$

4.1. Self-intersection local time: the branching tree heuristic

We first show how to establish the analogy between the q-fold self-intersection local times and the martingale limit of a Galton–Watson tree in the Böttcher case. The idea is to construct a nested family of random walks embedded into the Brownian path: the natural nesting of the embedded walks establishes the tree structure and a constant multiple of the total number of steps of the finest embedded walk approximates the q-fold self-intersection local times.

Let $(W(t) : t \geq 0)$ be a Brownian motion started at the origin, for each non-negative integer n, let

$$\mathfrak{D}_n := \left\{k2^{-n} : k \in \{-2^n, \ldots, 2^n\}\right\}$$

be the collection of dyadic points of the nth stage and let $0 = \tau_0^{(n)} < \tau_1^{(n)} < \cdots < \tau_{N(n)}^{(n)} = \sigma$ be the collection of stopping times defined for $j \geq 1$ by

$$\tau_j^{(n)} := \inf\left\{t > \tau_{j-1}^{(n)} : W(t) \in \mathfrak{D}_n,\ W(t) \neq W\left(\tau_{j-1}^{(n)}\right)\right\}.$$

Then $(X^{(n)}(j) : 0 \leq j \leq N(n))$, defined by

$$X^{(n)}(j) := 2^n W\left(\tau_j^{(n)}\right),$$

is the nth embedded random walk and $N(n)$ its length. We assign $N(1)$ offspring to the root so that the vertices in the first generation correspond to the steps of height $1/2$ which the path takes to reach level 1 or -1 for the first time. The number of children of each vertex in the first generation is then determined by the number of steps of height $1/4$ that the path makes during

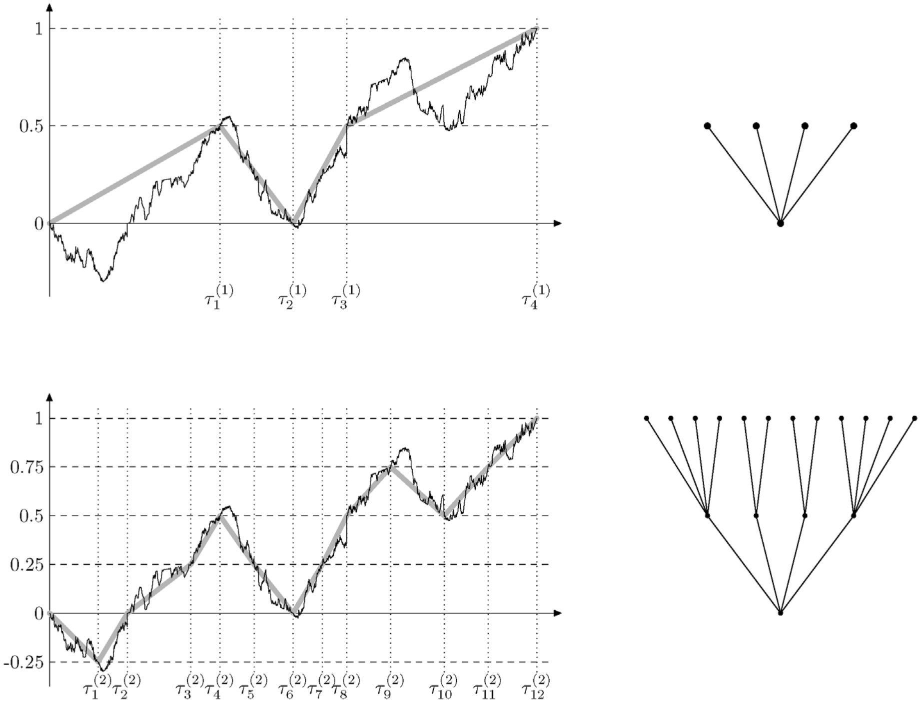

Figure 4. On the left, the first two embedded random walks with step sizes $\frac{1}{2}$ (resp., $\frac{1}{4}$); on the right, the corresponding first two generations of the associated tree.

the step of height $1/2$ corresponding to that vertex. This will be iterated ad infinitum to map the Brownian path to an infinite tree. Note that the resulting tree is a Galton–Watson tree and every vertex in this tree has at least two offspring so that we are in the Böttcher case (see Figure 4 for illustration).

4.2. Self-intersection local time: the lower bound

Recall from the last subsection the definition of the stopping times $0 = \tau_0^{(n)} < \tau_1^{(n)} < \cdots < \tau_{N(n)}^{(n)} = \sigma$ and that of $N(n)$. Note that $N(n) \geq 2^n$ and $\mathbb{P}\{N(n) = 2^n\} = 2(1/2)^{2^n}$. Hence, for any n and $\varepsilon > 0$,

$$\mathbb{P}\left\{\int_{-\infty}^{\infty} L^q(x,\sigma)\,\mathrm{d}x \leq \varepsilon\right\} \geq \mathbb{P}\left\{\int_{-\infty}^{\infty} L^q(x,\sigma)\,\mathrm{d}x \leq \varepsilon \,\Big|\, N(n) = 2^n\right\} \times 2(1/2)^{2^n}.$$

By scaling, there exists a positive constant $C(q)$ such that, for all $j \in \{1, \dots, N(n)\}$, the random variables

$$Y_j := C(q) 2^{n(1+q)} \int_{-\infty}^{\infty} L^q\big(x, \tau_{j-1}^{(n)}, \tau_j^{(n)}\big)\,\mathrm{d}x$$

have mean one. Given $\varepsilon > 0$, we choose the integer n such that $2^{-(n+1)q} \le C(q)2^{-2q}\varepsilon < 2^{-nq}$. Conditional on $N(n) = 2^n$, for every $x \in \mathbb{R}$, we know that in the decomposition

$$L(x,\sigma) = \sum_{j=1}^{2^n} L\big(x, \tau_{j-1}^{(n)}, \tau_j^{(n)}\big),$$

only two summands can be non-zero. Thus, using the convexity of $x \mapsto x^q$ for $q \ge 1$, we obtain

$$\int_{-\infty}^{\infty} L^q(x,\sigma)\,\mathrm{d}x \le 2^{q-1} \sum_{j=1}^{2^n} \int_{-\infty}^{\infty} L^q\big(x, \tau_{j-1}^{(n)}, \tau_j^{(n)}\big)\,\mathrm{d}x \le \varepsilon 2^{-1-n} \sum_{j=1}^{2^n} Y_j$$

and the summands on the right are independent, identically distributed random variables with mean one. Hence, by the law of large numbers,

$$\mathbb{P}\bigg\{\int_{-\infty}^{\infty} L^q(x,\sigma)\,\mathrm{d}x \le \varepsilon \,\bigg|\, N(n) = 2^n\bigg\} \ge \mathbb{P}\bigg\{2^{-n} \sum_{j=1}^{2^n} Y_j \le 2 \,\bigg|\, N(n) = 2^n\bigg\} \overset{n\uparrow\infty}{\longrightarrow} 1$$

and, altogether, for $c(q) := 4(\log 2)C(q)^{-1/q} > 0$ and all large values of n,

$$\mathbb{P}\bigg\{\int_{-\infty}^{\infty} L^q(x,\sigma)\,\mathrm{d}x \le \varepsilon\bigg\} \ge (1/2)^{2^n} \ge \exp(-c(q)\varepsilon^{-1/q}).$$

4.3. Self-intersection local time: the upper bound

Using the notation from the previous section, given $\varepsilon > 0$, we choose the integer n such that $2^{-(n+1)q} \le 2C(q)\varepsilon < 2^{-nq}$. Using the super-additivity of $x \mapsto x^q$ for $q \ge 1$, we obtain

$$\int_{-\infty}^{\infty} L^q(x,\sigma)\,\mathrm{d}x \ge \sum_{j=1}^{N(n)} \int_{-\infty}^{\infty} L^q\big(x, \tau_{j-1}^{(n)}, \tau_j^{(n)}\big)\,\mathrm{d}x \ge \varepsilon 2^{-n+1} \sum_{j=1}^{2^n} Y_j.$$

Hence, we obtain

$$\mathbb{P}\bigg\{\int_{-\infty}^{\infty} L^q(x,\sigma)\,\mathrm{d}x < \varepsilon\bigg\} \le \mathbb{P}\bigg\{2^{1-n} \sum_{j=1}^{2^n} Y_j < 1\bigg\} = \mathbb{P}\{S(2^n) > 0\},$$

where $S(k) := \sum_{j=1}^{k} X_j$ for $X_j := \frac{1}{2} - Y_j$. By the simple large deviation bound for the sum of bounded random variables with negative mean given in Section 2.5, we deduce the existence of a constant $0 < \varphi < 1$ such that

$$-\log \mathbb{P}\bigg\{\int_{-\infty}^{\infty} L^q(x,\sigma)\,\mathrm{d}x < \varepsilon\bigg\} \ge -\log \mathbb{P}\{S(2^n) \ge 0\} \ge (-\log\varphi)2^n \ge \tilde{c}(q)\varepsilon^{-1/q}$$

for the constant $\tilde{c}(q) := (-\log\varphi)(2^{-1-1/q}C(q)^{-1/q}) > 0$.

5. Outlook to future research

Small value probabilities for intersection local times of Brownian motions in dimensions two and three are considerably more difficult to handle, but, in principle, our method still applies. An analog of Theorem 2 for Brownian motions in dimensions two and three is proved, using the branching tree heuristic, in Mörters and Shieh (2007); see also Klenke and Mörters (2005) for partial results and their applications in multifractal analysis.

There is no direct analog to Theorem 3 for a higher-dimensional Brownian motion. However, our main results have natural analogs for random walks and in the random walk setting, problems analogous to Theorem 3 can also be tackled in higher dimensions. This research project, together with some applications to weakly self-avoiding walks, is currently ongoing.

Finally, it is natural to ask whether the main results of the present paper can be extended from Brownian motion to Lévy processes. It appears that the approach presented here may be suited to such an extension and further investigations of this problem appear promising.

Acknowledgements

We thank the Nuffield Foundation for awarding an Undergraduate Research Bursary which allowed us to study the first example. The first author would like to thank Greg Lawler and Wenbo Li for useful discussions and the EPSRC for support through grant EP/C500229/1 and an Advanced Research Fellowship.

References

Athreya, K.B. and Ney, P.E. (1972). *Branching Processes*. New York: Springer. MR0373040

Bingham, N.H. (1988). On the limit of a supercritical branching process. *J. Appl. Probab.* **25A** 215–228. MR0974583

Dereich, S., Fehringer, F., Matoussi, A. and Scheutzow, M. (2003). On the link between small ball probabilities and the quantization problem for Gaussian measures on Banach spaces. *J. Theoret. Probab.* **16** 249–265. MR1956830

Dubuc, S. (1971a). La densité de la loi-limite d'un processus en cascade expansif. *Z. Wahrsch. Verw. Gebiete* **19** 281–290. MR0300353

Dubuc, S. (1971b). Problèmes relatifs à l'itération de fonctions suggérés par les processus en cascade. *Ann. Inst. Fourier (Grenoble)* **21** 171–251. MR0297025

Graf, S., Luschgy, H. and Pagès, G. (2003). Functional quantization and small ball probabilities for Gaussian processes. *J. Theoret. Probab.* **16** 1047–1062. MR2033197

Hofstad, R. v. d., den Hollander, F. and König, W. (1997). Central limit theorem for the Edwards model. *Ann. Probab.* **25** 573–597. MR1434119

Klartag, B. and Vershynin, R. (2007). Small ball probability and Dvoretzky theorem. *Israel J. Math.* **157** 193–207. MR2342445

Klenke, A. and Mörters, P. (2005). The multifractal spectrum of Brownian intersection local times. *Ann. Probab.* **33** 1255–1301. MR2150189

Lawler, G. (1996). Hausdorff dimension of cut points for Brownian motion. *Electron. J. Probab.* **1** 1–20. MR1386294

Li, W. and Linde, W. (1999). Approximation, metric entropy and small ball estimates for Gaussian measures. *Ann. Probab.* **27** 1556–1578. MR1733160

Li, W. and Shao, Q. (2001a). Capture time of Brownian pursuits. *Probab. Theory Related Fields* **121** 30–48. MR1857107

Li, W. and Shao, Q. (2001b). Gaussian processes: Inequalities, small ball probabilities and applications. In *Stochastic Processes: Theory and Methods. Handbook of Statistics* **19** (C. Rao and D. Shanbhag, eds.) 533–597. Amsterdam: North-Holland. MR1861734

Lifshits, M. (2006). Bibliography of small deviation probabilities. Updated version downloadable from http://www.proba.jussieu.fr/pageperso/smalldev/biblio.pdf.

Mörters, P. and Shieh, N.-R. (2007). The exact packing measure of Brownian double points. *Probab. Theory Related Fields*. To appear.

Received January 2007