BRANCHING RANDOM WALKS AND GAUSSIAN FIELDS
Notes for Lectures

Ofer Zeitouni

Department of Mathematics
Weizmann Institute, Rehovot 76100, Israel
and
University of Minnesota, Minneapolis, MN 55455, USA

October 20, 2012. Version 3.3b

1 Introduction

These lecture notes are designed to accompany a mini-course on extrema of branching random walks (BRW) and Gaussian fields. The first half of this sentence, extrema of branching random walks, is of course a classical topic. The "modern" part of the theory, which can be traced back to the work of Mckean and of Bramson, is already quite old, and it is not unreasonable to wonder why would anybody be presenting a topics course based on this theory in 2012. My excuse for revisiting this topic is twofold. First, the methods developed for studying branching random walks have recently become relevant in seemingly unrelated problems, such as the study of the maximum of certain Gaussian fields (we will discuss in some detail such an application, to the study of the so called Gaussian Free Field in two dimensions, the 2DGFF); there are conjectured (and some proved) relations with other problems, like the cover time of graphs by simple random walk. Second, new results and questions have recently emerged even in the context of branching random walks. One can mention in particular the recent study of the point process generated by branching Brownian motion [ABBS11, ABK11], the construction of the limit law of the maximum of BRW [Ai11] or the study of BRW in some critical models [BDMM07, BBS10]. While these are not discussed in these notes, my hope that the latter serve as a quick introduction to the area which is a good enough basis on which one can explore these recent exciting developments. As a teaser toward some of these "new" topics, I did include some material related to time inhomogeneous branching random walks and the surprising phase transitions that they exhibit.

Of course, the presentation is skewed toward my second main goal, namely an explanation of recent results concerning the maximum of the 2DGFF. As it turns out, with the right point of view one can relate this a-priori hard question to a question concerning a class of (modified) branching random walks, which can be analysed using the tools presented in the first part. Many questions in this setup are still left open: study of the law of the maximum of the 2DGFF, structure of the process of near maxima, extensions to non Gaussian Ginzburg-Landau fields, tight relations with covering problems, etc. I hope that the exposition will convince some of the readers to look closer into these questions.

Finally, a caveat is in order: my goal was not to present the sharpest results, nor to strive for full generality. While I do sometimes point out to relevant extensions and generalizations in suitable remarks, these notes are not an encyclopedic treatment, and by necessity I have kept the bibliography rather limited. This should not be interpreted as a statement on the value of results that are not described or works that are not cited, but rather as a compromise reflecting what I hope can be reasonably covered in 7-10 hours.

I thank Amine Asselah, Ming Fang, Nicolas Flammarion, Pascal Maillard and Olivier Zindy for their comments on earlier versions of these notes.

2 Branching Random Walks

Branching random walks (BRWs), and their continuous time counterparts, branching Brownian motions (BBMs), form a natural model that describe the evolution of a population of particles where spatial motion is present. Groundbreaking work on this, motivated by biological applications, was done in the 1930's by Kolmogorov-Petrovsky-Piskounov and by Fisher. The model itself exhibit a rich mathematical structures; for example, rescaled limits of such processes lead to the study of superprocesses, and allowing interactions between particles creates many challenges when one wants to study scaling limits.

Our focus is slightly different: we consider only particles in $\mathbb{R}$, and are mostly interested in the atypical particles that "lead the pack". Surprisingly, this innocent looking question turns out to show up in unrelated problems, and in particular techniques developed to handle it show up in the study of the two dimensional Gaussian Free Field, through an appropriate underlying tree structure.

2.2 Definitions and models

We begin by fixing notation. Let $\mathcal{T}$ be a tree rooted at a vertex o, with vertex set V and edge set E. We denote by $|v|$ the distance of a vertex v from the root, i.e. the length of the geodesic (=shortest path, which is unique) connecting v to o, and we write $o \leftrightarrow v$ for the collection of vertices on that geodesic (including o and v). With some abuse of notation, we also write $o \leftrightarrow v$ for the collection of *edges* on the geodesic connecting o and v. Similarly, for $v, w \in V$, we write $\rho(v, w)$ for the length of the unique geodesic connecting v and w, and define $v \leftrightarrow w$ similarly. The nth generation of the tree is the collection $D_n := \{v \in V : |v| = n\}$, while for $v \in D_m$ and $n > m$, we denote by

$$D_n^v = \{w \in D_n : \rho(w,v) = n - m\}$$

the collection of descendants of v in D_n. Finally, the degree of the vertex v is denoted d_v.

Let $\{X_e\}_{e \in E}$ denote a family of independent (real valued throughout this course) random variables attached to the edges of the tree $\mathcal{T}$. For $v \in V$, set $S_v = \sum_{e \in o \leftrightarrow v} X_e$. The *Branching Random Walk* (BRW) is simply the collection of random variables $\{S_v\}_{v \in V}$. We will be interested in the *maximal displacement* of the BRW, defined as

$$M_n = \max_{v \in D_n} S_v \,.$$

Subsets of the following assumptions will be made throughout.

– [**A1**] The variables $\{X_e\}_{e \in E}$ are i.i.d., of common law μ.

– [**A2**] μ possesses super-exponential tails:

$$E_\mu(e^{\lambda X_e}) =: e^{\Lambda(\lambda)} < \infty, \quad \lambda \in \mathbb{R}. \tag{2.1.1}$$

– [**A3**] The tree $\mathcal{T}$ is a k-ary tree, with $k \geq 2$: $d_o = k$ and $d_v = k + 1$ for $v \neq o$.

Whenever assumptions **A1-A3** hold, introduce the large deviations rate function associated with Λ:

$$I(x) = \sup_{\lambda \in \mathbb{R}} (\lambda x - \Lambda(\lambda)) \,, \tag{2.1.2}$$

which is strictly convex and has compact level sets. Set x^* to be the unique point so that $x^* > E_\mu(X_e)$ and $I(x^*) = \log k$. We then have $I(x^*) = \lambda^* x^* - \Lambda(\lambda^*)$ where $x^* = \Lambda'(x^*)$ and $I'(x^*) = \lambda^*$.

2.2 Warm up: getting rid of dependence

We begin with a warm-up computation. Note that M_n is the maximum over a collection of k^n variables, that are not independent. Before tackling computations related to M_n, we first consider the same question when those k^n variables are independent. That is, let $\{\tilde{S}_v\}_{v \in D_n}$ be a collection of i.i.d. random variables, with $\tilde{S}_v$ distributed like S_v, and let $\tilde{M}_n = \max_{v \in D_n} \tilde{S}_v$. We then have the following, which we state for simplicity only in the non-lattice case.

Theorem 1. *With notation as above, and non-lattice assumption of the distribution of X_e, there exists a constant C so that*

$$P(\tilde{M}_n \leq \tilde{m}_n + x) \to \exp(-C e^{-I'(x^*) x}) \,, \tag{2.2.1}$$

where

$$\tilde{m}_n = n x^* - \frac{1}{2 I'(x^*)} \log n. \tag{2.2.2}$$

Proof. The key is the estimate, valid for $a_n = o(\sqrt{n})$,

$$P(\tilde{S}_v > nx^* - a_n) \sim \frac{C}{\sqrt{n}} \exp(-nI(x^* - a_n/n)) \,, \tag{2.2.3}$$

which can be proved, following Bahadur and Rao, by a change of measure (tilting) and the (local) CLT, see [DZ98, Proof of Theorem 3.7.4] for a similar argument (here, we use the non-lattice assumption). With our assumptions, I is smooth at x^* (since x^* is in the interior of the domain of I), and hence

$$nI(x^* - a_n/n) = nI(x^*) - I'(x^*)a_n + o(1) \,.$$

Therefore, recalling that $I(x^*) = \log k$,

$$\begin{aligned} P(\tilde{M}_n \le nx^* - a_n) &\sim \left(1 - \frac{C}{k^n \sqrt{n}} e^{I'(x^*)a_n + o(1)}\right)^{k^n} \\ &\sim \exp(-Ce^{I'(x^*)a_n + o(1)}/\sqrt{n}) \,. \end{aligned}$$

Choosing now $a_n = \log n / 2I'(x^*) - x$, one obtains

$$P(\tilde{M}_n \le m_n + x) \sim \exp(-Ce^{-I'(x^*)x} + o(1)) \,.$$

The claim follows. □

Remark 1. With some effort, the constant C can also be evaluated, but this will not be of interest to us. On the other hand, the constant in front of the $\log n$ term will play an important role in what follows.

Remark 2. Note the very different asymptotics of the right and left tails: the right tail decays exponentially while the left tail is doubly exponential. This is an example of extreme distribution of the Gumbel type.

2.3 BRW: the law of large numbers

As a further warm up, we will attempt to obtain a law of large numbers for M_n. Recall, from the results of Section 2.2, that $\tilde{M}_n/n \to x^*$. Our goal is to show that the same result holds for M_n.

Theorem 2 (Law of Large Numbers). *Under assumptions **A1**-**A3**, we have that*

$$\frac{M_n}{n} \to_{n\to\infty} x^* \,, \quad \textit{almost surely} \tag{2.3.1}$$

Proof. While we do not really need in what follows, we remark that the almost sure convergence can be deduced from the subadditive ergodic theorem. Indeed, note that each vertex in D_n can be associated with a word $a_1 \ldots a_n$ where $a_i \in \{1, \ldots, k\}$. Introduce an arbitrary (e.g., lexicographic) order on the vertices of D_n, and define

$$v_m^* = \min\{v \in D_m : S_v \geq \max_{w \in D_m} S_w\}\,.$$

For $n > m$, write

$$M_n^m = \max_{w \in D_n^{v_m^*}} S_w - S_{v_m^*}\,.$$

We then have, from the definitions, that $M_n \geq M_m + M_n^m$, and it is not hard to check that M_n possesses all moments (see the first and second moment arguments below). One thus concludes, by applying the subadditive ergodic theorem (check the stationarity and ergodicity assumptions, which here follow from independence!), that $M_n/n \to c$, almost surely, for some constant c. Our goal is now to identify c.

The upper bound Let $\bar{Z}_n = \sum_{v \in D_n} \mathbf{1}_{S_v > (1+\epsilon)x^* n}$ count how many particles, at the nth generation, are at location greater than $(1+\epsilon)nx^*$. We apply a first moment method: we have, for any $v \in D_n$, that

$$E\bar{Z}_n \leq k^n P(S_v > n(1+\epsilon)x^*) \leq k^n e^{-nI((1+\epsilon)x^*)}\,,$$

where we applied Chebychev's inequality in the last inequality. By the strict monotonicity of I at x^*, we get that $E\bar{Z}_n \leq e^{-nc(\epsilon)}$, for some $c(\epsilon) > 0$. Thus,

$$P(M_n > (1+\epsilon)nx^*) \leq E\bar{Z}_n \leq e^{-c(\epsilon)n}\,.$$

It follows that

$$\limsup_{n\to\infty} \frac{M_n}{n} \leq x^*\,, \quad \text{almost surely}\,.$$

The lower bound A natural way to proceed would have been to define

$$\underline{Z}_n = \sum_{v \in D_n} \mathbf{1}_{S_v > (1-\epsilon)x^* n}$$

and to show that with high probability, $\underline{Z}_n \geq 1$. Often, one handles this via the second moment method: recall that for any nonegative, integer valued random variable Z,

$$EZ = E(Z\mathbf{1}_{Z\geq 1}) \leq (EZ^2)^{1/2}(P(Z \geq 1))^{1/2}$$

and hence

$$P(Z \geq 1) \geq \frac{(EZ)^2}{E(Z^2)}. \tag{2.3.2}$$

In the case of independent summands, we obtain by this method that

$$P(\tilde{M}_n \geq (1-\epsilon)x^* n) \geq \frac{k^{2n} P(\tilde{S}_v \geq (1-\epsilon)x^* n)^2}{k^n(k^n-1)P(\tilde{S}_v \geq (1-\epsilon)x^* n)^2 + k^n P(\tilde{S}_v \geq (1-\epsilon)x^* n)}\,.$$

Since (e.g., by Cramer's theorem of large deviations theory),

$$\alpha_n := k^n P(\tilde{S}_v \geq (1-\epsilon)x^* n) \to \infty\,, \quad \text{exponentially fast}$$

one obtains that

$$P(\tilde{M}_n \geq (1-\epsilon)x^* n) \geq \frac{1}{\frac{k^n-1}{k^n} + 1/\alpha_n} \geq 1 - e^{-c'(\epsilon)n}\,,$$

implying that

$$\liminf \frac{\tilde{M}_n}{n} \geq x^*\,, \quad \text{almost surely.}$$

Any attempt to repeat this computation with M_n, however, fails, because the correlation between the events $\{S_v > nx^*(1-\epsilon)\}$ and $\{S_w > nx^*(1-\epsilon)\}$ with $v \neq w$ is too large (check this!). Instead, we will consider different events, whose probability is similar but whose correlation is much smaller. Toward this end, we keep track of the trajectory of the ancestors of particles at generation n. Namely, for $v \in D_n$ and $t \in \{0, \ldots, n\}$, we define the ancestor of v at levet t as $v_t := \{w \in D_t : \rho(v,w) = n-t\}$. We then set $S_v(t) = S_{v_t}$, noting that $S_v = S_v(n)$ for $v \in D_n$. We will later analyze in more detail events involving $S_v(t)$, but our current goal is only to prove a law of large number. Toward this end, define, for $v \in D_n$, the event

$$B_v^\epsilon = \{|S_v(t) - x^* t| \leq \epsilon n, t = 1, \ldots, n\}\,.$$

We now recall a basic large deviations result.

Theorem 3 (Varadhan, Mogulskii). *Under Assumption **A2**,*

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log P(B_v^\epsilon) = \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log P(B_v^\epsilon) = -I(x^*)\,.$$

Define now

$$Z_n = \sum_{v \in D_n} \mathbf{1}_{B_v^\epsilon}\,.$$

By theorem 3, we have that

$$EZ_n \geq e^{-c(\epsilon)n}\,. \tag{2.3.3}$$

To obtain an upper bound requires a bit more work. Fix a pair of vertices $v, w \in D_n$ with $\rho(v,w) = 2r$. Note that the number of such (ordered) pairs is $k^{n+r-1}(k-1)$. Now, using independence in the first equality, and homogenuity in the first inequality,

$$\begin{aligned}
P(B_v^\epsilon \cap B_w^\epsilon) &= P(|S_v(t) - x^* t| \leq \epsilon n, t = 1, \ldots, n-r) \\
&\quad \cdot E\left(P\left(|S_v(t) - x^* t| \leq \epsilon n, t = n-r+1, \ldots, n | S_v(n-r)\right)\right)^2 \\
&\leq P(|S_v(t) - x^* t| \leq \epsilon n, t = 1, \ldots, n-r) \\
&\quad \cdot P(|S_v(t) - x^* t| \leq 2\epsilon n, t = 1, \ldots, r)^2\,.
\end{aligned}$$

Using Theorem 3, we then get that

$$P(B_v^\epsilon \cap B_w^\epsilon) \le e^{-(n-r)I(x^*)-2rI(x^*)+c(\epsilon)n},$$

where $c(\epsilon) \to_{\epsilon\to 0} 0$. Therefore,

$$EZ_n^2 \le \sum_{r=0}^{n} k^{n+r} e^{-(n+r)I(x^*)+c(\epsilon)n} = e^{c(\epsilon)n}.$$

It follows from (2.3.2), (2.3.3) and the last display that, for any $\delta > 0$,

$$P(\exists v \in D_n : S_v \ge (1-\delta)x^* n) \ge e^{-o(n)}. \tag{2.3.4}$$

It seems that (2.3.4) is not quite enough to conclude. However, that turns out not to be the case. Indeed, fix $\epsilon > 0$, pick a value x so that $P(X_e > x) > 1/k$, and consider the tree $\mathcal{T}_\epsilon$ of depth ϵn which corresponds to independent bond percolation on $\mathcal{T}$ in levels $1, \ldots, \epsilon n$, keeping only those edges e with $X_e > x$. Because the percolation is supercritical (due to $kP(X_e > x) > 1$), there exists a constant C independent of ϵ such that the event $\mathcal{C}_\epsilon := \{|\mathcal{T}_\epsilon \cap D_{n\epsilon}| > e^{\epsilon C n}\}$ has probability at least $C_x > 0$, with $C_x \to_{x\to-\infty} 1$. By independence, we then conclude that

$$P(M_n \ge n(1-\epsilon)x^* + n\epsilon x) \ge C_x[1-(1-e^{-o(n)})^{e^{\epsilon C n}}] \to_{n\to\infty} C_x.$$

Taking now $n \to \infty$ followed by $\epsilon \to 0$ and finally $x \to -\infty$ yields (exercise!) that

$$\liminf_{n\to\infty} \frac{M_n}{n} \ge x^*, \quad \text{almost surely.} \quad \square$$

Exercise 1. Complete the proof that $EM_n/n \to x^*$.

2.4 A prelude to tightness: the Dekking-Host argument

The law of large number in Theorem 2 is weaker than the statement in Theorem 1 in two respects: first, no information is given in the latter concerning corrections from linear behavior, and second, no information is given, e.g., on the tightness of $M_n - EM_n$, let alone on its convergence in distribution. In this short section, we describe an argument, whose origin can be traced to [DH91], that will allow us to address the second point, once the first has been settled.

The starting point is the following recursion:

$$M_{n+1} \stackrel{d}{=} \max_{i=1}^{k}(M_{n,i} + X_i), \tag{2.4.1}$$

where $\stackrel{d}{=}$ denotes equality in distribution, $M_{n,i}$ are independent copies of M_n, and X_i are independent copies of X_e which are also independent of the collection $\{M_{n,i}\}_{i=1}^k$. Because of the independence and the fact that $EX_i = 0$, we have that

$$E\left(\max_{i=1}^{k}(M_{n,i}+X_i)\right) \geq E\left(\max_{i=1}^{k}(M_{n,i})\right)\,.$$

Therefore,

$$EM_{n+1} \geq E\left(\max_{i=1}^{k}(M_{n,i})\right) \geq E\left(\max_{i=1}^{2}(M_{n,i})\right)\,.$$

Using the identity $\max(a,b) = (a+b+|a-b|)/2$, we conclude that

$$E(M_{n+1}-M_n) \geq \frac{1}{2}E|M_n - M_n'|\,, \tag{2.4.2}$$

where M_n' is an independent copy of M_n.

The importance of (2.4.2) cannot be over-estimated. First, suppose that there exists $K<\infty$ such that $X_e < K$, almost surely (this was the setup for which Dekking and Host invented this argument). In that case, we have that $EM_{n+1} - EM_n \leq K$, and therefore, using (2.4.2), we immediately see that the sequence $\{M_n - EM_n\}_{n\geq 1}$ is tight (try to prove this directly to appreciate the power of (2.4.2)). In making this assertion, we used the easy

Exercise 2. Prove that for every $C>0$ there exists a function $f=f_C$ on R with $f(K)\to_{K\to\infty} 0$, such that if X,Y are i.i.d. with $E|X-Y|<C<\infty$, then $P(|X-EX|>K)\leq f(K)$.

(See also Exercise 6 below.)

However, (2.4.2) has implications even when one does not assume that $X_e < K$ almost surely for some K. First, it reduces the question of tightness to the question of computing an upper bound on $EM_{n+1}-EM_n$ (we will provide such a bound, of order 1, in the next section). Second, even without the work involved in proving such a bound, we have the following observation, due to [BDZ11].

Corollary 1. *For any $\delta>0$ there exists a deterministic sequence $\{n_j^\delta\}_{j\geq 1}$ with $\limsup(n_j^\delta/j)\leq(1+\delta)$, so that the sequence $\{M_{n_j^\delta} - EM_{n_j^\delta}\}_{j\geq 1}$ is tight.*

Proof. Fix $\delta\in(0,1)$. By Exercise 1, $EM_n/n\to x^*$. By (2.4.2), $EM_{n+1} - EM_n \geq 0$. Define $n_0^\delta = 0$ and $n_{j+1}^\delta = \min\{n\geq n_j^\delta : EM_{n+1} - EM_n \leq 2x^*/\delta\}$. We have that $n_{j+1}^\delta<\infty$ because otherwise we would have $\limsup EM_n/n \geq 2x^*/\delta$. Further, let $K_n = |\{\ell<n : \ell\notin\{n_j^\delta\}\}|$. Then, $EM_n\geq 2K_nx^*/\delta$, hence $\limsup K_n/n\leq\delta/2$, from which the conclusion follows. □

2.5 Tightness of the centered maximum

We continue to refine results for the BRW, in the spirit of Theorem 1; we will not deal with convergence in law, rather, we will deal with finer estimates on EM_n, as follows.

Theorem 4. *Under Assumptions **A1**-**A3**, we have*

$$EM_n = nx^* - \frac{3}{2I'(x^*)}\log n + O(1)\,. \tag{2.5.1}$$

Remark 3. It is instructive to compare the logarithmic correction term in (2.5.1) to the independent case, see (2.2.2): the constant $1/2$ coming from the Bahadur-Rao estimate (2.2.3) is replaced by $3/2$. As we will see, this change is due to extra constraints imposed by the tree structure, and ballot theorems that are close to estimates on Brownian bridges conditioned to stay positive.

Theorem 4 was first proved by Bramson [Br78] in the context of Branching Brownian Motions. The branching random walk case was discussed in [ABR09], who stressed the importance of certain ballot theorems. Recently, Roberts [Ro11] significantly simplified Bramson's original proof. The proof we present combines ideas from these sources. To reduce technicalities, we often consider only the case of Gaussian increments in the proofs; when we do so, we explicitely mention it.

Proof. We begin with a general ballot theorem; this version can be found in [ABR08, Theorem 1]. For simplicity, we assume throughout that $\mu\,(-1/2,1/2)) > 0$.

Theorem 5 (Ballot theorem). *Let X_i be iid random variables of zero mean, finite variance, with $P(X_1 \in (-1/2,1/2)) > 0$. Define $S_n = \sum_{i=1}^n X_i$. Then, for $0 \le k \le \sqrt{n}$,*

$$P(k \le S_n \le k+1, S_i > 0, 0 < i < n) = \Theta\left(\frac{k+1}{n^{3/2}}\right)\,, \tag{2.5.2}$$

and the upper bound in (2.5.2) *holds for any $k \ge 0$.*

Here, we write that $a_n = \Theta(b_n)$ if there exist constanst $c_1, c_2 > 0$ so that

$$c_1 \le \liminf_{n\to\infty} \frac{a_n}{b_n} \le \limsup_{n\to\infty} \frac{a_n}{b_n} \le c_2\,.$$

Theorem 5 is plausible in that its continuous limit is easily proved by using Brownian motion and the reflection principle, see [Br78], or Bessel processes, see [Ro11].

An easy corollary is the following.

Corollary 2. *Continue with the assumptions and notation of Theorem 5. For $y > 0$, define $S_n^y = S_n + y$. Then there exists a constant C independent of y so that, for $0 \le y \le k < \sqrt{n}/2$,*

$$P(k \le S_n^y \le k+1, S_i^y > 0, 0 < i < n) \le C\left(\frac{(k+1)(y+1)^2}{n^{3/2}}\right)\,. \tag{2.5.3}$$

Remark 4. At the cost of a more complicated proof, the estimate on the right side can be improved to $C(k+1)(y+1)n^{-3/2}$, which is the expression one obtains from the reflection principle.

Proof. For simplicity, we assume throughout the proof that X_1 is Gaussian of zero mean and variance 1. Denote by ρ its density. The modification needed for the general case is routine. Note that

$$\begin{aligned}
&\frac{k+1}{n^{3/2}} \geq P(k-y \leq S_{n+y^2} \leq k+1-y, S_i > 0, 0 < i < n+y^2) \\
&\geq P(0 < S_j, j=1,\ldots,y^2, S_{y^2} \in (y,y+1]) \qquad (2.5.4) \\
&\quad \times \min_{\theta\in[0,1]} \int_0^\infty \rho(x-\theta) P(x+S_i^y > 0, i=1,\ldots,n-1, x+S_{n-1}^y \in [k,k+1]) dx \\
&\geq \frac{c}{(y+1)^2} \min_{\theta\in[0,1]} \int_0^\infty \rho(x-\theta) P(x+S_i^y > 0, i=1,\ldots,n-1, x+S_{n-1}^y \in [k,k+1]) dx\,,
\end{aligned}$$

where in the second inequality we applied the lower bound in the ballot theorem for $n=y^2$ and $k=\sqrt{n}=y$. (Again, a direct computation is also possible by the reflection principle.) We further have that

$$\begin{aligned}
&P(S_i^y > 0, i=1,\ldots,n, S_n^y \in [k,k+1]) \\
&= \int_0^\infty \rho(x) P(x+S_i^y > 0, i=1,\ldots,n-1, x+S_{n-1}^y \in [k,k+1]) dx \\
&\leq c \min_{\theta\in[0,1]} \int_0^\infty \rho(x-\theta) P(x+S_i^y > 0, i=1,\ldots,n-1, x+S_{n-1}^y \in [k,k+1]) dx\,.
\end{aligned}$$

Combining the last display with (2.5.4) completes the proof of the corollary. □

A lower bound on the right tail of M_n Fix $y>0$ independent of n and set

$$a_n = a_n(y) = x^* n - \frac{3}{2I'(x^*)} \log n + y.$$

For $v \in D_n$, define the event

$$A_v = \{S_v \in [a_n - 1, a_n], S_v(t) \leq a_n t/n, t = 1,2,\ldots,n\},$$

and set

$$Z_n = \sum_{v \in D_n} \mathbf{1}_{A_v}\,.$$

In deriving a lower bound on EM_n, we first derive a lower bound on the right tail of the distribution of M_n, using a second moment method. For this, we need to compute $P(A_v)$. Recall that we have $I(x^*) = \lambda^* x^* - \Lambda(\lambda^*) = \log k$, with $\lambda^* = I'(x^*)$. Introduce the new parameter λ_n^* so that

$$\lambda_n^* \frac{a_n}{n} - \Lambda(\lambda_n^*) = I(a_n/n) \,.$$

Since $I'(a_n/n) = \lambda_n^*$, it is easy to check that $\lambda_n^* = \lambda^* - 3I''(x^*)\log n/(2nI'(x^*)) + O((y+1)/n)$.

Define a new probability measure Q on $\mathbb{R}$ by

$$\frac{d\mu}{dQ}(x) = e^{-\lambda_n^* x + \Lambda(\lambda_n^*)} \,,$$

and with a slight abuse of notation continue to use Q when discussing a random walk whose iid increments are distributed according to Q. Note that in the Gaussian case, Q only modifies the mean of P, not the variance.

We can now write

$$\begin{aligned} P(A_v) &= E_Q(e^{-\lambda_n^* S_v + n\Lambda(\lambda_n^*)} \mathbf{1}_{A_v}) \\ &\geq e^{-n[\lambda_n^* a_n/n - \Lambda(\lambda_n^*)]} E_Q(A_v) \\ &= e^{-nI(a_n/n)} P_Q(\tilde{S}_v \in [0,1], \tilde{S}_v(t) \geq 0, t = 1, 2, \ldots, n) \,. \end{aligned} \tag{2.5.5}$$

where $\tilde{S}_v(t) = a_n t/n - S_v(t)$ is a random walk with iid increments whose mean vanishes under Q. Again, in the Gaussian case, the law of the increments is Gaussian and does not depend on n.

Applying Theorem 5, we get that

$$P(A_v) \geq C \frac{1}{n^{3/2}} e^{-nI(a_n/n)} \,. \tag{2.5.6}$$

Since

$$I(a_n/n) = I(x^*) - I'(x^*)\left(\frac{3}{2I'(x^*)} \cdot \frac{\log n}{n} - \frac{y}{n}\right) + O\left(\left(\frac{\log n}{n}\right)^2\right) ,$$

we conclude that

$$P(A_v) \geq C k^{-n} e^{-I'(x^*) y} \,,$$

and therefore

$$EZ_n = k^n P(A_v) \geq c_1 e^{-I'(x^*) y} \,. \tag{2.5.7}$$

We next need to provide an upper bound on

$$EZ_n^2 = k^n P(A_v) + \sum_{v \neq w \in D_n} P(A_v \cap A_w) = EZ_n + k^n \sum_{s=1}^n k^s P(A_v \cap A_{v_s}) \,, \tag{2.5.8}$$

where $v_s \in D_n$ and $\rho(v, v_s) = 2s$.

The strategy in computing $P(A_v \cap A_{v_s})$ is to condition on the value of $S_v(n-s)$. More precisely, with a slight abuse of notation, writing $I_{j,s} = a_n(n-s)/n + [-j, -j+1]$, we have that

$$
\begin{aligned}
&P(A_v \cap A_{v_s}) \qquad (2.5.9)\\
&\leq \sum_{j=1}^{\infty} P(S_v(t) \leq a_n t/n, t=1,2,\ldots,n-s, S_v(n-s) \in I_{j,s})\\
&\quad \times \max_{z\in I_{j,s}} (P(S_v(s) \in [a_n-1,a_n], S_v(t) \leq a_n(n-s+t)/n, t=1,2,\ldots,s|S_v(0)=z))^2 .
\end{aligned}
$$

Repeating the computations leading to (2.5.6) (using time reversability of the random walk) we conclude that

$$
P(A_v \cap A_{v_s}) \leq \sum_{j=1}^{\infty} \frac{j^3}{s^3(n-s)^{3/2}} e^{-j\lambda^*} n^{3(n+s)/2n} k^{-(n+s)} e^{-(n+s)I'(x^*)y/n} . \qquad (2.5.10)
$$

Substituting in (2.5.8) and (2.5.9), and performing the summation over j first and then over s, we conclude that $EZ_n^2 \leq cEZ_n$, and therefore, using again (2.3.2),

$$
P(M_n \geq a_n - 1) \geq P(Z_n \geq 1) \geq cEZ_n \geq c_0 e^{-I'(x^*)y} . \qquad (2.5.11)
$$

This completes the evaluation of a lower bound on the right tail of the law of M_n.

An upper bound on the right tail of M_n A subtle point in obtaining upper bounds is that the first moment method does not work directly - in the first moment one cannot distinguish between the BRW and independent random walks, and the displacement for these has a different logarithmic corrections (the maximum of k^n independent particles is larger).

To overcome this, note the following: a difference between the two scenarios is that at intermediate times $0 < t < n$, there are only k^t particles in the BRW setup while there are k^n particles in the independent case treated in Section 2.2. Applying the first moment argument at time t shows that there cannot be any BRW particle at time t which is larger than $x^*t + C\log n$, while this constraint disappears in the independent case. One thus expect that imposing this constraint in the BRW setup (and thus, pick up an extra $1/n$ factor from the ballot theorem 5) will modify the correction term.

Carrying out this program thus involves two steps: in the first, we consider an upper bound on the number of particles that never cross a barrier reflecting the above mentioned constraint. In the second step, we show that with high probability, no particle crosses the barrier. The approach we take combines arguments from [Ro11] and [ABR09]; both papers build on Bramson's original argument.

Turning to the actual proof, fix a large constant $\kappa > 0$, fix $y > 0$, and define the function

$$
h(t) = \begin{cases} \kappa \log t, & 1 \leq t \leq n/2 \\ \kappa \log(n-t+1), & n/2 < t \leq n . \end{cases} \qquad (2.5.12)
$$

Recall the definition $a_n = a_n(y) = x^* n - \frac{3}{2I'(x^*)}\log n + y$ and let

$$\tau(v) = \min\{t > 0 : S_v(t) \geq (a_n t/n + h(t) + y - 1\} \wedge n\,,$$

and $\tau = \min_{v \in D_n} \tau(v)$. (In words, τ is the first time in which there is a particle that goes above the line $a_n t/n + h(t) + y$.)

Introduce the events

$$B_v = \{S_v(t) \leq a_n t/n + h(t) + y, 0 < t < n, S_v \in [a_n - 1, a_n]\}$$

and define $Y_n = \sum_{v \in D_n} \mathbf{1}_{B_v}$. We will prove the following.

Lemma 1. *There exists a constant c_2 independent of y so that*

$$P(B_v) \leq c_2 (y+1)^3 e^{-I'(x^*)y} k^{-n}\,. \tag{2.5.13}$$

Remark 5. The estimate in Lemma 1 is not optimal (in particular, using the improved estimate mentioned in Remark 4 would reduce the power of $(y+1)$ to 2), however it is sufficient for our needs.

Proof of Lemma 1 (Gaussian case). Let $\beta_i = h(i) - h(i-1)$ (note that β_i is of order $1/i$ and therefore the sequence β_i^2 is summable). Define parameters $\tilde{\lambda}_n^*(i)$ so that

$$\tilde{\lambda}_n^*(i)\left(\frac{a_n}{n} + \beta_i\right) - \Lambda(\tilde{\lambda}_n^*(i)) = I\left(\frac{a_n}{n} + \beta_i\right)\,.$$

Using that $I'(a_n/n + \beta_i) = \tilde{\lambda}_n^*(i)$, one has that

$$\tilde{\lambda}_n^*(i) = \lambda_n^* + I''(a_n/n)\beta_i + O\left(\beta_i^2\right)\,.$$

Define the new probability measures Q_i on $\mathbb{R}$ by

$$\frac{dP}{dQ_i}(x) = e^{-\tilde{\lambda}_n^*(i)x + \Lambda(\tilde{\lambda}_n^*(i))}\,,$$

and use $\tilde{Q}$ to denote the measure where X_i are independent of law Q_i. We have, similarly to (2.5.5),

$$P(B_v) = E_{\tilde{Q}}(e^{-\sum_{i=1}^n \tilde{\lambda}_n^*(i)X_i + \sum_{i=1}^n \Lambda(\tilde{\lambda}_n^*(i))}\mathbf{1}_{B_v})\,. \tag{2.5.14}$$

Expanding around λ_n^* and using that $\sum_{i=1}^n \beta_i = 0$, one gets that on the event $S_n \in [a_n - 1, a_n]$,

$$\sum_{i=1}^n \tilde{\lambda}_n^*(i)X_i - \sum_{i=1}^n \Lambda(\tilde{\lambda}_n^*(i)) = nI(a_n/n) + I''(a_n/n)\sum_{i=1}^n \beta_i X_i + \sum_{i=1}^n O(\beta_i^2)X_i + O(1)\,. \tag{2.5.15}$$

Substituting in (2.5.14), and using again that $\sum_{i=1}^n \beta_i = 0$, one gets

$$P(B_v) \le Cn^{3/2}k^{-n}e^{-I'(x^*)y}E_{\tilde Q}(e^{-I''(a_n/n)\sum_{i=1}^n \beta_i(X_i-a_n/n-\beta_i)+\sum_{i=1}^n \delta_i X_i}\mathbf{1}_{B_v})\,, \tag{2.5.16}$$

where $\delta_i = O(\beta_i^2)$. (The addition of the term $a_n/n+\beta_i$ is done because $E_{\tilde Q}(X_i) = a_n/n+\beta_i$.) Using again that $\sum_{i=1}^n \beta_i = 0$, integration by parts yields $\sum \beta_i(X_i - a_n/n - \beta_i) = -\sum \tilde S(i)\tilde\gamma_i$, where under $\tilde Q$, $\tilde S(i)$ is a random walk with standard Gaussian increments, and $\tilde\gamma_i = \beta_{i+1}-\beta_i$. We thus obtain that with $\theta_i = \delta_{i+1}-\delta_i$,

$$\begin{aligned} P(B_v) &\le Cn^{3/2}k^{-n}e^{-I'(x^*)y}E_{\tilde Q}(e^{I''(a_n/n)\sum_{i=1}^n \tilde S(i)\tilde\gamma_i + \sum_{i=1}^n + \sum_{i=1}^n \theta_i \tilde S_i}\mathbf{1}_{B_v}) \\ &\le Cn^{3/2}k^{-n}e^{-I'(x^*)y}E_{\tilde Q}(e^{\sum_{i=1}^n \tilde S(i)\gamma_i}\mathbf{1}_{B_v})\,, \end{aligned} \tag{2.5.17}$$

where $\gamma_i = O(1/i^2)$. In terms of $\tilde S_i$, we can write

$$B_v = \{\tilde S(t) \le y, \tilde S(n) \in [-1,0]\}\,.$$

Without the exponential term, we have, by Corollary 2, that

$$\tilde Q(B_v) \le c(y+1)^3 n^{-3/2}\,.$$

Our goal is to show that the exponential does not destroy this upper bound. Let

$$\mathcal{C}_- = \{\exists t \in [(\log n)^4, n/2] : \tilde S(t) < -t^{2/3}\},$$

$$\mathcal{C}_+ = \{\exists t \in (n-(\log n)^4, n] : \tilde S(n) - \tilde S(t) < -(n-t)^{2/3}\}.$$

Then,

$$P(\mathcal{C}_- \cup \mathcal{C}_+) \le 2\sum_{t=(\log n)^4}^{n/2} e^{-ct^{1/3}} \le 2e^{-c(\log n)^{4/3}}\,.$$

Since $\gamma_i \sim 1/i^2$, one has $\sum_{(\log n)^4}^{n-(\log n)^4} \gamma_i \to_{n\to\infty} 0$ and further $\sum \tilde S(i)\gamma_i$ is Gaussian of zero mean and bounded variance. We thus obtain

$$\begin{aligned} &E_{\tilde Q}(e^{I''(a_n/n)\sum_{i=1}^n \tilde S(i)\gamma_i}\mathbf{1}_{B_v}) \\ &\le \frac{1}{n^2} + E_{\tilde Q}(e^{I''(a_n/n)\sum_{i=1}^{(\log n)^4}(\tilde S(i)\gamma_i + \tilde S(n-i)\gamma_{n-i})}\mathbf{1}_{B_v\cap \mathcal{C}_-^c \cap \mathcal{C}_+^c})\,. \end{aligned}$$

Denote by

$$B(z,z',t) = \{\tilde S(i) \le z+y, i=1,\ldots,n-t, \tilde S(n-t) \in [z'-1,z']\,.$$

We have, for $0 \le z,z',t < (\log n)^4$, by Corollary 2,

$$E_{\tilde Q}(B(z,z',t)) \le C\frac{(z+z')^3+y^3}{n^{3/2}}\,.$$

We then get

$$\begin{aligned}
&E_{\tilde{Q}}(e^{I''(a_n/n)\sum_{i=1}^{(\log n)^4}(\tilde{S}(i)\gamma_i+\tilde{S}(n-i)\gamma_{n-i})}\mathbf{1}_{B_v\cap\mathcal{C}_-^c\cap\mathcal{C}_+^c})\\
\le & \sum_{z_-,z_+=0}^{(\log n)^4}\sum_{t_-,t_+=1}^{(\log n)^4} e^{c(z_-+z_+)I''(a_n/n)}e^{-cz_-^2t_-^{1/3}/2}e^{-cz_+^2t_+^{1/3}/2}\\
&\quad\times \max_{u,u'\in[0,1]} E_{\tilde{Q}}(B(u+z_-,u+z_++y,t_-+t_+))\\
\le & C\frac{(y+1)^3}{n^{3/2}}.
\end{aligned}$$

Combined with (2.5.17), this completes the proof of Lemma 1. □

We need to consider next the possibility that $\tau = t < n$. Assuming that κ is large enough ($\kappa > 3/2I'(x^*)$ will do), an application of the lower bound (2.5.11) to the descendants of the particle v with $\tau_v < n$ reveals that for some constant c_3 independent of y,

$$E[Y_n|\tau < n] \ge c_3\,.$$

We conclude that

$$P(\tau < n) \le \frac{E(Y_n)P(\tau<n)}{E(Y_n\mathbf{1}_{\tau<n})} = \frac{EY_n}{E(Y_n|\tau<n)} \le cEY_n\,.$$

One concludes from this and Lemma 1 that

$$P(M_n \ge a_n) \le P(\tau<n) + EY_n \le c_5(y+1)^3 e^{-I'(x^*)y}\,. \tag{2.5.18}$$

In particular, this also implies that

$$EM_n \le x^* n - \frac{3}{2I'(x^*)}\log n + O(1)\,. \tag{2.5.19}$$

We finally prove a complementary lower bound on the expectation. Recall, see (2.5.11), that for any $y > 0$,

$$P(M_n \ge a_n(y)) \ge ce^{-I'(x^*)y}\,.$$

In order to have a lower bound on EM_n that complements (2.5.19), we need only show that

$$\lim_{y\to-\infty}\limsup_{n\to\infty}\int_{-\infty}^{y} P(M_n \le a_n(y)) = 0\,. \tag{2.5.20}$$

Toward this end, fix $\ell > 0$ integer, and note that by the first moment argument used in the proof of the LLN (Theorem 2 applied to $\max_{w\in D_\ell}(-S_w)$), there exist positive constants c, c' so that

$$P(\min_{w\in D_\ell}(S_w) \le -c\ell) \le e^{-c'\ell}\,.$$

On the other hand, for each $v \in D_n$, let $w(v) \in D_\ell$ be the ancestor of v in generation ℓ. We then have, by independence,

$$P(M_n \le -c\ell + (n-\ell)x^* - \frac{3}{2I'(x^*)}\log(n-\ell)) \le (1-c_0)^{k^\ell} + e^{-c'\ell}\,,$$

where c_0 is as in (2.5.11). This implies (2.5.20). Together with (2.5.19), this completes the proof of Theorem 4. □

2.6 Time-varying profiles

Before dropping the study of BRW, we discuss some interesting phase transitions that occur when the increments are allowed to depend on time. In doing so, we follow closely [FZ11], to which we refer for more details on the proof than in the sketch that we provide below.

For simplicity, we continue to consider BRW with deterministic binary branching and Gaussian increments. The twist here is that the variance is a function of time and of the time-horizon. More precisely, for $\sigma > 0$, let $N(0, \sigma^2)$ denote the normal distribution with mean zero and variance σ^2. Let n be an integer, and let $\sigma_1^2, \sigma_2^2 > 0$ be given. We start the system with one particle at location 0 at time 0. Suppose that v is a particle at location S_v at time k. Then v dies at time $k+1$ and gives birth to two particles $v1$ and $v2$, and each of the two offspring ($\{vi, i = 1,2\}$) moves independently to a new location S_{vi} with the increment $S_{vi} - S_v$ independent of S_v and distributed as $N(0, \sigma_1^2)$ if $k < n/2$ and as $N(0, \sigma_2^2)$ if $n/2 \le k < n$. As before, let $M_n = \max_{v \in D_n} S_v$ denote the maximal displacement of the walk.

General tightness results, using recursions, establish the tightness of $\{M_n - EM_n\}$ [BZ09, Fa10]. We will however be more interested in the asymptotics of M_n. That is, we will show that

$$M_n = (\sqrt{2\log 2}\sigma_{\text{eff}})n - \beta \frac{\sigma_{\text{eff}}}{\sqrt{2\log 2}} \log n + O(1) \text{ a.s.} \tag{2.6.21}$$

were $\sigma_{\text{eff}}, \beta$ depend not only on the value of σ_1, σ_2 but more important, on their order. That is, we have the following.

Theorem 6. *In the setup above, the relation* (2.6.21) *holds, where*

$$\sigma_{\text{eff}} = \begin{cases} \sqrt{(\sigma_1^2+\sigma_2^2)/2}, & \sigma_1^2 < \sigma_2^2, \\ (\sigma_1+\sigma_2)/2, & \sigma_1^2 > \sigma_2^2, \\ \sigma_1, & \sigma_1^2 = \sigma_2^2\,, \end{cases} \qquad \beta = \begin{cases} 1/2, & \sigma_1^2 < \sigma_2^2, \\ 3, & \sigma_1^2 > \sigma_2^2, \\ 3/2, & \sigma_1^2 = \sigma_2^2. \end{cases} \tag{2.6.22}$$

Note that the function σ_{eff} is continuous in its arguments, whereas β exhibits a phase transition.

Proof of Theorem 6 (sketch) The first step consists of finding the typical behavior, that is, at the level of large deviations, the path followed by the maximal path. Indeed, the probability for a random walk (with variance profile as above) to follow a path $n\phi(nr), r \in [0,1]$, decays exponentially (in n) at rate $I(\phi,1)$, where

$$I(\phi,t)=\int_0^{1/2\wedge t}\frac{\phi'(s)^2}{2\sigma_1^2}+\int_{1/2\wedge t}^{1\wedge t}\frac{\phi'(s)^2}{2\sigma_2^2}\,.$$

A first moment computation (using that at time $k \le n$, there are only 2^k particles) yields the constraint

$$I(\phi,t)\le t\log 2, t\in[0,1]\,.$$

Of course, one aims at maximizing $\phi(1)$, subject to $\phi(0)=0$ and the constraints.

The solution of this variational problem differs according to whether $\sigma_1^2 < \sigma_2^2$ or $\sigma_1^2 > \sigma_2^2$: in either case, by convexity, the maximizing function $\bar\phi$ is piecewise linear, and we only need to compute $\bar\phi(1/2)$ and $\bar\phi(1)$. If $\sigma_1^2<\sigma_2^2$, one obtains

$$\bar\phi(s)=\begin{cases}\dfrac{2\sigma_1^2\sqrt{\log 2}}{\sqrt{(\sigma_1^2+\sigma_2^2)}}s, & 0\le s\le \dfrac12;\\[2ex] \dfrac{2\sigma_1^2\sqrt{\log 2}}{\sqrt{(\sigma_1^2+\sigma_2^2)}}\dfrac12+\dfrac{2\sigma_2^2\sqrt{\log 2}}{\sqrt{(\sigma_1^2+\sigma_2^2)}}(s-\dfrac12), & \dfrac12\le s\le 1,\end{cases} \tag{2.6.23}$$

whereas in case $\sigma_1^2>\sigma_2^2$ one has

$$\bar\phi(s)=\begin{cases}\sigma_1\sqrt{2\log 2}s, & 0\le s\le\dfrac12;\\[2ex] \sigma_1\sqrt{\log 2/2}+(s-1/2)\sigma_2\sqrt{2\log 2}, & \dfrac12\le s\le 1,\end{cases} \tag{2.6.24}$$

In particular, the optimal curve $\bar\phi(t)$ satisfies at all times the constraint with strict inequality (when $\sigma_1^2<\sigma_2^2$) and equality (if $\sigma_1^2\ge\sigma_2^2$). The computation of the LLN, that is of σ_{eff}, now follows the recipe described in Section 2.3, and yields the value for σ_{eff}.

In order to evaluate the logarithmic correction term, when $\sigma_1^2<\sigma_2^2$, for an upper bound one applies directly the first moment method, without worrying about the walk straying above the path $\bar\phi(t)$; one thus replaces the Ballot Theorem 5 with a local CLT. A similar phenomenon occurs for the lower bound, leading to $\beta=1/2$.

When $\sigma_1^2>\sigma_2^2$, the lower bound is very simple: one just consider the maximal descendant of the leading particle at generation $n/2$. The upper bound requires more work, but essentially, the constraint to consider is not just for a random walk to be below a (slightly curved bounday) for time $1,\ldots,n$. Instead, the random walk has to satisfy this *and* be in a small neighborhood of the path at time $n/2$; this results in two Brownian bridge factors, leading to $\beta=3$. □

We remark that by considering a decreasing variance profile with k distinct values, a repeat of the above argument leads to a logarithmic correction term with $\beta = 3k/2$. This immediately leads to the following natural question: what is the maximal slowdown (with respect to the LLN behavior) that can be achieved? Specifically, we conjecture the following.

Conjecture 1. Consider a strictly decreasing smooth function $\sigma : [0,1] \to [1,2]$, and consider binary BRW with increments, at generation k, which are normal of mean 0 and variance $\sigma(k/n)$. Then, that for large n,

$$EM_n = c_\sigma n - g_\sigma(n)\,,$$

where $g_\sigma(n) = O(n^{1/3})$.

In the rest of this section, we provide some evidence toward Conjecture 1. To avoid unimportant technical issues, we do so in the context of Branching Brownian Motions (BBM), with continuous binary branching. Translating this to the BRW setup requires only minimal, and straightforward, adaptations. Our treatment is borrowed from [FZ12].

Thus, consider the BBM model where at time t, all particles move independently as Brownian motions with variance $\sigma_T^2(t) = \sigma^2(t/T)$, and branch independently at rate 1. We consider a smooth strictly decreasing function σ. Let $\xi_i(t), i = 1, \ldots, N_t$ denote the set of particles at time t, let $M_t = \max_i \xi_i(t)$ denote the maximal displacement of the BBM. We are interested in M_T.

Theorem 7. *With $\sigma(\cdot)$ as above, we have that*

$$\mathrm{Med}(M_T) = v_\sigma T - g_\sigma(T)\,, \tag{2.6.25}$$

where v_σ is defined in (2.6.29)*, and*

$$0 < \liminf_{T\to\infty} \frac{g_\sigma(T)}{T^{1/3}} \le \limsup_{T\to\infty} \frac{g_\sigma(T)}{T^{1/3}} < \infty \tag{2.6.26}$$

We emphasize that it is already known a-priori that $\{M_T - EM_T\}$ is a tight sequence, so (2.6.25) would hold with $\mathrm{Med}(M_T)$ replaced by any quantile of M_T or by its mean.

Before bringing the proof of Theorem 7, we collect some preliminary information concerning the path of individual particles. With $W_\cdot$ and $\tilde{W}_\cdot$ denoting standard Brownian motions, let

$$X_t = \int_0^t \sigma_T(s)dW_s\,, \quad t \in [0,T].$$

Let $\tau(t) = \int_0^t \sigma_T^2(s)ds$. Clearly, $X_\cdot$ has the same law as $\tilde{W}_{\tau(t)}$. The following is a standard adaptation of Schilder's theorem, using the scaling properties of Brownian motion.

Theorem 8 (Schilder). *Define* $Z_t = \frac{1}{T} X_{t/T}, t \in [0,1]$*. Then* Z_t *satisfies a large deviation principle in* $C_0[0,1]$ *of speed* T *and rate function*

$$I(f) = \begin{cases} \int_0^1 \frac{f'(s)^2}{2\sigma^2(s)} ds \,, & f \in H_1[0,1]\,, \\ \infty\,, & \text{else} \end{cases} .$$

Here, $H_0[0,1]$ is the space of absolutely continuous function on $[0,1]$ that vanish at 0, whose (almost everywhere defined) derivative is square-integrable.

We now wish to define a barrier for the particle systems that is unlikely to be crossed. This barrier will also serve as a natural candidate for a change of measure. Recall that at time t, with overwhelming probability there are at most $e^{t+o(t)}$ particles alive in the system. Thus, it becomes unlikely that any particle crosses a boundary of the form $Tf(\cdot/T)$ if, at any time,

$$J_t(f) := \int_0^t \frac{f'(s)^2}{2\sigma^2(s)} ds > t\,.$$

This motivates the following lemma.

Lemma 2. *Assume* σ *is strictly decreasing. Then the solution of the variational problem*

$$v_\sigma := \sup\{f(1) : J_t(f) \le t, t \in [0,1]\} \tag{2.6.27}$$

exists, and the unique minimizing path is the function

$$\bar{f}(t) = \sqrt{2}\int_0^t \sigma(s) ds\,. \tag{2.6.28}$$

In particular,

$$v_\sigma = \sqrt{2}\int_0^1 \sigma(s) ds\,. \tag{2.6.29}$$

Proof of Lemma 2: We are going to prove that no other functions can do better than $\bar{f}$. That is, if some absolutely continuous function g satisfies $g(0) = 0$ and the constraint $J_t(g) \le t$ for all $0 \le t \le 1$, then $g(1) \le \bar{f}(1) = v_\sigma$. In fact, denote $\phi(t) = J_t(g) \le t$ for $0 \le t \le 1$, and then $\phi'(t) = \frac{g'(t)^2}{2\sigma^2(t)}$ a.e.. We can write $g^2(1)$ as

$$g^2(1) = \left(\int_0^1 g'(t)dt\right)^2 = \left(\int_0^1 \sqrt{2\phi'(t)}\sigma(t)dt\right)^2.$$

Using Hölder's inequality, we have

$$g^2(1) \le 2\left(\int_0^1 \phi'(t)\sigma(t)dt\right)\left(\int_0^1 \sigma(t)dt\right) = \sqrt{2}v_\sigma\left(\int_0^1 \phi'(t)\sigma(t)dt\right).$$

Using integration by parts, the above is equal to

$$\sqrt{2}v_\sigma \left(\phi(1)\sigma(1) - \int_0^1 \phi(t)\sigma'(t)dt\right).$$

Since $\phi(t) \le t$ and $\sigma'(t) \le 0$ for all $0 \le t \le 1$, the above is less than or equal to

$$\sqrt{2}v_\sigma \left(\sigma(1) - \int_0^1 t\sigma'(t)dt\right) = \sqrt{2}v_\sigma \int_0^1 \sigma(t)dt = v_\sigma^2,$$

where we apply integration by parts in the first equality. This completes the proof. □

Proof of Theorem 7: From Lemma 2 one immediately obtains that $\text{Med}(M_T)/T \le v_\sigma(1+o(1))$. To obtain an a-priori lower bound on EM_T (which will also give a weak form of the lower bound in (2.6.25)), we argue by comparison with a simpler BRW. A crucial role will be played by the following elementary lemma.

Lemma 3. *Let $\mathcal{T}$ be a finite set. Let $\{X_t\}_{t\in\mathcal{T}}, \{Y_t\}_{t\in\mathcal{T}}$ denote two independent collections of zero mean random variables indexed by $t \in \mathcal{T}$. Then*

$$E\max_{t\in\mathcal{T}}(X_t+Y_t) \ge E\max_{t\in\mathcal{T}} X_t\,. \tag{2.6.30}$$

Proof of Lemma 3 Let t^* be such that $\max_{t\in\mathcal{T}} X_t = X_{t^*}$. (If more that one such t^* exists, lake the minimal according to some a-priori order on $\mathcal{T}$.) We have

$$E\max_{t\in\mathcal{T}}(X_t+Y_t) \ge E(X_{t^*}+Y_{t^*}) = EX_{t^*} = E\max_{t\in\mathcal{T}} X_t\,,$$

where the first equality used the independence of $\{X_t\}$ and $\{Y_t\}$, and the fact that $EY_t = 0$. □

The promised (weak) lower bound on EM_T is obtained as follows. Divide the interval $[0,T]$ into $T^{1/3}$ intervals $I_j = [(j-1)T^{2/3}, jT^{2/3}]$ of length $T^{2/3}$. In the interval I_j, fix $\sigma_j = \sigma(jT^{-1/3})$, noting that $\sigma_j \le \sigma(x)$ for $x \in I_j$. Let $\tilde{M}_T$ denote the maximum of BBM where the variance of particles at time $t \in I_j$ is σ_j^2. By lemma 3, $EM_T \ge E\tilde{M}_T$. On the other hand,

$$E\tilde{M}_T \ge \sum_{j=1}^{T^{1/3}} E\tilde{M}^j_{T^{2/3}}\,,$$

where $\tilde{M}^j_{T^{2/3}}$ is the maximum, at time $T^{2/3}$, of BBM with particle variance σ_j^2. Applying Bramson's theorem for BBM (the continuous analogue of Theorem 4) we thus get

$$EM_T \ge \sum_{j=1}^{T^{1/3}} (T^{2/3}\sqrt{2}\sigma_j - C\log T)\,,$$

with the constant C independent of j (here we use that $\sigma_j \in [1,2]$). Since $\sigma_j = \sigma(jT^{-1/3})$ and the function σ is smooth, one has that

$$|T^{2/3} \sum_{j=1}^{T^{1/3}} \sigma_j - T \int_0^1 \sigma(s) ds|$$

is uniformly bounded by a multiple of $T^{2/3}$, and therefore

$$EM_T \geq T v_\sigma - C T^{2/3} ,$$

yielding a weaker form of the lower bound in (2.6.25).

Improving the lower bound to the form in (2.6.25) requires more work. Full details are provided in [FZ12]. We will sketch the proof of slightly weaker result, with a logarithmic correction. The key is to follow the strategy outlined above (dividing the interval $[0, T]$ to intervals of length $T^{2/3}$), but instead of comparison with a BBM with increments constant in time, we use the original BBM with variance that is (very slowly) varying in time. The key is the following lemma, where we set $\bar{\sigma}_j(t) = \sigma_T\left(t + (j-1)T^{2/3}\right)$ and $\bar{\sigma}_j = \int_0^{T^{2/3}} \bar{\sigma}_j(s) ds$.

Lemma 4. *Consider a BBM of binary branching, run over time* $[0, T^{2/3}]$*, with time varying variance* $\bar{\sigma}_j(t)$*, and let* M_j *denote its maximum. Then there exists a constant* A *so that*

$$EM_j \geq \sqrt{2} \bar{\sigma}_j - A \log T . \tag{2.6.31}$$

Indeed, given the lemma, we get

$$EM_T \geq v_\sigma T - A T^{1/3} \log T ,$$

proving (a weak form of) (2.6.26).

Before proceeding to the proof of the upper bound in (2.6.25), we provide the following.

Proof of Lemma 4 (sketch)

The proof uses the second moment method. Note that the path of a single particle can be represented as $W_{\tau(t)}$, where $W_\cdot$ is a Brownian motion and $\tau(t) = \int_0^t \sigma_j(s)^2 ds$. Note also that

$$\tau(T^{2/3}) = T^{2/3} (\bar{\sigma}_j / T^{2/3})^2 + O(1) =: T^{2/3} \eta_j / \sqrt{2} + O(1) .$$

Consider the path of particles, denoted $x_i(t)$, where $i = 1, \ldots, N(T^{2/3})$. Fix a constant C and call a particle *good* if

$$x_i(t) \leq \eta_j t, t \leq T^{2/3}, x_i(T^{2/3}) \geq \eta_j T^{2/3} - C \log T .$$

Set

$$\mathcal{A}_j = \sum_{i=1}^{N(T^{2/3})} \mathbf{1}_{\text{i is a good particle}} .$$

Using the time change $\tau(t)$ and standard computations for Brownian motion and Brownian bridge, one checks that for some C large enough, $E\mathcal{A}_j \geq 1$ and $E((\mathcal{A}_j)^2) \leq e^{B(C)\log T}$. Hence, with this value of C,

$$P(\text{there exists a good particle}) \geq T^{-C'}\,.$$

By truncating the tree of particles at depth $2C'\log T$ and using independence, one conclude that

$$P(\text{there exists } i \in N(T^{2/3}) \text{ with } x_i(T^{2/3}) > \eta_j T^{2/3} - C''\log T) > \frac{1}{2}\,.$$

Using the tightness of M_j one concludes that $EM_j \geq \eta_j T^{2/3} - C''\log T - O(1)$. This completes the proof of Lemma 4. $\square$

It remains to provide a proof of the upper bound in (2.6.25). The first step is to show that in fact, no particle will be found significantly above $T\bar{f}(t/T)$.

Lemma 5. *There exists C large enough such that, with*

$$\mathcal{A} = \{\exists t \in [0,T], i \in \{1,\ldots,N_t\}\xi_i(t) > T\bar{f}(t/T) + C\log T\}\,,$$

it holds that

$$P(\mathcal{A}) \to_{T\to\infty} 0\,. \tag{2.6.32}$$

Proof of Lemma 5: Recall the process $X.$ in $C_0[0,T]$, whose law we denote by P_0. Consider the change of measure with Radon–Nykodim derivative

$$\begin{aligned}\frac{dP_1}{dP_0}|_{\mathcal{F}_t} &= \exp\left(-\int_0^t \frac{\bar{f}'(s/T)}{\sigma^2(s/T)}dX_s - \frac{1}{2}\int_0^t \frac{(\bar{f}'(s/T))^2}{\sigma^2(s/T)}ds\right)\\ &= \exp\left(-\int_0^t \frac{\sqrt{2}}{\sigma(s/T)}dX_s - t\right)\,.\end{aligned} \tag{2.6.33}$$

The process $X.$ under P_0 is the same as the process $X. + T\bar{f}(\cdot/T)$ under P_1. Note that for any $t \leq T$,

$$\int_0^t \frac{\sqrt{2}}{\sigma(s/T)}dX_s = \frac{\sqrt{2}X_t}{\sigma(t/T)} + \frac{\sqrt{2}}{T}\int_0^t X_s\frac{\sigma'(s/T)}{\sigma^2(s/T)}ds\,. \tag{2.6.34}$$

We then have, with $\tau = \inf\{t \leq T : X_t \geq C\log T\}$, on the event $\tau \leq T$,

$$\begin{aligned}\int_0^\tau \frac{(\bar{f}'(s/T))}{\sigma^2(s/T)}dX_s &\geq \frac{\sqrt{2}C\log T}{\sigma(t/T)} + \frac{\sqrt{2}C\log T}{T}\int_0^t \frac{\sigma'(s/T)}{\sigma^2(s/T)}ds\\ &= \frac{\sqrt{2}C\log T}{\sigma(0)}\,,\end{aligned}$$

and therefore, with $\tau' = \inf\{t \leq T : X_t \geq T\bar{f}(t/T) + C\log T\}$, we have, for $k \leq T$,

$$P_0(\tau' \in [k-1,k)) = P_1(\tau \in [k-1,k)) = E_{P_0}\left(\frac{dP_1}{dP_0}\mathbf{1}_{\tau\in[k-1,k)}\right)$$
$$\leq E_{P_0}\left(\mathbf{1}_{\tau\in[k-1,k)}\exp\left(-\frac{\sqrt{2}C\log T}{\sigma(0)} - \tau\right)\right).$$

Define

$$\theta = \inf\{t \leq T : \text{there is } v \in \mathcal{N}_T \text{ so that } x_v(t) \geq T\bar{f}(t/T) + C\log T\},$$

and Z_k to be the number of particles $z \in \mathcal{N}_k$ such that $x_v(t) \leq T\bar{f}(t/T) + C\log T$ for all $t \leq k-1$ and $x_v(t) \geq T\bar{f}(t/T)+C\log T$ for some $k-1 \leq t \leq k$. Then,

$$P(\theta \leq T) \leq \sum_{k=1}^{T} P(\theta \in [k-1,k)) \leq P(Z_k \geq 1),$$

and, using a first moment computation, we obtain

$$P(Z_k \geq 1) \leq EZ_k \leq e^k P_0(\tau' \in [k-1,k)) \leq \exp\left(-\frac{\sqrt{2}C\log T}{\sigma(0)} + 1\right).$$

Therefore,

$$P(\theta \leq T) \leq T\exp\left(-\frac{\sqrt{2}C\log T}{\sigma(0)} + 1\right).$$

This completes the proof of Lemma 5. □

We need one more technical estimate.

Lemma 6. *With $X_\cdot$ and C as in Lemma 5, there exists a constant $C' \in (0,1)$ so that*

$$e^T P_0(X_t \leq T\bar{f}(t/T) + C\log T, t \in [0,T], X_T \geq T\bar{f}(1) - C'T^{1/3}) \to_{T\to\infty} 0. \quad (2.6.35)$$

Proof of Lemma 6: Fix $C' \in (0,1)$. We apply a change of measure similar to the one used in Lemma 6, whose notation we continue to use. We deduce the existence of positive constants c_1, c_2 (independent of T) such that

$$P_0(X_t \leq T\bar{f}(t/T) + C\log T, t \in [0,T], X_T \geq T\bar{f}(1) - C'T^{1/3})$$
$$\leq e^{-T} e^{c_1(C'T^{1/3}+\log T)}$$
$$\cdot E_{P_0}\left(\exp\left(\frac{c_2}{T}\int_0^T X_s ds\right)\mathbf{1}_{X_T\geq -C'T^{1/3}}\mathbf{1}_{X_t\leq 0, t\leq T}\right),$$

where here we used that $-\sigma'$ is bounded below by a positive constant and σ is bounded above. By representing $X_\cdot$ as a time-changed Brownian motion, the lemma will follows (for a small enough C') if we can show that for any constant c_3 there exists a $c_4 = c_4(c_3) > 0$ independent of $C' \in (0,1)$ such that

$$\mathtt{D} := E\left(\exp\left(\frac{c_3}{T}\int_0^T B_s ds\right)\mathbf{1}_{B_T \geq -C'T^{1/3}}\mathbf{1}_{B_t \leq 0, t \leq T}\right) \leq e^{-c_4 T^{1/3}}, \quad (2.6.36)$$

where $\{B_t\}_{t\geq 0}$ is a Brownian motion started at $-C\log T$. Note however that

$$\mathtt{D} \leq E\left(\exp\left(-\frac{c_3}{T}\int_0^T |B_s| ds\right)\mathbf{1}_{|B_T| \leq T^{1/3}}\right) e^{c_5 \log T} \leq e^{-c_4 T^{1/3}},$$

where here $B.$ is a Brownian motion started at 0 and the last inequality is a consequence of known estimates for Brownian motion, see e.g. [BS96, Formula 1.1.8.7, pg. 141]. □

We have completed all steps required for the proof of the upper bound in Theorem 7. Due to the strong tightness result in [Fa10] and Lemma 5, it is enough to show that

$$P(\{M_T \geq \bar{T}f(1) - C'T^{1/3}\} \cap \mathcal{A}^{\mathtt{C}}) \to 0.$$

This follows from the first moment method and Lemma 6. □

Exercise 3. Provide a direct proof of (2.6.36) using the following outline.
a) Divide the time interval $[0,T]$ to intervals of length $T^{2/3}$. Declare an interval *good* if throughout it, the path $B.$ is above $\delta T^{1/3}$ for at least half the length of the interval.
b) Note that regardless of the end point of the previous interval, the probability of an interval to be good is uniformly bounded below.
c) Apply now large deviations for Binomial random variables.

2.7 Additional remarks

We have of course not covered all that is known concerning branching random walks. In particular, we have not discussed results concerning fluctuations of the maximal displacement of BRW over time, where the 3/2 factor becomes 1/2 when lim sups are concerned. More precisely, one has the following result, due to Hu and Shi [HS09].

Proposition 1. *For BRW with increments possessing exponential moments of high enough order, one has*

$$\liminf_{n\to\infty} \frac{M_n - nx^*}{\log n / I'(x^*)} = -\frac{3}{2}, \quad \limsup_{n\to\infty} \frac{M_n - nx^*}{\log n / I'(x^*)} = -\frac{1}{2}.$$

For a modern proof in the case of BBM, see [Ro11]. A discussion of maximal *consistent* displacement is contained in [FZ10] and [FHS10].

We have also not discussed the study of the limit law of the maximal displacement of BRW (and BBM). A recent breakthrough in this direction is [Ai11].

3 The Discrete Gaussian Free Field

We discuss in this chapter maxima of certain Gaussian fields, called *Gaussian free fields*. As we are about to see, in certain cases these are closely related to the maxima of BRWs. We will first introduce the GFF on general graphs, discuss some general tools of Gaussian processes theory, and then specialize our discussion to the two dimensional GFF. For the initiated reader, we state our goal, which is the proof of the following theorem.

Theorem 9. *Let Θ_N denote the maximum of the two dimensional (discrete) GFF in a box of side N with Dirichlet boundary conditions. Then*

$$E\Theta_N = 2m_N + O(1)\,, \tag{3.1.1}$$

where

$$m_N = (2\sqrt{2/\pi})\log N - (3/4)\sqrt{2/\pi}\log\log N\,, \tag{3.1.2}$$

and the sequence $\{\Theta_N - E\Theta_N\}_N$ is tight.

(**Warning:** Be aware that what we call here the GFF differs by a scaling factor 2 from the definitions in [BDG01] and [BZ11]; the difference has to do with the relation between continuous and discrete time random walks. To be consistent, in the proof below we will follow the conventions in the latter papers; this is explained in the beginning of Section 3.3.)

Our exposition follows the following steps. We first introduce in Section 3.2 the GFF on arbitrary graphs, and discuss some of its properties. In Section 3.2 we bring, without proof, some basic tools of Gaussian fields theory. Section 3.3 introduces our main object of study, the two-dimensional (discrete) GFF (2D-GFF). Section 3.4 is devoted to the proof of a law of large numbers for the 2D-GFF, thus providing a (simple) proof to the main result of [BDG01]. Section 3.5 introduces an argument, borrowed from [BDZ11] and based on the Dekking-Host argument, that reduces tightness questions to the precise evaluation of the expectation of the maximum of the 2D-GFF. The latter is evaluated in Section 3.6, following [BZ11], and uses yet another Gaussian field, the modified BRW, that in an appropriate sense interpolates between the GFF and (Gaussian) BRW.

3.2 The Gaussian free field on graphs

We introduce first the Gaussian free field on an arbitrary (finite) graph. Let $G = (\mathbf{V}, \mathbf{E})$ be an (undirected, connected) graph with vertex set $\mathbf{V}$ and edge set $\mathbf{E}$. Let o be a specifed vertex in the graph. The *Gaussian Free Field* on G is the collection of Gaussian variables (Gaussian field) $\{X_v\}_{v\in\mathbf{V}}$, indexed by $\mathbf{V}$, with $X_o = 0$ and where the p.d.f. of the remaining variables is given by the formula (with $x_o = 0$)

$$p(\{x_v\}_{v\in\mathbf{V},v\neq o}) = \frac{1}{Z_G}\exp\left(\sum_{(v,w)\in\mathbf{E}} -(x_v-x_w)^2/2\right), \tag{3.1.1}$$

$$Z_G = \int \exp\left(\sum_{(v,w)\in\mathbf{V}} -(x_v-x_w)^2/2\right)\prod_{v\in V,v\neq o} dx_v\,.$$

As a convenient notation, for any random field $\{X_v\}_{v\in\mathbf{V}}$, we denote by

$$\mathbf{R}_X(v,w) = EX_vX_w$$

its covariance function.

A few comments are now in order. First, note that if G is a tree, the Gaussian free field corresponds to the field obtained by assigning to each edge e a standard zero mean unit variance gaussian variable Y_e, and setting $X_v = \sum_{e\in o\leftrightarrow v} Y_e$ where $o\leftrightarrow v$ denoted the geodesic connecting o and v. In particular, if G is the binary tree rooted at o with depth n, then the values of the GFF on G, denoted $\{X_v^{T,n}\}_v$, at vertices at the nth generation correspond to the location of particles in a BRW with standard Gaussian increments. Therefore, the maximum of the GFF on the binary tree of depth n is distributed like the maximum displacement of all particles. By combining Theorem 4 with (2.4.2), one easily deduces that in this situation,

$$\max_x X_v^{T,n} = \sqrt{2\log 2}n - \frac{3}{2\sqrt{2\log 2}}\log n + O(1)\,.$$

Second, the GFF is closely related to certain random walks. To see that, introduce the (rate) matrix

$$\mathcal{L}_{v,w} := \begin{cases} 1\,, & (v,w)\in\mathbf{E}\,,\\ -d_v\,, & v=w\neq o\,,\\ 0\,, & \text{otherwise}\,,\end{cases}$$

where d_v is the degree of a vertex v. From (3.1.1), the covariance of the GFF $\{X_v\}_{v\neq o}$ is given by $-\bar{\mathcal{L}}^{-1}$, where $\bar{\mathcal{L}}$ is the matrix obtained from $\mathcal{L}$ by striking the row and column corresponding to $v=o$. This coincides with the Green function of a (continuous time) simple random walk $\{B_t\}_{t\geq 0}$ on G which is killed at hitting o, that is, with $\tau=\min\{t: B_t=o\}$, we have

$$\mathbf{R}_X(v,w) = EX_vX_w = E^v(\int_0^\tau \mathbf{1}_{\{B_t=w\}}dt)\,. \tag{3.1.2}$$

In the special case where $d_v=d^*$ for all $v\neq o$, the same reasoning shows that this also corresponds to the Green function of a discrete time simple random walk $\{S_n\}_{n\geq 0}$ killed at hitting o:

$$\mathbf{R}_X(v,w) = EX_vX_w = E^v(\sum_{n=0}^{\tau-1}\mathbf{1}_{\{S_n=w\}})/d^*\,, \tag{3.1.3}$$

where again $\tau = \min\{n : S_n = o\}$ is the killing time for $\{S_n\}$. (Here and throughout, $E^v(\cdot)$ is shorthand for $E(\cdot|S_0 = v)$.) The relation between the GFF and random walks runs deeper, but we postpone for a while a discussion of that.

The last comment relates to the *Markov property* of the GFF. From the definition or the random walk representation, one readily obtains that for any subsets $A \subset B \subset \mathbf{V}$ and any measurable function $F : \mathbb{R}^{|A|} \to \mathbb{R}$,

$$E\left(F(\{X_v\}_{v\in A}) \mid \sigma(X_v, v \in B^{\complement})\right) = E\left(F(\{X_v\}_{v\in A}) \mid \sigma(X_v, v \in \partial B)\right), \tag{3.1.4}$$

where $\partial B = \{v \in B^{\complement} : \exists w \in B, (v,w) \in \mathbf{E}\}$. Further, using the fact that for a Gaussian variable X and Gaussian vector $\mathbf{Y}$,

$$E[X|\mathbf{Y}] = E(X \cdot \mathbf{Y}) R_{\mathbf{Y}}^{-1} \mathbf{Y}',$$

where $'$ denotes transpose, and using that the covariance of the GFF is a Green function, we have that for any $w \in A$,

$$E\left(X_w \mid \sigma(X_v, v \in \partial B)\right) = \sum_{v\in\partial B} u(w,v) X_v, \tag{3.1.5}$$

where $u(w,v) = P^w(X_{\tau_B} = v)$ and $\tau_B = \min\{t \geq 0 : B_t \in B^{\complement}\}$.

Exercise 4. Prove (3.1.5).

3.2 Estimates and comparisons for Gaussian fields

The first tool we introduce shows that for Gaussian fields, some tail estimates on maxima are available a-priori.

Theorem 10 (Borell's inequality). *Let $\mathcal{T}$ be a finite set and let $\{X_t\}_{t\in\mathcal{T}}$ be a zero mean Gaussian field indexed by $\mathcal{T}$. Let $X^*_{\mathcal{T}} = \max_{t\in\mathcal{T}} X_t$. Then*

$$P(|X^*_{\mathcal{T}} - EX^*_{\mathcal{T}}| > x) \leq 2e^{-x^2/2\sigma^2_{\mathcal{T}}}, \tag{3.2.1}$$

where $\sigma^2_{\mathcal{T}} = \max_{t\in\mathcal{T}} EX_t^2$.

The inequality (in a more general form suited to infinite $\mathcal{T}$s) is cited as Borell's inequality in the literature, because of [Bo75]. However, it appeared at the same time in the proceedings [TIS76]. For an accessible and fun discussion, see [Ad90] (and note the errata for the proof).

Theorem 10 already answers some questions related to maxima of GFFs. For example, consider the following.

Corollary 3. *Assume G is an infinite graph on which (continuous time) random walk is transient and possesses a uniformly bounded Green function. Let $\{G_N\}_{N\geq 1}$ denote the sequence of graphs obtained from G by keeping only* *those vertices at distance at most N from the root, and glueing all vertices at distance N as a single vertex o. Let X_N^* denote the maximum of the GFF on G_N with Dirichlet boundary condition at o. Then, the sequence $\{X_N^* - EX_N^*\}_N$ is tight.*

Remark 6. Corollary 3 shows that the maximum of the GFF defined on boxes in $\mathbb{Z}^d$ and shifted around its mean is tight, when $d \geq 3$.

The second tool that we will use extensively is a comparison theorem concerning the expectation of the maxima of (centered) Gaussian fields.

Theorem 11 (Sudakov–Fernique). *Let $\mathcal{T}$ be a finite set and let $\{X_t^i\}_{t\in\mathcal{T}}$, $i = 1, 2$, denote two zero mean Gaussian fields. Set $X_i^* = \max_{t\in\mathcal{T}} X_t^i$. If*

$$E(X_t^1 - X_s^1)^2 \geq E(X_t^2 - X_s^2)^2\,, \quad \text{for all } t, s \in \mathcal{T}\,, \tag{3.2.2}$$

then

$$EX_1^* \geq EX_2^*\,. \tag{3.2.3}$$

For a proof, see [Fe75]. Note that for Gaussian fields, Theorem 11 is a far reaching generalization of the (trivial) Lemma 3.

3.3 The two dimensional discrete GFF

Of course, one of the simplest sequence of graphs to consider is the sequence of boxes in the lattice $\mathbb{Z}^d$, that is we take $V_N = ([0, N-1] \cap \mathbb{Z})^d$. Set $V_N^o = ((0, N-1) \cap \mathbb{Z})^d$ and identify all vertices in $\partial V_N = V_N \setminus V_N^o$, calling the resulting vertex the root of V_N. The collection of vertices thus obtained is denoted $\mathbf{V}_N$, and we take as edge set $\mathbf{E}_N$ the collection of all the (unordered) pairs (x, y) where either $x, y \in V_N^o$ and $|x-y|_1 = 1$ or $x \in V_N^o$, $y = o$ and there exists $z \in \partial V_N$ so that $|x-z|_1 = 1$. We thus obtain a sequence of graphs G_N where all vertices, except for the root, have degree $d^* = 2d$. The GFF on G_N is then defined as in Section 3.2. Keeping in mind the relation between (3.1.2) and (3.1.3), we introduce the field $\{\mathcal{X}_z^N\}_{z\in\mathbf{V}_N}$ as the rescaling by $\sqrt{2d}$ of the GFF:

$$E\mathcal{X}_z^N \mathcal{X}_{z'}^N = E^z\Big(\sum_{k=0}^{\tau-1} \mathbf{1}_{\{S_k = z'\}}\Big)\,, \tag{3.3.1}$$

where $\{S_k\}$ is a simple random walk on G_N killed upon hitting o, with killing time τ. As before we set $\mathcal{X}_N^* = \max_{z\in\mathbf{V}_N} \mathcal{X}_z^N$.

Remark 7. As alluded to above, many authors, including the present one, refer to the field $\mathcal{X}_z^N$ as the GFF. I hope that this extra factor of $\sqrt{2d}$ will not cause too much confusion in what follows.

Recall from Remark 6 that for $d \geq 3$, the sequence $\{\mathcal{X}_N^* - E\mathcal{X}_N^*\}_N$ is tight. On the other hand, for $d = 1$, the GFF is simply a random walk with standard Gaussian steps, conditioned to hit 0 at time N. In particular, $\mathcal{X}_N^*/\sqrt{N}$ scales like the maximum of a Brownian bridge, and thus $\mathcal{X}_N^* - E\mathcal{X}_N^*$ fluctuates at order $\sqrt{N}$. This leads us immediately to the question:

For $d = 2$, what is the order of $\mathcal{X}_N^*$ and are the fluctuations of order $O(1)$?

The rest of this chapter is devoted to the study of that question. In the rest of this section, we provide some a-priori comparisons and estimates.

Lemma 7. *For any $d \geq 1$, the sequence $E\mathcal{X}_N^*$ is monotone increasing in N.*

Proof. Let $N' > N$. For $z \in V_N^o$, write

$$\mathcal{X}_z^{N'} = E[\mathcal{X}_z^{N'}|\mathcal{F}_N] + \Big(\mathcal{X}_z^{N'} - E[\mathcal{X}_z^{N'}|\mathcal{F}_N]\Big) := A_z + B_z\,,$$

where $\mathcal{F}_N = \sigma(\mathcal{X}_{N'}^z : z \in V_{N'} \setminus V_N^o)$ and $\{A_z\}_{z\in V_N^o}$ and $\{B_z\}_{z\in V_N^o}$ are independent zero mean Gaussian fields. By the Markov property (3.1.4), we have that $\{B_z\}_{z\in V_N^o}$ is distributed like $\{\mathcal{X}_z^N\}_{z\in V_N^o}$. Therefore, by Lemma 3, we conclude that $E\mathcal{X}_{N'}^* \geq E\mathcal{X}_N^*$. □

The next lemma is an exercise in evaluating hitting probabilities for simple random walk.

Lemma 8 (GFF covariance, $d = 2$). *Fix $d = 2$. For any $\delta > 0$ there exists a $C = C(\delta)$ such that for any $v, w \in \mathbf{V}_N$ with $d(v, \partial V_N), d(w, \partial V_N) \geq \delta N$, one has*

$$\left|\mathbf{R}_{X^N}(v,w) - \frac{2}{\pi}\left(\log N - (\log \|v-w\|_2)_+\right)\right| \leq C\,. \tag{3.3.2}$$

Further,

$$\max_{x\in\mathbf{V}_N} \mathbf{R}_{X^N}(x,x) \leq (2/\pi)\log N + O(1)\,. \tag{3.3.3}$$

The proof of Lemma 8 can be found in [BDG01, Lemma 1] or [BZ11, Lemma 2.2].

Exercise 5. Using hitting estimates for simple random walks, prove Lemma 8.

3.4 The LLN for the 2D-GFF

We prove in this short section the Bolthausen-Deuschel-Giacomin LLN; our proof is shorter than theirs and involves comparisons with BRW.

Theorem 12. *Fix $d \geq 2$. Then,*

$$E\mathcal{X}_N^* \leq m_N + O(1)\,, \tag{3.4.1}$$

and

$$\lim_{N\to\infty} \frac{E\mathcal{X}_N^*}{m_N} = 1\,, \tag{3.4.2}$$

where m_N *is defined in* (3.1.2). *Further, for any* $\epsilon > 0$ *there exists a constant* $c^* = c^*(\epsilon)$ *so that for all large enough* N,

$$P(|\mathcal{X}_N^* - m_N| \geq \epsilon m_N) \leq 2e^{-c^*(\epsilon)\log N} . \tag{3.4.3}$$

Proof. We note first that (3.4.3) follows from (3.4.2), (3.3.3) and Borell's inequality (Theorem 10). Further, because of the monotonicity statement in Lemma 7, in the proof of (3.4.1) and (3.4.2) we may and will consider $N = 2^n$ for some integer n.

We begin with the introduction of a BRW that will be useful for comparison purposes. For $k = 0, 1, \ldots, n$, let $\mathcal{B}_k$ denote the collection of subsets of $\mathbb{Z}^2$ consisting of squares of side 2^k with corners in $\mathbb{Z}^2$, let $\mathcal{BD}_k$ denote the subset of $\mathcal{B}_k$ consisting of squares of the form $([0, 2^k - 1] \cap \mathbb{Z})^2 + (i2^k, j2^k)$. Note that the collection $\mathcal{BD}_k$ partitions $\mathbb{Z}^2$ into disjoint squares. For $x \in V_N$, let $\mathcal{B}_k(x)$ denote those elements $B \in \mathcal{B}_k$ with $x \in B$. Define similarly $\mathcal{BD}_k(x)$. Note that the set $\mathcal{BD}_k(x)$ contains exactly one element, whereas $\mathcal{B}_k(x)$ contains 2^{2k} elements.

Let $\{a_{k,B}\}_{k\geq 0, B\in\mathcal{BD}_k}$ denote an i.i.d. family of standard Gaussian random variables. The BRW $\{\mathcal{R}_z^N\}_{z\in V_N}$ is defined by

$$\mathcal{R}_z^N = \sum_{k=0}^{n} \sum_{B\in\mathcal{BD}_k(z)} a_{k,B} .$$

We again define $\mathcal{R}_N^* = \max_{z\in V_N} \mathcal{R}_z^N$. Note that $\mathcal{R}_z^N$ is a Branching random walk (with 4 descendants per particle), see the discussion in Section 3.2. Further, the covariance structure of $\mathcal{R}_z^N$ respects a hierarchical structure on V_N^o: for $x, y \in V_N^o$, set $d_H(x, y) = \max\{k : y \notin \mathcal{BD}_k(x)\}$. Then,

$$\mathbf{R}_{\mathcal{R}_N}(x, y) = n - d_H(x, y) \leq n - \log_2 \|x - y\|_2 . \tag{3.4.4}$$

We remark first that, as a consequence of the Markov property (see the computation in Lemma 7),

$$E\mathcal{X}_N^* \leq E \max_{x\in(N/2,N/2)+V_N} \mathcal{X}_x^{2N} .$$

Combined with Lemma 8 and the Sudakov-Fernique (Theorem 11), we thus obtain that for some constant C independent of N,

$$E\mathcal{X}_N^* \leq \sqrt{\frac{2\log 2}{\pi}} E\mathcal{R}_N^* + C .$$

Together with computations for the BRW (the 4-ary version of Theorem 4), this proves (3.4.1).

To see (3.4.2), we dilute the GFF by selecting a subset of vertices in V_N. Fix $\delta > 0$. Define $V_N^{\delta,1} = V_N$ and, for $k = 2, \ldots, n - \log_2(1 - \delta)n - 1$, set

$$V_N^{\delta,k} = \{x \in V_N^{\delta,k-1} : |x-y|_\infty \ge \delta N/2^{n-k}, \forall y \in \cup_{B \in \mathcal{BD}_k} \partial B\}\,.$$

Note that $|V_N^{\delta,k}| \sim (1-\delta)^{2k}|V_N|$. We can now check that for $x, y \in V_N^{\delta,n(1-\log_2(1-\delta))}$, $\log_2 |x-y|_2$ is comparable to $d_H(x,y)$. Obviously,

$$E\mathcal{X}_N^* \ge E(\max_{x \in V_N^{\delta,n(1-\log_2(1-\delta))}} \mathcal{X}_x^N)\,.$$

Applying the same comparison as in the upper bound, the right side is bounded below by the maximum of a diluted version of the BRW, to which the second moment argument used in obtaining the LLN for the BRW can be applied. (Unfortunately, a direct comparison with the BRW is not possible, so one has to repeat the second moment analysis. We omit further details since in Section 3.6 we will construct a better candidate for comparison, that will actually allow for comparison up to order 1.) We then get that for some universal constant C,

$$E\mathcal{X}_N^* \ge E(\max_{x \in V_N^{\delta,n(1-\log_2(1-\delta)}} \mathcal{X}_x^N) \ge \sqrt{\frac{2\log 2}{\pi}} E\mathcal{R}^*_{N(1-\delta)^{n+C}/C}\,.$$

This yields (3.4.2) after taking first $N \to \infty$ and then $\delta \to 0$. $\square$

3.5 A tightness argument: expectation is king

Our goal in this short section is to provide the following prelude to tightness, based on the Dekking–Host argument. It originally appeared in [BDZ11].

Lemma 9. *With $\mathcal{X}_N^{*'}$ an independent copy of $\mathcal{X}_N^*$, one has*

$$E|\mathcal{X}_N^{*'} - \mathcal{X}_N^*| \le 2(E\mathcal{X}_{2N}^* - E\mathcal{X}_N^*)\,. \tag{3.5.1}$$

Note that by Lemma 7, the right side of (3.5.1) is positive. The estimate (3.5.1) reduces the issue of tightness of $\{\mathcal{X}_N^* - E\mathcal{X}_N^*\}_N$ to a question concerning precise control of $E\mathcal{X}_N^*$, and more specifically, to obtaining a lower bound on $E\mathcal{X}_{2N}^*$ which differs only by a constant from the upper bound (3.4.1) on $E\mathcal{X}_N^*$.

Exercise 6. Prove that if A_n is a sequence of random variables for which there exists a constant C independent of n so that $E|A_n - A_n'| \le C$, where A_n' is an independent copy of A_n, then EA_n exists and the sequence $\{A_n - EA_n\}_n$ is tight.

In fact, Lemma 9 already yields a weak form of tightness.

Exercise 7. Combine Lemma 9 with the monotonicity statement (Lemma 7 and the LLN (Theorem 12) to deduce the existence of a deterministic sequence $N_k \to \infty$ so that $\{\mathcal{X}_{N_k^*} - E\mathcal{X}_{N_k}^*\}_k$ is tight.

(We eventually get rid of subsequences, but this requires extra estimates, as discussed in Lemma 10 below. The point of Exercise 7 is that tightness on subsequences is really a "soft" property.)

Proof of Lemma 9. By the Markov property of the GFF and arguing as in the proof of Lemma 7 (dividing the square V_{2N} into four disjoint squares of side N), we have

$$E\mathcal{X}^*_{2N} \geq E\max_{i=1}^{4} \mathcal{X}_N^{*,(i)} \geq E\max_{i=1}^{2} \mathcal{X}_N^{*,(i)} \,,$$

where $\mathcal{X}_N^{*,(i)}$, $i=1,\ldots,4$ are four independent copies of $\mathcal{X}^*_N$. Using again that $\max(a,b)=(a+b+|a-b|)/2$, we thus obtain

$$E\mathcal{X}^*_{2N} \geq E\mathcal{X}^*_N + E|\mathcal{X}_N^{*,(1)} - \mathcal{X}_N^{*,(2)}|/2 \,.$$

This yields the lemma. □

3.6 Expectation of the maximum: the modified BRW

We will now prove the following lemma, which is the main result of [BZ11].

Lemma 10. *With m_N as in* (3.1.2)*, one has*

$$E\mathcal{X}^*_N \geq m_N + O(1) \,. \tag{3.6.1}$$

Assuming Lemma 10, we have everything needed in order to prove Theorem 9.

Proof of Theorem 9. Combining Lemma 10 and (3.4.1), we have that $E\mathcal{X}^*_N = m_N + O(1)$. Since $\Theta_N = 2\mathcal{X}^*_N$, this yields (3.1.1). The tightness statement is now a consequence of Lemma 9, Exercise 6 and the fact that $m_{2N}-m_N$ is uniformly bounded. □

We turn to the main business of this section.

Proof of Lemma 10 (sketch). The main step is to construct a Gaussian field that interpolates between the BRW and the GFF, for which the second moment analysis that worked in the BRW case can still be carried out. Surprisingly, the new field is a very small variant of $\mathcal{R}^N_z$. We therefore refer to this field as the *modified branching random walk*, or in short MBRW.

We continue to consider $N=2^n$ for some positive integer n and again employ the notation $\mathcal{B}_k$ and $\mathcal{B}_k(x)$. For $x,y\in\mathbb{Z}^2$, write $x\sim_N y$ if $x-y\in(N\mathbb{Z})^2$. Similarly, for $B,B'\subset V_N$, write $B\sim_N B'$ if there exist integers i,j so that $B'=B+(iN,jN)$. Let $\mathcal{B}^N_k$ denote the collection of subsets of $\mathbb{Z}^2$ consisting of squares of side 2^k with lower left corner in V_N. Let $\{b_{k,B}\}_{k\geq 0, B\in\mathcal{B}^N_k}$ denote a family of independent centered Gaussian random variables where $b_{k,B}$ has variance 2^{-2k}, and define

$$b^N_{k,B} = \begin{cases} b_{k,B}\,, & B\in\mathcal{B}^N_k\,,\\ b_{k,B'}\,, & B\sim_N B'\in\mathcal{B}^N_k\,. \end{cases}$$

The MBRW $\{\mathcal{S}_z^N\}_{z\in V_N}$ is defined by

$$\mathcal{S}_z^N = \sum_{k=0}^{n} \sum_{B\in\mathcal{B}_k(z)} b_{k,B}^N .$$

We will also need a truncated form of the MBRW: for any integer $k_0 \geq 0$, set

$$\mathcal{S}_z^{N,k_0} = \sum_{k=k_0}^{n} \sum_{B\in\mathcal{B}_k(z)} b_{k,B}^N .$$

We again define $\mathcal{S}_N^* = \max_{z\in V_N} \mathcal{S}_z^N$ and $\mathcal{S}_{N,k_0}^* = \max_{z\in V_N} \mathcal{S}_z^{N,k_0}$. The correlation structure of $\mathcal{S}$ respects a torus structure on V_N. More precisely, with $d^N(x,y) = \min_{z:\ z\sim_N y} \|x-z\|$, one easily checks that for some constant C independent of N,

$$|\mathbf{R}_{\mathcal{S}^N}(x,y) - (n - \log_2 d^N(x,y))| \leq C . \tag{3.6.2}$$

In particular, for points $x, y \in (N/2, N/2) + V_N$, the covariance of $\mathcal{S}_{2N}$ is comparable to that of $\mathcal{X}_{2N}$. More important, the truncated MBRW has the following nice properties. Define, for $x, y \in V_N$, $\rho_{N,k_0}(x,y) = E((\mathcal{S}_x^{N,k_0} - \mathcal{S}_y^{N,k_0})^2)$. The following are basic properties of ρ_{N,k_0}; verification is routine and omitted.

Lemma 11. *The function ρ_{N,k_0} has the following properties.*

$$\rho_{N,k_0}(x,y) \text{ decreases in } k_0. \tag{3.6.3}$$

$$\limsup_{k_0\to\infty} \limsup_{N\to\infty} \sup_{x,y\in V_N: d_N(x,y)\leq 2^{\sqrt{k_0}}} \rho_{N,k_0}(x,y) = 0 . \tag{3.6.4}$$

$$\begin{gathered}\text{There is a function } g : \mathbb{Z}_+ \to \mathbb{R}_+ \text{ so that } g(k_0) \to_{k_0\to\infty} \infty \\ \text{and, for } x, y \in V_N \text{ with } d^N(x,y) \geq 2^{\sqrt{k_0}}, \\ \rho_{N,k_0}(x,y) \leq \rho_{N,0}(x,y) - g(k_0) , \quad n > k_0.\end{gathered} \tag{3.6.5}$$

Equipped with Lemma 11, and using the Sudakov-Fernique comparison (Theorem 11), we have the following.

Corollary 4. *There exists a constant k_0 such that, for all $N = 2^n$ large,*

$$E\mathcal{X}_N^* \geq \sqrt{\frac{2\log 2}{\pi}} E\mathcal{S}_{N/4,k_0}^* . \tag{3.6.6}$$

Therefore, the proof of Lemma 10 reduces to the derivation of a lower bound on the expectation of the maximum of the (truncated) MBRW. This is contained in the following proposition, whose proof we sketch below.

Proposition 2. *There exists a function $f : \mathbb{Z}_+ \to \mathbb{R}_+$ such that, for all $N \geq 2^{2k_0}$,*

$$E\mathcal{S}^*_{N,k_0} \geq (2\sqrt{\log 2})n - (3/(4\sqrt{\log 2}))\log n - f(k_0)\,. \qquad (3.6.7)$$

The proposition completes the proof of Lemma 10. □

Proof of Proposition 2 (sketch). Set $V'_N = V_{N/2} + (N/4, N/4) \subset V_N$ and define

$$\tilde{\mathcal{S}}^*_{N,k_0} = \max_{z \in V'_N} \mathcal{S}^{N,k_0}_z\,, \quad \tilde{\mathcal{S}}^*_N = \tilde{\mathcal{S}}^*_{N,0}\,.$$

Set

$$A_n = m_N\sqrt{\pi/2\log 2} = (2\sqrt{\log 2})n - (3/(4\sqrt{\log 2}))\log n\,.$$

An application of the second moment method (similar to what was done for the BRW, and therefore omitted) yields the following.

Proposition 3. *There exists a constant $\delta_0 \in (0,1)$ such that, for all N,*

$$P(\tilde{\mathcal{S}}^*_N \geq A_n) \geq \delta_0\,. \qquad (3.6.8)$$

We now explain how to deduce Proposition 2 from Proposition 3. Our plan is to show that the left tail of $\tilde{\mathcal{S}}^*_N$ is decreasing exponentially fast; together with the bound (3.6.8), this will imply (3.6.7) with $k_0 = 0$. At the end of the proof, we show how the bound for $k_0 > 0$ follows from the case $k_0 = 0$. In order to show the exponential decay, we compare $\tilde{\mathcal{S}}^*_N$, after appropriate truncation, to four independent copies of the maximum over smaller boxes, and then iterate.

For $i = 1,2,3,4$, introduce the four sets $W_{N,i} = [0, N/32)^2 + z_i$ where $z_1 = (N/4, N/4)$, $z_2 = (23N/32, N/4)$, $z_3 = (N/4, 23N/32)$ and $z_4 = (23N/32, 23N/32)$. (We have used here that $3/4 - 1/32 = 23/32$.) Note that $\cup_i W_{N,i} \subset V_N$, and that these sets are $N/4$-separated, that is, for $i \neq j$,

$$\min_{z \in W_{N,i}, z' \in W_{N,j}} d^N_\infty(x,y) > N/4\,.$$

Recall the definition of $\mathcal{S}^N_z$ and define, for $n > 6$,

$$\bar{\mathcal{S}}^N_z = \sum_{k=0}^{n-6} \sum_{B \in \mathcal{B}_k(z)} b^N_{k,B}\,;$$

note that

$$\mathcal{S}^N_z - \bar{\mathcal{S}}^N_z = \sum_{j=0}^{5} \sum_{B \in \mathcal{B}_{n-j}(z)} b^N_{n-j,B}\,.$$

Our first task is to bound the probability that $\max_{z \in V_N}(\mathcal{S}^N_z - \bar{\mathcal{S}}^N_z)$ is large. This will be achieved by applying Fernique's criterion in conjunction with Borell's inequality (Theorem 10). We introduce some notation.

Let $m(\cdot) = m_N(\cdot)$ denote the uniform probability measure on V_N (i.e., the counting measure normalized by $|V_N|$) and let $g : (0,1] \to \mathbb{R}_+$ be the function defined by

$$g(t) = (\log(1/t))^{1/2} .$$

Set $\mathbf{G}_z^N = \mathcal{S}_z^N - \bar{\mathcal{S}}_z^N$ and

$$B(z,\epsilon) = \{z' \in V_N : E((\mathbf{G}_z^N - \mathbf{G}_{z'}^N)^2) \le \epsilon^2\} .$$

Then, Fernique's criterion, see [Ad90, Theorem 4.1], implies that, for some universal constant $K \in (1,\infty)$,

$$E(\max_{z \in V_N} \mathbf{G}_z^N) \le K \sup_{z \in V_N} \int_0^\infty g(m(B(z,\epsilon)))d\epsilon . \tag{3.6.9}$$

For $n \ge 6$, we have, in the notation of Lemma 11,

$$E((\mathbf{G}_z^N - \mathbf{G}_{z'}^N)^2) = \rho_{N,n-5}(z,z') .$$

Therefore, there exists a constant C such that, for $\epsilon \ge 0$,

$$\{z' \in V_N : d_\infty^N(z,z') \le \epsilon^2 N/C\} \subset B(z,\epsilon) .$$

In particular, for $z \in V_N$ and $\epsilon > 0$,

$$m(B(z,\epsilon)) \ge ((\epsilon^4/C^2) \vee (1/N^2)) \wedge 1 .$$

Consequently,

$$\int_0^\infty g(m(B(z,\epsilon)))d\epsilon \le \int_0^{\sqrt{C/N}} \sqrt{\log(N^2)}d\epsilon + \int_{\sqrt{C/N}}^{\sqrt{C}} \sqrt{\log(C^2/\epsilon^4)}d\epsilon < C_4 ,$$

for some constant C_4. Applying Fernique's criterion (3.6.9), we deduce that

$$E(\max_{z \in V_N}(\mathcal{S}_z^N - \bar{\mathcal{S}}_z^N)) \le C_4 K .$$

The expectation $E((\mathcal{S}_z^N - \bar{\mathcal{S}}_z^N)^2)$ is bounded in N. Therefore, using Borell's inequality (Theorem 10), it follows that, for some constant C_5 and all $\beta > 0$,

$$P(\max_{z \in V_N}(\mathcal{S}_z^N - \bar{\mathcal{S}}_z^N) \ge C_4 K + \beta) \le 2e^{-C_5\beta^2} . \tag{3.6.10}$$

We also note the following bound, which is obtained similarly: there exist constants C_5, C_6 such that, for all $\beta > 0$,

$$P(\max_{z \in V'_{N/16}}(\bar{\mathcal{S}}_z^N - \mathcal{S}_z^{N/16}) \ge C_6 + \beta) \le 2e^{-C_7\beta^2} . \tag{3.6.11}$$

The advantage of working with $\bar{\mathcal{S}}^N$ instead of $\mathcal{S}^N$ is that the fields $\{\bar{\mathcal{S}}_z^N\}_{z\in W_{N,i}}$ are independent for $i=1,\ldots,4$. For every $\alpha,\beta>0$, we have the bound

$$\begin{aligned} &P(\tilde{\mathcal{S}}_N^* \geq A_n-\alpha) \qquad (3.6.12)\\ &\geq P(\max_{z\in V_N'} \bar{\mathcal{S}}_z^N \geq A_n + C_4 - \alpha+\beta) - P(\max_{z\in V_N'}(\mathcal{S}_z^N - \bar{\mathcal{S}}_z^N) \geq C_4+\beta)\\ &\geq P(\max_{z\in V_N'} \bar{\mathcal{S}}_N \geq A_n + C_4-\alpha+\beta) - 2e^{-C_5\beta^2}\,, \end{aligned}$$

where (3.6.10) was used in the last inequality. On the other hand, for any $\gamma,\gamma'>0$,

$$\begin{aligned} &P(\max_{z\in V_N'} \bar{\mathcal{S}}_z^N \geq A_n-\gamma) \geq P(\max_{i=1}^{4} \max_{z\in W_{N,i}} \bar{\mathcal{S}}_z^N \geq A_n-\gamma)\\ &= 1-(P(\max_{z\in W_{N,1}} \bar{\mathcal{S}}_z^N < A_n-\gamma))^4\\ &\geq 1-\left(P(\max_{z\in V'_{N/16}} \mathcal{S}_z^{N/16} < A_n-\gamma+C_6+\gamma') + 2e^{-C_7(\gamma')^2}\right)^4\,, \end{aligned}$$

where (3.6.11) was used in the inequality. Combining this estimate with (3.6.12), we get that, for any $\alpha,\beta,\gamma'>0$,

$$\begin{aligned} &P(\tilde{\mathcal{S}}_N^* \geq A_n-\alpha) \qquad (3.6.13)\\ &\geq 1-2e^{-C_5\beta^2}\\ &-\left(P(\max_{z\in V'_{N/16}} \mathcal{S}_z^{N/16} < A_n+C_4+C_6+\beta+\gamma'-\alpha) + 2e^{-C_7(\gamma')^2}\right)^4\,. \end{aligned}$$

We now iterate the last estimate. Let $\eta_0 = 1-\delta_0 < 1$ and, for $j\geq 1$, choose a constant $C_8 = C_8(\delta_0) > 0$ so that, for $\beta_j = \gamma_j' = C_8\sqrt{\log(1/\eta_j)}$,

$$\eta_{j+1} = 2e^{-C_5\beta_j^2} + (\eta_j + 2e^{-C_7(\gamma_j')^2})^4$$

satisfies $\eta_{j+1} < \eta_j(1-\delta_0)$. (It is not hard to verify that such a choice is possible.) With this choice of β_j and γ_j', set $\alpha_0=0$ and $\alpha_{j+1} = \alpha_j + C_4 + C_6 + \beta_j + \gamma_j'$, noting that $\alpha_j \leq C_9\sqrt{\log(1/\eta_j)}$ for some $C_9 = C_9(\delta_0)$. Substituting in (3.6.13) and using Proposition 2 to start the recursion, we get that

$$P(\tilde{\mathcal{S}}_N^* \geq A_n - \alpha_{j+1}) \geq 1-\eta_{j+1}\,. \qquad (3.6.14)$$

Therefore,

$$\begin{aligned} E\tilde{\mathcal{S}}_N^* &\geq A_n - \int_0^\infty P(\tilde{\mathcal{S}}_N^* \leq A_n - \theta) d\theta \\ &\geq A_n - \sum_{j=0}^\infty \alpha_j P(\tilde{\mathcal{S}}_N^* \leq A_n - \alpha_j) \\ &\geq A_n - C_9 \sum_{j=0}^\infty \eta_j \sqrt{\log(1/\eta_j)}\,. \end{aligned}$$

Since $\eta_j \leq (1-\delta_0)^j$, it follows that there exists a constant $C_{10} > 0$ so that

$$E\mathcal{S}_N^* \geq E\tilde{\mathcal{S}}_N^* \geq A_n - C_{10}\,. \tag{3.6.15}$$

This completes the proof of Proposition 2 in the case $k_0 = 0$.

To consider the case $k_0 > 0$, define

$$\hat{\mathcal{S}}_{N,k_0}^* = \max_{z \in V_N' \cap 2^{k_0}\mathbb{Z}^2} \mathcal{S}_z^{N,k_0}\,.$$

Then, $\hat{\mathcal{S}}_{N,k_0}^* \leq \tilde{\mathcal{S}}_{N,k_0}^*$. On the other hand, $\hat{\mathcal{S}}_{N,k_0}^*$ has, by construction, the same distribution as $\tilde{\mathcal{S}}_{2^{-k_0}N,0}^* = \tilde{\mathcal{S}}_{2^{-k_0}N}^*$. Therefore, for any $y \in \mathbb{R}$,

$$P(\tilde{\mathcal{S}}_{N,k_0}^* \geq y) \geq P(\hat{\mathcal{S}}_{N,k_0}^* \geq y) \geq P(\tilde{\mathcal{S}}_{2^{-k_0}N}^* \geq y)\,.$$

We conclude that

$$E\mathcal{S}_{N,k_0}^* \geq E\tilde{\mathcal{S}}_{N,k_0}^* \geq E\tilde{\mathcal{S}}_{2^{-k_0}N}^*\,.$$

Application of (3.6.15) completes the proof of Proposition 2. □

3.7 Additional remarks.

The ideas presented in this section can be taken further. Some recent developments are summarized here. This is not intended as an exhaustive review, rather as a sample.

Tail estimates J. Ding [Di11a] has improved on the proof of tightness by providing the following tail estimates.

Proposition 4. *The variance of $\mathcal{X}_N^*$ is uniformly bounded. Further, there exist universal constants c, C so that for any $x \in (0, (\log n)^{2/3})$, and with $\bar{\mathcal{X}}_N^* = \mathcal{X}_N^* - E\mathcal{X}_N^*$,*

$$ce^{-Cx} \leq P(\bar{\mathcal{X}}_N^* \geq x) \leq Ce^{-cx}\,, \quad ce^{-Ce^{Cx}} \leq P(\bar{\mathcal{X}}_N^* \leq -x) \leq Ce^{-ce^{cx}}\,,$$

It is interesting to compare these bounds with the case of BRW: while the bounds on the upper tail are similar, the lower tail exhibits quite different behavior, since in the case of BRW, just modifying a few variables near the root of the tree can have a significant effect on the maximum.

Tightness of maxima for other graphs As discussed above, if a sequence of graphs G_N is obtained from an infinite graph G by truncation to a bounded box and imposing Dirichlet boundary conditions, the maximum of the resulting GFF fluctuates within $O(1)$ if the underlying random walk on G is transient and possesses a uniformly bounded Green function. The recurrent case, and especially the null-recurrent case, can have either fluctuations of the GFF maximum of order $O(1)$ or increasing (with N) fluctuations, compare the 1D and 2D lattice GFF. It is natural to conjecture that if the Green function blows up polynomially (in N), then the fluctuations of the GFF maximum blow up. For certain (random) fractal graphs, this is verified in work in progress with T. Kumagai.

Cover times Gaussian fields on graphs possess close relations with the cover time of these graphs by random walk, through *isomorphism theorems*. We do not discuss these in details but instead refer to [DLP10] and [Di11b] for a thorough discussion. For cover times of the binary tree, precise estimates are contained in [DiZ11].

First passage percolation The maximum of BRW can also be seen as a first passage percolation on trees, and the Dekking–Host argument then yields control on tightness of fluctuations of first passage percolation on regular trees or, more generally, on Galton–Watson trees. This can be extended to a large class of graphs that in some vague sense possess "tree-like" recursive structure, including some tilings of the hyperbolic plane, see [BeZ10] for details. Recall that proving that the fluctuations of first passage percolation are bounded for the d-dimensional Euclidean lattice ($d \geq 3$, or even d large) is an open problem, see [NP96, BKS03, PP94] for discussion.

References

[ABR08] L. Addario-Berry and B. Reed, Ballot theorems for random walks with finite variance. arXiv:0802.2491v2. (Updated version: http://www.math.mcgill.ca/louigi/papers/ballot.pdf)

[ABR09] L. Addario-Berry and B. Reed, Minima in branching random walks, *Annals Probab.* **37** (2009), pp. 1044–1079.

[Ad90] R. Adler, *An Introduction to Continuity, Extrema and Related Topics for General Gaussian Processes*, Lecture Notes - Monograph Series, Institute Mathematical Statistics, Hayward, CA (1990). Errata, http://webee.technion.ac.il/people/adler/borell.pdf.

[Ai11] E. Aïdékon, Convergence in law of the minimum of a branching random walk. arXiv:1101.1810v3. To appear, *Annals Probab.*

[ABBS11] E. Aïdékon, The branching Brownian motion seen from its tip, arXiv:1104.3738v2 (2011).

[ABK11] L.-P. Arguin, A. Bovier and N. Kistler, The extremal process of branching Brownian motion, arXiv:1103.2322v1 (2011).

[BKS03] I. Benjamini, G. Kalai and O. Schramm, First passage percolation has sublinear distance variance, *Annals Probab.* **31** (2003), pp. 1970–1978.

[BeZ10] I. Benjamini and O. Zeitouni, Tightness of Fluctuations of First Passage Percolation on Some Large Graphs, arXiv:1010.1412v2 (2010). To appear, *GAFA seminar notes, LNM, Springer*.

[BBS10] J. Berestycki, N. Berestycki and J. Schweinsberg, The genealogy of branching Brownian motion with absorption, arXiv:1001.2337v2 (2010).

[BDG01] E. Bolthausen, J.-D. Deuschel and G. Giacomin, Entropic repulsion and the maximum of the two-dimensional harmonic crystal, *Annals Probab.* **29** (2001), pp. 1670–1692.

[BDZ11] E. Bolthausen, J.-D. Deuschel and O. Zeitouni, Recursions and tightness for the maximum of the discrete, two dimensional Gaussian free field, *Elec. Comm. Prob.* **16** (2011), pp. 114–119.

[Bo75] C. Borell, The Brunn–Minkowski inequality in Gauss space, *Invent. Math.* **30** (1975), pp. 207–216.

[BS96] A. N. Borodin and O. Salminen, *Handbook of Brownian motion - facts and formulae*, Birkhauser, Basel (1996).

[Br78] M. Bramson, Maximal displacement of branching Brownian motion, *Comm. Pure Appl. Math.* **31** (1978), pp. 531–581.

[BZ09] M. Bramson and O. Zeitouni, Tightness for a family of recursion equations, *Annals Probab.* **37** (2009), pp. 615–653.

[BZ11] M. Bramson and O. Zeitouni, Tightness of the recentered maximum of the two–dimensional discrete Gaussian Free Field, *Comm. Pure App. Math* **65** (2011), pp. 1–20.

[BDMM07] E. Brunet, B. Derrida, A. H. Mueller, and S. Munier, Noisy traveling waves: effect of selection on genealogies, *Europhys. Lett.* **76** (2007), pp. 1–7.

[DH91] F. M. Dekking and B. Host, Limit distributions for minimal displacement of branching random walks, *Probab. Theory Relat. Fields* **90** (1991), pp. 403–426.

[DZ98] A. Dembo and O. Zeitouni, *Large deviation techniques and applications*, 2nd edition, Springer, New-York (1998).

[DLP10] J. Ding, J. R. Lee and Y. Peres, Cover times, blanket times, and majorizing measures. arXiv:1004.4371v5 (2010). To appear, *Annals Math.*

[Di11a] J. Ding, Exponential and double exponential tails for maximum of two-dimensional discrete Gaussian free field, arXiv:1105.5833v2 (2011).

[Di11b] J. Ding, Asymptotics of cover times via Gaussian free fields: bounded-degree graphs and general trees, arXiv:1103.4402v2 (2011).

[DiZ11] J. Ding and O. Zeitouni, A sharp estimate for cover times on binary trees. arXiv:1104.0434v1 (2011).

[Fa10] M. Fang, Tightness for maxima of generalized branching random walks (2010). arXiv:1012.0826v1 (2010).

[FZ10] M. Fang and O. Zeitouni, Consistent minimal displacement of branching random walks, *ECP* **15** (2010), pp. 106–118.

[FZ11] M. Fang and O. Zeitouni, Branching random walks in time inhomogeneous environments, arXiv:1112.1113v1 (2011).

[FZ12] M. Fang and O. Zeitouni, Slowdown for time inhomogeneous branching Brownian motion, arXiv:1205.1769v1 (2012).

[FHS10] G. Faraud, Y. Hu and Z. Shi, Almost sure convergence for stochastically biased random walks on trees. arXiv:1003.5505v5 (2010).

[Fe75] X. Fernique, Regularité des trajectoires des fonctions aléatoires gaussiennes, Ecole d'Eté de Probabilités de Saint-Flour, IV-1974, pp. 1–96. *Lecture Notes in Math.* **480**, Springer, Berlin, (1975).

[HS09] Y. Hu and Z. Shi, Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Annals Probab.* **37** (2009), pp. 742–789.

[NP96] C. Newmann and M. Piza, Divergence of shape fluctuations in two dimensions, *Annals Probab.* **23** (1995), 977–1005.

[PP94] R. Pemantle and Y. Peres, Planar first-passage percolation times are not tight, *Probability and phase transitions*, Cambridge University Press (1994), pp. 261–264.

[Ro11] M. Roberts, A simple path to asymptotics for the frontier of a branching Brownian motion, to appear, *Annals Probab.* (2012). arXiv:1106.4771v4.

[TIS76] B. S. Tsirelson, I. A. Ibragimov and V. N. Sudakov, Norms of Gaussian sample functions, *Lecture Notes Math.* **550**, Springer (1976).