Electronic Journal of Probability

Electron. J. Probab. **18** (2013), no. 53, 1–25.
ISSN: 1083-6489 DOI: 10.1214/EJP.v18-2082

An ergodic theorem for the frontier of branching Brownian motion*

Louis-Pierre Arguin† Anton Bovier‡ Nicola Kistler§

Abstract

We prove a conjecture of Lalley and Sellke [*Ann. Probab.* **15** (1987)] asserting that the empirical (time-averaged) distribution function of the maximum of branching Brownian motion converges almost surely to a double exponential, or Gumbel, distribution with a random shift. The method of proof is based on the decorrelation of the maximal displacements for appropriate time scales. A crucial input is the localization of the paths of particles close to the maximum that was previously established by the authors [*Comm. Pure Appl. Math.* **64** (2011)].

Keywords: Branching Brownian motion; ergodicity; extreme value theory; KPP equation and traveling waves.
AMS MSC 2010: 60J80; 60G70; 82B44.
Submitted to EJP on June 12, 2012, final version accepted on April 29, 2013.

1 Introduction

Branching Brownian Motion (BBM) on $\mathbb{R}$ is a continuous-time Markov branching process which plays an important role in the theory of partial differential equations [6, 7, 29], in particle physics [30], in the theory of disordered systems [10, 18], and in mathematical biology [21, 24]. It is constructed as follows on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Consider a standard Brownian motion $x(t)$, starting at 0 at time 0. We consider $x(t)$ to be the position of a *particle* at time t. After an exponential random time T of mean one and independent of x, the particle splits into k particles with probability p_k, where $\sum_{k=1}^\infty p_k = 1$, $\sum_{k=1}^\infty kp_k = 2$, and $\sum_k k(k-1)p_k < \infty$. (The choice of mean 2 is arbitrary and is fixed to lighten notation.) The positions of the k particles are independent

*Support: Discovery Grant of the Natural Sciences; Engineering Research Council of Canada; Établissement de Nouveaux Chercheurs - Fonds de recherche du Québec - Nature et technologies; German Research Foundation SFB 1060 ; Centre International de Rencontres Mathématiques (Jean Morlet Chair 2013).
†Université de Montréal, Canada;
E-mail: arguinlp@dms.umontreal.ca
‡Institut für Angewandte Mathematik, Rheinische Friedrich-Wilhelms-Universität Bonn, Germany;
E-mail: bovier@uni-bonn.de
§CIRM & Université Aix Marseille;
E-mail: kistler@cirm.univ-mrs.fr

Brownian motions starting at $x(T)$. Each of these processes have the same law as the first Brownian particle. Thus, after a time $t > 0$, there will be $n(t)$ particles located at $x_1(t), \ldots, x_{n(t)}(t)$, with $n(t)$ being the random number of offspring generated up to that time (note that $\mathbb{E}n(t) = e^t$).

An interesting link between BBM and partial differential equations was observed by McKean [29]. If one denotes by

$$u(t, x) \equiv \mathbb{P}\left[\max_{1 \le k \le n(t)} x_k(t) \le x\right] \tag{1.1}$$

the law of the maximal displacement, a renewal argument shows that $u(t, x)$ solves the Kolmogorov-Petrovsky-Piscounov equation [KPP], also referred to as the Fisher-KPP equation,

$$\begin{aligned} u_t &= \frac{1}{2}u_{xx} + \sum_{k=1}^{\infty} p_k u^k - u, \\ u(0, x) &= \begin{cases} 1, \text{ if } x \ge 0, \\ 0, \text{ if } x < 0. \end{cases} \end{aligned} \tag{1.2}$$

This equation has raised a lot of interest, in part because it admits traveling wave solutions: there exists a unique solution satisfying

$$u\big(t, m(t) + x\big) \to \omega(x), \quad \text{uniformly in } x, \text{ as } t \uparrow \infty, \tag{1.3}$$

where the centering term, the *front* of the wave, is given by

$$m(t) = \sqrt{2}t - \frac{3}{2\sqrt{2}} \ln t, \tag{1.4}$$

and $\omega(x)$ is the unique solution (up to translation) of the o.d.e.

$$\frac{1}{2}\omega_{xx} + \sqrt{2}\omega_x + \sum_{k=1}^{\infty} p_k \omega^k - \omega = 0. \tag{1.5}$$

The leading order of the front has been established by Kolmogorov, Petrovsky, and Piscounov [25]. The logarithmic corrections have been obtained by Bramson [11], using the probabilistic representation given above. (See also the recent contribution by Roberts [26] for a different derivation through spine techniques).

Equations (1.1) and (1.3) show the weak convergence of the distribution of the recentered maximum of BBM.

Let

$$M(t) \equiv \max_{k \le n(t)} x_k(t) - m(t) \ , \tag{1.6}$$

$$y_k(t) \equiv \sqrt{2}t - x_k(t), \quad z_k(t) \equiv y_k(t) e^{-y_k(t)}, \tag{1.7}$$

and finally

$$Y(t) \equiv \sum_{k \le n(t)} e^{-\sqrt{2} y_k(t)} \qquad \text{and} \qquad Z(t) \equiv \sum_{k \le n(t)} z_k(t) \ . \tag{1.8}$$

In 1987, Lalley and Sellke [27] proved that

$$\lim_{t \uparrow \infty} Y(t) = 0 \text{ a.s.} \qquad \text{and} \qquad \lim_{t \uparrow \infty} Z(t) \quad = Z \text{ a.s.}, \tag{1.9}$$

where Z is a strictly positive, almost surely finite random variable (with infinite mean).

This paper is concerned with the large time limit of the empirical (time-averaged) distribution of the maximal displacement

$$F_T(x) \equiv \frac{1}{T}\int_0^T \mathbb{1}_{\{M(s)\leq x\}}\, ds,\ x \in \mathbb{R}\,. \tag{1.10}$$

The main result is that F_T converges almost surely as $T \uparrow \infty$ to a random distribution function. The limit is the double exponential (Gumbel) distribution that is shifted by the random variable $\frac{1}{\sqrt{2}}\ln Z$:

Theorem 1 (Ergodic Theorem)**.** *For any $x \in \mathbb{R}$,*

$$\lim_{T\uparrow\infty} F_T(x) = \exp\left(-CZe^{-\sqrt{2}x}\right), \quad \textit{a.s.,} \tag{1.11}$$

where $C > 0$ is a positive constant.

The derivative martingale Z encodes the dependence on the early evolution of the system. The mechanism for this is subtle, and we shall provide first some intuition in the next section.

Theorem 1 was conjectured by Lalley and Sellke in [27]. They showed that, despite the weak convergence (1.3), the empirical distribution $F_T(x)$ *cannot* converge to $\omega(x)$ in the limit of large times (for any $x \in \mathbb{R}$), and proved that the latter is recovered when Z is integrated, i.e.

$$\omega(x) = \mathbb{E}\left[\exp\left(-CZ\mathrm{e}^{-\sqrt{2}x}\right)\right]. \tag{1.12}$$

(A similar representation for the law of the branching random *walk* has been recently obtained by Aïdékon [1]). The issue of ergodicity of BBM has also been discussed by Brunet and Derrida in [15]. Ergodic results similar to Theorem 1 can be proved for statistics of extremal particles of BBM other than the distribution of the maximum. This will be detailed in a separate work.

A description of the law of the statistics of extremal particles has been obtained in a series of papers by the present authors [3, 4, 5] and in the work of Aïdékon, Berestycki, Brunet, and Shi [2]: it is now known that the joint distribution of extremal particles re-centered by $m(t)$ converges weakly to a randomly shifted Poisson cluster process; the positions of the clusters is a random shift of a Poisson point process with exponential density, and the law of the individual clusters is also known, but has a different description in each work. We refer the reader to the aforementioned papers for details.

We point out that the interest in the properties of BBM stems also from its alleged universality: it is conjectured, and in some instances also proved, that different models of probability and of statistical mechanics share many structural features with the extreme values of BBM. A partial list includes the two-dimensional Gaussian free field [8, 9, 13], the cover times of graphs by random walks [19, 20], and in general, log-correlated Gaussian fields, see e.g. [17, 22].

2 Outline of the proof

Consider a compact interval $\mathcal{D} = [d, D]$ with $-\infty < d < D < \infty$. It is clear that almost sure convergence of the empirical distribution on these sets implies almost sure convergence of the distribution function $F_T(x)$. As a first step in the proof of Theorem 1, we introduce a "cutoff" $\varepsilon > 0$ and split the integration over the sets $[0, T\varepsilon]$ and $(T\varepsilon, T]$:

$$F_T(D) - F_T(d) = \frac{1}{T}\int_{\varepsilon T}^T \mathbb{1}_{\{M(s)\in\mathcal{D}\}}\, ds + \frac{1}{T}\int_0^{\varepsilon T} \mathbb{1}_{\{M(s)\in\mathcal{D}\}}\, ds. \tag{2.1}$$

The second term on the r.h.s. above does not contribute in the limit when $T \uparrow \infty$ first, and $\varepsilon \downarrow 0$ next. It thus suffices to compute the double limit for the first term.

To this aim, we introduce the time $R_T > 0$, which will play the role of the *early evolution*. For the moment we only require that $R_T \uparrow \infty$, but $R_T/\sqrt{T} \downarrow 0$, as $T \uparrow \infty$. A particular choice will be made later. We rewrite the empirical distribution as

$$\begin{aligned}\frac{1}{T}\int_{\varepsilon T}^{T} \mathbb{1}_{\{M(s)\in\mathcal{D}\}} ds &= \frac{1}{T}\int_{\varepsilon T}^{T} \mathbb{P}\left[M(s)\in\mathcal{D} \mid \mathcal{F}_{R_T}\right] ds \\ &+ \frac{1}{T}\int_{\varepsilon T}^{T} \Big(\mathbb{1}_{\{M(s)\in\mathcal{D}\}} - \mathbb{P}\left[M(s)\in\mathcal{D} \mid \mathcal{F}_{R_T}\right]\Big) ds\ .\end{aligned} \tag{2.2}$$

We now state two theorems which immediately imply Theorem 1: Theorem 2 below addresses the first term on the r.h.s of (2.2), while Theorem 3 addresses the second term.

Theorem 2. *Let $R_T \uparrow \infty$ as $T \uparrow \infty$ but with $R_T = o(\sqrt{T})$. Then for any $s \in [\varepsilon, 1]$,*

$$\lim_{T\uparrow\infty} \mathbb{P}\left[M(T\cdot s)\in\mathcal{D} \mid \mathcal{F}_{R_T}\right] = \int_{\mathcal{D}} d\left(\exp\left(-CZe^{-\sqrt{2}x}\right)\right), \text{a.s.}. \tag{2.3}$$

The above statement is an improvement of [27, Theorem 1], where the probability was conditioned on a *fixed* time that only subsequently was let to infinity. The proof closely follows this case and relies on precise estimates of the law of the maximal displacement obtained by Bramson [12].

Theorem 2 together with a change of variables and using dominated convergence imply

$$\lim_{\varepsilon\downarrow 0}\lim_{T\uparrow\infty} \frac{1}{T}\int_{\varepsilon T}^{T} \mathbb{P}\left[M(s)\in\mathcal{D} \mid \mathcal{F}_{R_T}\right] ds = \int_{\mathcal{D}} d\left(\exp -CZe^{-\sqrt{2}x}\right) \text{ a.s.}, \tag{2.4}$$

which is the r.h.s. of (1.11).

The integrand of the second term on the r.h.s of (2.2) has mean zero. Therefore, to prove Theorem 1 we need only the following strong law of large numbers.

Theorem 3. *For $\varepsilon > 0$, $\mathcal{D}$ as above, and R_T as in Theorem 2,*

$$\lim_{T\uparrow\infty} \frac{1}{T}\int_{\varepsilon T}^{T} \left(\mathbb{1}_{\{M(s)\in\mathcal{D}\}} - \mathbb{P}\left[M(s)\in\mathcal{D} \mid \mathcal{F}_{R_T}\right]\right) ds = 0, \text{a.s.} \tag{2.5}$$

The short proof of Theorem 2 is given in Section 3. The proof of Theorem 3 turns out to be quite delicate. Due to the possibly strong correlations among the Brownian particles, it is perhaps surprising that a law of large numbers holds at all. Let T be large and consider two times $s, s' \in [0, T]$. It is clear that if the distance between s and s' is of order one, say, then the extremal particles at s are strongly correlated with the ones at s', since the children of extremal particles are very likely to remain extremal for some time. Therefore, s and s' need to be *well separated* for the correlations to be weak. On the other hand, and this is the crucial point, it is generally not true that the correlations between the extremal particles at time s and s' decay as the distance between s and s' increases. As shown by Lalley and Sellke [27, Theorem 2 and corollary], *"every particle born in a branching Brownian motion has a descendant particle in the lead at some future time"*. Hence, if s and s' are too far from each other (for example, if s is of order one with respect to T and s' is of order T), correlations build up again and mixing fails. Therefore, weak correlations between the frontiers at two different times only set in at precise time scales. It turns out that if s and s' are both of order T, $s, s' \in [\varepsilon T, T]$ and well separated, i.e. $|s - s'| > T^{\xi}$ for some $0 < \xi < 1$, then the correlations between

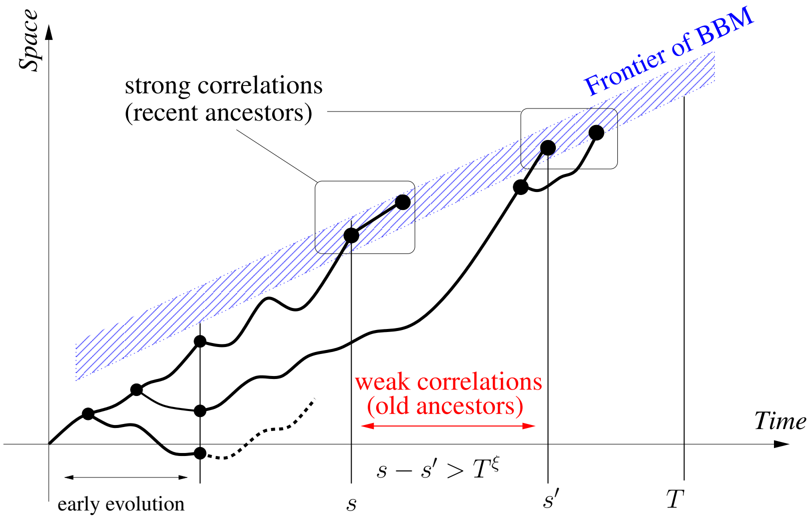

Figure 1: Leaders and their ancestors.

the frontiers are weak enough to provide a law of large numbers. By weak enough, we understand a summability condition on the correlations that lead to a SLLN by a theorem of Lyons, see Theorem 8 below. See Figure 1 for a graphical representation. A precise control on the correlations is achieved by controlling the paths of extremal particles in the spirit of [3] (see Section 4 below for precise statements).

3 Almost sure convergence of the conditional maximum

We start with some elementary facts that will be of importance. First, observe that for $t, s > 0$ such that $s = o(t)$ for $t \uparrow \infty$, the level of the maximum (1.4) satisfies

$$\begin{aligned} m(t) &= m(t-s) + \sqrt{2}s + \frac{3}{2\sqrt{2}} \ln\left(\frac{t-s}{t}\right) \\ &= m(t-s) + \sqrt{2}s + o(1). \end{aligned} \tag{3.1}$$

Here and henceforth, we write $\overline{x_k(t)}$ for the position of particle k at time t *shifted* by the level of the maximum, i.e. $\overline{x_k(t)} \equiv x_k(t) - m(t)$.

Second, let $\{x_j(s), j \le n(s)\}$ and, for $j = 1 \dots n(s)$, $\{x_k^{(j)}(t-s), k \le n^{(j)}(t-s)\}$ be all independent, identically distributed BBMs. The Markov property of BBM implies

$$\{x_k(t), k \le n(t)\} \stackrel{law}{=} \{x_j(s) + x_k^{(j)}(t-s), j \le n(s),\, k \le n^{(j)}(t-s)\}, \tag{3.2}$$

In particular, if $\mathcal{F}_s$ denotes the σ-algebra generated by the process up to time s, the combination of (3.1) and (3.2) yields for $X \in \mathbb{R}$

$$\mathbb{P}\left[\forall_{k \le n(t)} : \overline{x_k(t)} \le X \mid \mathcal{F}_s\right] = \prod_{k \le n(s)} \mathbb{P}\left[\forall_{j \le n(t-s)} : \overline{x_j(t-s)} \le X + x_k(s) + o(1) \mid \mathcal{F}_s\right] . \tag{3.3}$$

We will typically deal with situations where only a subset of $\{k : k = 1, \dots, n(t)\}$ appears. In all such cases, the generalization of (3.3) is straightforward.

A key ingredient to the proof of Theorem 2 is a precise estimate on the right-tail of the distribution of the maximal displacement. It is related to [5, Proposition 3.3], which heavily relies on the work by Bramson [12].

Lemma 4. *Consider* $t \geq 0$ *and* $X(t) \geq 0$ *such that* $\lim_{t\uparrow\infty} X(t) = +\infty$ *and* $X(t) = o(\sqrt{t})$ *in the considered limit. Then, for* $X(t)$ *and* t *both greater than* $8r$*,*

$$C\gamma(r)^{-1}X(t)e^{-\sqrt{2}X(t)}\left(1-\frac{X(t)}{t-r}\right) \leq \mathbb{P}\left[M(t) \geq X(t)\right] \leq C\gamma(r)X(t)e^{-\sqrt{2}X(t)} \tag{3.4}$$

for some $\gamma(r) \downarrow 1$ *as* $r \uparrow \infty$ *and* C *as in* (1.12)*.*

Proof. Let us denote by $\overline{u}(t,x) \equiv 1-u(t,x)$, with u the distribution of the maximal displacement defined in (1.1). We define

$$\begin{aligned}\psi(r,t,x+\sqrt{2}t) \equiv \frac{e^{-\sqrt{2}x}}{\sqrt{t-r}}\int_0^\infty \frac{dy'}{\sqrt{2\pi}}\cdot \overline{u}(r,y'+\sqrt{2}r)\cdot e^{y'\sqrt{2}} \\ \times\left\{1-\exp\left(-2y'\frac{x+\frac{3}{2\sqrt{2}}\ln t}{t-r}\right)\right\}\exp\left(-\frac{(y'-x)^2}{2(t-r)}\right).\end{aligned} \tag{3.5}$$

According to [5, Proposition 3.3], for $t \geq 8r$, and $x \geq 8r - \frac{3}{2\sqrt{2}}\ln(t)$, the following bounds hold:

$$\gamma(r)^{-1}\psi(r,t,x+\sqrt{2}t) \leq \overline{u}(t,x+\sqrt{2}t) \leq \gamma(r)\psi(r,t,x+\sqrt{2}t) \tag{3.6}$$

for some $\gamma(r) \downarrow 1$ as $r \uparrow \infty$.

As $\sqrt{2}t = m(t) + \frac{3}{2\sqrt{2}}\ln(t)$, by putting $\overline{x} \equiv x + \frac{3}{2\sqrt{2}}\ln(t)$, we reformulate the above as

$$\gamma(r)^{-1}\psi(r,t,\overline{x}+m(t)) \leq \overline{u}(t,\overline{x}+m(t)) \leq \gamma(r)\psi(r,t,\overline{x}+m(t)). \tag{3.7}$$

(The bounds in (3.7) hold for $\overline{x} \geq 8r$).

Setting

$$G(t,r;\overline{x},y') \equiv \overline{u}(r,y'+\sqrt{2}r)\cdot e^{y'\sqrt{2}}\cdot\exp\left(-\frac{(y'-\overline{x}+\frac{3}{2\sqrt{2}}\ln t)^2}{2(t-r)}\right), \tag{3.8}$$

we can rewrite (3.7) as

$$\begin{aligned}\psi(r,t,\overline{x}+m(t)) &= \frac{t^{3/2}e^{-\overline{x}\sqrt{2}}}{\sqrt{t-r}}\int_0^\infty \frac{dy'}{\sqrt{2\pi}}\cdot\left\{1-e^{-2y'\frac{\overline{x}}{t-r}}\right\}\cdot G(t,r;\overline{x},y') \\ &= t(1+o(1))e^{-\overline{x}\sqrt{2}}\int_0^\infty \frac{dy'}{\sqrt{2\pi}}\cdot\left\{1-e^{-2y'\frac{\overline{x}}{t-r}}\right\}\cdot G(t,r;\overline{x},y').\end{aligned} \tag{3.9}$$

By a dominated convergence argument [12, Prop. 8.3 and its proof], one shows that

$$C(r) \equiv \lim_{t\uparrow\infty}\int_0^\infty 2y'G(t,r;\overline{x},y')\frac{dy'}{\sqrt{2\pi}}, \tag{3.10}$$

exists, *uniformly* for $\overline{x}$ in compacts. In fact, Bramson's argument easily extends to the case where $\overline{x} = o(\sqrt{t})$ (to see this, one simply expands the quadratic term in the Gaussian density appearing in the definition of the function G). Moreover, $C(r) \to C$ as $r \uparrow \infty$, with C as in (1.12), see [12, p. 145-146]. An elementary estimate on the exponential function yields

$$2y'\frac{\overline{x}}{t-r} - \frac{2(y')^2\overline{x}^2}{(t-r)^2} + \frac{f(t,r;x,y')}{(t-r)^3} \leq 1-e^{-2y'\frac{\overline{x}}{t-r}} \leq 2y'\frac{\overline{x}}{t-r}\,, \tag{3.11}$$

for some function $f(t,r;x,y')$ which is integrable with respect to $G(t,r;\overline{x},y')dy'$. Inserting (3.11) into (3.9), we get the bounds

$$\begin{aligned} &\overline{x}e^{-\overline{x}\sqrt{2}}\int_0^\infty 2y'G(t,r;\overline{x},y')\frac{dy'}{\sqrt{2\pi}} \\ &\geq \overline{u}(t,\overline{x}+m(t)) \\ &\geq \overline{x}e^{-\overline{x}\sqrt{2}}\left(1+\frac{\overline{x}}{t-r}\right)\int_0^\infty 2y'G(t,r;\overline{x},y')\frac{dy'}{\sqrt{2\pi}}+O((t-r)^{-2}), \end{aligned} \tag{3.12}$$

for large enough t. The assertion of the Lemma follows by taking $\overline{x}\equiv X(t)$ in (3.12) and using(3.10). □

Proof of Theorem 2. The proof of Theorem 2 is a straightforward application of Lemma 4 and the convergence of the derivative martingale. First we write

$$\mathbb{P}\left[M(T\cdot s)\in\mathcal{D}\mid\mathcal{F}_{R_T}\right]=\mathbb{P}\left[M(T\cdot s)\leq D\mid\mathcal{F}_{R_T}\right]-\mathbb{P}\left[M(T\cdot s)\leq d\mid\mathcal{F}_{R_T}\right]. \tag{3.13}$$

We show only the almost sure convergence of the first term, the treatment of the second is identical. Since s is in $(\varepsilon,1)$, we have $R_T=o(T\cdot s)$ for $T\uparrow\infty$. Therefore, by (3.1) and (3.2),

$$\begin{aligned} &\mathbb{P}\left[M(T\cdot s)\leq D\mid\mathcal{F}_{R_T}\right]= \\ &\quad=\prod_{k\leq n(R_T)}\mathbb{P}\left[M(Ts-R_T)\leq D+y_k(R_T)\mid\mathcal{F}_{R_T}\right] \\ &\quad=\prod_{k\leq n(R_T)}\left\{1-\mathbb{P}\left[M(Ts-R_T)>D+y_k(R_T)\mid\mathcal{F}_{R_T}\right]\right\} \\ &\quad=\exp\left(\sum_{k\leq n(R_T)}\ln\left(1-\mathbb{P}\left[M(Ts-R_T)>D+y_k(R_T)\mid\mathcal{F}_{R_T}\right]\right)\right). \end{aligned} \tag{3.14}$$

By (1.9), $\lim_{R_T\uparrow\infty}\min_{k\leq n(R_T)}y_k(R_T)=+\infty$ a.s. Therefore, we may use Lemma 4 to establish upper and lower bounds for the probability of the maximum being larger than $D+y_k(R_T)$, namely

$$\begin{aligned} &C\gamma(r)^{-1}\big\{D+y_k(R_T)\big\}\exp\Big\{-\sqrt{2}(D+y_k(R_T)\Big\}\left(1-\frac{(D+y_k(R_T))}{Ts-R_T-r_T}\right)\leq \\ &\qquad\leq\mathbb{P}\left[M(Ts-R_T)\geq D+y_k(R_T)\mid\mathcal{F}_{R_T}\right]\leq \\ &\qquad\qquad\leq C\gamma(r)\big\{D+y_k(R_T)\big\}\exp\Big\{-\sqrt{2}(D+y_k(R_T)\Big\}, \end{aligned} \tag{3.15}$$

for $Ts-R_T\geq 8r>0$. Now write (3.15) as

$$\begin{aligned} C\gamma(r)^{-1}e^{-\sqrt{2}D}z_k(R_T)+\omega_k(R_T) &\leq \mathbb{P}\left[M(Ts-R_T)\geq D+y_k(R_T)\mid\mathcal{F}_{R_T}\right] \\ &\leq C\gamma(r)e^{-\sqrt{2}D}z_k(R_T)+\Omega_k(R_T)\,, \end{aligned} \tag{3.16}$$

where

$$\begin{aligned} \omega_k(R_T) \;\equiv\; & C\,D\,\gamma(r)^{-1}e^{-\sqrt{2}D}e^{-\sqrt{2}y_k(R_T)}\left(1-\frac{D+y_k(R_T)}{Ts-R_T-r_T}\right) \\ & -C\,\gamma(r)^{-1}e^{-\sqrt{2}D}z_k(R_T)\frac{D+y_k(R_T)}{Ts-R_T-r_T}, \end{aligned} \tag{3.17}$$

and

$$\Omega_k(R_T)\equiv C\,D\,\gamma(r)e^{-\sqrt{2}D}e^{-\sqrt{2}y_k(R_T)}\;. \tag{3.18}$$

Using that $-a \le \ln(1-a) \le -a + a^2/2$ (valid for $0 < a < 1/2$) together with the bounds (3.16), we obtain

$$\exp\left(-C\gamma(r)e^{-\sqrt{2}D}Z(R_T) - \sum_{k\le n(R_T)} \Omega_k(R_T)\right) \le \mathbb{P}\left[M(T\cdot s) \le D \mid \mathcal{F}_{R_T}\right]$$
$$\le \exp\Bigg(-C\gamma(r)^{-1}e^{-\sqrt{2}D}Z(R_T) - \sum_{k\le n(R_T)} \omega_k(R_T)$$
$$+\frac{1}{2}\sum_{k\le n(R_T)} \left\{C\gamma(r)e^{-\sqrt{2}D}z_k(R_T) + \Omega_k(R_T)\right\}^2\Bigg). \tag{3.19}$$

Next we show that the only contribution in the limit of large times in the above upper and lower bounds comes from the Z-terms. Regarding the terms involving $\Omega_k(R_T)$ in the lower bound, we note that

$$0 \le \sum_{k\le n(R_T)} \Omega_k(R_T) = CD\gamma(r)e^{-\sqrt{2}D}Y(R_T), \tag{3.20}$$

which is indeed vanishing by (1.9) in the limit $T \uparrow \infty$ To control the term involving $\omega_k(R_T)$ in the upper bound, we first observe that

$$\begin{aligned}\Big|\sum_{k\le n(R_T)} \omega_k(R_T)\Big| &\le \sum_{k\le n(R_T)} \Big|\omega_k(R_T)\Big| \\ &\le CD\gamma(r)^{-1}e^{-\sqrt{2}D}\left\{Y(R_T)+Z(R_T)\right\}\frac{\sup_k(D+y_k(R_T))}{T\varepsilon - R_T - r_T}.\end{aligned} \tag{3.21}$$

But this term vanishes in the large time limit, since $Y(R_T) \to 0$ and $Z(R_T) \to Z$, a.s., as $T \uparrow \infty$, again by (1.9). Moreover, one easily sees that one can choose $\kappa < \infty$ such that $\sup_k |y_k(R_T)| \le \kappa \log(R_T)$, a.s., and therefore

$$\frac{\sup_k(D+y_k(R_T))}{T\varepsilon - R_T - r_T} \to 0. \tag{3.22}$$

Thus, the $\omega_k(R_T)$ term in the upper bound vanishes in the limit $T \uparrow \infty$.

It remains to control the third term in the exponential on the r.h.s. of (3.19). Using that $(a+b)^2 \le 2a^2 + 2b^2$, one gets

$$\begin{aligned}&\frac{1}{2}\sum_k \left\{C\gamma(r)e^{-\sqrt{2}D}z_k(R_T) + \Omega_k(R_T)\right\}^2 \\ &\quad\le \left(C\gamma(r)e^{-\sqrt{2}D}\right)^2\left(\sum_k z_k(R_T)^2 + \sum_k e^{-2\sqrt{2}y_k(R_T)}\right).\end{aligned} \tag{3.23}$$

Clearly,

$$\sum_k z_k(R_T)^2 \le \sup_k \left(y_k(R_T)^2 e^{-\sqrt{2}y_k(R_T)}\right) Y(R_T), \tag{3.24}$$

which vanishes by (1.9). A similar reasoning shows that $\sum_k e^{-2\sqrt{2}y_k(R_T)} \to 0$.

To summarize, the non-trivial contributions in (3.19) come from the terms involving the random variable Z: taking the limit $T \uparrow \infty$ first and $r \uparrow \infty$ next (so that $\gamma(r) \downarrow 1$), implies that

$$\lim_{T\uparrow\infty} \mathbb{P}\left[M(T\cdot s) \le D \mid \mathcal{F}_{R_T}\right] = \exp\left(-CZe^{-\sqrt{2}D}\right), \text{ a.s.} \tag{3.25}$$

This concludes the proof of Theorem 2. □

4 The strong law of large numbers

This section is divided into two subsections. In Subsection 4.1 we analyse *localization* properties of the *paths* of extremal particles. Localization of the paths has played a fundamental role in [3] in the context of the genealogies of extremal particles. The details of the proof of the law of large numbers are given in Subsection 4.2.

4.1 Preliminaries and localization of the paths

The following fundamental result by Bramson [11] provides bounds to the right tail of the maximal displacement. (See also Roberts [26] for a different derivation). These bounds are not optimal (in fact, Lemma 4 is an improvement), they are however sufficient for our purposes here, and simpler.

Lemma 5. *[11, Section 5] Consider a branching Brownian motion $\{x_j(t)\}_{j\leq n(t)}$. Then, for $0 \leq y \leq t^{1/2}$ and $t \geq 2$,*

$$\mathbb{P}\left[M(t) \geq y\right] \leq \gamma (y+1)^2 e^{-\sqrt{2}y}, \tag{4.1}$$

where γ is independent of t and y.

Next we recall a property of the paths of extremal particles established in [3]. This requires some notation. For $t \in \mathbb{R}_+$ and $\gamma > 0$, we define

$$f_{\gamma,t}(s) \equiv \begin{cases} s^\gamma & 0 \leq s \leq t/2, \\ (t-s)^\gamma & t/2 \leq s \leq t. \end{cases} \tag{4.2}$$

Choose

$$0 < \alpha < 1/2 < \beta < 1, \tag{4.3}$$

and introduce the *time-t entropic* envelope, and the *time-t lower* envelope respectively:

$$F_{\alpha,t}(s) \equiv \frac{s}{t} m(t) - f_{\alpha,t}(s), \quad 0 \leq s \leq t, \tag{4.4}$$

and

$$F_{\beta,t}(s) \equiv \frac{s}{t} m(t) - f_{\beta,t}(s), \quad 0 \leq s \leq t, \tag{4.5}$$

($m(t)$ is the level of the maximum of a BBM of length t). By definition,

$$F_{\beta,t}(s) < F_{\alpha,t}(s), \quad 0 < s < t, \tag{4.6}$$

and

$$F_{\beta,t}(0) = F_{\alpha,t}(0) = 0, \quad F_{\beta,t}(t) = F_{\alpha,t}(t) = m(t). \tag{4.7}$$

The space/time region between the entropic and lower envelopes will be denoted throughout as the *time-t tube*, or simply the *tube*.

Given a particle $k \leq n(t)$ which is at position $x_k(t)$ at time t, we denote by $x_k(t,s)$ the position of its ancestor at time $s \in (0,t)$. We refer to the map $s \mapsto x_k(t,s)$ as the *path* of the particle k. We say that a particle k is *localized* in the time t-tube during the interval $(r, t-r)$ if and only if

$$F_{\beta,t}(s) \leq x_k(t,s) \leq F_{\alpha,t}(s), \forall s \in (r, t-r)\ . \tag{4.8}$$

Otherwise, we say that it is *not localized*. The following proposition gives strong bounds on the probability of finding particles which are, at given times, close to the level of the maximum, but not localized.

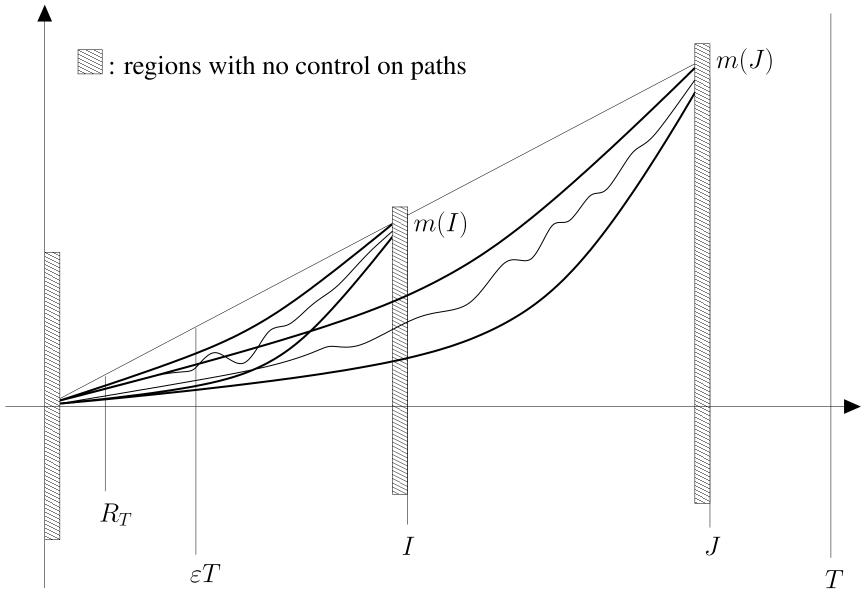

Figure 2: Maxima at different times I, J are localized. The first shaded region is the interval $(0, r_T)$, the second is $(I - r_T, I)$, the third is $(J - r_T, J)$.

Proposition 6. *Let the subset $\mathcal{D} = [d, D]$ be given, with $-\infty < d < D \leq \infty$. There exist $r_o, \delta > 0$ depending on α, β and $\mathcal{D}$ such that for $r \geq r_o$*

$$\sup_{t \geq 3r} \mathbb{P}\left[\exists_{k \leq n(t)}\, \overline{x_k(t)} \in \mathcal{D} \text{ but } x_k(s,t) \text{ not localized during } (r, t-r)\right] \leq \exp\left(-r^{\delta}\right). \quad (4.9)$$

Proof. The bound (4.9) is obtained combining equations (5.5), (5.54), (5.62) and (5.63) in [3, Corollary 2.6]. □

What lies behind the Proposition is a phenomenon of "energy vs. entropy" which is fundamental for the whole picture. This is explained in detail in [3], but for the convenience of the reader we briefly sketch the argument here.

As it turns out, at any given time $s \in (r, t-r)$ well inside the lifespan of a BBM, there are simply not enough particles lying above the entropic envelope for their offspring to make the jumps which eventually bring them to the edge at time t. On the other hand, although there are plenty of ancestors lying below the lower envelope, their position is so low that again none of their offspring will make it to the edge at time t. A delicate balance between number and positions of ancestors has to be met, and this feature is fully captured by the tubes.

Take $\delta = \delta(\alpha, \beta, \mathcal{D})$ as in Proposition 6. We pick $r \geq r_o(\alpha, \beta, \mathcal{D})$: we choose

$$r_T \equiv (20 \ln T)^{1/\delta}. \quad (4.10)$$

We now consider the maximum of the particles at a given time $s \in (R_T, T)$ that are also localized during the interval $(r_T, s - r_T)$. See Figure 2 for a graphical representation. We denote this maximum by $M_{\text{loc}}(s)$. With this notation, by Proposition 6 and the choice (4.10),

$$0 \leq \mathbb{P}\left[M(s) \in \mathcal{D}\right] - \mathbb{P}\left[M_{\text{loc}}(s) \in \mathcal{D}\right] \leq \frac{1}{T^{20}}. \quad (4.11)$$

We pick $R_T \equiv 40 \cdot r_T$, with r_T as in (4.10). This choice clearly satisfies $R_T = o(\sqrt{T})$ as required in Theorem 2. The choice of the prefactor is arbitrary: we only need that

$R_T > r_T$ (for the localization of path on $[R_T, T - R_T]$) and a choice of R_T that ensures summability in T (as seen, e.g. in (4.11)). We assume henceforth without loss of generality that both T and εT are integers.

4.2 Implementing the strategy

Recall that Theorem 3 asserts that

$$\mathrm{Rest}_{\varepsilon,\mathcal{D}}(T) \equiv \frac{1}{T}\int_{\varepsilon T}^{T} \left(\mathbb{1}_{\{M(s)\in\mathcal{D}\}} - \mathbb{P}\left[M(s)\in\mathcal{D} \mid \mathcal{F}_{R_T}\right]\right) ds \tag{4.12}$$

tends to zero a.s. for $T \uparrow \infty$. In order to prove the claim, we consider $\mathrm{Rest}^{\mathrm{loc}}_{\varepsilon,\mathcal{D}}(T)$, defined as $\mathrm{Rest}_{\varepsilon,\mathcal{D}}(T)$ but with the requirement that all particles in $\mathcal{D}$ are localized:

$$\mathrm{Rest}^{\mathrm{loc}}_{\varepsilon,\mathcal{D}}(T) \equiv \frac{1}{T}\int_{\varepsilon T}^{T} \left(\mathbb{1}_{\{M_{\mathrm{loc}}(s)\in\mathcal{D}\}} - \mathbb{P}\left[M_{\mathrm{loc}}(s)\in\mathcal{D} \mid \mathcal{F}_{R_T}\right]\right) ds\ . \tag{4.13}$$

We now claim that the large T-limit of $\mathrm{Rest}^{\mathrm{loc}}_{\varepsilon,\mathcal{D}}(T)$ and that of $\mathrm{Rest}_{\varepsilon,\mathcal{D}}(T)$ coincide (provided one of the two exists, but this will become apparent below).

Lemma 7. *With the notation introduces above,*

$$\lim_{T\uparrow\infty}\Big(\mathrm{Rest}_{\varepsilon,\mathcal{D}}(T) - \mathrm{Rest}^{\mathrm{loc}}_{\varepsilon,\mathcal{D}}(T)\Big) = 0,\ \textit{a.s.} \tag{4.14}$$

Proof of Lemma 7. We have

$$\begin{aligned}\mathrm{Rest}_{\varepsilon,\mathcal{D}}(T) - \mathrm{Rest}^{\mathrm{loc}}_{\varepsilon,\mathcal{D}}(T) &= \frac{1}{T}\int_{\varepsilon T}^{T}\left(\mathbb{1}_{\{M(s)\in\mathcal{D}\}} - \mathbb{1}\{M_{\mathrm{loc}}(s)\in\mathcal{D}\}\right) ds \\ -\frac{1}{T}\int_{\varepsilon T}^{T}&\Big(\mathbb{P}\left[M(s)\in\mathcal{D}\mid\mathcal{F}_{R_T}\right] - \mathbb{P}\left[M_{\mathrm{loc}}(s)\in\mathcal{D}\mid\mathcal{F}_{R_T}\right]\Big) ds \\ \equiv \boldsymbol{(1)_{T,\varepsilon}} - \boldsymbol{(2)_{T,\varepsilon}}. \end{aligned} \tag{4.15}$$

The proofs that $\lim_{T\uparrow\infty}\boldsymbol{(1)_{T,\varepsilon}} = 0$ and $\lim_{T\uparrow\infty}\boldsymbol{(2)_{T,\varepsilon}} = 0$ (almost surely) are identical and relies on an application of the Borel-Cantelli lemma. We thus prove only the first limit. Let $\epsilon > 0$. By Markov's inequality,

$$\begin{aligned}\mathbb{P}\left[\boldsymbol{(1)_{T,\varepsilon}} > \epsilon\right] &\leq \frac{1}{T\epsilon}\int_{\varepsilon T}^{T}\Big(\mathbb{P}\left[M(s)\in\mathcal{D}\right] - \mathbb{P}\left[M_{\mathrm{loc}}(s)\in\mathcal{D}\right]\Big) ds \leq \\ &\overset{(4.11)}{\leq} \frac{1-\varepsilon}{\epsilon}T^{-20},\end{aligned} \tag{4.16}$$

which is summable in T (recalling that we assume $T \in \mathbb{N}$). Therefore, by Borel-Cantelli,

$$\mathbb{P}\left[\{\boldsymbol{(1)_{T,\varepsilon}} > \epsilon\} \text{ infinitely often}\right] = 0. \tag{4.17}$$

As the above holds for all $\epsilon > 0$ we have that $\boldsymbol{(1)_{T,\varepsilon}}$ converges to 0 as $T \uparrow \infty$ almost surely, and concludes the proof of Lemma 7. □

The following result is the major tool to establish the SLLN for the term $\mathrm{Rest}^{\mathrm{loc}}_{\varepsilon,\mathcal{D}}(T)$. (By Lemma 7, this will then imply that the same is true for $\mathrm{Rest}_{\varepsilon,\mathcal{D}}(T)$). The result is a small extension of a theorem of Lyons [28, Theorem 1], where the statement is given for the sum of random variables.

Theorem 8. *Consider a process $\{X_s\}_{s\in\mathbb{R}_+}$ such that $\mathbb{E}[X_s] = 0$ for all s. Assume furthermore that the random variables are uniformly bounded almost surely. If*

$$\sum_{T=1}^{\infty}\frac{1}{T}\mathbb{E}\Big[\Big|\frac{1}{T}\int_0^T X_s\, ds\Big|^2\Big] < \infty, \tag{4.18}$$

then

$$\lim_{T\uparrow\infty} \frac{1}{T}\int_0^T X_s\, ds = 0,\ \textit{a.s.} \tag{4.19}$$

Proof. We suppose without loss of generality that $\sup_s |X_s| \le 1$. The extension to integrals is straightforward. By the summability assumption, we can find a subsequence $T_k \in \mathbb{N}$ of times such that

$$\sum_{k=1}^{\infty} \mathbb{E}\Big[\Big|\frac{1}{T_k}\int_0^{T_k} X_t\, dt\Big|^2\Big] < \infty \tag{4.20}$$

where $T_k \uparrow \infty$ and $T_{k+1}/T_k \to 1$. (In our case, the expectation in (4.18) will decay faster than $e^{-(\ln T)^\epsilon}$ for some $\epsilon > 0$, see Theorem 9 below. In particular, one can take the subsequence $T_k = \exp(k^{1/2})$. For the general case, we refer to [28, Lemma 2] and references therein.) Therefore by Fubini, the sum without the expectation is almost surely finite, and we must have

$$\lim_{k\uparrow\infty} \frac{1}{T_k}\int_0^{T_k} X_t\, dt \to 0 \text{ a.s.} \tag{4.21}$$

It remains to show this is true for all $T \in \mathbb{N}$. This is easy since the variables are bounded. For any T, there exists k such that $T_k \le T \le T_{k+1}$. Thus

$$\Big|\frac{1}{T}\int_0^T X_t\, dt\Big| \le \Big|\frac{1}{T_k}\int_0^{T_k} X_t\, dt\Big| + \sup_{1\le s\le T_{k+1}-T_k} \Big|\frac{1}{T_k}\int_{T_k}^{T_k+s} X_t\, dt\Big| . \tag{4.22}$$

The first term goes to zero by the previous argument. The second term goes to zero since

$$\sup_{1\le s\le T_{k+1}-T_k} \Big|\frac{1}{T_k}\int_{T_k}^{T_k+s} X_t\, dt\Big| \le \frac{T_{K+1}-T_k}{T_k} , \tag{4.23}$$

and $T_{k+1}/T_k \to 1$. □

Note that

$$\begin{aligned} \mathrm{Rest}^{\mathrm{loc}}_{\varepsilon,\mathcal{D}}(T) &= \frac{1}{T}\int_{\varepsilon T}^T \left(\mathbb{1}_{\{M_{\mathrm{loc}}(s)\le D\}} - \mathbb{P}\left[M_{\mathrm{loc}}(s)\le D \mid \mathcal{F}_{R_T}\right]\right) ds \\ &\quad - \frac{1}{T}\int_{\varepsilon T}^T \Big(\mathbb{1}\{M_{\mathrm{loc}}(s)\le d\} - \mathbb{P}\left[M_{\mathrm{loc}}(s)\le d \mid \mathcal{F}_{R_T}\right]\Big) ds \\ &\equiv \frac{1}{T}\int_{\varepsilon T}^T X_s^{\{D\}} ds - \frac{1}{T}\int_{\varepsilon T}^T X_s^{\{d\}} ds, \end{aligned} \tag{4.24}$$

with obvious notations. The goal is thus to prove that both integrals satisfy the assumptions of Theorem 8. We address the first integral, the proof for the second being identical. By construction, $\left|X_s^{\{D\}}\right| \le 2$ a.s. for all s, and

$$\mathbb{E}\left[X_s^{\{D\}}\right] = 0 . \tag{4.25}$$

It therefore suffices to check the assumption concerning the summability of correlations. Let

$$\widehat{C}_T(s,s') \equiv \mathbb{E}\left[X_s^{(D)} \cdot X_{s'}^{(D)}\right] . \tag{4.26}$$

Note that by the properties of conditional expectation

$$\begin{aligned}\widehat{C}_T(s,s') &= \mathbb{E}\Bigg[\left(\mathbb{1}_{\{M_{\rm loc}(s)\leq D\}} - \mathbb{P}\left[M_{\rm loc}(s)\leq D \mid \mathcal{F}_{R_T}\right]\right)\\ &\qquad\qquad \times \left(\mathbb{1}_{\{M_{\rm loc}(s')\leq D\}} - \mathbb{P}\left[M_{\rm loc}(s')\leq D \mid \mathcal{F}_{R_T}\right]\right)\Bigg]\\ &= \mathbb{E}\Bigg[\Bigg(\mathbb{P}\left[M_{\rm loc}(s)\leq D, M_{\rm loc}(s')\leq D \mid \mathcal{F}_{R_T}\right]\\ &\qquad\qquad - \mathbb{P}\left[M_{\rm loc}(s)\leq D \mid \mathcal{F}_{R_T}\right]\times \mathbb{P}\left[M_{\rm loc}(s')\leq D \mid \mathcal{F}_{R_T}\right]\Bigg)\Bigg].\end{aligned} \tag{4.27}$$

We claim that

$$\sum_T \frac{1}{T}\mathbb{E}\left[\Big|\frac{1}{T}\int_{\varepsilon T}^T X_s^{(D)}\, ds\Big|^2\right] = 2\sum_T \frac{1}{T^3}\int_{\varepsilon T}^T ds \int_s^T ds' \widehat{C}_T(s,s') < \infty\ . \tag{4.28}$$

In order to see this, and proceeding with the program outlined at the end of Section 2, we now specify the concept of times *well separated* from each other. Choose $0<\xi<1$ and split the integration according to the distance between s and s':

$$\frac{1}{T^3}\int_{\varepsilon T}^T ds \int_s^T ds'(\cdot) = \frac{1}{T^3}\int_{\varepsilon T}^T ds \int_s^{s+T^\xi} ds'(\cdot) + \frac{1}{T^3}\int_{\varepsilon T}^T ds \int_{s+T^\xi}^T ds'(\cdot). \tag{4.29}$$

The contribution of the first term on the r.h.s. above is negligible due to the uniform boundedness of the integrand and to the choice $0<\xi<1$. We are thus left to prove that the contribution to (4.28) of the second term in (4.29) is finite. The following is the key estimate.

Theorem 9. *There exists a finite T_o such that the following holds for $T\geq T_o$: for some $\epsilon>0$* not *depending on T (but on the other underlying parameters), the bound*

$$\widehat{C}_T(s,s') \leq e^{-(\ln T)^\epsilon} \tag{4.30}$$

holds uniformly for all s,s' such that $\varepsilon T\leq s<s'\leq T$ and $s'-s>T^\xi$.

Theorem 9 controls the decay of correlations at *specific* timescales. We have not tried to derive optimal bounds. There is in fact a certain freedom in the choice of the timescales, and certain choices are likely to yield better estimates. For the purpose of checking the conditions in Lyons Theorem, the bounds established are more than sufficient: they imply that the second term in (4.29) is at most $T^{-1}e^{-(\ln T)^\epsilon}$, which is summable over T by comparing to $(T\ln T)^{-1}$ (recall that T is assumed to be an integer). Theorem 3 therefore follows as soon as we prove Theorem 9. The proof of the latter is somewhat lengthy, and done in the next section.

5 Uniform bounds for the correlations.

We use here I and J to denote the two times s,s' from the statement of Theorem 9. $\widehat{C}_T(I,J)$ is the expectation of the random variable

$$\begin{aligned}\hat{c}_T(I,J) \equiv \mathbb{P}\left[M_{\rm loc}(I)\leq D, M_{\rm loc}(J)\leq D \mid \mathcal{F}_{R_T}\right]&\\ - \mathbb{P}\left[M_{\rm loc}(I)\leq D \mid \mathcal{F}_{R_T}\right]\times \mathbb{P}\left[M_{\rm loc}(J)\leq D \mid \mathcal{F}_{R_T}\right].&\end{aligned} \tag{5.1}$$

We rewrite these conditional probabilities using the Markov property of BBM, considering independent BBM's starting at their respective position at time R_T and shifting the time by R_T. This requires some additional notation. Take

$$I_T \equiv I - R_T, J_T \equiv J - R_T \ ,$$

and note that $m(I) = m(I_T) + \sqrt{2} I_T + o(1)$ as $T \uparrow \infty$. We consider the collection $\{y_k(R_T) \equiv \sqrt{2} R_T - x_k(R_T)\}_{k \le n(R_T)}$ where the $\{x_k(R_T)\}$ are the positions of the particles of the original BBM at time R_T.

Let $\{\tilde{x}_l(J_T), l \le n(J_T)\}$ be a BBM starting at zero, of length J_T, and of law $\tilde{\mathbb{P}}$ independent of $\mathbb{P}$. We write $\tilde{M}_{\text{loc}}(J_T)$ for the maximum shifted by $m(J_T)$ of this collection, restricted to l's whose *paths* (recall the notation introduced in Section 4.1) satisfy

$$y_k(R_T) + \frac{s'}{J} m(J) - f_{\beta,J}(R_T + s') \le \tilde{x}_l(J_T, s') \le y_k(R_T) + \frac{s'}{J} m(J) - f_{\alpha,J}(R_T + s') \ , \tag{5.2}$$

for $0 \le s' \le J_T - r_T$, the "shifted" J-tube.

Similarly, $\tilde{M}^k_{\text{loc}}(I_T)$ is the maximum shifted by $m(I_T)$ of the positions of the particles at time I_T with the localization condition

$$y_k(R_T) + \frac{s'}{I} m(I) - f_{\beta,I}(R_T + s') \le \tilde{x}_l(I_T, s') \le y_k(R_T) + \frac{s'}{J} m(J) - f_{\alpha,J}(R_T + s'), \tag{5.3}$$

for $0 \le s' \le I_T - r_T$, the "shifted" I-tube.

(Note that the localization depends on k through the random variable $y_k(R_T)$).

By the Markov property, the first conditional probability in $\hat{c}_T(I, J)$ can be written in terms of the shifted process just defined:

$$\begin{aligned} &\mathbb{P}\left[M_{\text{loc}}(I) \le D, M_{\text{loc}}(J) \le D \mid \mathcal{F}_{R_T}\right] \\ &= \prod_{k \le n(R_T)}^{\star} \tilde{\mathbb{P}}\left[\tilde{M}^k_{\text{loc}}(I_T) \le D + y_k(R_T), \tilde{M}^k_{\text{loc}}(J_T) \le D + y_k(R_T)\right], \end{aligned} \tag{5.4}$$

where the product runs over all the particles k's at time R_T whose path is localized in the intersection of the $I-$ and $J-$tubes during the interval (r_T, R_T). The restriction to localized positions at time R_T is weaker and sufficient for our purpose: we thus introduce the set of particles

$$\triangle \equiv \left\{k = 1, \dots, n(R_T) : \ y_k(R_T) \in \left(R_T^{\alpha} + \Omega_T, R_T^{\beta} + \Omega_T\right)\right\}. \tag{5.5}$$

(Here and henceforth, we will use Ω_T to denote a negligible term, which is not necessarily the same at different occurrences. In the above case it holds $\Omega_T = O(\ln \ln T)$ by definition of the tubes). We thus get that (5.4) is at most

$$\prod_{k \in \triangle} \tilde{\mathbb{P}}\left[\tilde{M}^k_{\text{loc}}(I_T) \le D + y_k(R_T), \tilde{M}^k_{\text{loc}}(J_T) \le D + y_k(R_T)\right] \ . \tag{5.6}$$

Let

$$\wp(I_T; y_k(R_T)) \equiv \tilde{\mathbb{P}}\left[\tilde{M}^k_{\text{loc}}(I_T) > D + y_k(R_T)\right], \tag{5.7}$$

(analogously for J_T) and

$$\wp(I_T, J_T; y_k(R_T)) \equiv \tilde{\mathbb{P}}\left[\tilde{M}_{\text{loc}}(I_T) > D + y_k(R_T) \text{ and } \tilde{M}_{\text{loc}}(J_T) > D + y_k(R_T)\right]. \tag{5.8}$$

Finally, define

$$\widehat{Z}(\,\cdot\,;R_T) \equiv \sum_{k\in\triangle}\wp(\,\cdot\,;y_k(R_T)), \tag{5.9}$$

$$\mathcal{R}_T \equiv \frac{1}{2}\sum_{k\in\triangle}\left\{\wp(I_T;y_k(R_T))+\wp(J_T;y_k(R_T))-\wp(I_T,J_T;y_k(R_T)\right\}^2. \tag{5.10}$$

Proposition 10. *With the above definitions,*

$$0\le \hat{c}_T(I,J)\le \widehat{Z}(I_T,J_T;R_T)+\mathcal{R}_T, \tag{5.11}$$

almost surely, for T large enough.

Proof. To simplify the notation, we drop here and henceforth the dependence on k in $\tilde{M}^k_{\mathrm{loc}}$. We have:

$$\begin{aligned}&\prod_{k\in\triangle}\tilde{\mathbb{P}}\left[\tilde{M}_{\mathrm{loc}}(I_T)\le D+y_k(R_T),\tilde{M}_{\mathrm{loc}}(J_T)\le D+y_k(R_T)\right]\\&\quad=\exp\left\{\sum_{k\in\triangle}\ln\left[1-\wp(I_T;y_k(R_T))-\wp(J_T;y_k(R_T))+\wp(I_T;J_T;y_k(R_T))\right]\right\}\ .\end{aligned} \tag{5.12}$$

Note that for all $k\in\triangle$

$$\wp(I_T;y_k(R_T))\le\tilde{\mathbb{P}}\left[\tilde{M}(I_T)\ge D+y_k(R_T)\right]\le\gamma(1+y_k(R_T)+D)^2e^{-\sqrt{2}(y_k(R_T)+D)}\ . \tag{5.13}$$

The first inequality holds by dropping the localization condition, the second follows from (5). Therefore, this probability can be made arbitrarily small (uniformly in k) by choosing T large enough. The same obviously holds for $\wp(J_T;y_k(R_T))$ and $\wp(I_T,J_T;y_k(R_T))$. Choose T large enough so that

$$\sup_{\triangle}\max\{\wp(I_T;y_k(R_T)),\wp(J_T;y_k(R_T)),\wp(I_T,J_T;y_k(R_T))\}\le 1/6. \tag{5.14}$$

Coming back to (5.12) and using that

$$-a\le\ln(1-a)\le -a+a^2/2\quad(0\le a\le 1/2), \tag{5.15}$$

(with $a\equiv\wp(I_T;y_k(R_T))+\wp(J_T;y_k(R_T))-\wp(I_T;J_T;y_k(R_T))$, for $k\in\triangle$), we get that (5.12) is at most

$$\exp\left(-\widehat{Z}(I_T;R_T)-\widehat{Z}(J_T;R_T)+\widehat{Z}(I_T,J_T;R_T)+\mathcal{R}_T\right). \tag{5.16}$$

This is an upper bound for the first conditional probability in the definition of $\hat{c}_T(I,J)$. A similar reasoning, using this time the first inequality in (5.15), yields a lower bound for the second term in $\hat{c}_T(I,J)$, i.e, the product of the conditional probabilities. This gives

$$\hat{c}_T(I,J)\le e^{-\widehat{Z}(I_T;R_T)-\widehat{Z}(J_T;R_T)}\left\{e^{\widehat{Z}(I_T,J_T;R_T)+\mathcal{R}_T}-1\right\}, \tag{5.17}$$

almost surely for T large enough. Using that for $a\ge 0$, $e^a-1\le a\cdot e^a$ shows that the right-hand side of (5.17) is bounded from above by

$$e^{-\widehat{Z}(I_T;R_T)-\widehat{Z}(J_T;R_T)}\left(\widehat{Z}(I_T,J_T;R_T)+\mathcal{R}_T\right)e^{\widehat{Z}(I_T,J_T;R_T)+\mathcal{R}_T}. \tag{5.18}$$

By construction, $\widehat{Z}(I_T, J_T; R_T) \leq \min\left\{\widehat{Z}(I_T; R_T); \widehat{Z}(I_T; R_T)\right\}$. This implies that

$$\widehat{Z}(I_T, J_T; R_T) - \frac{1}{2}\widehat{Z}(I_T; R_T) - \frac{1}{2}\widehat{Z}(J_T; R_T) \leq 0, \tag{5.19}$$

and therefore

$$\hat{c}_T(I, J) \leq \left(\widehat{Z}(I_T, J_T; R_T) + \mathcal{R}_T\right) e^{\mathcal{R}_T - \frac{1}{2}\widehat{Z}(I_T;R_T) - \frac{1}{2}\widehat{Z}(J_T;R_T)}. \tag{5.20}$$

To arrive at the claim of Proposition 10, it remains to get rid of the exponential on the r.h.s. of the preceeding inequality. Using the bound (5.26) together with the definition of $\widehat{Z}$ and rearranging terms, we arrive at

$$\begin{aligned} \mathcal{R}_T - \frac{1}{2}\widehat{Z}(I_T; R_T) - \frac{1}{2}\widehat{Z}(J_T; R_T) \quad &\leq \quad \sum_{k\in\triangle} \wp(I_T; y_k(R_T)) \left(3\wp(I_T; y_k(R_T)) - \frac{1}{2}\right) \\ &+ \quad \sum_{k\in\triangle} \wp(J_T; y_k(R_T)) \left(3\wp(J_T; y_k(R_T)) - \frac{1}{2}\right) \end{aligned} \tag{5.21}$$

In view of (5.13), there exists $T < \infty$ such that for all $k \in \triangle$:

$$3\wp(I_T; y_k(R_T)) - \frac{1}{2} \leq 0, \qquad 3\wp(J_T; y_k(R_T)) - \frac{1}{2} \leq 0. \tag{5.22}$$

Thus, for such T, *all* terms appearing in (5.21) are negative, and this implies that

$$\hat{c}_T(I, J) \leq \widehat{Z}(I_T, J_T; R_T) + \mathcal{R}_T, \tag{5.23}$$

concluding the proof of Proposition 10. □

Proof of Theorem 9. Taking expectation in Proposition 10, we first show that the expectation of $\mathcal{R}_T$ by its expectation yields the the desired bound in Theorem 9. Indeed, using that $(a + b + c)^2 \leq 4a^2 + 4b^2 + 4c^2$, we get the upper bound

$$\begin{aligned} \mathcal{R}_T &= \frac{1}{2} \sum_{k\in\triangle} \left(\wp(I_T; y_k(R_T)) + \wp(J_T; y_k(R_T)) - \wp(I_T, J_T; y_k(R_T)\right)^2 \\ &\leq 2 \sum_{k\in\triangle} \left(\wp(I_T; y_k(R_T))^2 + \wp(J_T; y_k(R_T))^2 + \wp(I_T, J_T; y_k(R_T))^2\right). \end{aligned} \tag{5.24}$$

Moreover,

$$\wp(I_T, J_T; y_k(R_T)) \leq \frac{1}{2}\wp(I_T; y_k(R_T)) + \frac{1}{2}\wp(J_T; y_k(R_T)). \tag{5.25}$$

Inserting this estimate into (5.24), we get

$$\mathcal{R}_T \leq \sum_{k\in\triangle} \left(3\wp(I_T; y_k(R_T))^2 + 3\wp(J_T; y_k(R_T))^2\right). \tag{5.26}$$

By (5.26), (5.13) and (5.5), and using the density of branching Brownian motion at time R_T, we get that there is a constant $\kappa < \infty$ such that for sufficiently large T,

$$\begin{aligned} \mathbb{E}\left[\mathcal{R}_T\right] &\leq \kappa \mathbb{E}\left[\sum_{k\in\triangle} y_k(R_T)^2 e^{-2\sqrt{2} y_k(R_T)}\right] \\ &\leq \kappa e^{R_T} \int_{R_T^\alpha + \Omega_T}^{R_T^\beta + \Omega_T} y^2 e^{-2\sqrt{2}y} e^{-\frac{(y-\sqrt{2}R_T)^2}{2R_T}} \frac{dy}{\sqrt{2\pi R_T}} \\ &\leq \kappa R_T^2 e^{-\sqrt{2}R_T^\alpha} = \kappa(\ln T)^{2/\delta} e^{-\kappa(\ln T)^{\alpha/\delta}}. \end{aligned} \tag{5.27}$$

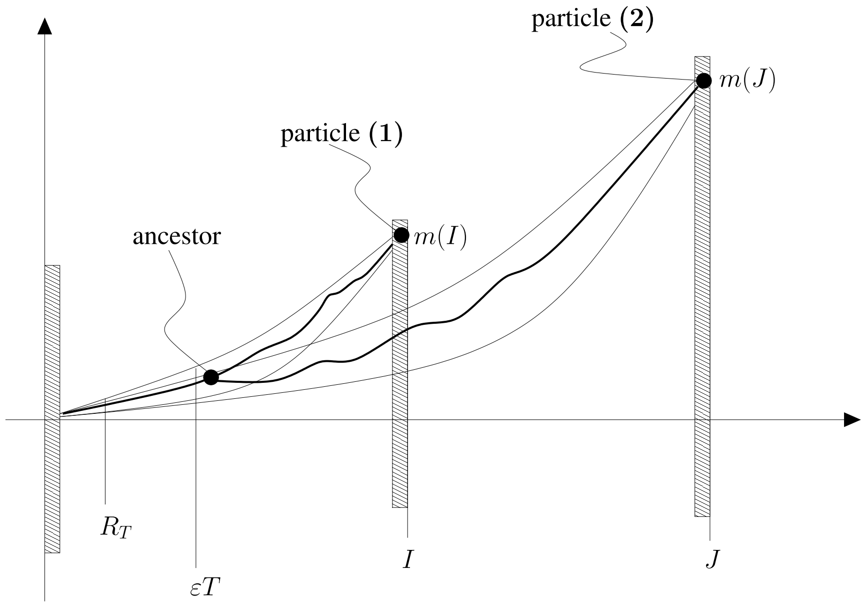

Figure 3: Time of branching before I. The first shaded region depicts the interval $(0, r_T)$, the second $(I - r_T, I)$, etc.

It remains to prove similar estimates for $\mathbb{E}[\widehat{Z}(I_T, J_T; R_T)]$ in order to prove Theorem 9.

Recall that

$$\widehat{Z}(I_T, J_T; R_T) = \sum_{k \in \triangle} \wp(I_T, J_T; y_k(R_T)), \tag{5.28}$$

and

$$\wp(I_T, J_T; y_k(R_T)) = \tilde{\mathbb{P}}\left[\tilde{M}_{\text{loc}}(I_T) > D + y_k(R_T) \text{ and } \tilde{M}_{\text{loc}}(J_T) > D + y_k(R_T)\right]. \tag{5.29}$$

By definition, (5.29) is the probability to find a particle of the BBM which has two extremal descendants, particle $(\mathbf{1})$ say, whose position is above $m(I_T) + D + y_k(R_T)$ at time I_T, and particle $(\mathbf{2})$, which lies above $m(J_T) + D + y_k(R_T)$ at time J_T. These two particles also satisfy localization conditions on their paths. In other words, this is the probability that the same ancestor k, with (relative) position $y_k(R_T)$, produces children $(\mathbf{1})$ and $(\mathbf{2})$ which are extremal at time I and J. As these generations are *well separated* in time, that is $J - I > T^\xi$ (and thus also $J_T - I_T > T^\xi$), we may expect this probability to be very small.

In order to see that this is indeed the case, split the probabilities according to whether the *most recent* common ancestor of particles $(\mathbf{1})$ and $(\mathbf{2})$ has branched *before* time $I_T - r_T$ (with r_T as in (4.10)), or *after*. We write this as

$$\begin{aligned}\wp(I_T, J_T; y_k(R_T)) = \wp(I_T, J_T; y_k(R_T); \text{split before } I_T - r_T)\\ + \wp(I_T, J_T; y_k(R_T); \text{split after } I_T - r_T).\end{aligned}$$

(Figure 3 illustrates the first case). The second probability is in fact zero. Indeed, the condition (5.2) implies that the ancestor of $(\mathbf{2})$ at time $I - r_T$ lies at height which is at most the level of the entropic envelope associated with J. Since $J - I > T^\xi$ and this is easily seen to be way lower than the *lower* envelope of particle $(\mathbf{1})$ associated with time I. In other words, the localization tubes of particles $(\mathbf{1})$ and $(\mathbf{2})$ are disjoint if their ancestor split after $I - r_T$. Hence, the splitting of the ancestor of particles $(\mathbf{1})$ and $(\mathbf{2})$

can only happen before time $I_T - r_T$:

$$\widehat{Z}(I_T, J_T; R_T) = \sum_{k\in\triangle} \wp(I_T, J_T; y_k(R_T); \text{split before } I_T - r_T), \text{ a.s.} \tag{5.30}$$

Proposition 11. *For some $\epsilon > 0$ and T large enough,*

$$\begin{aligned} &\wp(I_T, J_T; y_k(R_T); \textit{split before } I_T - r_T) \\ &\quad \leq e^{-(\ln T)^\epsilon} y_k(R_T) e^{-\sqrt{2}(y_k(R_T)}, \end{aligned} \tag{5.31}$$

uniformly for all $k \in \triangle$ and I_T, J_T as considered, almost surely.

The proof of this proposition is technical and postponed to Section 5.1. We show how this provides the last piece for the proof of Theorem 9. This is straightforward: by similar computations as in (5.27),

$$\mathbb{E}\left[\sum_{k\in\triangle} y_k(R_T) e^{-\sqrt{2} y_k(R_T)}\right] \leq \kappa R_T = \kappa' \ln(T)^{1/\delta}, \tag{5.32}$$

for large enough $\kappa' > 0$ and recalling that by definition $R_T = 40(\ln T)^{1/\delta}$. This, together with (5.31) implies

$$\mathbb{E}\left[\sum_{k\in\triangle} \wp(I_T, J_T; y_k(R_T); \text{split before } I_T - r_T)\right] \leq \kappa' e^{-(\ln T)^\epsilon}, \tag{5.33}$$

where ϵ has been adjusted to absorb the log-term. Combining this estimate with (5.27), the claim of Theorem 9 follows. □

5.1 Proof of Proposition 11

The claim is that

$$\begin{aligned} &\wp(I_T, J_T; y_k(R_T); \text{split before } I_T - r_T) \\ &\quad \leq e^{-(\ln T)^\epsilon} y_k(R_T) e^{-\sqrt{2} y_k(R_T)}, \end{aligned} \tag{5.34}$$

holds uniformly for $k \in \triangle$. In order to prove this, we use a formula by Sawyer [31] concerning the *expected number of pairs of particles* whose ancestor branched in the interval $(0, I_T - r_T)$ and whose paths satisfy certain localization conditions, say $T^{(1)}$ and $T^{(2)}$ respectively. The expected number of such pairs is given by

$$\begin{aligned} &K e^{I_T} \int_0^{I_T - r_T} ds \cdot e^{J_T - s} \int d\mu_s(y) \mathbb{P}\left[x \in T^{(1)}_{(0,s)} \cap T^{(2)}_{(0,s)} \mid x(s) = y\right] \\ &\times \mathbb{P}\left[x \in T^{(1)}_{(s,I_T)} \mid x(s) = y\right] \times \mathbb{P}\left[x \in T^{(2)}_{(s,J_T)} \mid x(s) = y\right]. \end{aligned} \tag{5.35}$$

Here the probability $\mathbb{P}$ is the law of a Brownian motion x, and $K = \sum_j p_j j(j-1)$ (with $\{p_j\}$ the offspring distribution). The time s is the branching time of the common ancestor, and μ_s is the Gaussian measure with variance s. $T^{(\cdot)}_{(a,b)}$ denotes the condition on the path during the time interval (a, b).

A proof of this formula is given in [31, p. 664 and 686]. Sawyer counts the pairs of particles for the *same* time, whereas our case concerns particles for two different times: particle $(\mathbf{1})$ at time I_T, and particle $(\mathbf{2})$ at time J_T. The generalization of Sawyer's formula is straightforward, although a formal derivation is somewhat involved. The reader is referred to the intuitive construction of the formula provided by Bramson [11, p. 564].

Dropping the condition $T^{(2)}$ in the first probability of (5.35) yields a simpler bound:

$$\begin{aligned}(5.35) \leq & K e^{I_T} \int_0^{I_T - r_T} ds \cdot e^{J_T - s} \int d\mu_s(y) \mathbb{P}\left[x \in T^{(1)}_{(0,I_T)} \mid x(s) = y\right] \\ & \times \mathbb{P}\left[x \in T^{(2)}_{(s,J_T)} \mid x(s) = y\right] .\end{aligned} \tag{5.36}$$

Note that $\wp(I_T, J_T; y_k(R_T); \text{split before } I_T - r_T)$ is by Markov's inequality at most the expected number of pairs $\{\mathbf{(1)}, \mathbf{(2)}\}$ of particles which satisfy their respective localization conditions with the common ancestor branching before time $I_T - r_T$. By (5.36), it holds that

$$\begin{aligned}&\wp(I_T, J_T; y_k(R_T); \text{split before } I_T - r_T) \\ &\leq K e^{I_T} \int_0^{I_T - r_T} ds \cdot e^{J_T - s} \int d\mu_s(y) \mathbb{P}\left[x \in T^{(1)}_{(0,I_T)} \mid x(s) = y\right] \\ &\qquad \times \mathbb{P}\left[x \in T^{(2)}_{(s,J_T)} \mid x(s) = y\right]\end{aligned} \tag{5.37}$$

with $T^{(1)}$ and $T^{(2)}$ being the shifted tubes defined in (5.2) and (5.3).

The idea is now to bound the second probability appearing in (5.37) uniformly in y. This procedure has been introduced in Bramson [11, Lemma 11], and proved useful also in [3, Theorem 2.1].

Lemma 12. *It holds that*

$$\begin{aligned}&\mathbb{P}\left[x \in T^{(2)}_{(s,J_T)} \mid x(s) = y\right] \\ &\leq \Omega_T^2 e^{-(J_T - s)} \exp\left(-\sqrt{2} f_{\alpha,J}(R_T + s) - \frac{3}{2} \ln\left(\frac{J_T - s}{J_T}\right) - \frac{3}{2} \frac{s}{J_T} \ln J_T\right),\end{aligned} \tag{5.38}$$

where $\Omega_T = O((\ln T)^{1/2\delta})$ *as* $T \uparrow \infty$.

For the proof of Lemma 12 some facts concerning the Brownian bridge are needed. Denoting a standard Brownian motion by x, the Brownian bridge of length t starting and ending at zero, is the Gaussian process

$$\mathfrak{z}_t(s) \equiv x(s) - \frac{s}{t} x(t), \qquad 0 \leq s \leq t. \tag{5.39}$$

The Brownian bridge is a Markov process, and it has the property that $\mathfrak{z}_t(s), 0 \leq s \leq t$ is independent of $x(t)$. This construction generalizes to the case where the endpoints of the bridge are $a, b \neq 0$; we denote by $\mathfrak{z}_t^{(a,b)}(s)$ such a process. The following is also well known:

$$\mathfrak{z}_t^{(a,b)}(s) \overset{(d)}{=} \mathfrak{z}_t(s) + \left(1 - \frac{s}{t}\right) a + \left(\frac{s}{t}\right) b, \qquad 0 \leq s \leq t, \tag{5.40}$$

with equality holding in distribution.

We now recall [3, Lemma 3.4] which deals with probabilities that a Brownian bridge stays below linear functions; the proof is elementary and will not be given here.

Lemma 13. *Let* $z_1, z_2 \geq 0$ *and* $r_1, r_2 \geq 0$. *Then for* $t > r_1 + r_2$,

$$\begin{aligned}&\mathbb{P}\left[\mathfrak{z}_t(s) \leq \left(1 - \frac{s}{t}\right) z_1 + \frac{s}{t} z_2,\, r_1 \leq s \leq t - r_2\right] \\ &\qquad \leq \frac{2}{t - r_1 - r_2} \prod_{i=1,2} \left\{z(r_i) + \sqrt{r_i}\right\},\end{aligned} \tag{5.41}$$

where $z(r_1) \equiv \left(1 - \frac{r_1}{t}\right) z_1 + \frac{r_1}{t} z_2$ *and* $z(r_2) \equiv \frac{r_2}{t} z_1 + \left(1 - \frac{r_2}{t}\right) z_2$.

Proof of Lemma 12. We begin by first writing explicitly the underlying conditions on the paths. For $f : \mathbb{R}_+ \to \mathbb{R}, t \mapsto f(t)$ a generic function, we denote by $f^S(\cdot) \equiv f(S+\cdot)$ its time-shift by $S > 0$. We also shorten $y(s) \equiv \sqrt{2}s - x(s)$, where $x(s) = y$ as in (5.37), and $J_{T,s} \equiv J_T - s$. We also set $\Omega_T \equiv O(\ln\ln T)$. Elementary manipulations lead to

$$\mathbb{P}\left[x \in T^{(2)}_{(s,J_T)} \mid x(s) = y\right] = \mathbb{P}\left[(\boldsymbol{E})\right], \tag{5.42}$$

where $(\boldsymbol{E})$ is the event

$$(\boldsymbol{E}) = \begin{cases} x(J_{T,s}) \geq m(J_{T,s}) + y(s) + \frac{3}{2\sqrt{2}}\ln\left(\frac{J_{T,s}}{J_T}\right) + D + y_k(R_T) + \Omega_T & (\boldsymbol{E_1}) \\ F_2(t) \leq x(t) \leq F_1(t), \qquad 0 \leq t \leq J_{T,s} - r_T & (\boldsymbol{E_2}) \end{cases} \tag{5.43}$$

where F_1, F_2 are the entropic (resp. lower) envelopes of (5.2) shifted by s:

$$\begin{aligned} F_1(t) &\equiv y_k(R_T) + y(s) + \frac{t}{J_T}m(J_T) + \frac{3}{2\sqrt{2}}\frac{s}{J_T}\ln(J_T) - f^{R_T+s}_{\alpha,J}(t) + \Omega_T, \\ F_2(t) &\equiv y_k(R_T) + y(s) + \frac{t}{J_T}m(J_T) + \frac{3}{2\sqrt{2}}\frac{s}{J_T}\ln(J_T) - f^{R_T+s}_{\beta,J}(t) + \Omega_T, \end{aligned} \tag{5.44}$$

with $\Omega_T = O(\ln\ln T)$. By the very same localization, we also have a condition on $x(s)$. This reads

$$x(s) \in \Big(-f^{R_T}_{\beta,J}(s); -f^{R_T}_{\alpha,J}(s)\Big) + y_k(R_T) + \sqrt{2}s - \frac{3}{2\sqrt{2}}\frac{s}{J_T}\ln(J_T). \tag{5.45}$$

For later use, we reformulate (5.45) into a condition on $y_k(R_T) + y(s)$, namely:

$$y_k(R_T) + y(s) \in \Big(f^{R_T}_{\alpha,J}(s); f^{R_T}_{\beta,J}(s)\Big) + \frac{3}{2\sqrt{2}}\frac{s}{J_T}\ln(J_T). \tag{5.46}$$

We now construct an event $(\boldsymbol{E'}) \supsetneq (\boldsymbol{E})$. First, we drop the condition that the Brownian path is required to stay *above* F_2. Second, we replace the condition on F_1 by the condition that the x-path remains, on the interval $(0, J_{T,s} - r_T)$, *below* the line segment interpolating between $(0, F_1(0))$ and $(J_{T,s}, F_1(J_{T,s}))$, see Figure 4 for a graphical representation. We consider

$$(\boldsymbol{E'}) = \begin{cases} x(J_{T,s}) \geq m(J_{T,s}) + y(s) + \frac{3}{2\sqrt{2}}\ln\left(\frac{J_{T,s}}{J_T}\right) + D + y_k(R_T) + \Omega_T & (\boldsymbol{E'_1}) \\ x(t) \leq \left(1 - \frac{t}{J_{T,s}}\right)F_1(0) + \frac{t}{J_{T,s}}F_1(J_{T,s}) \qquad 0 \leq t \leq J_{T,s} - r_T & (\boldsymbol{E'_2}) \end{cases} \tag{5.47}$$

By construction,

$$\mathbb{P}\left[(\boldsymbol{E})\right] \leq \mathbb{P}\left[(\boldsymbol{E'})\right]. \tag{5.48}$$

Let us put

$$\begin{aligned} X(s, J_T) &\equiv m(J_{T,s}) + y(s) + \frac{3}{2\sqrt{2}}\ln\left(\frac{J_{T,s}}{J_T}\right) + D + y_k(R_T) + \Omega_T \\ &= \sqrt{2}J_{T,s} - \frac{3}{2\sqrt{2}}\ln J_{T,s} + \left\{\frac{3}{2\sqrt{2}}\ln\left(\frac{J_{T,s}}{J_T}\right) + y(s) + y_k(R_T) + \Omega_T\right\}. \end{aligned} \tag{5.49}$$

We write

$$\mathbb{P}\left[(\boldsymbol{E'})\right] = \int_0^\infty \mathbb{P}\left[(\boldsymbol{E'_2})\middle| x(J_{T,s}) = X(s,J_T) + X\right]\tilde{\mu}(dX), \tag{5.50}$$

where $\tilde{\mu}$ is a Gaussian with variance $J_{T,s}$ and mean $-X(s, J_T)$, i.e.

$$\tilde{\mu}(dX) = \exp\left(-\frac{(X + X(s,J_T))^2}{2J_{T,s}}\right)\frac{dX}{\sqrt{2\pi J_{T,s}}}. \tag{5.51}$$

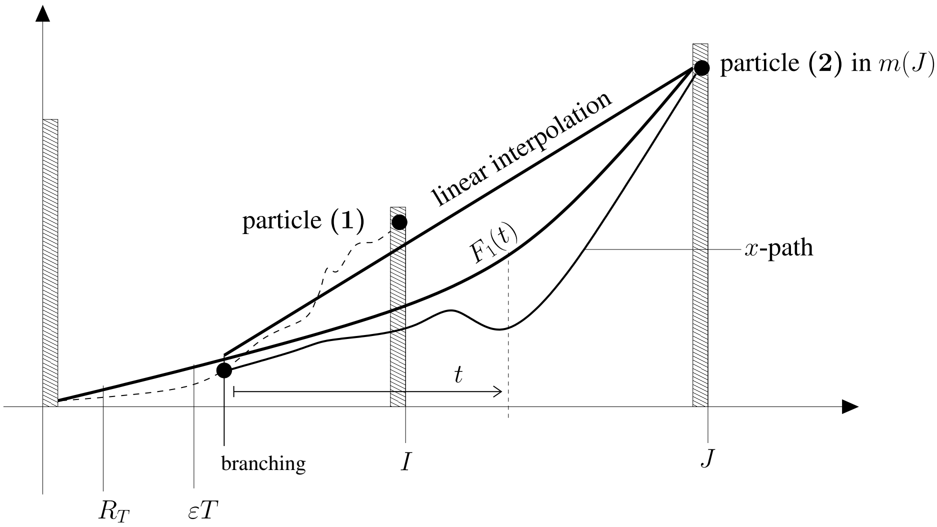

Figure 4: The path of an extremal particle at J stays below the linear interpolation.

We now make some observations concerning the Gaussian density and the conditional probability appearing in (5.50).

For the Gaussian density, we recall that $J_{T,s} = J_T - s$ for $0 \leq s \leq I_T - r_T \leq I_T$. Moreover, since $J_T - I_T > T^\xi$ and $J_T \geq \varepsilon T$, we see that

$$-(1-\xi)\ln T - \ln \varepsilon \leq \ln \left(\frac{J_{T,s}}{J_T}\right) \leq 0. \tag{5.52}$$

And,

$$y(s) + y_k(R_T) = o(J_{T,s}) \quad (T \uparrow \infty), \tag{5.53}$$

by (5.46). Therefore, combining (5.52) and (5.53) we have that $X(s, J_T) = \sqrt{2}J_{T,s} + o(J_{T,s})$ as $T \uparrow \infty$. The Gaussian density can thus be developed as follows

$$\tilde{\mu}(dX) = J_{T,s} e^{-J_{T,s}} e^{-\sqrt{2}\Delta_T(s)} g_T(X) dX, \tag{5.54}$$

where

$$\Delta_T(s) \equiv y(s) + \frac{3}{2\sqrt{2}} \ln\left(\frac{J_{T,s}}{J_T}\right) + y_k(R_T), \tag{5.55}$$

and

$$g_T(X) \equiv \frac{e^{-X^2/2J_{T,s}}}{\sqrt{2\pi}} e^{-\sqrt{2}(1+\omega_T)X} \left(1 + \Omega_T\right), \tag{5.56}$$

$\omega_T = o(1)$ as $T \uparrow \infty$, and $\Omega_T = O(\ln \ln T)$.

For the conditional probability appearing in (5.50), we observe that conditioning on the event $\{x(J_{T,s}) = X\}$, turns the Brownian motion involved in the definition of $\boldsymbol{E_2'}$ into a Brownian bridge ending at the conditioning point:

$$\mathbb{P}\left[(\boldsymbol{E_2'}) \mid x(J_{T,s}) = X(s, J_T) + X\right] = \mathbb{P}\left[(\boldsymbol{E''})\right], \tag{5.57}$$

where

$$(\boldsymbol{E''}) \equiv \left\{ \forall_{0\leq t\leq J_{T,s}-r_T} : \mathfrak{z}_{J_{T,s}}(t) \leq \left(1-\frac{t}{J_{T,s}}\right) F_1(0) + \frac{t}{J_{T,s}} \left(F_1(J_{T,s}) - X(s,J_T) - X\right) \right\}$$
$$= \left\{ \forall_{0\leq t\leq J_{T,s}-r_T} : \mathfrak{z}_{J_{T,s}}(t) \leq \left(1-\frac{t}{J_{T,s}}\right) F_1(0) + \frac{t}{J_{T,s}} \left(\Omega_T - X\right) \right\}, \tag{5.58}$$

since by (5.44) one has $F_1(J_{T,s}) = \Omega_T = O(\ln\ln T)$. We easily compute an upper bound to the probability of the $(\boldsymbol{E''})$-event. By Lemma 13, putting there $z_1 \equiv F_1(0)$ and $z_2 \equiv \max\{\Omega_T - X; 0\}$), it holds:

$$\mathbb{P}\left[(\boldsymbol{E''})\right] \leq \frac{2}{J_{T,s}-r_T} F_1(0) \left(\frac{r_T}{J_{T,s}} F_1(0) + \left(1-\frac{r_T}{J_{T,s}}\right)\max\{\Omega_T - X; 0\} + \sqrt{r_T}\right). \tag{5.59}$$

Since $F_1(0) = y_k(R_T) + y(s) - f_{\beta,J}(R_T+s) \leq \Omega_T$ by the localization (5.53), and $r_T \ll J_{T,s} = O(T)$, as $T\uparrow\infty$,

$$\mathbb{P}\left[(\boldsymbol{E''})\right] \leq \frac{2\max\{\Omega_T - X; 0\} + \sqrt{r_T}}{J_{T,s}}. \tag{5.60}$$

Inserting the bounds (5.60) and (5.54) into (5.50), perform the integral over dX, we immediately get that

$$\mathbb{P}\left[(\boldsymbol{E'})\right] \leq \Omega_T^2 e^{-J_{T,s}-\sqrt{2}\Delta_T(s)}, \tag{5.61}$$

for some $\Omega_T = O\left((\ln T)^\epsilon\right)$. By (5.46) we may now bound $\Delta_T(s)$ from below, *uniformly* in $y(s)$: the upshot is

$$\mathbb{P}\left[(\boldsymbol{E'})\right] \leq \Omega_T^2 e^{-J_{T,s}} \exp\left(-\sqrt{2} f_{\alpha,J}(R_T+s) - \frac{3}{2}\ln\left(\frac{J_{T,s}}{J_T}\right) - \frac{3}{2}\frac{s}{J_T}\ln J_T\right). \tag{5.62}$$

This is the uniform bound we were looking for and concludes the proof of Lemma 12. □

We finally give the

Proof of Proposition 11. Using the uniform bound provided by Lemma 12 in (5.37) and integrating over $\mu_s(dy)$ we obtain

$$\wp(I_T, J_T; y_k(R_T); \text{split before } I_T - r_T) \leq \kappa\cdot\Omega_T\cdot e^{I_T}\cdot\mathbb{P}\left[x\in T^{(1)}_{(0,I_T)}\right]$$
$$\times \int_0^{I_T-r_T} ds\cdot\exp\left(-\sqrt{2} f_{\alpha,J}(R_T+s) - \frac{3}{2}\ln\left(\frac{J_T-s}{J_T}\right) - \frac{3}{2}\frac{s}{J_T}\ln J_T\right). \tag{5.63}$$

The term $e^{I_T}\mathbb{P}\left[x\in T^{(1)}_{(0,I_T)}\right]$ can be handled by considerations similar to those in the proof of Lemma 12. The condition $T^{(1)}_{(0,I_T)}$ gives rise to the event

$$\begin{cases} x(I_T) \geq m(I_T) + D + y_k(R_T), & \\ F_2(t) \leq x(t) \leq F_1(t), & 0\leq t\leq I_T - r_T. \end{cases} \tag{5.64}$$

where

$$F_1(t) \equiv y_k(R_T) + \frac{t}{I_T} m(I_T) - f^{R_T}_{\alpha,I}(t) + \Omega_T$$
$$F_2(t) \equiv y_k(R_T) + \frac{t}{I_T} m(I_T) - f^{R_T}_{\beta,I}(t) + \Omega_T. \tag{5.65}$$

(For some $\Omega_T = O(\ln\ln T)$). In particular, the probability of the event is bounded by the probability that a Brownian motion stays below the linear interpolation of the points

$(0, F_1(0))$ and $(I_T - r_T, F_1(I_T - r_T))$ during the interval of time $(0, I_T - r_T)$ intersected with the event $x(I_t - r_T) \geq F_2(I_T - r_T)$, that is:

$$\mathbb{P}\left[x(t) \leq \frac{t}{I_T - r_T} F_1(I_T - r_T) + \left(1 - \frac{t}{I_T - r_T}\right) F_1(0),\ \forall 0 \leq t \leq I_T - r_T, \right. \\ \left. x(I_t - r_T) \geq F_2(I_T - r_T)\right] \tag{5.66}$$

Subtracting $\frac{t}{I_T - r_T} x(I_T - r_T)$ and using the fact that $x(I_t - r_T) \geq F_2(I_T - r_T)$, the above can be bounded above by $\mathbb{P}\Big[x(I_t - r_T) \geq F_2(I_T - r_T)\Big]$ times the Brownian bridge probability:

$$\mathbb{P}\left[\mathfrak{z}_{I_T - r_T}(t) \leq \frac{t}{I_T - r_T}(F_1(I_T - r_T) - F_2(I_T - r_T)) \right. \\ \left. + \left(1 - \frac{t}{I_T - r_T}\right) F_1(0),\ \forall 0 \leq t \leq I_T - r_T\right]. \tag{5.67}$$

Now $F_1(I_T - r_T) - F_2(I_T - r_T) \leq \kappa R_T^\beta$, for some $\kappa > 0$. Therefore the probability in (5.67) can be bounded using Lemma 13 by

$$\frac{2\kappa}{I_T - r_t} R_T^\beta\, F_1(0) = \frac{2\kappa}{I_T - r_t} R_T^\beta\, \left(y_k(R_T) + D - R_T^\alpha + \Omega_T\right). \tag{5.68}$$

Now, note that $m(I_T) - m(I_T - r_T) = \sqrt{2}\, r_T + o(1)$. Therefore, for some $\kappa > 0$,

$$F_2(I_T - r_T) - m(I_T - r_T) \geq y_k(R_T) + \kappa\, r_T\,. \tag{5.69}$$

A standard Gaussian estimate thus yields for some $\epsilon > 0$,

$$\mathbb{P}\Big[x(I_t - r_T) - m(I_T - r_T) \geq F_2(I_T - r_T) - m(I_T - r_T)\Big] \\ \leq \kappa(I_T - r_T) e^{-\sqrt{2} y_k(R_T)} e^{-(\ln T)^\epsilon}. \tag{5.70}$$

A combination of the above equation and (5.68) gives a bound of the desired form (5.31).

It remains to provide a similar bounds for the integral in (5.63). We first write

$$\int_0^{I_T - r_T} = \int_0^{I_T/2} + \int_{I_T/2}^{I_T - r_T}\ . \tag{5.71}$$

For the first integral, since $s \leq I_T/2$, we have

$$\ln\left(\frac{J_T - s}{J_T}\right) = \ln\left(1 - \frac{s}{J_T}\right) \geq \ln\left(\frac{1}{2}\right) \tag{5.72}$$

hence, up to irrelevant numerical constant, the contribution of the first integral is at most

$$\Omega_T^2 \int_0^{I_T/2} ds e^{-\sqrt{2}(R_T + s)^\alpha} \leq \Omega_T^2 \int_{R_T}^\infty ds e^{-\sqrt{2} s^\alpha} \leq e^{-\sqrt{2} R_T^\epsilon} \tag{5.73}$$

for some $\epsilon > 0$ small enough. The contribution of the second integral is sub-exponentially small (in T). To see this, recall that $J_T - I_T > T^\xi$ and $s \in [I_T/2, I_T - r_T]$, thus for some $\kappa_1 < 0 < \kappa_2$,

$$\kappa_1 \ln T \leq \ln\left(\frac{J_T - s}{J_T}\right) \leq \kappa_2 \ln T \tag{5.74}$$

implying that the second integral is, for some $\kappa > 0$, at most

$$T^\kappa \int_{I_T/2}^{I_T - r_T} e^{-\sqrt{2} f_{\alpha,J}(R_T+s)} ds \quad \le T^\kappa e^{-T^{\epsilon^{(6)}}} \le e^{-T^{\epsilon^{(7)}}} \tag{5.75}$$

for some $\epsilon^{(6)}, \epsilon^{(7)} > 0$. This is obviously much smaller than the first contribution (5.73). Therefore, summing this up,

$$\begin{aligned} &\wp(I_T, J_T; y_k(R_T); \text{split before } I_T - r_T) \\ &\quad \le (\ln T)^\epsilon e^{-(\ln T)^\epsilon} y_k(R_T) e^{-\sqrt{2} y_k(R_T)}. \end{aligned} \tag{5.76}$$

We can now adjust the value of ϵ to absorb the log-term. This concludes the proof of Proposition 11.

□

References

[1] E. Aidekon, *Convergence in law of the minimum of a branching random walk*, Ann. Probab., *to appear*

[2] E. Aidekon, J. Berestycki, E. Brunet, and Z. Shi. *The branching Brownian motion seen from its tip*, Probab. Theory Relat. Fields, DOI 10.1007/s00440-012-0461-0.

[3] L.-P. Arguin, A. Bovier, and N. Kistler, *The genealogy of extremal particles of branching Brownian motion*, Comm. Pure Appl. Math. 64: 1647-1676 (2011) MR-2838339

[4] L.-P. Arguin, A. Bovier, and N. Kistler, *Poissonian statistics in the extremal process of branching Brownian motion*, Ann. Appl. Prob., 22: 1693-1711 (2012). MR-2985174

[5] L.-P. Arguin, A. Bovier, and N. Kistler, *The extremal process of branching Brownian motion*, Probab. Theory Relat. Fields, DOI 10.1007/s00440-012-0464-x.

[6] D.G. Aronson and H.F. Weinberger, *Nonlinear diffusion in population genetics, combustion and nerve propagation*, in *Partial Differential Equations and Related Topics*, ed. J.A. Goldstein; Lecture Notes in Mathematics 446: 5-49, Springer, New York (1975). MR-0427837

[7] D.G. Aronson and H.F. Weinberger, *Multi-dimensional nonlinear diffusions arising in population genetics*, Adv. in Math., 30: 33-76 (1978). MR-0511740

[8] E. Bolthausen, J.-D. Deuschel, and G. Giacomin, *Entropic repulsion and the maximum of the two- dimensional harmonic crystal*, Ann. Probab. 29: 1670–1692 (2001) MR-1880237

[9] E. Bolthausen, J.-D. Deuschel, and O. Zeitouni, *Recursions and tightness for the maximum of the discrete, two dimensional Gaussian Free Field*, Electron. Commun. Probab. 16: 114-119 (2011) MR-2772390

[10] A. Bovier and I. Kurkova, *Derrida's generalized random energy models. 2. Models with continuous hierarchies.* Ann. Inst. H. Poincare. Prob. et Statistiques (B) Prob. Stat. 40: 481-495 (2004). MR-2070335

[11] M. Bramson, *Maximal displacement of branching Brownian motion*, Comm. Pure Appl. Math. 31: 531-581 (1978) . MR-0494541

[12] M. Bramson, *Convergence of solutions of the Kolmogorov equation to travelling waves*, Mem. Amer. Math. Soc. 44, no. 285, iv+190 pp. (1983) MR-0705746

[13] M. Bramson and O. Zeitouni, *Tightness of the recentered maximum of the two-dimensional discrete Gaussian Free Field*, Comm. Pure Appl. Math. 65: 1-20 (2012) MR-2846636

[14] E. Brunet and B. Derrida, *Statistics at the tip of a branching random walk and the delay of traveling waves*, Eurphys. Lett. 87: 60010 (2009).

[15] E. Brunet and B. Derrida, *A branching random walk seen from the tip*, J. Stat. Phys. 143: 420-446 (2010). MR-2799946

[16] B. Chauvin and A. Rouault, *Supercritical Branching Brownian Motion and K-P-P Equation in the Critical Speed-Area*, Math. Nachr. 149: 41-59 (1990). MR-1124793

[17] D. Carpentier and P. Le Doussal, *Glass transition of a particle in a random potential, front selection in nonlinear renormalization group, and entropic phenomena in Liouville and sinh-Gordon models*, Physical Review E63:026110 (2001); Erratum-ibid.E73:019910 (2006).

[18] B. Derrida and H. Spohn, *Polymers on disordered trees, spin glasses, and travelling waves*, J. Statist. Phys. 51: 817-840 (1988). MR-0971033

[19] A. Dembo, *Simple random covering, disconnection, late and favorite points*, Proceedings of the International Congress of Mathematicians, Madrid, Volume III, pp. 535-558 (2006) MR-2275695

[20] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni, *Cover times for Brownian motion and random walks in two dimensions*, Ann. Math. 160: 433-464 (2004) MR-2123929

[21] R.A. Fisher, *The wave of advance of advantageous genes*, Ann. Eugen. 7: 355-369 (1937).

[22] Y.V. Fyodorov and J.-P. Bouchaud, *Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential*, J. Phys. A 41 372001 (2008). MR-2430565

[23] S.C. Harris, *Travelling-waves for the FKPP equation via probabilistic arguments*, Proc. Roy. Soc. Edin., 129A: 503-517 (1999). MR-1693633

[24] D. A. Kessler, H. Levine, D. Ridgway, and L. Tsimring, *Evolution on a Smooth Landscape*, J. Stat. Phys. 87: 519-544 (1997).

[25] A. Kolmogorov, I. Petrovsky, and N. Piscounov, *Etude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique*, Moscou Universitet, Bull. Math. 1: 1-25 (1937).

[26] M. I. Roberts, *A simple path to asymptotics for the frontier of a branching Brownian motion*, Ann. Probab., to appear

[27] S.P. Lalley and T. Sellke, *A conditional limit theorem for the frontier of a branching Brownian motion*, Ann. Probab. 15: 1052-1061 (1987). MR-0893913

[28] R. Lyons, *Strong laws of large numbers for weakly correlated random variables*, Michigan Math. J. 35: 353-359 (1988) MR-0978305

[29] H.P. McKean, *Application of Brownian Motion to the equation of Kolmogorov-Petrovskii-Piskunov*, Comm. Pure. Appl. Math. 28: 323-331 (1976). MR-0400428

[30] S. Munier and R. Peschanski, *Traveling wave fronts and the transition to saturation*, Phys. Rev. D 69 034008 (2004).

[31] S. Sawyer, *Branching diffusion processes in population genetics*, Adv. Appl. Prob. 8: 659-689 (1976) MR-0432250

Acknowledgments. Nicola Kistler gratefully acknowledges discussions with Eric Brunet and Bernard Derrida.