More on the structure of extreme level sets in branching Brownian motion

Aser Cortines*   Lisa Hartung†   Oren Louidor‡

Abstract

This work is a continuation of [6], in which the same authors studied the fine structure of the extreme level sets of branching Brownian motion, namely the sets of particles whose height is within a finite distance from the global maximum. It is well known that such particles congregate at large times in clusters of order-one genealogical diameter around local maxima which form a Cox process in the limit. Our main finding here is that most of the particles in an extreme level set come from only a small fraction of the clusters, which are atypically large.

Keywords: Branching Brownian motion; extreme value theory.
MSC2020 subject classifications: 60J80; 60G70; 60G15.
Submitted to ECP on March 4, 2019, final version accepted on December 27, 2020.
Supersedes `arXiv:1902.08197`.

1 Introduction and results

1.1 Setup and state of the art

This work is a continuation of [6], in which the fine structure of the extreme values of branching Brownian motion (BBM) was studied. Let us first recall the definition of BBM and some of the state-of-the-art concerning its extreme value statistics. Let L_t be the set of particles alive at time $t \geq 0$ in a continuous time Galton-Watson process with binary branching at rate 1. The entire genealogy can be recorded via the metric space (T, d), consisting of the elements $T := \cup_{t \geq 0} L_t$ and equipped with the *genealogical distance*,

$$
\mathrm{d}_t(x, x') := \inf \left\{ \tfrac{(t-s)+(t'-s)}{2} \; : \; s \geq 0, \; x, x' \text{ share an ancestor in } L_s \right\} \quad ; x \in L_t, \, x' \in L_{t'} \,,
$$

for any $t, t' \geq 0$.

Conditional on (T, d), let $h = (h(x) : x \in T)$ be a mean-zero Gaussian process with covariance function given by $\mathbb{E}h(x)h(x') = (t + t')/2 - d(x, x')$ for $x \in L_t$ and $x' \in L_{t'}$ and $t, t' \geq 0$. Equivalently, $\mathbb{E}h(x)h(x')$ is equal to the largest $s \geq 0$ such that both x and x' share a common ancestor at time s. Then the triplet (h, T, d) (or just h for short) forms a standard BBM and $h(x)$ for $x \in L_t$ is interpreted as the height of particle x at time t. The restriction of T to all particles born up-to time t, will be denoted by $T_t := \cup_{s \leq t} L_s$, with d_t

---

*Universität Zürich, Switzerland. E-mail: `aserpeixoto@gmail.com`
†Universität Mainz, Germany. E-mail: `lhartung@uni-mainz.de`
‡Technion, Israel. E-mail: `oren.louidor@gmail.com`.

and h_t the corresponding restrictions of d and h, respectively. The natural filtration of the process $(\mathcal{F}_t : t \geq 0)$ can then be defined via $\mathcal{F}_t = \sigma(h_t, T_t, \mathrm{d}_t)$ for all $t \geq 0$.

The study of extreme values of h dates back to works of McKean [9], Bramson [3, 4] and Lalley and Sellke [8] who derived asymptotics for the law of the maximal height $h_t^* = \max_{x \in L_t} h_t(x)$. Introducing the centering function

$$m_t := \sqrt{2}t - \frac{3}{2\sqrt{2}} \log^+ t, \quad \text{where} \qquad \log^+ t := \log(t \vee 1), \tag{1.1}$$

and writing $\widehat{h}_t$ for the centered process $h_t - m_t$ and $\widehat{h}_t^* := h_t^* - m_t$ for its maximum, these works show that $\widehat{h}_t^*$ converges in law to $G + (1/\sqrt{2}) \log Z$ as $t \to \infty$, where G is a Gumbel distributed random variable and Z, which is independent of G, is the almost sure limit as $t \to \infty$ of (a multiple of) the so-called *derivative martingale*:

$$Z_t := C_\diamond \sum_{x \in L_t} \left(\sqrt{2}t - h_t(x)\right) \mathrm{e}^{\sqrt{2}(h_t(x) - \sqrt{2}t)}, \tag{1.2}$$

for some properly chosen $C_\diamond > 0$.

Other extreme values of h can be studied simultaneously by considering the *extremal process* associated with it. To describe the latter, given $t \geq 0$, $x \in L_t$ and $r > 0$, we let $\mathcal{C}_{t,r}(x)$ denote the *cluster* of *relative heights* of particles in L_t, which are at genealogical distance at most r from x. This is defined formally as the point measure,

$$\mathcal{C}_{t,r}(x) := \sum_{y \in \mathrm{B}_r(x)} \delta_{h_t(y) - h_t(x)}, \quad \text{where} \quad \mathrm{B}_r(x) := \{y \in L_t : \mathrm{d}_t(x, y) < r\}. \tag{1.3}$$

Fixing any positive function $t \mapsto r_t$ such that both r_t and $t - r_t$ tend to ∞ as $t \to \infty$ and letting $L_t^* = \{x \in L_t : h_t(x) \geq h_t(y), \forall y \in \mathrm{B}_{r_t}(x)\}$, the *structured extremal process* is then given as

$$\widehat{\mathcal{E}}_t := \sum_{x \in L_t^*} \delta_{h_t(x) - m_t} \otimes \delta_{\mathcal{C}_{t,r_t}(x)}. \tag{1.4}$$

That is, $\widehat{\mathcal{E}}_t$ is a point process on $\mathbb{R} \times \mathbb{M}$, where $\mathbb{M}$ denote the space of all point measures on $(-\infty, 0]$, which records *the-centered-height-of* (u) and *the-cluster-around* ($\mathcal{C}$) – all r_t-local maxima of h. We shall sometimes refer to the pair $(u, \mathcal{C})$ as a *cluster-pair*. It was then shown in [1, 2] that

$$\left(\widehat{\mathcal{E}}_t, Z_t\right) \overset{t \to \infty}{\Longrightarrow} \left(\widehat{\mathcal{E}}, Z\right), \quad \text{where} \qquad \widehat{\mathcal{E}}|Z \sim \mathrm{PPP}(Z\mathrm{e}^{-\sqrt{2}u}\mathrm{d}u \otimes \nu). \tag{1.5}$$

Above Z_t and Z are as before and ν is a deterministic distribution on $\mathbb{M}$, which we will call the *cluster distribution*. As two consequences, one gets the convergence of the *standard* extremal process of h:

$$\mathcal{E}_t \overset{t \to \infty}{\Longrightarrow} \mathcal{E} \quad \text{where} \quad \mathcal{E}_t := \sum_{x \in L_t} \delta_{h_t(x) - m_t}, \quad \mathcal{E} := \sum_{(u, \mathcal{C}) \in \widehat{\mathcal{E}}} \mathcal{C}(\cdot - u), \tag{1.6}$$

as well as the convergence of the *extremal process of local maxima*:

$$\mathcal{E}_t^* \overset{t \to \infty}{\Longrightarrow} \mathcal{E}^* \quad \text{where} \quad \mathcal{E}_t^* := \sum_{x \in L_t^*} \delta_{h_t(x) - m_t}, \quad \mathcal{E}^* := \sum_{(u, \mathcal{C}) \in \widehat{\mathcal{E}}} \delta_u \sim \mathrm{PPP}(Z\mathrm{e}^{-\sqrt{2}u}). \tag{1.7}$$

Henceforth, we shall use the unified notation $\mathcal{E}_{(t)}$ (also $\mathcal{E}_{(t)}^*$, $\widehat{\mathcal{E}}_{(t)}$) to mean $\mathcal{E}$ or $\mathcal{E}_t$.

The asymptotic growth of the number of extreme values which are also r_t-local maxima can then be read off directly from (1.7). Indeed, a simple application of the weak law of large numbers combined with the convergence statement in (1.7) yield,

$$\frac{\mathcal{E}^*([-u, \infty))}{Z\mathrm{e}^{\sqrt{2}u}/\sqrt{2}} \overset{\mathbb{P}}{\to} 1 \text{ as } u \to \infty, \qquad \frac{\mathcal{E}_t^*([-u, \infty))}{Z\mathrm{e}^{\sqrt{2}u}/\sqrt{2}} \overset{\mathbb{P}}{\to} 1 \text{ as } t \to \infty \text{ then } u \to \infty. \tag{1.8}$$

The asymptotic growth of the number of *all* extreme values, which is arguably the more interesting quantity, is not however a straightforward consequence of (1.6). This is because the limiting process $\mathcal{E}$ is now a superposition of i.i.d. clusters $\mathcal{C}$, and the law ν of the latter will determine the number of points inside any given set in the overall process.

To address this question, a study of the cluster law ν was carried out in [6] and then used to show (Proposition 1.5) that

$$\mathbb{E}\mathcal{C}([-v,0]) \sim C_\star \mathrm{e}^{\sqrt{2}v} \text{ as } v \to \infty \quad ; \quad \mathcal{C} \sim \nu, \tag{1.9}$$

for some $C_\star > 0$. This was combined with (1.5) and (1.6) to derive (Theorem 1.1 in [6])

$$\frac{\mathcal{E}([-v,\infty))}{C_\star Z v \mathrm{e}^{\sqrt{2}v}} \xrightarrow{\mathbb{P}} 1 \text{ as } v \to \infty \quad , \quad \frac{\mathcal{E}_t([-v,\infty))}{C_\star Z v \mathrm{e}^{\sqrt{2}v}} \xrightarrow{\mathbb{P}} 1 \text{ as } t \to \infty \text{ then } v \to \infty, \tag{1.10}$$

The above shows that points coming from the clusters around extreme local maxima account for an additional multiplicative linear prefactor in the overall growth rate of extreme values.

Next, it is natural to ask what are the *typical height* $u \in \mathbb{R}$ and *typical cluster configuration* $\mathcal{C}$ of those cluster-pairs $(u,\mathcal{C}) \in \widehat{\mathcal{E}}_{(t)}$ which *carry* the level set $\mathcal{E}_{(t)}|_{[-v,\infty)}$. The question of the typical height was addressed in [6]. To give a precise statement, for a Borel set $B \subseteq \mathbb{R} \times \mathbb{M}$ let us define (compare with (1.6)):

$$\mathcal{E}_t(\,\cdot\,;B) := \sum_{(u,\mathcal{C}) \in \widehat{\mathcal{E}}_t} \mathcal{C}(\,\cdot - u)1_{\{(u,\mathcal{C})\in B\}} \quad \text{and} \quad \mathcal{E}(\,\cdot\,;B) := \sum_{(u,\mathcal{C}) \in \widehat{\mathcal{E}}} \mathcal{C}(\,\cdot - u)1_{\{(u,\mathcal{C})\in B\}}. \tag{1.11}$$

These processes record all extreme values coming from cluster pairs $(u,\mathcal{C})$ in B. Theorem 1.2 from [6] then showed that for any $\alpha \in (0,1]$,

$$\frac{\mathcal{E}\big([-v,\infty);\ [-\alpha v,\infty)\times\mathbb{M}\big)}{\mathcal{E}\big([-v,\infty)\big)} \xrightarrow[v\to\infty]{\mathbb{P}} \alpha \quad \text{and} \quad \frac{\mathcal{E}_t\big([-v,\infty);\ [-\alpha v,\infty)\times\mathbb{M}\big)}{\mathcal{E}_t\big([-v,\infty)\big)} \xrightarrow[\substack{t\to\infty\\v\to\infty}]{\mathbb{P}} \alpha. \tag{1.12}$$

By the second limit we mean that the probability that difference between the right and left sides exceeds any $\epsilon > 0$, goes to zero when one first takes a limit-superior as $t \to \infty$ followed by a limit as $v \to \infty$. In other words, the typical height u of those extreme local maxima of which clusters contribute to $\mathcal{E}_{(t)}|_{[-v,\infty)}$ is uniformly chosen in $\big[-v, O(1)\big)$. This left the open question of the typical cluster configurations carrying $\mathcal{E}_{(t)}|_{[-v,\infty)}$, and this is precisely the focus of this manuscript.

1.2 New results

Our first observation is that $\mathcal{C}([-v,0])$ is not concentrated around its mean. This could be readily concluded from the results in [6], by comparing the first and second moment of this quantity (see also Lemma 2.7 below). The next proposition provides a quantitative version for this assertion.

Proposition 1.1. *Let $\mathcal{C} \sim \nu$. For all $\epsilon > 0$, there exists $\delta > 0$, such that for all v large enough,*

$$\mathbb{E}\Big(\mathcal{C}([-v,0]);\ \mathcal{C}([-v,0]) \leq \delta v \mathrm{e}^{\sqrt{2}v}\Big) \leq \epsilon \mathrm{e}^{\sqrt{2}v}. \tag{1.13}$$

Moreover, there exists $C > 0$ such that for all $\delta > 0$ and $v > 0$,

$$\mathbb{P}\Big(\mathcal{C}([-v,0]) > \delta v \mathrm{e}^{\sqrt{2}v}\Big) \leq \frac{C}{\delta}\, v^{-1}. \tag{1.14}$$

The above proposition shows that the asymptotic mean on the right hand side of (1.9) is the result of an unlikely event of probability $O(v^{-1})$, in which the number of cluster

points above $-v$ is of unusually high order $v \mathrm{e}^{\sqrt{2}v}$. Let us call a cluster satisfying the event in (1.14) a δ-*fat cluster for the height* $-v$, or a $(-v, \delta)$-*fat* cluster, for short.

Since $\mathcal{E}_{(t)}$ is a superposition of many i.i.d. clusters $\mathcal{C}$, the above proposition should translate, by virtue of the law of large numbers, to the assertion that most extreme values come from only a small fraction of the clusters – the fat ones. To phrase this assertion in formal terms, if $v \geq 0$ and $\delta > 0$, we set

$$F_\delta(v) := \Big\{ (u, \mathcal{C}) \in \mathbb{R} \times \mathbb{M} : \; u \geq v \,, \; \mathcal{C} \text{ is } (v - u, \delta)\text{-fat} \Big\} . \tag{1.15}$$

The cluster $\mathcal{C}$ in the pair $(u, \mathcal{C}) \in F_\delta(v)$ is fat precisely for the height that is relevant for its contribution to $\mathcal{E}|_{[v, \infty)}$. Given a Borel set $B \subseteq \mathbb{R} \times \mathbb{M}$ we also define in analog to (1.11),

$$\mathcal{E}_t^*(\,\cdot\,; B) := \sum_{(u, \mathcal{C}) \in \widehat{\mathcal{E}}_t} \delta_u 1_{\{(u, \mathcal{C}) \in B\}} \quad \text{and} \quad \mathcal{E}^*(\,\cdot\,; B) := \sum_{(u, \mathcal{C}) \in \widehat{\mathcal{E}}} \delta_u 1_{\{(u, \mathcal{C}) \in B\}} . \tag{1.16}$$

Then,

Theorem 1.2. *For all $\epsilon > 0$ there exist $\delta > 0$, such that for all $\alpha \in (0, 1)$ with probability tending to 1 as $v \to \infty$,*

$$\frac{\mathcal{E}\big([-v, \infty); \; [-\alpha v, \infty) \times \mathbb{M}, \; F_\delta(-v)\big)}{\mathcal{E}\big([-v, \infty); \; [-\alpha v, \infty) \times \mathbb{M}\big)} \geq 1 - \epsilon \,. \tag{1.17}$$

Moreover, there exists $C > 0$ such that for all $\delta > 0$ and $\alpha \in (0, 1)$, with probability tending to 1 as $v \to \infty$,

$$\frac{\mathcal{E}^*\big([-\alpha v, \infty); \; F_\delta(-v)\big)}{\mathcal{E}^*\big([-\alpha v, \infty)\big)} < \frac{C}{\delta(1 - \alpha)} \, v^{-1} \,. \tag{1.18}$$

Both statements hold in the limit when $t \to \infty$ followed by $v \to \infty$ if we replace $\mathcal{E}$ and $\mathcal{E}^*$ by $\mathcal{E}_t$ and $\mathcal{E}_t^*$ respectively.

As a corollary we get,

Corollary 1.3. *Let $\epsilon > 0$ be arbitrarily small. Then for all v large enough, we may find $G_\epsilon(-v) \subseteq \mathbb{R} \times \mathbb{M}$, such that with probability at least $1 - \epsilon$,*

$$\frac{\mathcal{E}\big([-v, \infty); \; G_\epsilon(-v)\big)}{\mathcal{E}\big([-v, \infty)\big)} \geq 1 - \epsilon \quad \text{and} \quad \frac{\mathcal{E}^*\big([-v, \infty); \; G_\epsilon(-v)\big)}{\mathcal{E}^*\big([-v, \infty)\big)} < \epsilon \tag{1.19}$$

The same holds for all t large enough (depending on v and ϵ), if we replace $\mathcal{E}$ and $\mathcal{E}^*$ by $\mathcal{E}_t$ and $\mathcal{E}_t^*$ respectively.

1.3 Proof idea and heuristic picture

The proof of Theorem 1.2 is a rather standard application of Chebyshev's inequality, using Proposition 1.1 (along with a second moment bound on $\mathcal{C}([-v, 0])$ from [6]) and the explicit (conditional) Poissonian structure of $\widehat{\mathcal{E}}$. We therefore omit further explanations of this proof and the straightforward one for Corollary 1.3, focusing instead on the argument for Proposition 1.1. As in [6], the key ingredient is a handle on the cluster distribution ν. This was carried out in Section 3 of [6] and is presented in this manuscript again in Section 2. Let us therefore first briefly describe this derivation and recall how it was used in [6] to show (1.9). The reader is referred to [6] for more details.

Thanks to (1.5) and the product structure of the intensity measure in the definition of $\widehat{\mathcal{E}}$, the law ν can be obtained as the limiting distribution of the cluster around a
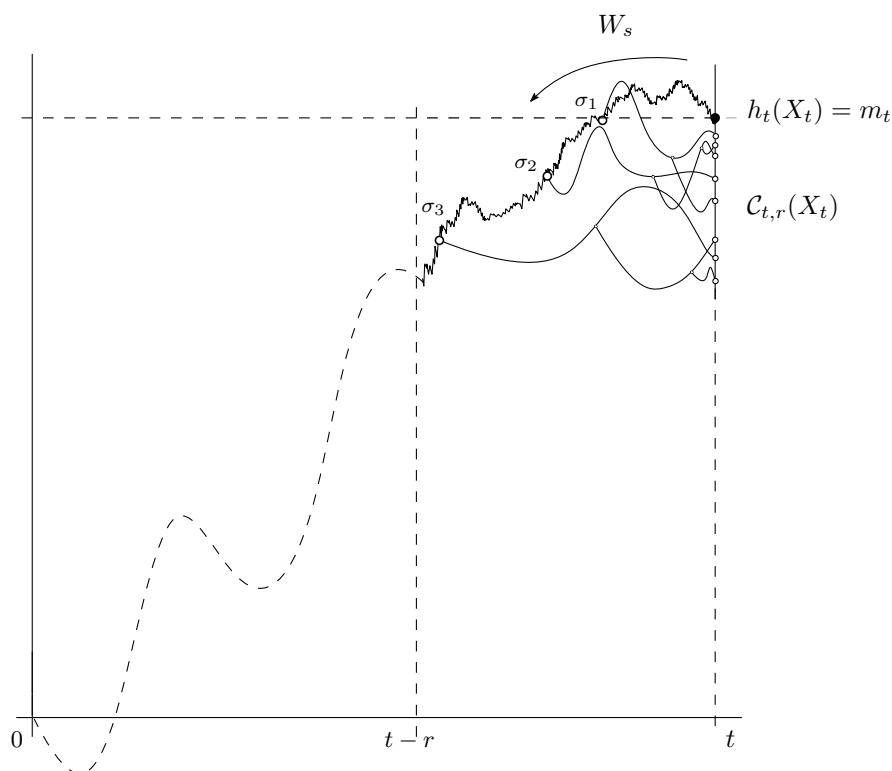
Figure 1: The cluster $\mathcal{C}_{t,r}(X_t)$ around the spine X_t, conditioned to be the maximum and at height m_t. Before conditioning, the process W_s is a Brownian bridge from $(0, m_t)$ to $(t, 0)$ and $\sigma_1, \sigma_2, \dots$ are the branching times.

uniformly chosen particle X_t in L_t, conditioned to be the global maximum at time t and having height, say, m_t. Tracing the trajectory of this distinguished particle backwards in time and accounting for the random genealogical structure leading to size-biasing (see Subsection 2.1) can be described as follows. One sees a (distinguished) particle, which is also called the spine particle, performing a standard Brownian motion $W = (W_s)_{s \geq 0}$ from m_t at time 0 to 0 at time t and giving birth at random Poissonian times (at an accelerated rate 2, see Subsection 2.1) to independent standard branching Brownian motions, which then evolve back to time 0. The entire particle population at time 0 is then conditioned not exceed m_t. The cluster distribution at genealogical distance r around X_t is therefore determined by the relative heights of particles of those BBMs which branched off before time r (see Figure 1).

Formally, denoting by $0 \leq \sigma_1 < \sigma_2 < \dots$ the points of a Poisson point process $\mathcal{N}$ on $\mathbb{R}_+$ with rate 2 and letting $H = (h_t^s(x) : t \geq 0, \; x \in L_t^s)_{s \geq 0}$ be a collection of independent branching Brownian motions (with $W, \mathcal{N}$ and H independent), the cluster distribution ν can then be written as the weak limit (Lemma 2.2):

$$\nu(\cdot) = \lim_{t \to \infty} \mathbb{P}\Big(\mathcal{C}_{t,r_t}(X_t) \in \cdot \,\Big|\, h_t^* = h_t(X_t) = m_t\Big)$$

$$= \lim_{t \to \infty} \mathbb{P}\Big(\sum_{\sigma_k \leq r_t} \sum_{x \in L_{\sigma_k}^{\sigma_k}} \delta_{h_{\sigma_k}^{\sigma_k}(x) + W_{\sigma_k} - m_t} \in \cdot \,\Big|\, \max_{\sigma_k \in [0,t]} \big(W_{\sigma_k} + h_{\sigma_k}^{\sigma_k *}\big) \leq m_t, \, W_0 = m_t, \, W_t = 0 \Big)$$

$$\tag{1.20}$$

where r_t is as in (1.4) and h_t^* is as above (1.1). Tilting by $s \mapsto -m_t(1 - s/t)$ and setting

$\widehat{W}_{t,s} := W_s - \gamma_{t,s}$, with $\gamma_{t,s} := m_t(s/t) - m_s = 3/(2\sqrt{2})(\log^+ s - \frac{s}{t}\log^+ t)$, we obtain,

$$\nu(\cdot) = \lim_{t\to\infty} \mathbb{P}\Big(\sum_{\sigma_k \leq r_t} \mathcal{E}_{\sigma_k}^{\sigma_k}\big(\cdot - \widehat{W}_{t,\sigma_k}\big) \in \cdot \Big| \max_{k:\sigma_k \in [0,t]} \big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}\big) \leq 0, \widehat{W}_{t,0} = \widehat{W}_{t,t} = 0\Big),$$

(1.21)

where $\mathcal{E}_t^s$ is the extremal process associated with h_t^s and $\widehat{h}_t^s = h_t^s - m_t$. In this paper we refer to the triplet $(\widehat{W}, \mathcal{N}, H)$ as a decorated random-walk-like (DRW) process (see Subsection 2.2).

Now let $\mathcal{C} \sim \nu$ and pick $v > 0$. Provided that we can exchange limit and integration, it follows form (1.21) and Palm calculus that $\mathbb{E}\mathcal{C}([-v, 0])$ is equal to the limit as $t \to \infty$ of

$$\frac{\int_0^{r_t} 2\mathrm{d}s \int \mathbb{E}\Big(\mathcal{E}_s^s\big([-v,0]-z\big); z+\widehat{h}_s^{s*} \leq 0\Big)\mathbb{P}\Big(\max_{\sigma_k \leq t}\big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}\big) \leq 0, \widehat{W}_{t,s} \in \mathrm{d}z\Big|\widehat{W}_{t,0} = \widehat{W}_{t,t} = 0\Big)}{\mathbb{P}\Big(\max_{k:\,\sigma_k \in [0,t]}\big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}\big) \leq 0\Big|\widehat{W}_{t,0} = \widehat{W}_{t,t} = 0\Big)}$$

(1.22)

where the inner integral is over $z = O(1)$ and is the result of the total probability formula, after conditioning on $\{\widehat{W}_{t,s} = z\}$.

The left most term in the integrand is the first moment of the size of the (global) extreme level set of h_s^s, subject to a truncation event restricting the height of its global maximum. This was estimated in [6] (Lemma 4.2) as

$$\mathbb{E}\big(\mathcal{E}_t([-v,0] - z); \widehat{h}_t^* \leq -z\big) \approx Czv\mathrm{e}^{\sqrt{2}v - v^2/(2t)}.$$

(1.23)

(In this heuristics subsection, we use $\approx$ for the loose notion of "roughly equal to".) The remaining terms are of the form

$$\mathbb{P}\Big(\max_{k:\sigma_k \in [s_1,s_2]}\big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}\big) \leq 0 \Big| \widehat{W}_{t,s_1} = x, \widehat{W}_{t,s_2} = y\Big),$$

(1.24)

namely the probability that a DRW process stays negative on $[s_1, s_2]$.

For standard Brownian motion, the well known reflection principle gives

$$\mathbb{P}\Big(\max_{s \in [s_1,s_2]} W_s \leq 0 \Big| W_{s_1} = x, W_{s_2} = y\Big) \sim \frac{2xy}{s_2 - s_1} \quad \text{as } s_2 - s_1 \to \infty.$$

(1.25)

It was shown in [6] (Subsection 2.1; see also Lemma 2.5 here), that similar estimates hold for (1.24) as well. (This is because the drift function $\gamma_{t,s}$ is bounded by $1 + \log^+(s \wedge (t-s))$, the random decorations $(h_s^{s*} : s \geq 0)$ have (at least) exponential tails and the random sampling times $(\sigma_k : k \geq 1)$ arrive at a Poissonian rate.) We can therefore estimate the probability in the denominator in (1.22) by Ct^{-1} and the probability in the numerator by

$$Cz^2(s(t-s))^{-1}\mathbb{P}(\widehat{W}_{t,s} \in \mathrm{d}z|\widehat{W}_{t,0} = \widehat{W}_{t,t} = 0).$$

(1.26)

Using also (1.23) and performing the integration over $z = O(1)$ in (1.22), we obtain,

$$\mathbb{E}\mathcal{C}([-v,0]) \approx Cv\mathrm{e}^{\sqrt{2}v}\int_{s=0}^\infty s^{-3/2}\mathrm{e}^{-v^2/(2s)}\mathrm{d}s = C\mathrm{e}^{\sqrt{2}v}\int_{r=0}^\infty r^{-3/2}\mathrm{e}^{-1/(2r)}\mathrm{d}r = C_\star\mathrm{e}^{\sqrt{2}v}.$$

(1.27)

This explains (1.9) and is precisely the derivation in [6] (Proposition 1.5).

Now, looking more carefully at (1.27), we see that for large v, "most" of the contribution to the first integral comes from $s \in [\eta v^2, \eta^{-1}v^2]$. Rolling back the derivation all the way to (1.22), the latter is equivalent to saying that most of the contribution to $\mathbb{E}\mathcal{C}([-v,0])$ comes from trajectories in which $\max_{s\in[\eta v^2, \eta^{-1}v^2]}|\widehat{W}_{t,s}| = O(1)$. Such trajectories have very small probability:

$$\mathbb{P}\Big(\max_{s/v^2\in[\eta,\frac{1}{\eta}]}|\widehat{W}_{t,s}| = O(1) \Big| \max_{\sigma_k \leq t}\big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}\big) \leq 0, \widehat{W}_{t,0} = \widehat{W}_{t,t} = 0\Big) \approx \int_{\eta v^2}^{v^2/\eta} Cs^{-\frac{3}{2}}\mathrm{d}s = \frac{C'}{v}$$
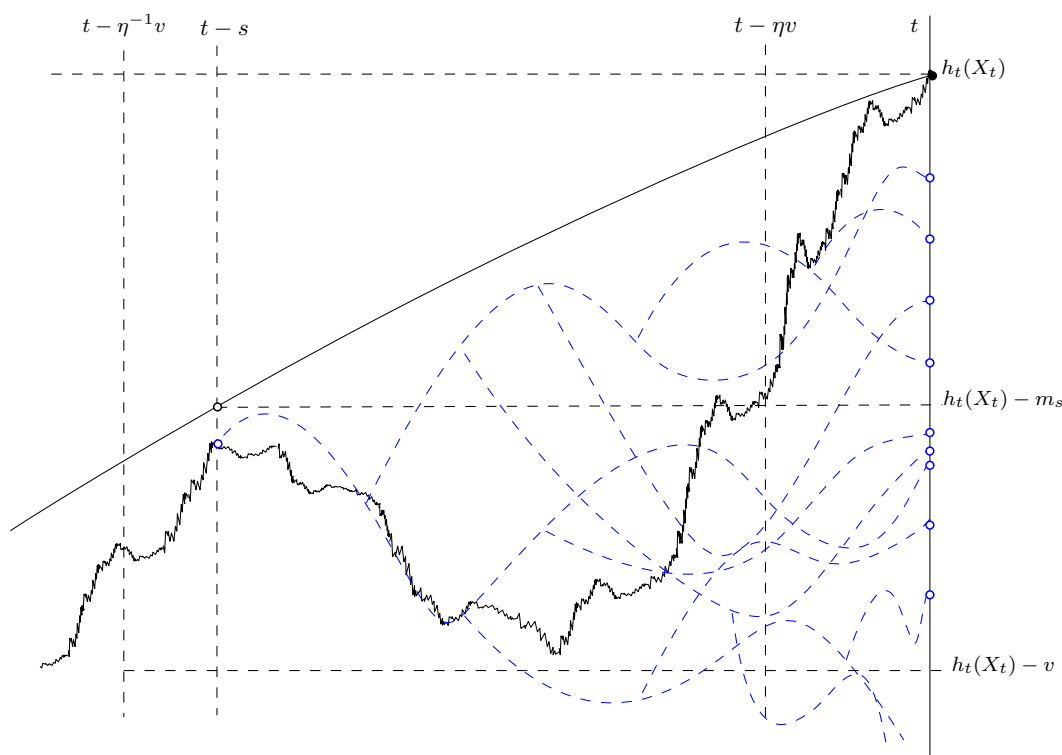
Figure 2: The contribution to the cluster level set $\mathcal{C}_{t,r_t}(X_t)|_{(-v,0]}$ on an atypical event when the cluster maximum X_t ascends from $h_t(X_t) - m_s + O(1)$ at time $t - s$ to $h_t(X_t)$ at time t, where $s \in [\eta v^2, \eta^{-1} v^2]$.

But when they occur, it follows from (1.23) with $s \in [\eta v^2, \eta^{-1} v^2]$ (and an additional concentration argument) that $\mathcal{C}([-v, 0]) \approx C\mathcal{E}_s^s([-v, 0] - z) 1_{\{\widehat{h}_s^{s*} \leq -z\}} \approx C'v\mathrm{e}^{\sqrt{2}v}$ (see Figure 2).

Reversing further the two reduction steps (1.20) and (1.21), we can rephrase the last statement as follows (see Lemma 3.1 and Lemma 3.2): Having $C_\star \mathrm{e}^{\sqrt{2}v}$ mean number of particles above $-v$ (for large v) in a cluster is the result of an atypical $O(v^{-1})$ probability event, in which the number of such particles is of order $v\mathrm{e}^{\sqrt{2}v}$. Such a cluster is realized at a large time t when its local maximum X_t has ascended (atypically slowly) to $h_t(X_t)$ from $h_t(X_t) - m_s + O(1)$ at time $t - s$, where $s \in [\eta v^2, \eta^{-1} v^2]$.

2   A handle on the cluster distribution

As mentioned in the introduction, understanding the cluster distribution ν is key to proving the statements in the manuscript. As in [6], a handle on this distribution is obtained by first identifying ν with the law of the cluster around a distinguished spine particle, conditioned to be the global maximum of the process. Then, by tracing the trajectory of the spine particle backwards in time, the events involved can be recast in terms of a decorated random-walk-like process conditioned to stay negative. The two reduction steps are summarized in the next two subsections. Estimates for such random walks are given in the succeeding subsection. Finally, some upper bounds from [6], derived using the statements in the first three subsections, are stated in the last subsection. For all proofs see [6].

2.1 Reduction to the cluster of the spine, conditioned to be the maximum

We begin by recalling the useful technique of spinal decomposition (c.f., [7]). The (one) spine branching Brownian motion (SBBM) is defined as the original process (h, T, d), only that at any given time one of the particles is designated as the *spine particle*. Particles which are not the spine branch and diffuse exactly as before. The spine particle also diffuses as before, but branches (by two) at rate 2 not 1. When the spine branches, one of its children, chosen uniformly at random, is designated the new spine. We shall use the same notation (T, h, d) for the SBBM process and distinguish this process from the original one by renaming the underlying probability measure to $\widetilde{\mathbb{P}}$ (with $\widetilde{\mathbb{E}}$ the corresponding expectation). The identity of the spine at time $t \geq 0$ will be recorded via the random variable $X_t \in L_t$. The genealogical line of decent of the spine particle, namely the function $t \mapsto X_t$, will be referred to as the spine of the process.

The following is known as the Many-To-One Lemma. To avoid integrability issues, we state it for bounded functions. Recall that $\mathcal{F}_t$ is the sigma-algebra generated by h_t, T_t and d_t (but not X_t).

Lemma 2.1 (Many-To-One Lemma). *Let $F = (F(x) : x \in L_t)$ be a bounded $\mathcal{F}_t$-measurable real-valued random function on L_t. Then,*

$$\mathbb{E}\Big(\sum_{x \in L_t} F(x) \Big) = \mathrm{e}^t \, \widetilde{\mathbb{E}} F(X_t) \,. \tag{2.1}$$

We write $\widehat{h}_t = h_t - m_t$ and $\widehat{h}_t^* = \max_{x \in L_t} \widehat{h}_t(x)$. Recalling (1.3), let now $\mathcal{C}_{t,r}^* := \mathcal{C}_{t,r_t}(X_t)$ denote the cluster around the spine. Thanks to Lemma 2.1 and the convergence in (1.5), it is then not difficult to show,

Lemma 2.2 (Lemma 5.1 in [6]). *Let $\mathcal{C} \sim \nu$ be distributed according to the cluster law. Then for any ν-continuity set $B \subseteq \mathbb{M}$,*

$$\mathbb{P}(\mathcal{C} \in B) = \lim_{t \to \infty} \widetilde{\mathbb{P}}\big(\mathcal{C}_{t,r_t}^* \in B \,\big|\, \widehat{h}_t(X_t) = \widehat{h}_t^* = 0 \big) \,. \tag{2.2}$$

2.2 Reduction to a decorated random-walk conditioned to stay negative

Next, we introduce the decorated random-walk-like process using-which, probabilities such as the ones on the right hand side of (2.2), can be handled. Let $W = (W_s : s \geq 0)$ be a standard Brownian motion, whose initial position we leave free to be determined according to the conditional statements we make. For $0 \leq s \leq t$, we fix

$$\gamma_{t,s} := \tfrac{3}{2\sqrt{2}} \big(\log^+ s - \tfrac{s}{t} \log^+ t \big) \quad \text{and} \quad \widehat{W}_{t,s} := W_s - \gamma_{t,s} \,. \tag{2.3}$$

Let us also define the collection $H = \big(h^s : s \geq 0 \big)$ of independent copies of h, that we will assume to be independent of W as well. Finally, let $\mathcal{N}$ be a Poisson point process with intensity $2\mathrm{d}x$ on $\mathbb{R}_+$, independent of H and W and denote by $\sigma_1 < \sigma_2 < \dots$ its ordered atoms. The triplet $(\widehat{W}, \mathcal{N}, H)$ forms what we shall call a *decorated random-walk-like process* (DRW). The underlying probability measure will still be denoted by $\mathbb{P}$ and the corresponding expectation by $\mathbb{E}$.

The following reduction statements appear in Subsection 3.1 of [6]. We refer the reader to that reference for their straightforward proofs. Recall that $\mathrm{B}_r(x)$ is the ball of radius r around x in the genealogical distance d. For $A \subseteq L_t$ set also $\widehat{h}_t^*(A)$ for $\max_{x \in A} \widehat{h}_t(x)$. The first lemma concerns the event that particles in a genealogical neighborhood of the spine (including the spine itself) stay below a given height.

Lemma 2.3. *For all $0 \leq r \leq t$ and $u, w \in \mathbb{R}$,*

$$\widetilde{\mathbb{P}}\Big(\widehat{h}_t^*\big(\mathrm{B}_r^c(X_t)\big) \leq u \,\Big|\, \widehat{h}_t(X_t) = w \Big) = \mathbb{P}\Big(\max_{\sigma_k \in [r,t]} \big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k*} \big) \leq 0 \Big| \widehat{W}_{t,r} = w - u, \, \widehat{W}_{t,t} = -u \Big)$$

In particular for all $t \geq 0$ and $v, w \in \mathbb{R}$,

$$\widetilde{\mathbb{P}}\big(\widehat{h}_t^* \leq u \,\big|\, \widehat{h}_t(X_t) = w\big) = \mathbb{P}\Big(\max_{k:\sigma_k \in [0,t]} \big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}\big) \leq 0 \,\Big|\, \widehat{W}_{t,0} = w - u, \widehat{W}_{t,t} = -u \Big).$$

In a similar way, we can express the distribution of the cluster around the spine particle, given that it reaches height m_t. For what follows $\mathcal{E}_t^s$ denotes the extremal process of h_t^s, defined as $\mathcal{E}_t$ in (1.6) only with respect to h_t^s in place of h_t. Then,

Lemma 2.4. *For all $0 \leq r \leq t$,*

$$\widetilde{\mathbb{P}}\Big(\big(\mathcal{C}_{t,r}(X_t), (h_{t-s}(X_{t-s}) - m_t)_{s \leq r}\big) \in \cdot \,\Big|\, \widehat{h}_t^* = \widehat{h}_t(X_t) = 0 \Big)$$
$$= \mathbb{P}\Big(\Big(\sum_{\sigma_k \leq r} \mathcal{E}_{\sigma_k}^{\sigma_k}\big(\cdot - \widehat{W}_{t,\sigma_k}\big), (\widehat{W}_{t,s} - m_s)_{s \leq r}\Big) \in \cdot \,\Big|\, \max_{k:\sigma_k \in [0,t]} \big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}\big) \leq 0, \widehat{W}_{t,0} = \widehat{W}_{t,t} = 0 \Big)$$

2.3 Estimates for a decorated random-walk conditioned to stay negative

Having reduced the analysis of the cluster distribution to the study of the DRW conditioned to stay negative, we need estimates for probabilities involving the latter. Such probabilities are treated in general in [5], which is the supplement material to [6]. Specialized to our setup here, they take the form of the two lemmas below. The first lemma was already included in [6] (see third part of Lemma 3.4).

Lemma 2.5 (Lemma 3.4 in [6])**.** *There exists non-increasing functions $g : \mathbb{R} \to (0, \infty)$ and $f : \mathbb{R} \to (0, \infty)$ such that*

$$\mathbb{P}\Big(\max_{k:\sigma_k \in [0,t]} \big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}\big) \leq 0 \,\Big|\, \widehat{W}_{t,0} = v\,, \ \widehat{W}_{t,t} = w \Big) \sim 2\frac{f(v)g(w)}{t}, \tag{2.4}$$

as $t \to \infty$ uniformly in v, w satisfying $v, w < 1/\epsilon$ and $(v^- + 1)(w^- + 1) \leq t^{1-\epsilon}$ for any $\epsilon > 0$.

Next we have the following sharp entropic-repulsion result. This was not needed in [6] and hence left out from that work. Nevertheless, it is an immediate consequence of Proposition 1.5 from the supplement material [5] of [6].

Lemma 2.6. *For all $M \geq 1$,*

$$\sup_{s \geq 1} \limsup_{t \to \infty} \sqrt{s}\, \mathbb{P}\Big(\max_{u \in [s, t-s]} \widehat{W}_{t,u} \geq -M \,\Big|\, \max_{\sigma_k \in [0,t]} \big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}\big) \leq 0\,, \ \widehat{W}_{t,0} = \widehat{W}_{t,t} = 0 \Big) < \infty\,.$$

Proof. Recalling the definition of $\widehat{W}_{t,u}$ the desired statement is precisely that in Proposition 1.5 in [5] with $x = y = 0$. We just need to verify that the assumptions in that proposition hold. Assumptions (1.1), (1.2) and (1.3) from [5] (with $\lambda = 2$ and sufficiently small $\delta > 0$) were already verified in the proof of Lemma 2.5 in [6]. It remains therefore to check that

$$\gamma_{t,u} - \frac{u}{r}\gamma_{t,r} \leq \delta^{-1}\big(1 + (\wedge^r(u))^{1/2-\delta}\big)\,, \quad \gamma_{t,u'} - \frac{t-u'}{t-r}\gamma_{t,r} \leq \delta^{-1}\big(1 + (\wedge^{t-r}(u'-r))^{1/2-\delta}\big)\,, \tag{2.5}$$

for all $0 < u < r < u' < t$ and some $\delta > 0$. To this end, a simple computation gives

$$\gamma_{t,u} - \gamma_{t,r}(u/r) = \tfrac{3}{2\sqrt{2}}\Big(\log^+ r - \tfrac{r}{u}\log^+ u \Big)$$
$$\gamma_{t,u'} - \tfrac{t-u'}{t-r}\gamma_{t,r} = \tfrac{3}{2\sqrt{2}}\Big(\log^+ u - \big(\tfrac{t-u'}{t-r}\log^+ r + \tfrac{u'-r}{t-r}\log^+ t\big) \Big)\,. \tag{2.6}$$

Then the upper bound of Lemma 3.3 in [6], once with $(u, r, 0)$ and once with $(t, u' - r, r)$ in place of (t, s, r), gives (2.5) with any $\delta \leq 2\sqrt{\log 2}/3$. $\qquad \square$

2.4 Moment bounds for the number of cluster points

Next we state several moment bounds for the number of cluster points. These were derived in [6] using the statements in the last three subsections. We start with a second moment bound for $\mathcal{C}([-v,0])$. It is part of Proposition 1.5 in [6].

Lemma 2.7. *Let* $\mathcal{C} \sim \nu$. *Then there exists* $C > 0$ *such that for all* $v \geq 0$,

$$\mathbb{E}\big[\mathcal{C}([-v,0])^2\big] \leq C(v+1)\mathrm{e}^{2\sqrt{2}v}. \tag{2.7}$$

Next, if $v \geq 0$, $0 \leq s \leq t$ and $M \geq 0$, we define:

$$J_{t,v}^{\geq M}(s) := \mathcal{E}_s^s\big([-v,0] - \widehat{W}_{t,s}\big) \times 1_{\{|\widehat{W}_{t,s}| \geq M\}} 1_{\{\widehat{h}_s^{s*} \leq -\widehat{W}_{t,s}\}} 1_{\{\max_{k:\sigma_k \in [0,t]} (\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k *}) \leq 0\}}$$

$$j_{t,v}^{\geq M}(s) := \mathbb{E}\Big(J_{t,v}^{\geq M}(s)\,\Big|\,\widehat{W}_{t,0} = \widehat{W}_{t,t} = 0\Big), \tag{2.8}$$

and set $J_{t,v}(s) \equiv J_{t,v}^{\geq 0}(s)$ and $j_{t,v}(s) \equiv j_{t,v}^{\geq 0}(s)$. The following is part of Lemma 5.5 in [6].

Lemma 2.8. *There exists* $C, C' > 0$ *such that for all* $t \geq 0$, $0 \leq s \leq t/2$, $v \geq 0$ *and* $M \geq 0$,

$$j_{t,v}^{\geq M}(s) \leq C\frac{\mathrm{e}^{\sqrt{2}v}(v+1)}{t(s+1)\sqrt{s}} \times \mathrm{e}^{-C'M}\left(\mathrm{e}^{-\frac{v^2}{16s}} + \mathrm{e}^{-\frac{v}{2}}\right). \tag{2.9}$$

3 Proof of main results

3.1 Proof of Proposition 1.1

The proof of Proposition 1.1 will be based on the following two lemmas.

Lemma 3.1. *The following quantity*

$$\mathrm{e}^{-\sqrt{2}v}\widetilde{\mathbb{E}}\Big(\mathcal{C}_{t,r_t}^*\big([-v,0]\big);\ \max_{s \in [\eta v^2, \eta^{-1} v^2]} \big(h_{t-s}(X_{t-s}) - m_t + m_s\big) \leq -M\,\Big|\,\widehat{h}_t^* = \widehat{h}_t(X_t) = 0\Big) \tag{3.1}$$

tends to 0 *when* $t \to \infty$, *followed by* $v \to \infty$, *then* $M \to \infty$ *and finally* $\eta \to 0$.

Lemma 3.2. *For all* $\eta > 0$ *and* $M > 0$, *there exists* $\delta > 0$ *such that,*

$$\widetilde{\mathbb{P}}\Big(\max_{s \in [\eta v^2, \eta^{-1} v^2]} \big(h_{t-s}(X_{t-s}) - m_t + m_s\big) > -M\,\Big|\,\widehat{h}_t^* = \widehat{h}_t(X_t) = 0\Big) \leq \frac{\delta^{-1}}{v}.$$

for all v *large enough and then* t *large enough.*

Let us first prove Proposition 1.1.

Proof of Proposition 1.1. The second statement in the proposition follows trivially from (1.9) and Markov's inequality. It therefore remains to show (1.13). Given $t \geq 0$, $v \geq 0$, $M > 0$ and $\eta > 0$, define the event

$$\mathcal{B} := \Big\{\max_{s \in [\eta v^2, \eta^{-1} v^2]} \big(h_{t-s}(X_{t-s}) - m_t + m_s\big) > -M\Big\}. \tag{3.2}$$

Thanks to Lemma 3.1 and Lemma 3.2, for any $\epsilon > 0$ there exist $\eta > 0$, $M > 0$ and $\delta' > 0$ such that for all v and then t large enough,

$$\widetilde{\mathbb{E}}\Big(\mathcal{C}_{t,r_t}^*\big([-v,0]\big);\ \mathcal{B}^c\,\Big|\,\widehat{h}_t^* = \widehat{h}_t(X_t) = 0\Big) \leq \frac{\epsilon}{2}\,\mathrm{e}^{\sqrt{2}v}\ ,\qquad \widetilde{\mathbb{P}}\Big(\mathcal{B}\,\Big|\,\widehat{h}_t^* = \widehat{h}_t(X_t) = 0\Big) \leq \frac{1}{2\delta' v}. \tag{3.3}$$

Next let $\delta := \epsilon\delta'$ and define also the event $\mathcal{A} := \big\{\mathcal{C}_{t,r_t}^*\big([-v,0]\big) > \delta v\mathrm{e}^{\sqrt{2}v}\big\}$. Then,

$$\widetilde{\mathbb{E}}\Big(\mathcal{C}_{t,r_t}^*\big([-v,0]\big);\ \mathcal{A}^c\,\Big|\,\widehat{h}_t^* = \widehat{h}_t(X_t) = 0\Big)$$

$$\leq \widetilde{\mathbb{E}}\Big(\mathcal{C}_{t,r_t}^*\big([-v,0]\big);\ \mathcal{B}^c\,\Big|\,\widehat{h}_t^* = \widehat{h}_t(X_t) = 0\Big) + \widetilde{\mathbb{E}}\Big(\mathcal{C}_{t,r_t}^*\big([-v,0]\big);\ \mathcal{B} \cap \mathcal{A}^c\,\Big|\,\widehat{h}_t^* = \widehat{h}_t(X_t) = 0\Big)$$

$$\leq \frac{\epsilon\mathrm{e}^{\sqrt{2}v}}{2} + \frac{1}{2\delta' v}\delta'\epsilon v\mathrm{e}^{\sqrt{2}v} \leq \epsilon\mathrm{e}^{\sqrt{2}v}.$$

Now take $t \to \infty$ in the last display. Then by Lemma 2.2 and the bounded convergence theorem, (1.13) follows for all large enough v such that ν-almost-surely the point $-v$ is not charged by $\mathcal{C}$. Removing this stochastic continuity restriction requires a standard argument, the kind of which was used in many of the proofs in [6] (e.g., the proof of Proposition 1.5). We therefore omit further details. $\qquad \square$

Proof of Lemma 3.1. Using Lemma 2.3 and Lemma 2.4 we can write the product of the expectation in (3.1) with $\widetilde{\mathbb{P}}\big(\widehat{h}_t^* \leq 0 \,\big|\, \widehat{h}_t(X_t) = 0\big)$ as

$$\mathbb{E}\Big(\int_{s=0}^{r_t} J_{t,v}(s)\mathcal{N}(\mathrm{d}s) \ ; \ \max_{s \in [\eta v^2, \eta^{-1}v^2]} \widehat{W}_{t,s} \leq -M \,\Big|\, \widehat{W}_{t,0} = \widehat{W}_{t,t} = 0\Big)$$

$$\leq \mathbb{E}\Big(\int_{s=0}^{r_t} \big(J_{t,v}(s)1_{[\eta v^2, \eta^{-1}v^2]^c}(s) + J_{t,v}^{\geq M}(s)1_{[\eta v^2, \eta^{-1}v^2]}\big)\mathcal{N}(\mathrm{d}s) \,\Big|\, \widehat{W}_{t,0} = \widehat{W}_{t,t} = 0\Big), \quad (3.4)$$

where $J_{t,v}^{\geq M}(s)$ is defined in (2.8). Then, by the Palm-Campbell Theorem, the right hand side above is equal to $2\int_{s=0}^{r_t} \big(j_{t,v}(s)1_{[\eta v^2, \eta^{-1}v^2]^c} + j_{t,v}^{\geq M}(s)1_{[\eta v^2, \eta^{-1}v^2]}\big)\mathrm{d}s$. Using Lemma 2.8 we can upper bound the above integral by $Ct^{-1}\mathrm{e}^{\sqrt{2}v}(v+1)$ times

$$\int_{s=0}^{\infty} \frac{\mathrm{e}^{-v^2/(16s)} + \mathrm{e}^{-v/2}}{\sqrt{s}(s+1)}\Big(1_{\{s \in [\eta v^2, \eta^{-1}v^2]^c\}} + \mathrm{e}^{-C'M}1_{\{s \in [\eta v^2, \eta^{-1}v^2]\}}\Big)\mathrm{d}s$$

$$\leq \int_0^{\infty} \frac{\mathrm{e}^{-\frac{v}{2}}}{\sqrt{s}(s+1)}\mathrm{d}s + \int_0^{\eta v^2} \frac{\mathrm{e}^{-v/(16s)}}{s^{3/2}}\mathrm{d}s + \mathrm{e}^{-C'M}\int_{\eta v^2}^{\infty} s^{-\frac{3}{2}}\mathrm{d}s + \int_{\eta^{-1}v^2}^{\infty} s^{-\frac{3}{2}}\mathrm{d}s \quad (3.5)$$

$$\leq C\mathrm{e}^{-\frac{v}{2}} + v^{-1} + \mathrm{e}^{-C'M}/(v\sqrt{\eta}) + \sqrt{\eta}/v.$$

Altogether (3.1) is at most $Ct^{-1}\mathrm{e}^{\sqrt{2}v}\big(\mathrm{e}^{-v/4} + \mathrm{e}^{-C'M}\eta^{-\frac{1}{2}} + \eta^{\frac{1}{2}}\big)$. Since $\widetilde{\mathbb{P}}\big(\widehat{h}_t^* \leq 0 \,\big|\, \widehat{h}_t(X_t) = 0\big) \sim Ct^{-1}$ as $t \to \infty$, thanks to the second part of Lemma 2.3 with $u = w = 0$ and Lemma 2.5, the desired statement follows. $\qquad \square$

Proof of Lemma 3.2. Thanks to Lemma 2.4 (ignoring the distribution of $\mathcal{C}_{t,r}(X_t)$), the conditional probability in the statement of the lemma is equal to

$$\mathbb{P}\Big(\max_{s \in [\eta v^2, \eta^{-1}v^2]} \widehat{W}_{t,s} > -M \,\Big|\, \max_{k:\sigma_k \in [0,t]} \big(\widehat{W}_{t,\sigma_k} + \widehat{h}_{\sigma_k}^{\sigma_k*}\big) \leq 0 \,, \, \widehat{W}_{t,0} = \widehat{W}_{t,t} = 0\Big).$$

Invoking Lemma 2.6, then gives the desired upper bound with some $\delta > 0$, depending on η, M and all v and then t large enough. $\qquad \square$

3.2 Proofs of Theorem 1.2 and Corollary 1.3

The proofs of Theorem 1.2 will be based on the following lemma.

Lemma 3.3. *For any $\epsilon > 0$, there exists $\delta > 0$ such that for all $\alpha \in (0,1)$,*

$$\lim_{v \to \infty} \mathbb{P}\left(\frac{\mathcal{E}\big([-v, \infty); \, [-\alpha v, \infty) \times \mathbb{M}, \, F_\delta(-v)^c\big)}{C_\star Z \alpha v \mathrm{e}^{\sqrt{2}v}} \leq \epsilon \right) = 1. \quad (3.6)$$

Moreover, there exists $C > 0$ such that for any $\delta > 0$ and $\alpha \in (0,1)$,

$$\lim_{v \to \infty} \mathbb{P}\left(\frac{\widehat{\mathcal{E}}\big([-\alpha v, \infty) \times \mathbb{M} \cap F_\delta(-v)\big)}{Z\mathrm{e}^{\sqrt{2}\alpha v}/\sqrt{2}} < \frac{C}{\delta(1-\alpha)}v^{-1} \right) = 1. \quad (3.7)$$

Proof. Let $\alpha \in (0,1)$. Given $-\infty < -v < w < z \leq \infty$ and $\delta > 0$, define:

$$H_v(w,z) := \sum_{(u,\mathcal{C}) \in \widehat{\mathcal{E}}} \mathcal{C}\big([-v-u, 0]\big) \times 1_{[w,z]}(u), \quad (3.8)$$

$$Q_{v,\delta}(w,z) := \sum_{(u,\mathcal{C}) \in \widehat{\mathcal{E}}} \mathcal{C}\big([-v-u, 0]\big) \times 1_{[w,z)}(u) \times 1_{\{\mathcal{C}([-v-u,0]) \leq \delta(v+u)\mathrm{e}^{\sqrt{2}(v+u)}\}}. \quad (3.9)$$

Clearly $Q_{v,\delta}(w,z) \leq H_v(w,z)$. Therefore, setting $z := \sqrt{\log v}$ and $w := -\alpha v$, we can bound the numerator in (3.6) by $Q_{v,\delta}(w,z) + H_v(z,\infty)$.

Since conditional on Z, the intensity measure governing the law of $\widehat{\mathcal{E}}$ is finite on $[0,\infty) \times \mathbb{M}$ almost surely, we must have $\widehat{\mathcal{E}}([z,\infty) \times \mathbb{M}) = 0$ for all large enough z. This shows that

$$H_v(z,\infty)\big(Zve^{\sqrt{2}v}\big)^{-1} \longrightarrow 0 \quad \text{as } v \to \infty \text{ a.s.} \tag{3.10}$$

At the same time, for any $\epsilon > 0$, thanks to the first part of Proposition 1.1, we may find $\delta > 0$ such that

$$\mathbb{E}\big(Q_{v,\delta}(w,z) \,\big|\, Z\big) = \int_{u=w}^{z} \mathbb{E}\Big(\mathcal{C}([-v-u,0]); \, \mathcal{C}([-v-u,0]) \leq \delta(v+u)e^{\sqrt{2}(v+u)}\Big) Ze^{-\sqrt{2}u}\mathrm{d}u$$

$$\leq \int_{u=w}^{z} \epsilon e^{\sqrt{2}(v+u)} Ze^{-\sqrt{2}u}\mathrm{d}u \leq 2\epsilon Z\alpha v e^{\sqrt{2}v}.$$

In addition, thanks to Lemma 2.7,

$$\mathbb{V}\mathrm{ar}\big(Q_{v,\delta}(w,z)|Z\big) = \int_{w}^{z} \mathbb{E}\Big(\mathcal{C}([-v-u,0]); \, \mathcal{C}([-v-u,0]) > \delta(v+u)e^{\sqrt{2}(v+u)}\Big)^2 Ze^{-\sqrt{2}u}\mathrm{d}u$$

$$\leq \int_{w}^{z} \mathbb{E}\Big(\mathcal{C}([-v-u,0])\Big)^2 Ze^{-\sqrt{2}u}\mathrm{d}u \leq CZve^{2\sqrt{2}v+\sqrt{2}z}.$$

It follows by Chebyshev's inequality that

$$\mathbb{P}\Big(Q_v(w,z) > 4\epsilon Z\alpha v e^{\sqrt{2}v} \,\Big|\, Z\Big) \overset{v\to\infty}{\longrightarrow} 0\,, \tag{3.11}$$

almost surely. By the bounded convergence theorem, the above limit holds also for the unconditional probability. Together with (3.10) and the union bound, this shows that

$$\mathbb{P}\Big(Q_v(w,z) + H_v(z,\infty) > 8\epsilon Z\alpha v e^{\sqrt{2}v}\Big) \longrightarrow 0 \quad \text{as } v \to \infty\,. \tag{3.12}$$

Since the quantity on the left hand side in the probability above dominates the one in the numerator of (3.6), this gives (3.6) with $8\epsilon/C_\star$ in place of ϵ.

Turning to the second statement of the lemma, by the definition of $\widehat{\mathcal{E}}$, for any $\delta > 0$ and $\alpha \in (0,1)$, the law of the first numerator in (3.7), conditional on Z, is Poisson with parameter given by

$$\lambda_{\alpha,\delta}(v) = \int_{u=-\alpha v}^{\infty} \mathbb{P}\Big(\mathcal{C}([-v-u,0]) > \delta(v+u)e^{\sqrt{2}(v+u)}\Big) Ze^{-\sqrt{2}u}\mathrm{d}u\,. \tag{3.13}$$

Therefore by the second part of Proposition 1.1,

$$\lambda_{\alpha,\delta}(v) \leq \int_{u=-\alpha v}^{\infty} \frac{C}{\delta(v+u)} Ze^{-\sqrt{2}u}\mathrm{d}u \leq \frac{C'}{\delta(1-\alpha)} v^{-1} Ze^{\sqrt{2}\alpha v}/\sqrt{2}\,, \tag{3.14}$$

for some $C' > 0$. Then by Chebyshev's inequality, conditional on Z, the probability of the complement of the event in (3.7) with $C := 2C'$ is at most

$$\frac{\lambda_{\alpha,\delta}(v)}{\Big(\frac{2C'}{\delta(1-\alpha)} v^{-1} Ze^{\sqrt{2}\alpha v}/\sqrt{2} - \lambda_{\alpha,\delta}(v)\Big)^2}\,, \tag{3.15}$$

which goes to 0 as $v \to \infty$, for $\mathbb{P}$-almost every Z. The same then also holds for the unconditional probability, thanks again to the bounded convergence theorem. $\qquad\square$

We can now prove Theorem 1.2.

Proof of Theorem 1.2. Given $\epsilon > 0$, we use Lemma 3.3 to find $\delta > 0$ such that for any $\alpha \in (0, 1)$ (3.6) holds. We then use (1.8), (1.10) and (1.12) to claim that also

$$\frac{\mathcal{E}\big([-v, \infty);\ [-\alpha v, \infty) \times \mathbb{M}\big)}{C_\star Z \alpha v \mathrm{e}^{\sqrt{2}v}} > \frac{3}{4} \quad \text{and} \quad \frac{\mathcal{E}^*\big([-\alpha v, \infty)\big)}{Z \mathrm{e}^{\sqrt{2}\alpha v}/\sqrt{2}} > \frac{3}{4} \qquad (3.16)$$

holds with probability tending to 1 as $v \to \infty$.

Dividing both the numerator and denominator of (1.17) by $C_\star Z \alpha v \mathrm{e}^{\sqrt{2}v}$, we now observe that whenever the event in (3.6) and the first event in (3.16) hold, we must also have (1.17) with $4\epsilon/3$ in place of ϵ. At the same time, observing that $\mathcal{E}^*\big([-\alpha v, \infty);\ F_\delta(-v)\big)$ is equal to $\widehat{\mathcal{E}}\big(([-\alpha v, \infty) \times \mathbb{M}) \cap F_\delta(-v)\big)$ which follows by definition, we now divide both the numerator and denominator of (1.18) by $Z \mathrm{e}^{\sqrt{2}\alpha v}/\sqrt{2}$. We then see that whenever the event in (3.7) and the second event in (3.16) hold, we must also have (1.18) with $4C/3$ in place of C. Renaming ϵ and C, and using the union bound, we complete the proof of the theorem for $\mathcal{E}$ and $\mathcal{E}^*$.

To obtain the finite t analogs (with $t \to \infty$), it is sufficient to argue that all random quantities on the left hand sides of (1.17) and (1.18) are the joint weak limits of their respective finite time analogs as $t \to \infty$. This in turn follows from (1.5) and standard arguments, the kind of which were used in many of the proofs in [6] (e.g., the proof of Theorem 1.1). We therefore omit further details. $\qquad\square$

Proof of Corollary 1.3. We shall only show that statement for $\mathcal{E}^*$ and $\mathcal{E}$ as the argument for the case involving $\mathcal{E}_t$ and $\mathcal{E}_t^*$ is almost identical. Given $\epsilon > 0$, we let δ be given by Theorem 1.2 for $\epsilon/2$, choose $\alpha := 1 - \epsilon/3$ and set $G_\epsilon(-v) := \big([-\alpha v, \infty) \times \mathbb{M}\big) \cap F_\delta(-v)$. Then, the first part of Theorem 1.2, together with (1.12) and the union bound show that

$$\mathcal{E}\big([-v, \infty);\ G_\epsilon(-v)\big) \geq (1 - \epsilon/2)\mathcal{E}\big([-v, \infty);\ [-\alpha v, \infty) \times \mathbb{M}\big) \geq (1 - \epsilon)\mathcal{E}\big([-v, \infty)\big), \quad (3.17)$$

with probability at least $1 - \epsilon/2$ whenever v is large enough. At the same time, the second part of Theorem 1.2 shows that

$$\mathcal{E}^*\big([-v, \infty);\ G_\epsilon(-v)\big) \leq \epsilon \, \mathcal{E}^*\big([-\alpha v, \infty)\big) < \epsilon \, \mathcal{E}^*\big([-v, \infty)\big), \qquad (3.18)$$

with probability at least $1 - \epsilon/2$, again whenever v is large enough. A final application of the union bound then completes the proof. $\qquad\square$

References

[1] E. Aïdékon, J. Berestycki, É. Brunet, and Z. Shi, *Branching Brownian motion seen from its tip*, Probab. Theory Relat. Fields **157** (2013), 405–451. MR-3101852

[2] Louis-Pierre Arguin, Anton Bovier, and Nicola Kistler, *The extremal process of branching Brownian motion*, Probab. Theory Relat. Fields **157** (2013), 535–574. MR-3129797

[3] Maury Bramson, *Convergence of solutions of the Kolmogorov equation to travelling waves*, Mem. Amer. Math. Soc. (1983), no. 285. MR-0705746

[4] Maury D. Bramson, *Maximal displacement of branching Brownian motion*, Comm. Math. Phys. **31** (1978), no. 5, 531–581. MR-0494541

[5] Aser Cortines, Lisa Hartung, and Oren Louidor, *Decorated random walk restricted to stay below a curve*, Ann. Probab. **47** (2019), no. 4, supplement material to "the structure of extreme level sets in branching Brownian motion". MR-3980921

[6] Aser Cortines, Lisa Hartung, and Oren Louidor, *The structure of extreme level sets in branching Brownian motion*, Ann. Probab. **47** (2019), no. 4, 2257–2302. MR-3980921

[7] Simon C. Harris and Matthew I. Roberts, *The many-to-few lemma and multiple spines*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 1, 226–242. MR-3606740

[8] S. P. Lalley and T. Sellke, *A conditional limit theorem for the frontier of a branching Brownian motion*, Ann. Probab. **15** (1987), no. 3, 1052–1061. MR-0893913

[9] H. P. McKean, *Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov*, Comm. Pure Appl. Math. **28** (1975), no. 3, 323–331. MR-0400428