From Derrida's random energy model to branching random walks: from 1 to 3

Nicola Kistler[*]   Marius A. Schmidt[†]

Abstract

We study the extremes of a class of Gaussian fields with in-built hierarchical structure. The number of scales in the underlying trees depends on a parameter $\alpha \in [0, 1]$: choosing $\alpha = 0$ yields the random energy model by Derrida (REM), whereas $\alpha = 1$ corresponds to the branching random walk (BRW). When the parameter α increases, the level of the maximum of the field decreases smoothly from the REM- to the BRW-value. However, as long as $\alpha < 1$ strictly, the limiting extremal process is always Poissonian.

1 Introduction and main result

The Gaussian fields we consider are constructed as follows. Let $\alpha \in [0, 1]$ and $N \in \mathbb{N}$. We refer to the parameter N as the *size of the system*. For $j = 1 \ldots N^\alpha$ and $\sigma_j = 1 \ldots 2^{\left(N^{1-\alpha}\right)}$, consider the vectors $\sigma = (\sigma_1, \ldots, \sigma_{N^\alpha})$. (We assume, without loss of generality, that N and α are such that N^α and $N^{1-\alpha}$ are both integers). We refer to the indices $j = 1 \ldots N^\alpha$ as *scales*, and to the labels σ as *configurations*. The space of configurations is denoted by $\Sigma_N^{(\alpha)}$. Remark that, by construction, $\sharp\Sigma_N^{(\alpha)} = 2^N$. For scales $j \leq N^\alpha$ and $(\sigma_1, \ldots, \sigma_j)$, consider independent centered Gaussian random variables $X_{\sigma_1, \ldots, \sigma_j}^{(\alpha, j)}$ with variance $N^{1-\alpha}$ defined on some common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. To given configuration $\sigma \in \Sigma_N^{(\alpha)}$ we associate the *energies*

$$X_\sigma^{(\alpha, N)} \equiv \sum_{j=1}^{N^\alpha} X_{\sigma_1, \ldots, \sigma_j}^{(\alpha, j)} \tag{1.1}$$

The collection $X^{(\alpha, N)} \equiv \left\{ X_\sigma^{(\alpha, N)}, \sigma \in \Sigma_N^{(\alpha)} \right\}$ defines a centered Gaussian field with

$$\operatorname{var}\left[X_\sigma^{(\alpha, N)} \right] = N, \quad \text{and} \quad \operatorname{cov}\left[X_\sigma^{(\alpha, N)}, X_\tau^{(\alpha, N)} \right] = (\sigma \wedge \tau) \, N^{1-\alpha},$$

where $\sigma \wedge \tau \equiv \inf \left\{ j \leq N^\alpha : (\sigma_1, \ldots, \sigma_j) = (\tau_1, \ldots, \tau_j) \text{ and } \sigma_{j+1} \neq \tau_{j+1} \right\}$. In spin glass terminology, $\sigma \wedge \tau$ is the *overlap* of the configurations σ and τ. In other words, the

---

[*]J.W. Goethe-Universität Frankfurt, Germany. E-mail: kistler@math.uni-frankfurt.de
[†]J.W. Goethe-Universität Frankfurt, Germany. E-mail: mschmidt@math.uni-frankfurt.de

Gaussian field $X^{(\alpha,N)}$ is *hierarchically* correlated. The parameter α governs the number of scales in the underlying "trees". The choice $\alpha = 0$ yields the celebrated REM of Derrida [12]; in this case the tree consists of a single scale (only for this boundary case is the field uncorrelated). The choice $\alpha = 1$ yields the (classical) BRW, also known as the directed polymer on Cayley trees [15]: in this model, the number of scales grows linearly with the size of the system. In this sense, the fields $X^{(\alpha,N)}$ interpolate between REM and BRW (remark that these boundary cases are, within our class, the least resp. the most correlated fields). See Figure 1 below for a graphical representation.
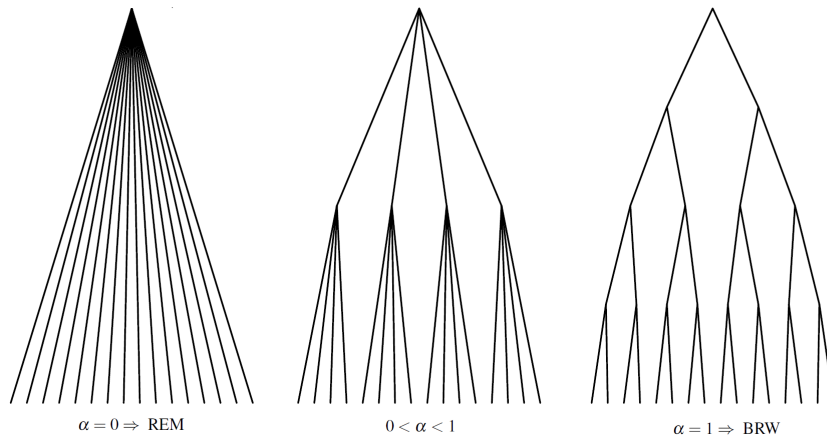

Figure 1: Trees interpolating between REM and BRW

A fundamental question in the study of random fields concerns the behavior of the extreme values in the limit of large system-size. The case of independent random variables is simple, and completely understood, see e.g. the classic [21]. On the other hand, the study of the extremes of *correlated* random fields is a much harder question. There is good reason to develop an extreme value theory for Gaussian fields defined on trees: besides being typically amenable to a detailed analysis (see e.g. [3, 5, 7, 8, 9, 10, 16, 22]), Gaussian hierarchical fields should be some sort of "universal attractors" in the limit of large system-size; this claim is a major pillar of the Parisi theory [24] which has remained to these days rather elusive (see however [19] and references therein for some recent advances). Our main result provides a characterization of the weak limit of the extremes of the hierarchical field (1.1).

Theorem 1.1. *Assume $\alpha \in [0,1)$. Let*

$$a_N^{(\alpha)} \equiv \beta_c N - \frac{1+2\alpha}{2\beta_c} \log N, \quad \text{where} \ \ \beta_c \equiv \sqrt{2\log 2},$$

and consider the random Radon measure on the real line

$$\Xi_N^{(\alpha)} \equiv \sum_{\sigma \in \Sigma_N^{(\alpha)}} \delta_{X_\sigma^{(\alpha,N)} - a_N^{(\alpha)}} \, .$$

Then $\Xi_N^{(\alpha)}$ converges weakly to a Poisson process Ξ of intensity $\mu(A) \equiv \int_A e^{-\beta_c x} dx / \sqrt{2\pi}$.

The weak limits of the extremes of Gaussian hierarchical fields with a *fixed* number of scales, the generalized random energy models by Derrida [13], have been rigorously derived in [10]. On the other hand, apart from the case $\alpha = 0$, the picture depicted in Theorem 1.1 seems to be new. There is good reason to leave out the case $\alpha = 1$: to clarify this, and to shed further light on our main result, let us spend a few words. First, the theorem implies that $a_N^{(\alpha)}$ is the *level of the maximum* of the random field $X^{(\alpha, N)}$, and $\Xi_N^{(\alpha)}$ is then the *extremal process*. It steadily follows from the convergence of the extremal process that the maximum of the field, recentered by its level, weakly converges to a Gumbel distribution. As expected under the light of (say) Slepian's Lemma, the level of the maximum *decreases* when α (hence the amount of correlations) *increases*. However, this feature is only detectable at the level of the second order, logarithmic corrections; curiously, the pre-factor $1 + 2\alpha$ interpolates smoothly between the REM- and the BRW-values ("from 1 to 3"). Notwithstanding, as long as $\alpha < 1$ strictly, and in spite of what might look at first sight as severe correlations, all our models fall into the universality class of the REM, which is indeed characterized by convergence towards Poissonian extremal processes. In the boundary case of the BRW, the picture is only partially correct: the logarithmic correction is still given by $a_N^{(\alpha)}$ with $\alpha = 1$, see [1, 2, 11], yet the weak limit of the maximum is no longer a Gumbel distribution [20], nor is the limiting extremal process a simple Poisson process [3, 5, 14, 22].

We conclude this section with a sketch of the proof of our main result. A natural approach would be to choose $a_N^{(\alpha)}$ such that the expected number of extremal configurations in any given compact $A \subset \mathbb{R}$ is of order one in the large N-limit. However, with the level of the maximum as given by Theorem 1.1, classical Gaussian estimates steadily yield

$$\mathbb{E}\left[\Xi_N^{(\alpha)}(A)\right] = 2^N \int_A \exp\left[-\left(x - a_N^{(\alpha)}\right)^2 / (2N)\right] \frac{dx}{\sqrt{2\pi N}} = N^\alpha (1 + o(1)) \qquad (N \to \infty),$$

which is exploding as soon as $\alpha > 0$ strictly. The reason for this is easily identified: by linearity of the expectation, we are completely omitting correlations, but these turn out to be strong enough to affect the level of the maximum. To overcome this problem, we rely on the multi-scale analysis which has emerged in the study of the extremes of branching Brownian motion (see e.g. [19]). To formalize, we need some notation. First, for a given $\sigma \in \Sigma_N^{(\alpha)}$, we refer to the process

$$S^\sigma = (S_k^\sigma, k \leq N^\alpha), \quad S_k^\sigma \equiv \sum_{j \leq k} X_{\sigma_1, \ldots \sigma_j}^{(\alpha, j)},$$

as the *path* of a configuration. (The process S^σ is a random walk with Gaussian increments, i.e. a discrete Brownian motion). We refer to any function $F_N : \{0 \ldots N^\alpha\} \to \mathbb{R}$, $k \mapsto F_N(k)$, as *barrier*. Given a barrier F_N, we denote by

$$\Xi_{N, F_N}^{(\alpha)} \equiv \sum_{\sigma \in \Sigma_N} \delta_{X_\sigma^{(\alpha, N)} - a_N^{(\alpha)}} \mathbf{1}_{\left\{S_k^\sigma \leq F_N(k) \text{ for all } k \in \{1, \ldots, N^\alpha\}\right\}}$$

the *modified (extremal) process*. A key step in the proof is to identify a barrier E_N, see (2.8) below for its explicit form, such that for any compact $A \subset \mathbb{R}$,

$$\lim_{N \to \infty} \mathbb{P}\left[\Xi_N^{(\alpha)}(A) = \Xi_{N, E_N}^{(\alpha)}(A)\right] = 1. \tag{1.2}$$

This naturally entails that the weak limit of the extremal process and that of the modified process must coincide (provided one of the two exists). We will thus focus our attention

on the modified process $\Xi_{N,E_N}^{(\alpha)}$, thereby proving that *mean of the process* as well as its *avoidance functions* converge to the Poissonian limit as given by Theorem 1.1, to wit:

$$\lim_{N\to\infty} \mathbb{E}\left[\Xi_{N,E_N}^{(\alpha)}(A)\right] = \mu(A) \qquad \text{(Convergence of mean)} \qquad (1.3)$$

and

$$\lim_{N\to\infty} \mathbb{P}\left(\Xi_{N,E_N}^{(\alpha)}(A) = 0\right) = \mathbb{P}\left(\Xi(A) = 0\right) \qquad \text{(Avoidance functions)} \qquad (1.4)$$

By (1.3) and (1.4), it follows from Kallenberg's theorem on Poissonian convergence [18], that the modified process weakly converges to the Poisson point process Ξ; but by (1.2), the same must be true for the extremal process, settling the proof of Theorem 1.1.

The rest of the paper is devoted to the proof of (1.2), (1.3) and (1.4). Since $\alpha \in [0,1)$ is fixed throughout, we lighten notations by dropping the α-dependence whenever no confusion can possibly arise, writing e.g. Σ_N for $\Sigma_N^{(\alpha)}$, X_σ for $X_\sigma^{(\alpha,N)}$, a_N for $a_N^{(\alpha)}$, etc.

2 Barriers, and the modified processes

The goal of this section is to construct the barrier E_N to which we alluded in the introduction, and to give a proof of (1.2) and (1.3). In a first step, we construct a barrier which is not "optimal", but which provides important *a priori* information:

Lemma 2.1. *Consider the barrier*

$$U_N(k) \equiv \beta_c k N^{1-\alpha} + \ln(N), \ k = 0, .., N^\alpha.$$

It then holds:

$$\lim_{N\to\infty} \mathbb{P}\left(S_k^\sigma \le U_N(k) \quad \forall k \in \{1, .., N^\alpha\}, \sigma \in \Sigma_N\right) = 1.$$

Proof. By Markov inequality, and simple counting, it holds:

$$\mathbb{P}\left(\exists \sigma \in \Sigma_N : \sum_{i \le j} X_{\sigma_1,..,\sigma_i}^{(i)} > U_N(j), \text{ for some } j \le N^\alpha\right)$$

$$\le \sum_{j \le N^\alpha} \exp\left(jN^{1-\alpha}\ln 2\right) \mathbb{P}\left(\sum_{i \le j} X_{1,...,1}^{(j)} > \beta_c jN^{1-\alpha} + \ln N\right).$$
(2.1)

By classical Gaussian estimates, the probability on the r.h.s. above is at most

$$\frac{\sqrt{jN^{1-\alpha}}}{\sqrt{2\pi}\left(\beta_c jN^{1-\alpha} + \ln N\right)} \exp\left[-\frac{\left(\beta_c jN^{1-\alpha} + \ln N\right)^2}{2jN^{1-\alpha}}\right].$$

Using this, and straightforward estimates, we get

$$(2.1) \le \exp\left[\left(\frac{3\alpha - 1}{2} - \beta_c\right)\ln N\right],$$

which is evidently vanishing in the large N-limit, since $\frac{3\alpha-1}{2} < \beta_c$. $\qquad\square$

The above Lemma immediately implies that the weak limit of the extremal process Ξ_N and the weak limit of the modified process Ξ_{N,U_N} must necessarily coincide (provided one of the two exists). We now identify conditions under which this remains true for barriers which lie even lower than U_N.

Lemma 2.2. *Consider a barrier F_N with the following properties:*

  i) $F_N \leq U_N$, *i.e.* $F_N(k) \leq U_N(k)$ *for all k;*

  ii) for $A \subset \mathbb{R}$ compact, it holds:

$$\lim_{N\to\infty} \mathbb{E}\left[\Xi_{N,F_N}(A)\right] = \lim_{N\to\infty} \mathbb{E}\left[\Xi_{N,U_N}(A)\right]$$

Then the weak limits of Ξ_{N,F_N} and Ξ_{N,U_N} coincide (provided one of the two exists).

Proof. The Lemma steadily follows from the claim

$$\mathbb{P}\left(\Xi_{N,U_N}(A) = \Xi_{N,F_N}(A)\right) \geq 1 - \mathbb{E}\left[\Xi_{N,U_N}(A) - \Xi_{N,F_N}(A)\right]. \tag{2.2}$$

The proof of (2.2) is straightforward. Simple rearrangements and subadditivity imply that the for probability of the complementary event, it holds:

$$
\begin{aligned}
&\mathbb{P}\left(\Xi_{N,U_N}(A) \neq \Xi_{N,F_N}(A)\right) \\
&= \mathbb{P}\left(\exists \sigma \in \Sigma_N : X_\sigma - a_N \in A,\ \forall_{j=1\dots N^\alpha} : S_j^\sigma \leq U_N(j) \text{ but } \exists_{j=1\dots N^\alpha} : S_j^\sigma > F_{j,N}\right) \\
&\leq \sum_{\sigma \in \Sigma_N} \mathbb{P}\left(X_\sigma - a_N \in A, \forall_{j=1\dots N^\alpha} : S_j^\sigma \leq U_N(j) \text{ but } \exists_{j=1\dots N^\alpha} : S_j^\sigma > F_{j,N}\right) \\
&= 2^N \mathbb{P}\left(X_\sigma - a_N \in A, \forall_{j=1\dots N^\alpha} : S_j^\sigma \leq U_N(j) \text{ but } \exists_{j=1\dots N^\alpha} : S_j^\sigma > F_{j,N}\right) \\
&= 2^N \mathbb{P}\left(X_\sigma - a_N \in A,\ S^\sigma \leq U_N\right) - 2^N \mathbb{P}\left(X_\sigma - a_N \in A,\ S^\sigma \leq F_N\right) \\
&= \mathbb{E}\left[\Xi_{N,U_N}(A) - \Xi_{N,F_N}(A)\right].
\end{aligned}
$$

Building the complement, (2.2) immediately follows. □

By the previous Lemma, and in view of a proof of the main theorem, it is crucial to identify conditions for which the mean(s) of the modified process(es) converge to a finite limit. This is done by

Proposition 2.3. *Consider a barrier of the form $F_N = U_N + f_N$, where f_N is such that*

  i) $f_N(0) = f_N(N^\alpha) = 0$

  ii) $\displaystyle\sup_{k\in\{1,..,N^\alpha\}} |f_N(k)| = o\left(N^{\frac{1-\alpha}{2}}\right)$ *for $N \uparrow \infty$.*

For $A \subset \mathbb{R}$ compact, and μ as in Theorem 1.1, it holds:

$$\lim_{N\to\infty} \mathbb{E}\left[\Xi_{N,F_N}(A)\right] = \mu(A).$$

Proof. By linearity of the expectation, and by conditioning on the "terminal event",

$$
\begin{aligned}
\mathbb{E}\left[\Xi_{N,F_N}(A)\right] &= \\
&= 2^N \int_A \mathbb{P}\left(\forall_{k\in\{1,..,N^\alpha\}} : S_k^\sigma \leq F_N(k) \,\Big|\, X_\sigma - a_N = x\right) \mathbb{P}\left(X_\sigma - a_N \in dx\right). \quad (2.3)
\end{aligned}
$$

Let us focus on the conditional probability: we first write this as

$$
\begin{aligned}
&\mathbb{P}\left(\forall_{k\in\{1,..,N^\alpha\}} : S_k^\sigma \leq F_N(k) \,\Big|\, X_\sigma - a_N = x\right) \\
&= \mathbb{P}\left(\forall_{k\in\{1,..,N^\alpha\}} : S_k^\sigma - \frac{k}{N^\alpha}X_\sigma \leq F_N(k) - \frac{k}{N^\alpha}(a_N + x) \,\Big|\, X_\sigma - a_N = x\right). \quad (2.4)
\end{aligned}
$$

Inspection of the covariances shows that the Gaussian vector $\left(S_k^\sigma - \frac{k}{N^\alpha} X_\sigma, k = 1 \ldots N^\alpha\right)$ is, in fact, *independent* of X_σ. Using this, and rescaling by $N^{-\frac{1-\alpha}{2}}$ yields

$$(2.4) = \mathbb{P}\left(\forall_{k \in \{1,..,N^\alpha\}} : \ N^{-\frac{1-\alpha}{2}}\left[S_k^\sigma - \frac{k}{N^\alpha} S_{N^\alpha}^\sigma\right] \leq N^{-\frac{1-\alpha}{2}}\left[F_N(k) - \frac{k}{N^\alpha}\left(a_N + x\right)\right]\right).$$
$$(2.5)$$

Again by inspection of the covariances, one immediately realizes that the law of the Gaussian vector $\left(N^{-\frac{1-\alpha}{2}}\left[S_k^\sigma - \frac{k}{N^\alpha} S_{N^\alpha}^\sigma\right], k = 0 \ldots N^\alpha\right)$ is that of a (discrete) Brownian bridge of lifespan N^α, starting and ending in 0. To lighten notations, let $(B_{N^\alpha}(k), k \leq N^\alpha)$ be such a Brownian bridge, and shorten

$$\widetilde{F}_N(k, x) \equiv N^{-\frac{1-\alpha}{2}}\left[F_N(k) - \frac{k}{N^\alpha}\left(a_N + x\right)\right].$$

It thus holds:

$$(2.4) = \mathbb{P}\left(\forall_{k \in \{1,..,N^\alpha\}} : \ B_{N^\alpha}(k) \leq \widetilde{F}_N(k, x)\right).$$

One immediately checks that within our choice of the barrier F_N, and since $\alpha < 1$ strictly,

$$\lim_{N \uparrow \infty} \sup_{k \leq N^\alpha, x \in A} \left|\widetilde{F}_N(k, x)\right| = 0.$$

in which case it follows from the lemmata in the Appendix that

$$\mathbb{P}\left(\forall_{k \in \{1,..,N^\alpha\}} : \ B_{N^\alpha}(k) \leq \widetilde{F}_N(k, x)\right) = \mathbb{P}\left(\forall_{k \in \{1,..,N^\alpha\}} : \ B_{N^\alpha}(k) \leq 0\right)(1 + o(1))$$
$$= N^{-\alpha}(1 + o(1)),$$
$$(2.6)$$

uniformly for x in compacts, and for $N \uparrow \infty$. Plugging this into (2.3), we have

$$\mathbb{E}\left[\Xi_{N,F_N}(A)\right] = 2^N N^{-\alpha}(1 + o(1)) \int_A \mathbb{P}\left(X_\sigma - a_N \in dx\right).$$

The claim of the Proposition then immediately follows by straightforward estimates on the Gaussian density. $\qquad\square$

Remark 2.4. The proof of Proposition 2.3 breaks down in the limiting case $\alpha = 1$: for this choice of the parameter, $F_N(k) - \frac{k}{N^\alpha}(a_N + x)$ is of logarithmic size, hence $\widetilde{F}_N(k, x)$ does not vanish in the large N limit. As technical as it may look, this is in fact a structural issue: for $\alpha = 1$ the extremal process cannot be a simple Poisson point process.

We can finally specify our choice of the barrier E_N alluded to in the introduction. The optimal choice is (by far) not unique, and depends on an additional free parameter γ. The only requirement is that

$$0 < \gamma < \frac{1-\alpha}{2}.$$
$$(2.7)$$

With any γ satisfying (2.7), and U_N as in Lemma 2.1, we set

$$E_N(k) \equiv U_N(k) - N^\gamma \mathbf{1}_{k \neq 0, N^\alpha}$$
$$(2.8)$$

This choice of a barrier clearly satisfies the assumptions of Proposition 2.3 and also Lemma 2.2. This has two fundamental consequences: first, the weak limit of the modified process $\Xi_{N,E_N}^{(\alpha)}$ and that of extremal process $\Xi_N^{(\alpha)}$ must necessarily coincide (provided one of the two exists); second, the mean of the modified process converges to the alleged limit, i.e. (1.3) holds with E_N as a barrier. Theorem 1.1 will thus follow as soon as we prove that avoidance functions (1.4) also converge with the very same choice for the

barrier. This will be done in the next section by means of the Chen-Stein method. Before that, let us spend a few words "on what really stands behind" the choice of the barrier (2.8). (The discussion is intentionally informal: for details, the reader is referred e.g. to [19].)

First, we remark that by Lemma 2.1, the path of extremal configurations (the process $k \mapsto S_k^\sigma$ for σ s.t. $X_\sigma \approx a_N$) must necessarily satisfy the "U_N-barrier condition". As we have seen in Proposition 2.3, conditioning onto the terminal event turns the path into a Brownian bridge which is required to stay below 0 during its lifespan. It is well known that in order to achieve this, the bridge will behave within good approximation as the path of its modulus in the negative, $k \mapsto - \mid S_k^\sigma \mid$, which is typically much lower than the shift $-N^\gamma \mathbf{1}_{k \neq 0, N^\alpha}$ for $\gamma < (1 - \alpha)/2$ (this is the so-called entropic repulsion, see e.g. [4]). In other words, requiring that the paths stay below E_N is no stricter requirement than asking them to stay below U_N. However, and crucially, the E_N-barrier is low enough to force the expected number of correlated extremal pairs to vanish in the large N-limit: it is this specific feature which stands behind the Chen-Stein method which we implement below.

3 Convergence of the avoidance functions

The goal of this section is to prove (1.4), which we recall reads

$$\lim_{N \to \infty} \mathbb{P} \left(\Xi_{N, E_N}(A) = 0 \right) = \mathbb{P} \left(\Xi(A) = 0 \right), \tag{3.1}$$

where E_N is given by (2.8), A is any compact set, and Ξ is a Poisson point process with density $\mu(A) = \int_A e^{-\beta_c x} dx / \sqrt{2\pi}$. To do so, we will use the so-called Chen-Stein method [6, Theorem 1A]. We begin with a warm-up computation. In what follows, we write $\mathcal{E}_N(\sigma)$ for the event that a configuration σ satifies the "E_N-barrier condition", more precisely:

$$\mathcal{E}_N(\sigma) \equiv \left\{ \omega \in \Omega : S_k^\sigma(\omega) \leq E_N(k), k = 1 \dots N^\alpha \right\}.$$

Recall that for two configurations $\sigma, \tau \in \Sigma_N^{(\alpha)}$, we denote by $\sigma \wedge \tau$ their overlap, namely the first scale at which the two configurations do not coincide.

Lemma 3.1 (Extremal pairs)**.** *Let $A \subset \mathbb{R}$ be compact. With the above notations, it holds:*

$$\mathbb{E} \left[\sharp \left\{ \sigma, \tau : \ \sigma \wedge \tau \neq 0, N^\alpha, \ \text{and } X_\sigma - a_N \in A, \mathcal{E}_N(\sigma); \ X_\tau - a_N \in A, \mathcal{E}_N(\tau) \right\} \right] = o(1),$$

as $N \to \infty$.

It follows from Lemma 3.1 that energies of extremal configurations are, in fact, *independent* random variables. It will come hardly as a surprise that this feature stands behind the onset of the Poisson point process in the large N-limit.

Proof of Lemma 3.1. By linearity of the expectation, and re-arranging the ensuing sum according to the possible overlap-values, it holds:

$$\mathbb{E} \left[\sharp \left\{ \sigma, \tau : \ \sigma \wedge \tau \neq 0, N^\alpha, \ \text{and } X_\sigma - a_N \in A, \mathcal{E}_N(\sigma); \ X_\tau - a_N \in A, \mathcal{E}_N(\tau) \right\} \right]$$
$$= \sum_{K=1}^{N^\alpha - 1} \# \left\{ (\sigma, \tau) | \sigma \wedge \tau = K \right\} \mathbb{P} \left(X_\sigma - a_N \in A, \mathcal{E}_N(\sigma), \ X_\tau - a_N \in A, \mathcal{E}_N(\tau) \right) \tag{3.2}$$

Let us focus on the probability on the r.h.s. above: since σ and τ coincide up to scale K, by conditioning on the "trunk" which is shared by σ and τ, we get

$$\mathbb{P} \left(X_\sigma - a_N \in A, \mathcal{E}_N(\sigma); X_\tau - a_N \in A, \mathcal{E}_N(\tau) \right) = \int\limits_{-\infty}^{E_{K,N}} (P) \times \mathbb{P} \left(S_K^\sigma \in dx \right), \tag{3.3}$$

where

$$(P) \equiv \mathbb{P}\left(x + (S_{N^\alpha}^\sigma - S_K^\sigma) - a_N \in A, \mathcal{E}_N(\sigma)\, ; \; x + (S_{N^\alpha}^\tau - S_K^\tau) - a_N \in A, \mathcal{E}_N(\tau)\, \Big|\, S_K^\sigma = x\right).$$

On the event appearing in (P) we drop the $\mathcal{E}$-requirements: by independence of the paths after the "branching point", this leads to

$$(3.3) \leq \int_{-\infty}^{E_{K,N}} \mathbb{P}\left(x + (S_{N^\alpha}^\sigma - S_K^\sigma) - a_N \in A\right)^2 \mathbb{P}\left(S_K^\sigma \in dx\right). \qquad (3.4)$$

This steadily implies that the r.h.s. of (3.4) is at most

$$\int_{-\infty}^{E_{K,N}} \left[\int_{A+a_N-x} \exp\left(-\frac{z^2}{2N^{1-\alpha}(N^\alpha - K)}\right) dz\right]^2 \exp\left(-\frac{x^2}{2N^{1-\alpha}K}\right) dx$$

$$\leq 2^{-2N+KN^{1-\alpha}}\lambda(A) \times$$

$$\times \sup_{a\in A} \int_{-\infty}^{0} \exp\left(-\frac{\left(a_N^{(\alpha)} - E_{K,N} - x + a\right)^2}{N^{1-\alpha}(N^\alpha - K)} - \frac{(x + E_{K,N})^2}{2N^{1-\alpha}K} + \left(2N - KN^{1-\alpha}\right)\ln 2\right) dx,$$

$$(3.5)$$

where λ denotes Lebesgue measure. The argument of the exponential in (3.5) is easily seen to be bounded by $\beta_c\left(3\ln N - N^\gamma + x - 2a\right)$, hence

$$(3.3) \leq 2^{-2N+KN^{1-\alpha}}\lambda(A) \sup_{a\in A} \int_{-\infty}^{0} \exp\left[\beta_c\left(3\ln N - N^\gamma + x - 2a\right)\right] dx$$

$$\leq 2^{-2N+KN^{1-\alpha}} C_A \exp\left[\beta_c\left(3\ln N - N^\gamma\right)\right], \qquad (3.6)$$

with $C_A \equiv \frac{1}{\beta_c}\lambda(A)\exp\left[-2\beta_c\inf\{A\}\right]$. Plugging (3.6) into (3.2), and using that

$$\#\left\{(\sigma,\tau)\,|\,\sigma \wedge \tau = K\right\} \times 2^{-2N+KN^{1-\alpha}} \leq 1,$$

we get

$$(3.2) \leq \sum_{K=1}^{N^\alpha - 1} C_A \exp\left[\beta_c\left(3\ln N - N^\gamma\right)\right],$$

which is evidently vanishing in the large N-limit. $\qquad\square$

We can now finally move to the last missing piece, namely a proof of convergence of the avoidance functions (3.1). As mentioned, the main technical device here will be the so-called Chen-Stein method, [6, Theorem 1A]. To implement this, we need to introduce some notation. For compact $A \subset \mathbb{R}$, we shorten

$$\mu_N(A) \equiv \mathbb{E}\left[\Xi_{N,E_N}(A)\right]$$

and denote by $\mathcal{L}_N(A)$ the *law* of the random variable $\Xi_{N,E_N}(A)$. For a (sigma-finite) measure ν on $\mathbb{R}$, we denote by $\mathrm{Pois}_{\nu(A)}$ the law of a Poisson random variable with mean $\nu(A)$. For $\rho, \rho' \in \mathcal{M}_1(\mathbb{R})$ two probability measures on $\mathbb{R}$ we denote by $d_{TV}(\rho, \rho')$ their distance in total variation. In order to closely stick to the notation in [6], we write

$$\Xi_{N,E_N}(A) = \sum_{\sigma\in\Sigma_N^{(\alpha)}} I_\sigma, \quad I_\sigma \equiv \mathbf{1}_{\{X_\sigma - a_N \in A, \mathcal{E}_N(\sigma)\}},$$

From REM to BRW

and define, for given $\sigma \in \Sigma_N$,

$$Z_\sigma \equiv \sum_{\tau \in \Sigma_N, \tau \wedge \sigma \neq 0, N^\alpha} I_\tau.$$

For a last piece of notation, we shorten $p_\sigma \equiv \mathbb{E}[I_\sigma]$.

Coming back to our main task of proving (3.1), with $\mu(A)$ as in Theorem 1.1, it holds:

$$\left| \mathbb{P}\left(\Xi_{N,E_N}(A) = 0\right) - \mathbb{P}\left(\Xi(A) = 0\right) \right| \leq d_{TV}\left(\mathcal{L}_N(A), \mathrm{Pois}_{\mu(A)}\right)$$
$$\leq d_{TV}\left(\mathcal{L}_N(A), \mathrm{Pois}_{\mu_N(A)}\right) + d_{TV}\left(\mathrm{Pois}_{\mu_N(A)}, \mathrm{Pois}_{\mu(A)}\right) \tag{3.7}$$

The convergence of $\mu_N(A)$ towards $\mu(A)$ is guaranteed by Proposition 2.3; in virtue of simple properties of Poisson random variables, this convergence implies that the second term on the r.h.s. above vanishes in the limit of large N. Concerning the first term on the r.h.s. of (3.7): the Chen-Stein method [6, Theorem 1A] yields the bound

$$d_{TV}\left(\mathcal{L}_N(A), \mathrm{Pois}_{\mu_N(A)}\right) \leq \sum_{\sigma \in \Sigma_N} \left(p_\sigma^2 + p_\sigma \mathbb{E}\left[Z_\sigma\right] + \mathbb{E}\left[I_\sigma Z_\sigma\right]\right) \tag{3.8}$$

Since for any $\sigma \in \Sigma_N$, $p_\sigma = 2^{-N}\mu_N(A)$, one immediately gets

$$\sum_{\sigma \in \Sigma_N} p_\sigma^2 = 2^{-N}\mu_N(A)^2, \tag{3.9}$$

and by simple counting,

$$\sum_{\sigma \in \Sigma_N} p_\sigma \mathbb{E}\left[Z_\sigma\right] \leq 2^{-N^{1-\alpha}}\mu_N(A)^2. \tag{3.10}$$

Plugging (3.9) and (3.10) in (3.8) we get

$$d_{TV}\left(\mathcal{L}_N(A), \mathrm{Pois}_{\mu_N(A)}\right) \leq 2^{-N}\mu_N(A)^2 + 2^{-N^{1-\alpha}}\mu_N(A)^2 + \sum_{\sigma \in \Sigma_N^{(\alpha)}} \mathbb{E}\left[I_\sigma Z_\sigma\right]$$
$$= 2^{-N}\mu_N(A)^2 + 2^{-N^{1-\alpha}}\mu_N(A)^2 + \sum_{\sigma \wedge \tau \neq 0, N^\alpha} \mathbb{E}\left[I_\sigma I_\tau\right], \tag{3.11}$$

the last equality by definition of Z_σ. Since $\mu_N(A)$ converges to a finite limit (by Proposition 2.3), the first two terms in the last display of (3.11) vanish in the limit of large N; the third term is exactly what was analyzed in Lemma 3.1, and therefore also vanishing. All in all, (3.7) is vanishing, hence (3.1) holds and the proof of Theorem 1.1 is concluded.

Appendix

A fundamental ingredient in the proof of Theorem 1.1 are the estimates (2.6) on Brownian bridge probabilities appearing in the proof of Proposition 2.3; these are somewhat classical [17], sometimes going under the name of "ballot theorems". For the reader's convenience, we give here a short proof of the estimates as needed in our framework.

Lemma 3.2. *Let $(\Delta_i)_{i \in \{0,..,n-1\}}$ be i.i.d random variables having a density with respect to the Lebesgue measure and $(B_n(j), j \in \{1,..,n\})$ the related bridge, i.e.*

$$B_n(j) \equiv \sum_{i=0}^{j-1} \Delta_i - \frac{j}{n}\sum_{i=0}^{n-1}\Delta_i,$$

then it holds:

$$\mathbb{P}\left(B_n(j) \leq 0 \text{ for all } j \in \{1,..,n\}\right) = \frac{1}{n} \tag{3.12}$$

Proof. We refer to $(\Delta_i)_{i \in \{0,..,n-1\}}$ as *increments*. The event in (3.12) is equivalent to the maximum of the bridge being lower than zero. Let $m \in \{0,..,n-1\}$ be the position of the maximum; remark that this is almost surely unique by the density-assumption. One steadily checks that applying a cyclic permutation, say π, to the increments of the bridge, shifts the position of the maximum to $\pi^{-1}m$. There is one cyclic permutation only, say $\hat{\pi}$, which shifts the position of the maximum to the origin, i.e. for which $\hat{\pi}^{-1}m = 0$. On the other hand, the distribution of the bridge is not affected by any permutation, hence $\hat{\pi}$ must be uniformly distributed among the n possible cyclic permutations: since the event in (3.12) is equivalent to $\hat{\pi}$ being the identity, the Lemma follows. $\square$

In other words, the probability that a discrete bridge stays below zero during its lifetime decays as the inverse of the length of the bridge. On the other hand, since our bridges have "square-root fluctuations", one expects that whether the bridge is required to stay below zero or below a straight line shouldn't alter (much) the asymptotic behavior of these probabilities. This is indeed the case:

Lemma 3.3. *Let $(B_n(j), j = \ldots n)$ be a (discrete) Brownian bridge of length n. Then there exists $c > 0$ independent of n such that for any n and $|\varepsilon| \leq c^{-1}$, it holds:*

$$\left| \mathbb{P}\left(B_n(j) \leq 0, j = 1 \ldots n-1\right) - \mathbb{P}\left(B_n(j) \leq \varepsilon, j = 1 \ldots n-1\right) \right| \leq c \frac{|\varepsilon|}{n}.$$

Proof. We proceed by induction on the length of the bridge.

Base case. For $n = 2$, it clearly holds:

$$\mathbb{P}\left(B_n(j) \leq \varepsilon, j \in \{1\}, \exists j \in \{1\} : B_n(j) > 0\right) = \mathbb{P}\left(B_n(1) \in [0, \varepsilon]\right) \leq \frac{2}{\sqrt{\pi}} \frac{|\varepsilon|}{n} \quad \text{for } \varepsilon > 0.$$

$$\mathbb{P}\left(B_n(j) \leq 0, j \in \{1\}, \exists j \in \{1\} : B_n(j) > \varepsilon\right) = \mathbb{P}\left(B_n(1) \in [\varepsilon, 0]\right) \leq \frac{2}{\sqrt{\pi}} \frac{|\varepsilon|}{n} \quad \text{for } \varepsilon < 0.$$

The proof in the cases $\varepsilon > 0$ and $\varepsilon < 0$ are similar, we thus consider only the first case.

Induction step. For $n \geq 3$, assume the claim is true for all $k \leq n-1$. By Markov inequality,

$$\mathbb{P}\left(B_n(j) \leq \varepsilon, j = 1 \ldots n-1 \text{ but } \exists i = 1 \ldots n-1 : B_n(i) > 0\right)$$

$$\leq \sum_{i=1}^{n-1} \mathbb{P}\left(B_n(j) \leq \varepsilon, j = 1 \ldots n-1 \text{ but } B_i > 0\right) \tag{3.13}$$

$$= \sum_{i=1}^{n-1} \int_0^{\varepsilon} \mathbb{P}\left(B_n(j) \leq \varepsilon, \ j = 1 \ldots n-1 \middle| B_n(i) = x\right) \mathbb{P}\left(B_n(i) \in dx\right).$$

By the Markov property of Brownian bridges, the conditional probability above reads

$$\mathbb{P}\left(\forall_{j \in \{1,..,i\}} : \ B_n(j) \leq \varepsilon \middle| B_n(i) = x\right) \mathbb{P}\left(\forall_{j \in \{i,..,n-1\}} B_n(j) \leq \varepsilon \middle| B_n(i) = x\right)$$

$$\leq \mathbb{P}\left(\forall_{j \in \{1,..,i\}} : \ B_n(j) \leq \varepsilon \middle| B_n(i) = 0\right) \mathbb{P}\left(\forall_{j \in \{i,..,n-1\}} B_n(j) \leq \varepsilon \middle| B_n(i) = 0\right)$$

where the inequality follows by monotonicity in x. Using this, and Gaussian estimates,

$$(3.13) \leq \sum_{i=1}^{n-1} \mathbb{P}\left(\forall_{j \in \{1,..,i\}} : B_n(j) \leq \varepsilon | B_n(i) = 0\right) \times$$

$$\times \mathbb{P}\left(\forall_{j \in \{i,..,n-1\}} : \ B_n(j) \leq \varepsilon | B_n(i) = 0\right) \frac{\varepsilon}{\sqrt{2\pi \left(i - \frac{i^2}{n}\right)}}.$$

Given that $B_n(i) = 0$, the process $(B_n(j), j = 1 \dots n)$ is a Brownian bridge of length $i \leq n - 1$; analogously, the second probability involves a Brownian bridge of length $n - i$. The assumption therefore applies, and the above is at most

$$\sum_{i=1}^{n-1} \left(\mathbb{P} \left(\forall_{j \in \{1,..,i\}} : B_i(j) \leq 0 \right) + c\frac{\varepsilon}{i} \right) \times$$
$$\times \left(\mathbb{P} \left(\forall_{j \in \{1,..,n-i\}} : B_{n-i}(j) \leq 0 \right) + c\frac{\varepsilon}{n-i} \right) \frac{\varepsilon}{\sqrt{2\pi \left(i - \frac{i^2}{n} \right)}}.$$
(3.14)

It then holds:

$$(3.14) \leq \varepsilon \frac{4\sqrt{n}}{\sqrt{2\pi}} \sum_{i=1}^{n-1} \left(\frac{1}{i(n-i)} \right)^{3/2} \leq \varepsilon \frac{8\sqrt{n}}{\sqrt{2\pi}} \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{1}{i(n-i)} \right)^{3/2} \leq c\frac{\varepsilon}{n},$$

the last inequality using the bound $[i(n-i)]^{-3/2} \leq [i(n/2)]^{-3/2}$, for $i \leq \lfloor n/2 \rfloor$. $\qquad \square$

References

[1] L. Addario-Berry and B. Reed, *Minima in branching random walks*, Ann. of Prob. 37, 1044-1079 (2009) MR-2537549

[2] E. Aidekon, *Convergence in law of the minimum of a branching random walk*, Ann. of Prob. 41, 362-1426 (2013) MR-3098680

[3] E. Aidekon, J. Berestycki, E. Brunet, and Z. Shi. *Branching Brownian motion seen from its tip*, Probab. Theor. Rel. Fields, 157, 405-451 (2013). MR-3101852

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

[5] L.-P. Arguin, A. Bovier, and N. Kistler, *The extremal process of branching Brownian motion*, Probab. Theor. Rel. Fields 157, 535-574 (2013). MR-3129797

[6] A.D. Barbour, L. Holst, and S. Janson, *Poisson Approximation*, Oxford: Clarendon Press, (1992). MR-1163825

[7] A. Bovier and L. Hartung, *The extremal process of two-speed branching Brownian motion*, Elect. J. Probab. 19, No. 18, 1-28 (2014) MR-3164771

[8] A. Bovier and L. Hartung, *Variable speed branching Brownian motion 1. Extremal processes in the weak correlation regime*, arXiv e-prints (2014)

[9] A. Bovier and L. Hartung, *Extended Convergence of the Extremal Process of Branching Brownian Motion*, arXiv e-prints (2014).

[10] A. Bovier, I. Kurkova. *Derrida's generalized random energy models. I. Models with finitely many hierarchies*. Ann. Inst. H. Poincare. Prob. et Statistiques (B) Prob. Stat. 40, 439-480 (2004). MR-2070334

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

[12] B. Derrida, *Random-energy model: An exactly solvable model of disordered systems*, Phys. Rev. B (3), 24 (1981) MR-0627810

[13] B. Derrida, *A generalization of the random energy model which includes correlations between energies*, Jour. de Phys. Lett. 46, 401-407 (1985).

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

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

[16] M. Fang, and O. Zeitouni, *Slowdown for time inhomogeneous branching Brownian motion*, Jour. Stat. Phys. 149, 1-9 (2012). MR-2981635

[17] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I, Wiley (1968) MR-0228020

[18] O. Kallenberg, *Random measures*, Akademie-Verlag, Berlin (1983). MR-0818219

[19] N. Kistler, *Derrida's random energy models. From spin glasses to the extremes of correlated random fields*, Springer Lecture Notes, to appear (2015+)

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

[21] M. R. Leadbetter, G. Lindgren, and H. Rootzen. *Extremes and related properties of random sequences and processes*. Vol. 21. New York: Springer-Verlag (1983). MR-0691492

[22] T. Madaule, *Convergence in law for the branching random walk seen from its tip*, arXiv e-prints (2011).

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

[24] M. Mezard, M.A. Virasoro, and G. Parisi, *Spin glass theory and beyond*. World scientific (1987). MR-1026102