

## On the overlap distribution of Branching Random Walks\*

Aukosh Jagannath<sup>†</sup>

### Abstract

In this paper, we study the overlap distribution and Gibbs measure of the Branching Random Walk with Gaussian increments on a binary tree. We first prove that the Branching Random Walk is 1 step Replica Symmetry Breaking and give a precise form for its overlap distribution, verifying a prediction of Derrida and Spohn. We then prove that the Gibbs measure of this system satisfies the Ghirlanda-Guerra identities. As a consequence, the limiting Gibbs measure has Poisson-Dirichlet statistics. The main technical result is a proof that the overlap distribution for the Branching Random Walk is supported on the set  $\{0, 1\}$ .

**Keywords:** Branching Random Walk; Ghirlanda-Guerra identities; spin glasses.

**AMS MSC 2010:** 60K35; 82B44; 82D30; 82D60.

Submitted to EJP on April 12, 2016, final version accepted on July 10, 2016.

## 1 Introduction

In this paper, we study the Branching Random Walk (BRW), or directed polymer, on a binary tree. To fix notation, let  $\mathcal{T}_N$  be the binary tree of depth  $N$  and let  $\{g_v\}_{v \in \mathcal{T}_N \setminus \emptyset}$  be a collection of i.i.d. standard Gaussian random variables indexed by this tree without its root. We define the Branching Random Walk by

$$H(v) = \sum_{\beta \in p(v)} g_\beta$$

where  $p(v)$  is the root-leaf path to  $v$  excluding the root. Viewed as a gaussian process on  $\partial \mathcal{T}_N$ ,  $(H(v))_{v \in \partial \mathcal{T}_N}$  is centered and has covariance structure

$$\mathbb{E}H_N(v)H_N(w) = |v \wedge w|,$$

where  $v \wedge w$  denotes the least common ancestor of  $v$  and  $w$  and  $|\alpha|$  is the depth of  $\alpha \in \mathcal{T}_N$ . In particular,

$$\mathbb{E}H_N(v)H_N(v) = N.$$

---

\*This research was supported in part by the National Science Foundation and New York University

<sup>†</sup>Department of Mathematics, University of Toronto, Canada E-mail: aukosh@math.toronto.edu

One can think of the root-leaf paths on the tree as polymer configurations and  $H_N$  as an energy. We denote the partition function corresponding to this polymer model by

$$Z_N(\beta) = \sum e^{\beta H_N(v)}$$

and the free energy by

$$F_N(\beta) = \frac{1}{N} \mathbb{E} \log Z_N(\beta).$$

If we denote the Gibbs measure by

$$G_{N,\beta}(v) = \frac{e^{\beta H_N(v)}}{Z}$$

then this induces a (random) probability measure on the leaves  $\Sigma_N = \partial \mathcal{T}_N$ . In the following  $\langle \cdot \rangle_N$  denotes integration with respect to this measure or the corresponding product measures. We will drop the subscript  $N$  when it is unambiguous. Finally, let  $R(v, w) = \frac{1}{N} |v \wedge w|$ , and let  $R_{12} = R(v_1, v_2)$ , which we call the overlap between two polymers. An important object in the study of mean field spin glasses is the (mean) overlap distribution

$$\mu_{N,\beta}(A) = \mathbb{E} G_{N,\beta}^{\otimes 2}(R_{12} \in A). \tag{1.1}$$

The Branching Random Walk, was introduced to the mean field spin glass community in [17]. There, Derrida and Spohn argued that the statistical physics of this model should be similar to the Random Energy Model (REM). They predicted that the overlap distribution should consist of one atom at high temperature and two atoms at low temperature. In the language of Replica theory it should be Replica Symmetric (RS) at high temperature and one step Replica Symmetry breaking (1RSB) at low temperature. Furthermore, they predicted that, as with the REM, the limiting Gibbs measure of the system should be a Ruelle Probability Cascade (see the discussion preceding Corollary 3.6 for a definition). As a consequence, it was suggested [15, 17] that the BRW should serve as an intermediate toy model for spin glass systems, between the REM and the Sherrington-Kirkpatrick (SK) model, as it is still analytically tractable, while having a key feature of SK that the REM lacks: a strong local correlation structure.

The study of replica symmetry breaking in its various forms is the subject of major research in the mathematical spin glass community. As such it is of interest to have a few simple, but non-trivial examples in which Replica Symmetry Breaking can be seen essentially “by hand”. In this paper, we give proofs of the predictions of Derrida and Spohn described above using a combination of arguments that are basic to both fields. In particular, we avoid the use and analysis of the extremal process.

We begin first with the study of Replica Symmetry and Replica Symmetry Breaking. Replica Symmetry above the critical temperature was proved by Chauvin and Rouault in [15]. Our contribution is proving Replica Symmetry Breaking below the critical temperature, and in particular obtaining the mass on the atom at 1.

**Theorem 1.1.** *Let  $\beta_c = \sqrt{2 \log 2}$ . Then*

$$\mathbb{E} G_{N,\beta}^{\otimes 2}(R_{12} \in \cdot) \rightarrow \mu_\beta(\cdot) = \begin{cases} \delta_0 & \beta < \beta_c \\ \frac{\beta_c}{\beta} \delta_0 + (1 - \frac{\beta_c}{\beta}) \delta_1 & \beta \geq \beta_c \end{cases} \tag{1.2}$$

*weakly as measures*

We now turn toward the characterization of the Gibbs measure for this system. Our next result is to prove that the Gibbs measure satisfies a class of identities called the Approximate Ghirlanda-Guerra Identities which will imply the Ruelle Probability Cascade Structure described above. To this end, let  $(v_i)$  be *i.i.d.* draws from  $G_{N,\beta}$ , let  $R_{ij} = R(v_i, v_j)$  and define  $R^n = (R_{ij})_{i,j \in [n]}$ . The doubly infinite array  $R = (R_{ij})_{i,j \geq 1}$  is called the *overlap array* corresponding to these draws.

**Theorem 1.2.** *The Branching Random Walk satisfies the Approximate Ghirlanda-Guerra identities. That is, if  $f$  is a bounded measurable function  $[0, 1]^{n^2}$  then for every  $p$ ,*

$$\lim_{N \rightarrow \infty} |\mathbb{E} \langle f(R^n) R_{1,n+1}^p \rangle_{G_{N,\beta}^{\otimes n+1}} - \frac{1}{n} \left( \mathbb{E} \langle f(R^n) \rangle_{G_{N,\beta}^{\otimes n}} \mathbb{E} \langle R_{12}^p \rangle_{G_{N,\beta}^{\otimes n}} + \sum_{k=2}^n \mathbb{E} \langle f(R^n) R_{1k}^p \rangle_{G_{N,\beta}^{\otimes n}} \right)| = 0$$

**Remark 1.3.** Note that by Theorem 1.1, only the case  $\beta > \beta_c$  is interesting in the above theorem.

Our proof is a version of the technique introduced by Bovier and Kurkova in [11, 12] (see [10] for a textbook presentation) and is analogous to [5, 6]. An immediate consequence of this is that the overlap array distribution for these systems converges to a Ruelle Probability Cascade, see Corollary 3.6. This also implies a mode of convergence of Gibbs measures and the convergence of the weights of balls in support a Poisson-Dirichlet process, which was first proved by Barral, Rhodes, and Vargas in greater generality by different methods [8]. This is explained in the discussion surrounding Corollary 3.6.

For experts in Branching Random Walks, we emphasize here the following point. Just as in the work of Arguin-Zindy, these methods allow us to obtain Poisson-Dirichlet statistics for the system without an analysis of the extremal process. In particular, we can avoid an analysis of the decoration (see, e.g., [27, 21] for this terminology) thereby side-stepping a major technical hurdle.

The Approximate Ghirlanda Guerra Identities (AGGI) have emerged as a unifying principle in spin glasses. Due to the characterization-by-invariance theory [24], we know that the limiting overlap distribution is an order parameter for models that satisfy the AGGIs, as originally predicted in the Replica Theoretic literature [22]. As such, it has become very important to find models that satisfy these identities in the limit. This has proved to be very difficult.

They are known to hold exactly for the generic mixed  $p$ -spin glass models [24], the REM and GREM [11, 12]. These ideas have extended to the 2D Gaussian Free Field and a class of Log-Correlated fields [5, 6]. For many other models, however, we only know these results in a perturbative sense [16, 19, 24, 25]. A contribution of this paper is the observation that the Branching Random Walk falls in to the class of models for which these identities hold exactly.

We finally turn to the main technical step involved in the proofs of the above results. Just as with the REM, both of these predictions can be shown to follow from standard concentration and integration-by-parts arguments provided one can show that the model is at most 1RSB and that the top of the support is at 1 when it is 1RSB. To our knowledge this result is thought of as folklore in the Branching Random Walk community. For example, such a result follows from similar ideas to those in [4, 5, 6, 21]. The proof of this result for similar models can be seen in [4] and [18]. In our setting, this is the content of the following proposition.

**Proposition 1.4.** *The mean overlap distribution is supported on the set  $\{0, 1\} \subsetneq [0, 1]$ . That is, for any weak limit we get that*

$$\mathbb{E} G_{N,\beta}^{\otimes 2}(R_{12} \in \cdot) \rightarrow m\delta_0 + (1 - m)\delta_1$$

for some  $m \in [0, 1]$ .

In the Random Energy Model, this is a consequence of the second moment method combined with a large deviations estimate. In our setting, however, this argument breaks down due to the correlation structure of the Branching Random walk. This is often explained by the seemingly innocuous observation that the sub-leading correction to

the expected maximum (see the definition of  $m_N$  in Section 2) is of the same order of magnitude as in the REM, but the pre-factor is  $3/2$  as opposed to  $1/2$ . We point the reader to [3, Section 3.2] for a discussion of how this small change is a signature of a profound structural difference.

In the study of such models, this issue is dealt with by a truncated second moment method approach. In our setting, this takes the form of the tilted barrier estimates of Bramson, see Section 2.1.

Before turning to the proofs of the above, we make the following remarks.

**Remark 1.5.** In our setting, one does not need the full power of the Ghirlanda-Guerra identities to obtain the aforementioned characterization of the Gibbs measure. In particular, as a consequence of Proposition 1.4, the Approximate Ghirlanda-Guerra Identities are equivalent to an approximate form of Talagrand’s identities [24], which characterize the Poisson-Dirichlet process through its moments. This is explained in more detail in the discussion surrounding Corollary 3.6.

**Remark 1.6.** These arguments do not depend greatly on the Gaussian nature of the problem. In particular, the main technical tool, Proposition 1.4, holds in fairly large generality (see Remark 2.1). The remaining results are essentially consequences of the sub-Gaussian tails of the model and an applications of integration-by-parts. These results should extend to increments that have sub-Gaussian tails. For experts, we also note that if the decoration process has enough moments, the first two results follow by an application of the Bolthausen-Sznitman invariance (see [9][24, Sect. 2.2]). As a study of the extendability of these results are not within the scope of this paper we do not examine these questions further.

**Acknowledgments.** The author would like to thank Ofer Zeitouni for many helpful discussions regarding Branching Random Walks and for a careful reading of an early draft of this paper. This research was conducted while the author was supported by an NSF Graduate Research Fellowship DGE-1342536, and NSF Grants DMS-1209165 and OISE-0730136. Preparation of this manuscript was partially supported by NSF OISE-1604232.

## 2 The support of the overlap distribution

In this section, we will prove Proposition 1.4, namely that the support of the overlap distribution is the set  $\{0, 1\}$ . To this end, we introduce the following notation. Let

$$S_v(l) = \sum_{\substack{w \in P(v) \\ \text{depth}(w) \leq l}} g_w$$

denote the walk corresponding to the BRW at vertex  $v$ . In this notation,  $H_N(v) = S_v(N)$ . Let  $S(l)$  denote a random walk with standard Gaussian increments and let  $P^z$  denote its law conditioned to start at  $z$ .

We think of the collection of walkers  $(S_v)_{v \in \partial T_N}$  as a pack of walkers that branch at each time step. We call  $M_N = \max_{v \in \partial T_N} S_v(N)$  the *leader* of the walkers. With slight—or, depending on your taste, great—exaggeration, we call  $m_N = \beta_c N - \frac{3}{2\beta_c} \log N$ . To justifying this simplification, we remind the reader of the result from [2] that the family of random variables  $(M_N - m_N)_{N \geq 0}$  is tight. In particular, if  $\mathcal{M}_K = \{|M_N - m_N| \leq K\}$ , there is a function  $\epsilon_1(K)$  with

$$\lim_{K \rightarrow \infty} \epsilon_1(K) = 0$$

such that

$$P(\mathcal{M}_K^c) \leq \epsilon_1(K).$$

Of course this does not necessarily show that  $m_N$  the true location of  $M_N$ . The true location will be an order 1 correction from this. Finally let  $\lambda_N = \frac{m_N}{N}$ , and  $\Lambda(x) = \frac{x^2}{2}$ . In the following we will drop the subscript  $N$  whenever possible for readability, and we say  $f \lesssim_a g$  if  $f \leq C(a)g$  where  $C$  is a constant that depends at most on  $a$ .

This section is organized as follows. First we prove a basic estimate that will be used through out the section. We then prove the main estimates required to prove Proposition 1.4. We finally turn to the proof of Proposition 1.4. Before we start we make a brief remark regarding the extension of these results to more general increments than Gaussian.

**Remark 2.1.** The results of this section hold in more generality than we study here. For the interested reader, note that in the following we use sub-Gaussian tails (we believe that one can relax this, however computing the free energy in this setting becomes delicate), that the increments are *i.i.d.* and have support on  $(-1/2, 1/2)$ , and finite moment generating function and rate functions, and that the tree is  $k$ -ary. In this setting,  $m_N = Nx + O(\log N)$  where  $x$  uniquely solves  $\Lambda^*(x) = \log k$  and  $x > \mathbb{E}X$ , and  $\lambda$  uniquely achieves the equality in  $\Lambda^*(x) + \Lambda(\lambda) = \lambda x$ . For more on this see, e.g., [13]. To avoid un-necessary notation and technicalities, we will stick to the Gaussian case where we have the self-duality of the moment generating function,  $\Lambda = \Lambda^*$ .

### 2.1 Tilted barrier estimates

In the following, we will repeatedly use of a class of estimates called *tilted-barrier estimates*. These estimates are used frequently in the study of Branching Random Walks and Log-Correlated fields and were, to our knowledge, introduced by Bramson [14]. The goal of these estimates is to bound probabilities of the form

$$P^z(S(l) \leq \lambda l + K \forall l \in [T]; S(T) \in \lambda T + [a, b]).$$

We think of the underlying event as follows: there is a random walker,  $S(l)$ , which starts at  $z$ , and two lines,  $\lambda l$  and  $\lambda l + K$ , which are barriers. Our goal is to compute the probability that the walker stays below the farther barrier,  $\lambda l + K$ , for the duration of its walk but ends in a window near the nearer barrier,  $\lambda l$ .

The idea of the estimate is to tilt the law of the walker,  $P^z$ , to a new measure,  $Q^{-z}$ , so that under  $Q^{-z}$ , the walk  $\tilde{S}(t) = \lambda \cdot t - S(t)$  will be centered. The result will then follow by an application of the ballot theorem applied to  $\tilde{S}$ . We will make this precise in Proposition 1.4. Before proving this estimate, we first prove the relevant ballot theorem.

**Lemma 2.2. (Ballot-type theorem)** Let  $(S(t))_{t=0}^n$  be given by  $S(t) = X_0 + \sum_{i=1}^t X_i$  be a random walk with  $(X_i)_{i=1}^n$  *i.i.d.* standard Gaussian and  $X_0$  is the starting position. For any  $A, B \in \mathbb{Z}$  with  $0 \leq A < A + 1 \leq B$ , and  $z \geq 0$ , we have that

$$P^z(S(t) \geq 0, 0 < t < n; S(n) \in [A, B]) \lesssim \frac{\max\{z, 1\}B(B - A)}{n^{3/2}}.$$

**Remark 2.3.** This proof is a minor modification of the proof of [1, Theorem 1]. This modification is explained, for example, in [13, Lemma 2.1] we include its proof for the convenience of the reader.

*Proof.* Let  $\tau_h = \min\{0 < k \leq n : S(t) < -h\}$  with the convention that if the condition never happens,  $\tau_h = n$ . Define the time reversed, reflected walk,  $S^r$ ,

$$\begin{cases} S^r(t) = S^r(t - 1) - X_{n-(t-1)} \\ S^r(0) = 0 \end{cases}.$$

That is,  $S^r(t) = S(t) - S(n)$ . Let  $\tau_h^r$  be the same hitting time as before for the reversed walk.

Observe that if  $S(t) \geq 0$  for  $t \in [n]$  and  $S(n) \in [A, B]$ , we must have that  $\tau_0 \geq \lfloor \frac{n}{4} \rfloor$ ,  $\tau_B^r \geq \lfloor \frac{n}{4} \rfloor$ , and  $S(n) \in [A, B]$ . This is for the following reasons. The first condition follows immediately from the positivity. The second condition follows from the fact that if  $\tau_B^r \leq \lfloor \frac{n}{4} \rfloor < \frac{n}{2}$ ,  $S(n)$  splits as

$$S(n) = S(n - \tau_B^r) - S^r(\tau_B^r) \geq -S^r(\tau_B^r) > B$$

by positivity of  $S$  which contradicts  $S(n) \leq B$ . As a result, since  $\{\tau_0 \geq \lfloor \frac{n}{4} \rfloor\}$ ,  $\{\tau_B^r \geq \lfloor \frac{n}{4} \rfloor\}$ , and  $(X_i)_{i=\lfloor \frac{n}{4} \rfloor+1}^{\lceil \frac{3n}{4} \rceil}$  are independent, it follows that

$$\begin{aligned} P^z(S(t) \geq 0, 0 < t < n; S(n) \in [A, B]) &\leq P^z\left(\tau_0 \geq \lfloor \frac{n}{4} \rfloor, \tau_B^r \geq \lfloor \frac{n}{4} \rfloor, S(\frac{3n}{4}) - S^r(\lfloor \frac{n}{4} \rfloor) \in [A, B]\right) \\ &\leq P^0(\tau_z \geq \lfloor \frac{n}{4} \rfloor) P^0(\tau_B \geq \lfloor \frac{n}{4} \rfloor) \sup_x P^0\left(S(\frac{n}{2}) \in x + [A, B]\right) \\ &\lesssim \frac{\max\{z, 1\} \max\{B, 1\} (B - A)}{(n/4)^{1/2} (n/4)^{1/2} \sqrt{n}}. \end{aligned}$$

The first inequality follows by set containment. The second inequality follows by a conditioning argument. The second equality follows from independence, the fact that  $S^r(t), t \in [n/4]$  does not depend on  $X_0 = z$ , and symmetry. The last inequality follows from [1, Lemma 3] and the fact that the last term is at most  $(B - A)/\sqrt{n/2}$  by [1, Theorem 2] combined with a union bound.  $\square$

We now turn to the proof of the Tilted Barrier Estimate.

**Lemma 2.4.** (Tilted Barrier Estimate) *Let  $K, z \geq 0$ , and let  $a$  and  $b$  be such that  $a + 1 \leq b \leq K$ . Then we have*

$$P^z(S(l) \leq \lambda l + K \forall l \in [T]; S(T) \in \lambda T + [a, b]) \lesssim \frac{e^{\lambda(z-a)} \max\{K + z, 1\} \cdot (K - a) \cdot (b - a)}{2^{-T} T^{3/2}}. \tag{2.1}$$

*Proof.* Define the measure  $Q^{-z}$  by the tilting

$$dQ^{-z}(S) = e^{\lambda(S(T)-z) - T\Lambda(\lambda)} dP^z(S).$$

Observe that under  $Q^{-z}(S)$ , the walk  $\tilde{S}$  has no drift, i.e.,

$$\mathbb{E}_{Q^{-z}}(\tilde{S}) = 0,$$

and starts at  $-z$ . Note that

$$\begin{aligned} P^z(S(l) \leq \lambda l + K \forall l \in [T]; S(T) \in \lambda T + [a, b]) &= \mathbb{E}_{Q^{-z}}\left(e^{-\lambda(S(T)-z) + T\Lambda(\lambda)} \mathbb{1}(S(l) \leq \lambda l + K \forall l \in [T]; S(T) \in \lambda T + [a, b])\right) \\ &\leq e^{\lambda(z-a) - \frac{\lambda^2}{2} T} Q^{K-z}\left(\tilde{S} \geq 0 \forall l \in [T]; \tilde{S} \in K - [a, b]\right). \end{aligned}$$

By Lemma 2.2 and the choice of  $\lambda$ , we have that

$$P^z(S(l) \leq \lambda \cdot l + K \forall l \in [T]; S(T) \in \lambda T + [a, b]) \lesssim 2^{-T} e^{\lambda(z-a)} \frac{\max\{K + z, 1\} (K - a) (b - a)}{T^{3/2}}. \square$$

**2.2 Applications of concentration and tilted barrier estimates**

In this subsection, we prove three estimates regarding the probability of the Branching Random Walker having walkers that behave pathologically. Before we begin, we remind the reader of the interpretation of  $H_N$  in terms of the walkers  $S_v$ , and the interpretation of  $m_N$  as (essentially) the location of the leader  $M_N = \max_{v \in \partial \mathcal{T}_N} S_v$  discussed at the beginning of this section.

In our first estimate, we will show that there is a barrier beyond which it is unlikely for any walker to cross. In order for this probability to go to zero, we will need that the barrier drifts off to infinity logarithmically in  $N$ . This will follow from the Gaussian tails of the increments. To make this precise, define the event

$$\Gamma_L^N = \{\exists v \in \partial \mathcal{T}_N : \exists l \in [N] : S_v(l) \geq \lambda l + K\} \tag{2.2}$$

which is the event that there is a leaf  $v$  whose corresponding walker  $S_v$  crosses the barrier  $L^K$  at some time  $l \leq N$ . We then have the following lemma.

**Lemma 2.5.** *For every  $\kappa > \frac{5}{2(\beta_c - \frac{3}{2e})}$ , for  $N$  sufficiently large and  $K \geq \kappa \log N$ ,*

$$P(\Gamma_K^N) = o_N(1).$$

*Proof.* By the union bound and the Gaussian tail inequality, we see that

$$P(\Gamma_L^N) \leq \sum_{l=1}^N \exp \left[ l \frac{\beta_c^2}{2} - \frac{1}{2l} \left( \frac{m_N}{N} l + K \right)^2 \right].$$

Now,

$$l \frac{\beta_c^2}{2} - \frac{1}{2l} \left( \frac{m_N}{N} l + K \right)^2 \leq \frac{3 \log N}{2} l - \left( \beta_c - \frac{3 \log N}{2\beta_c} \right) K$$

so that

$$P(\Gamma_K^N) \leq e^{-(\beta_c - \frac{3 \log N}{2\beta_c})K} \sum N^{\frac{3}{2} \frac{l}{N}} \leq N^{-\kappa(\beta_c - \frac{3 \log N}{2\beta_c})} N^{\frac{5}{2}}.$$

Thus provided

$$\kappa > \frac{5}{2 \left( \beta_c - \frac{3 \log N}{2\beta_c} \right)}$$

this is  $o_N(1)$ . The result follows by optimizing  $\frac{3 \log N}{2}$  in  $N$ . □

The next estimate we will need is a control on the probability that there is a walker that enters the strip  $\lambda l + [0, K]$  on a time of order  $tN$  and ends near the leader at time  $N$ , i.e. near  $m_N - x$  (recall the interpretation of  $m_N$  as the location of the leader from the beginning of this section). To this end, we define the event

$$\Xi_{\epsilon, K}^N = \{\exists v \in \partial \mathcal{T}_N, T \in N[\epsilon, 1 - \epsilon] : S_v(l) \leq \lambda l + K \forall l \in [N], S_v(T) \geq \lambda T, S_v(N) \geq m_N - x\} \tag{2.3}$$

which is the event that there is a leaf,  $v$ , whose corresponding walker,  $S_v$ , enters the window  $\lambda l + [0, K]$  on the time scale of  $N[\epsilon, 1 - \epsilon]$ . The probability of this event is bounded as follows.

**Lemma 2.6.** *For all  $x, K \in \mathbb{Z}$ ,  $x > 0, K \geq 1$ , and  $\epsilon \in (0, 1/2)$ , we have that*

$$P(\Xi_{\epsilon, K}) \lesssim_{\epsilon} \frac{K^4(x + K)^2}{N^{1/2}} e^{\lambda_N x}.$$

*Proof.* By a union bound

$$P(\Xi_{\epsilon,K}) \leq 2^N \sum_{T \geq \epsilon N}^{(1-\epsilon)N} P(S(l) \leq \lambda l + K \forall t \in [N], S(T) \geq \lambda T, S(N) \geq m_N - x). \quad (2.4)$$

The summand satisfies

$$\begin{aligned} &P(S(l) \leq \lambda l + K \forall t \in [N], S(T) \geq \lambda T, S(N) \geq m_N - x) \\ &\leq P(S(l) \leq \lambda l + K \forall l \in [N]; S(T) \in \lambda T + [0, K], S(N) \in m_N + [-x, K]) \\ &\leq \sum_{i=0}^{K-1} P(S(l) \leq \lambda l + K, \forall l \in [T]; S(T) \in \lambda T + [i, i+1]) \\ &\quad \cdot \max_{z \in [i, i+1]} P(S(l) \leq \lambda l + K, \forall l \in [T, N]; S(N) \in m_N + [-x, K] | S(T) = \lambda T + z). \end{aligned} \quad (2.5)$$

We now compute the multiplicands in the summand.

The first multiplicand can be controlled by the tilted barrier estimate (2.1) to get

$$P^0(S(l) \leq \lambda l + K \forall l \in [T]; S(T) \in \lambda T + [i, i+1]) \lesssim 2^{-T} e^{-\lambda i} \frac{K \cdot (K-i)}{T^{3/2}}. \quad (2.6)$$

To bound the second multiplicand, first let  $T' = N - T$ . Observe then that for all  $z \in [0, K-1]$ , we have  $K - z \geq 1$ , so that by the titled barrier estimate yields

$$\begin{aligned} &P(S(l) \leq \lambda l + K \forall l \in [T, N]; S(N) \in m_N + [-x, K] | S(T) = \lambda T + z) \\ &= P^z(S(l) \leq \lambda l + K \forall l \in [T']; S(T') \in \lambda T' + [-x, K]) \\ &\lesssim 2^{-T'} e^{\lambda(x+z)} \frac{(K-z)(x+K)^2}{(T')^{3/2}} \end{aligned} \quad (2.7)$$

Plugging (2.6)-(2.7) into (2.5) and plugging this into (2.4), yields the desired bound.  $\square$

Our last estimate (once combined with the above two estimates) shows that it is unlikely for there to be two walkers who branch on the genealogical time scale  $T$ —that is, that here are two  $v, w \in \partial T_N$  with  $|v \wedge w| = T$ —and both end only order 1 away from the leader  $m_N$  (recall again the interpretation of  $m_N$  from the beginning of the section). We make this precise in the following lemma and corollary.

First we fix a pair  $v, w \in \partial T_N$  with  $|v \wedge w| = T$  and bound the probability of this pathological event. To this end, define for such a pair  $v, w$  the event

$$E_\epsilon^N(T, K, v, w) = \{S_v(l), S_w(l) \leq \lambda l + K \forall l \in [N]; S_v(l), S_w(l) \leq \lambda l \forall l \in N[\epsilon, (1-\epsilon)]; S_v(N), S_w(N) \in m_N + [-x, K]\}. \quad (2.8)$$

This is the event that the walkers corresponding to  $v$  and  $w$  stay below the barrier  $\lambda l + K$  for all time, stay below the barrier  $\lambda l$  on the time scale  $tN$ , and end in the window  $[m_N - x, m_N + K]$ . We control the probability of this event as follows

**Lemma 2.7.** *Let  $v, w \in \partial T_N$  be such that  $R(v, w) = t, T = Nt, T' = N - T$ , and  $t' = 1 - t$ . Let  $K \geq 1x > 0$ . Then*

$$P(E_\epsilon^N(T, K)) \lesssim 2^{-(N+T')} \frac{e^{2\lambda_N x}}{N^{\frac{3}{2}(1+t)} t^{3/2} (1-t)^3} (x+K)^4 K^4$$

*Proof.* Observe that

$$P(E_\epsilon^N(T, K, v, w))$$

$$\leq \sum_{j=0}^{\infty} P(S(l) \leq \lambda l + K \forall l \in [T]; S_N(T) \in \lambda l + [-j - 1, -j]) \cdot \max_{z \in [-j-1, -j]} P(S(l) \leq \lambda l + K \forall l \in [T, N]; S(N) \in m_N + [-x, K] | S(T) = \lambda T + z)^2$$

We bound the two multiplicands by application of the tilted barrier estimate. Observe that the first multiplicand satisfies

$$P(S(l) \leq \lambda l + K \forall l \in [T]; S_N(T) \in \lambda l + [-j - 1, -j]) \lesssim 2^{-T} e^{\lambda \cdot j} \frac{K(K + j + 1)}{T^{3/2}}$$

and that the second satisfies

$$P(S(l) \leq \lambda l + K \forall T \in [T, N]; S(N) \in m_N + [-x, K] | S_v(T') = \lambda T + z) \lesssim 2^{-T'} e^{\lambda_N(x+z)} \frac{\max\{K + |z|, 1\}(x + K)^2}{T'^{3/2}}.$$

Here we used that  $K - z \geq 0$  (for us  $z \leq 0$ ). Combining the results then yields the desired estimate, namely

$$P(E_\epsilon^N(T, K, v, w)) \lesssim 2^{-(N+T')} \frac{e^{2\lambda x}}{N^{\frac{3}{2}(1+t)} t^{3/2} (1-t)^3} (x + K)^4 K^4. \quad \square$$

By an application of the union bound, we see that the previous estimate implies that it is rare for there to be any pairs of leaves  $|v \wedge w| = T$  that have this behavior.

**Corollary 2.8.** *Let  $v_T, w_T \in \partial T_N$  be a pair satisfying  $|v \wedge w| = T$ . Under the conditions of Lemma 2.7, we have that for  $\epsilon \in (0, \frac{1}{2})$*

$$\sum_{T \geq \epsilon N}^{(1-\epsilon)N} 2^{N+T'} P(E_\epsilon(T, K, v_T, w_T)) \lesssim_\epsilon \frac{e^{2\lambda x}}{N^{\frac{1}{2} + \frac{3}{2}\epsilon}} (x + K)^4 K^4.$$

### 2.3 Proof of Proposition 1.4

The proof of Proposition 1.4 will now follow immediately from an application of the following two estimates. The first estimate says that below the critical temperature, the Gibbs measure gives no mass to points that more than a large, but order 1, distance from the leader. This follows more or less immediately from the sub-gaussian tails of the increments and the tightness of the (centered) leader,  $M_N - m_N$ .

**Lemma 2.9.** *Let  $\beta > \beta_c$ . Then for each  $x$ ,*

$$\lim_{x \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{E} G_N(H \leq m_N - x) = 0.$$

*Proof.* Before we begin, we make the following useful definitions. For readability, we suppress dependence on  $N$  whenever it is unambiguous. Let

$$V_y^N = \{v \in \partial T_N : H_N(v) \in m_N - [y, y + 1]\}$$

and  $\mathcal{N}_y^N = \text{card}(V_y^N)$ . Choose  $\alpha = \alpha(\beta)$  such that  $\beta > \beta_c(1 + \alpha)$ . Let  $u_n = u_n(\alpha) = (1 + \alpha)^{-1} N^{1/2} 2^{-1/2}$ .

Now, recall from [18, Prop. 3.3] that for all  $y \geq 0$ ,

$$\mathbb{E} \mathcal{N}_y^N \lesssim N e^{-\beta_c y - \frac{y^2}{2N}}$$

and there is a universal constant  $c_1$  such that for all  $y, u$  with  $0 \leq u + y \leq \sqrt{N}$ ,  $u \geq -y$ ,

$$P(\mathcal{N}_y^N \geq e^{\beta_c(y+u)}) \lesssim e^{-\beta_c u + C \log_+(y+u)}.$$

Let  $x_0 = x_0(\alpha, \beta)$  be such that for all  $y \geq x_0$ ,  $c_1 \log((1 + \alpha)y) \leq \frac{\beta\alpha}{2}y$ . Then for all  $y \in [x, u_N]$ ,

$$P(\mathcal{N}_y^N \geq e^{\beta_c(1+\alpha)y}) \lesssim e^{-\frac{\beta_c\alpha}{2}y}.$$

This implies that

$$\mathbb{E}G_N(H_N \leq m_N - x) \leq \epsilon_1(K) + \mathbb{E}G_N(H_N \leq m_N - x) \mathbb{1}_{\mathcal{T}_K}.$$

Let

$$B_N^\alpha(x) = \left\{ \exists y \in [x, u_N] \cap \mathbb{N} : \mathcal{N}_y^N \geq e^{\beta_c(1+\alpha)y} \right\}.$$

Note that on the event  $\mathcal{M}_K$ , we have  $Z \geq e^{\beta m_N - \beta K}$  from which it follows that

$$\begin{aligned} \mathbb{E}G_N(H_N(v) \leq m_N - x) \mathbb{1}_{\mathcal{T}_K} &\leq P(B_N^\alpha(x)) + \mathbb{E} \sum_{y=x}^{\infty} \sum_{H \in V_y^N} \frac{e^{\beta H}}{Z} \mathbb{1}_{\mathcal{T}_K} \mathbb{1}_{B_N^\alpha(x)^c} \\ &\lesssim \sum_{y=x}^{u_N} e^{-\frac{\beta_c\alpha}{2}y} + e^{\beta K} \left( \sum_{y=x}^{u_N} e^{-(\beta - \beta_c(1+\alpha))y} + \sum_{u_N}^{\infty} N e^{-(\beta - \beta_c)y} \right) \\ &\lesssim_{\alpha, \beta} e^{-c(\alpha)x} + e^{\beta K} \left( e^{-c'(\alpha)x} + N e^{-c''(\beta)u_N} \right). \end{aligned}$$

Putting these together yields

$$\limsup_{N \rightarrow \infty} \mathbb{E}G_N(H \leq m_N - x) \lesssim_{\alpha, \beta} \epsilon_1(K) + e^{-c(\alpha)x} + e^{\beta K - c'(\alpha)x}.$$

Since  $K$  was arbitrary, we may take  $K = \frac{c'(\alpha)x}{2\beta}$ , and the result follows.  $\square$

We now show that it is unlikely to have two points that are both order 1 away from the leader, but overlap strictly in  $(0, 1)$ . The idea of this estimate is that for this to happen there must be two walkers,  $S_{v_1}$  and  $S_{v_2}$ , whose branching time is of order  $N$  and both land order 1 away from the leader. This event is rare by the above.

**Lemma 2.10.** *Let  $\beta > \beta_c$ . For all  $x$  and all  $A \subsetneq (0, 1)$ , we have that*

$$\limsup \mathbb{E}G^{\otimes 2}(R_{12} \in A; H(v_1), H(v_2) \geq m_N - x) = 0.$$

*Proof.* Recall the events  $\Gamma_L$ ,  $\Xi_{L, \epsilon}^N$ , and  $E_\epsilon^N(T, L, v_T, w_T)$  from (2.2)(2.3), and (2.8) respectively where  $v_T, w_T$  are as per Corollary 2.8. Observe that

$$\begin{aligned} \mathbb{E}G_N^{\otimes 2}(R_{12} \in A; H_N(v_1), H_N(v_2) \geq m_N - x) &\leq P(\Gamma_L^N) + P(\Xi_{L, \epsilon}^N) + \mathbb{E}G_N^{\otimes 2}(R_{12} \in A; H_N(v_1), H_N(v_2) \geq m_N - x) \mathbb{1}_{(\Gamma_L^N)^c, (\Xi_{L, \epsilon}^N)^c} \\ &\leq P(\Gamma_L^N) + P(\Xi_{L, \epsilon}^N) + \mathbb{E} \sum_{R_{12} \in (\epsilon, 1 - \epsilon)} \mathbb{1}_{H_N(v_1), H_N(v_2) \in m_N + [-x, K]} \mathbb{1}_{(\Gamma_L^N)^c, (\Xi_{L, \epsilon}^N)^c} \\ &\leq P(\Gamma_L^N) + P(\Xi_{L, \epsilon}^N) + \sum_{T \geq \epsilon N}^{(1-\epsilon)N} 2^{N+T} P(E_\epsilon^N(T, L, v_T, w_T)) \\ &\lesssim_\epsilon P(\Gamma_L^N) + \frac{L^4(x+L)}{N^{1/2}} e^{\lambda_N x} + \frac{e^{2\lambda_N x}}{N^{\frac{1}{2} + \frac{3}{2}\epsilon}} (x+L)^4 L^4 \end{aligned}$$

where the last inequality follows by Lemma 2.6 and Corollary 2.8. Choosing  $L = c \log N$  for  $c$  large enough, and sending  $N \rightarrow \infty$  then yields the result after applying Lemma 2.6.  $\square$

We now prove Proposition 1.4.

*Proof of Proposition 1.4.* By [15], it suffices to take  $\beta > \beta_c$ . Note that it suffices to show that  $\mu_N(\epsilon, 1 - \epsilon) \rightarrow 0$ . Let  $A = (\epsilon, 1 - \epsilon)$ . Observe that

$$\begin{aligned} \mu_N(A) &= \mathbb{E}G_N^{\otimes 2}(R_{12} \in A; H(v_1) \wedge H(v_2) \leq m_N - x) \\ &\quad + \mathbb{E}G_N^{\otimes 2}(R_{12} \in A; H(v_1), H(v_2) \geq m_N - x) \\ &\leq 2\mathbb{E}G_N(H_N \leq m_N - x) + \mathbb{E}G_N^{\otimes 2}(R_{12} \in A; H(v_1), H(v_2) \geq m_N - x). \end{aligned}$$

The result then follows from Lemma 2.9 and Lemma 2.10 by taking  $N \rightarrow \infty$  and then  $x \rightarrow \infty$ . □

### 3 The Derrida-Spohn conjecture and the Ghirlanda-Guerra identities

In this section we prove the Derrida-Spohn conjecture and show that the Branching Random Walk satisfies the Ghirlanda-Guerra Identities.

#### 3.1 Derrida-Spohn conjecture

The proof of the Derrida-Spohn conjecture will follow immediately after the following technical preliminaries. Recall first the following result of Chauvin and Rouault.

**Proposition 3.1.** (Chauvin-Rouault [15]) *The free energy satisfies*

$$F(\beta) = \lim \frac{1}{N} \mathbb{E} \log Z_N = \begin{cases} \log 2 + \beta^2/2 & \beta < \beta_c \\ \beta_c \beta & \beta \geq \beta_c \end{cases}. \tag{3.1}$$

Recall the following integration-by-parts. Let  $\Sigma$  denote an at most countable set;  $(x(\sigma))_{\sigma \in \Sigma}$  and  $(y(\sigma))_{\sigma \in \Sigma}$  be centered Gaussian processes with uniformly bounded variances and mutual covariance  $C(\sigma^1, \sigma^2) = \mathbb{E}x(\sigma^1)y(\sigma^2)$ ;  $G'$  be a finite measure on  $\Sigma$ ;  $Z = \sum e^{y(\sigma)}$ ; and  $G(\sigma) = e^{y(\sigma)}G'(\sigma)/Z$ .

**Lemma 3.2.** [24, Lemma 2](Gibbs-Gaussian Integration-by-parts). *We have the identity*

$$\mathbb{E} \langle x(\sigma) \rangle_G = \mathbb{E} \langle C(\sigma^1, \sigma^1) - C(\sigma^1, \sigma^2) \rangle_{G^{\otimes 2}}.$$

Furthermore, for any bounded measurable  $f$  on  $\Sigma^n$ ,

$$\mathbb{E} \langle f(\sigma^1, \dots, \sigma^n)x(\sigma^1) \rangle_{G^{\otimes n}} = \mathbb{E} \left\langle f(\sigma^1, \dots, \sigma^n) \left( \sum_{k=1}^n C(\sigma^1, \sigma^k) - nC(\sigma^1, \sigma^{n+1}) \right) \right\rangle_{G^{\otimes n+1}}.$$

As a consequence we have the following

**Corollary 3.3.** *We have*

$$F'(\beta) = \beta \int (1 - x)d\mu$$

where  $\mu$  is a limit point of the mean overlap measure.

*Proof.* Notice that Lemma 3.2 gives

$$F'_N(\beta) = \frac{1}{N} \mathbb{E} \langle H_N \rangle = \beta \frac{1}{N} \mathbb{E} \langle NR_{11} - NR_{12} \rangle = \beta \mathbb{E} \langle 1 - R_{12} \rangle = \beta \int (1 - x)d\mu_{N,\beta}.$$

Where  $\mu_{N,\beta}$  is as per (1.1). Since  $F_N$  and  $F$  are convex and  $C^1$ , and  $F_N \rightarrow F$  point-wise on  $\mathbb{R}_+$ , we have that  $F'_N \rightarrow F'$ . This gives us the lefthand side of the desired equality. Furthermore, since  $f(x) = 1 - x$  is in  $C([0, 1])$ , and the sequence  $\mu_{N,\beta} \in \text{Pr}([0, 1])$  is necessarily tight, weak convergence applied to the last term yields the righthand side of the desired equality. □

Finally we observe that the limiting overlap distribution has  $\text{supp}\mu \subset \{0, 1\}$ . We now turn to the proof of Theorem 1.1.

*Proof of Theorem 1.1.* By Corollary 3.3 and Proposition 1.4, we know that for any such weak limit, we get

$$F'(\beta) = \beta \int (1 - x)d\mu = \beta m.$$

Differentiating (3.1), equating, and solving for  $m$ , we get

$$m = \begin{cases} 1 & \beta \leq \beta_c \\ \frac{\beta_c}{\beta} & \beta \geq \beta_c \end{cases}. \quad \square$$

### 3.2 Ghirlanda-Guerra identities

We now turn to the proof of the Ghirlanda-Guerra Identities for these models. We need the following preliminary lemmas. Observe that a standard application of Gaussian concentration yields the following.

**Lemma 3.4.** *The free energy concentrates about its mean:*

$$\mathbb{P}(|F_N - \mathbb{E}F_N| > \epsilon) \leq 2e^{-\frac{N\epsilon^2}{2\beta^2}}$$

As a consequence of this, we find the Gibbs measure concentrates around a fixed energy level to order  $N$ .

**Lemma 3.5.** *The intensive energy concentrates. In particular,*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \langle |H_N - \mathbb{E} \langle H_N \rangle| \rangle_\beta = 0$$

for each  $\beta$ .

*Proof.* The result then follows from Lemma 3.4 after a modification of the proof of [24, Theorem 3.8].  $\square$

We now turn to the proof of Theorem 1.2.

*Proof of Theorem 1.2.* Observe first that for  $\beta \leq \beta_c$ ,  $\mu = \delta_0$  by Theorem 1.1 so that the identities are trivial in this setting. It suffices to study  $\beta > \beta_c$ . Furthermore, the limiting overlap distribution is supported on  $\{0, 1\}$  by Proposition 1.4, so it suffices to show the Approximate Ghirlanda-Guerra Identities for  $p = 1$ , since  $R_{12}^p = R_{12}$  when  $R_{12} \in \{0, 1\}$ . The result then follows by a standard integration-by-parts argument.

Notice that if we apply Lemma 3.2 with  $\Sigma_N = \partial T_N$ ,  $x(\sigma) = H_N(\sigma)$ ,  $y(\sigma) = \beta H_N(\sigma)$ , and  $C(\sigma^1, \sigma^2) = \beta N R_{12}$  it follows that

$$\frac{1}{N} \mathbb{E} \langle f(R^n) H_N(\sigma^1) \rangle = \beta \mathbb{E} \left\langle f(R^n) \left( \sum_{k=1}^n R_{1k} - n R_{1,n+1} \right) \right\rangle$$

and

$$\frac{1}{\beta N} \mathbb{E} \langle H_N(\sigma^1) \rangle = \mathbb{E} \langle 1 - R_{12} \rangle.$$

As a result,

$$\begin{aligned} & \frac{1}{\beta N} \mathbb{E} \langle f(R^n) (H_N(\sigma^1) - \mathbb{E} \langle H_N(\sigma) \rangle) \rangle \\ &= \frac{1}{\beta N} (\mathbb{E} \langle f(R^n) H_N(\sigma^1) \rangle - \mathbb{E} \langle f(R^n) \rangle \mathbb{E} \langle H_N(\sigma) \rangle) \end{aligned}$$

$$\begin{aligned}
 &= \left( \mathbb{E} \langle f(R^n) \rangle + \sum_{k=2}^n \mathbb{E} \langle f(R^n) R_{1k} \rangle - n \mathbb{E} \langle f(R^n) R_{1,n+1} \rangle \right) \\
 &\quad - (\mathbb{E} \langle f(R^n) \rangle - \mathbb{E} \langle f(R^n) \rangle \mathbb{E} \langle R_{12} \rangle) \\
 &= \mathbb{E} \langle f(R^n) \rangle \mathbb{E} \langle R_{12} \rangle + \sum_{k=2}^n \mathbb{E} \langle f(R^n) R_{1k} \rangle - n \mathbb{E} \langle f(R^n) R_{1,n+1} \rangle.
 \end{aligned}$$

This implies that

$$\begin{aligned}
 & \left| \mathbb{E} \langle f(R^n) \cdot R_{1,n+1} \rangle - \frac{1}{n} \left( \mathbb{E} \langle f(R^n) \rangle \mathbb{E} \langle R_{12} \rangle + \sum_{k=2}^n \mathbb{E} \langle f(R^n) \cdot R_{1k} \rangle \right) \right| \\
 &= \frac{1}{\beta n N} \left| \mathbb{E} \langle f(R^n) (H_N - \mathbb{E} \langle H_N \rangle) \rangle \right| \\
 &\leq \|f\|_{L^\infty([0,1]^{n^2})} \frac{1}{\beta n N} \mathbb{E} \langle |H_N - \mathbb{E} \langle H_N \rangle| \rangle \rightarrow 0
 \end{aligned}$$

by Lemma 3.5. □

The Approximate Ghirlanda-Guerra identities have many deep consequences. We highlight one simple consequence regarding the limit of the overlap array distribution in these systems. Let  $Q_N^\beta$  be the overlap array distribution corresponding to  $G_N^\beta$ . As  $\{Q_N\}$  is a sequence of measures on the compact polish space  $[0, 1]^{N^2}$ , it is tight. Let  $Q^\beta$  be any limit point of this sequence. We will show that it is the unique limit point and is given by what is called a 1RSB Ruelle Probability Cascade which we define presently.

To this end, define  $RPC(\zeta)$  for  $\zeta = \theta\delta_0 + (1-\theta)\delta_1$  with  $\theta \in (0, 1]$  as follows. If  $\theta \in (0, 1)$ , let  $\{e_n\}$  be the standard basis for  $\ell_2$ , let  $(w_n)$  be the ranked points of a Poisson-Dirichlet,  $PD(\theta, 0)$  process [26], and let

$$G = \sum w_n \delta_{e_n}.$$

If  $\theta = 1$ , let  $G = \delta_0$ . Then  $RPC(\zeta)$  is the overlap array distribution induced by  $G$ . One important result to note is that  $RPC(\zeta)$  satisfies the Ghirlanda-Guerra identities [24, Section 2]. This is a consequence of a standard invariance property of the Poisson-Dirichlet process/Poisson point processes of Gumbel type.

What we will show now is that as a consequence of Theorem 1.2, we will have that  $Q = RPC(\mu_\beta)$ . This an immediate consequence of the characterization-by-invariance theory used in spin glasses [24, 28, Section 15.13]. In this setting, the proof is fairly elementary and does not require the full machinery of this theory. Furthermore, it illustrates some essential ideas for this method. For these reasons, and to make this presentation self-contained, we include the proof.

**Corollary 3.6.** *We have the limit  $Q_N^\beta \xrightarrow{(d)} Q^\beta$  where  $Q^\beta$  is the overlap distribution corresponding to  $RPC(\mu_\beta)$  where  $\mu_\beta$  is as in (1.2).*

*Proof.* (For readability, we drop the dependence on  $\beta$  and denote  $RPC(\zeta)$  by  $\tilde{Q}$ .) It suffices to check that

$$Q(R^n = A) = \tilde{Q}(R^n = A)$$

for every  $n \times n$  Gram matrix,  $A$ , with entries in  $\{0, 1\}$ . This is done by a direct computation, which reduces both sides to the same polynomial in  $\mu(R_{12} = 1)$ . Before, we begin the computation, observe that  $Q$  satisfies:

- Weak Exchangeability: for every  $\pi : \mathbb{N} \rightarrow \mathbb{N}$  a permutation of finitely many indices

$$Q((R_{ij})_{i,j \geq 1} \in A) = Q((R_{\pi(i)\pi(j)})_{i,j \geq 1} \in A),$$

- Ultrametricity:

$$Q(R_{12} \geq \min\{R_{13}, R_{23}\}) = 1,$$

- The Ghirlanda-Guerra identities: for every  $f$  and  $\psi$  bounded measurable, we have

$$\int f(R^n)\psi(R_{1,n+1})dQ = \frac{1}{n} \left[ \int f dQ \cdot \int \psi d\mu_\beta + \sum_{k=2}^n \int f(R^n)\psi(R_{1,k})dQ \right],$$

- The diagonal is almost surely 1:  $Q(R_{11} = 0) = 0$ .

The probability  $Q(R^n = A)$  is computed by recursively applying the following cases:

**Case 1.**  $n = 2$ .

The  $Q$ -probability of this event is either  $\mu_\beta(R_{12} = 1)$  or  $\mu_\beta(R_{12} = 0)$ .

**Case 2.**  $n \geq 3$  and  $A$  is not the identity.

By weak exchangeability, we can assume that  $A$  is block-diagonal, and that the blocks are arranged in decreasing size. Take the first block and suppose that it is of length  $m \geq 2$ . Let  $R^n(m)$  denote the  $n - 1 \times n - 1$  matrix obtained by deleting the  $m$ -th row and column of  $R^n$  and similarly for  $A(m)$ . By ultrametricity and the Ghirlanda-Guerra Identities combined with weak exchangeability, it follows that

$$\begin{aligned} Q(R^n = A) &= Q(R^n(m) = A(m), R_{1m} = 1) \\ &= \frac{1}{n} Q(R^{n-1} = A(m)) [\mu(R_{12} = 1) + m - 1] \end{aligned}$$

It thus suffices to compute  $Q(R^{n-1} = A(m))$ .

**Case 3.**  $n \geq 3$  and  $A = Id$ .

By ultrametricity, it follows that

$$Q(R^n = Id) = Q(R^{n-1} = Id, R_{1,n} = 0) - \sum_{k=2}^{n-1} Q(R^{n-1} = Id, R_{k,n} = 1)$$

To see why, note that if  $R^{n-1} = Id$  and  $R_{k,n} = 1$  for some  $k \neq n$ , then the remaining must all be zero. By the Ghirlanda-Guerra identities, the first event is

$$Q(R^{n-1} = Id, R_{1,n} = 0) = Q(R^{n-1} = Id) \left[ 1 - \frac{1}{n-1} Q(R_{12} = 1) \right].$$

The first term is then computed by Case 3. The second event is computed by Case 2.

By applying repeatedly these cases, the probability of any such event is reduced to a polynomial in  $\mu(R_{12} = 1)$ . Evidently, the same argument applied to  $\tilde{Q}$  yields same polynomials, and thus the desired result.  $\square$

We now make the following remarks regarding how this can be understood to imply certain modes of convergence for the Gibbs measures.

**Remark 3.7.** Observe that by the uniqueness portion of the Dovbysh-Sudakov theorem [23], this shows that the Gibbs measure  $G_N$  sampling converges to  $G$  as above in the sense of Austin [7].

**Remark 3.8.** We also observe that for  $\beta > \beta_c$  we can recover a result like that from [8] mentioned in the introduction. In particular, if we partition  $\partial T_N$  in to groups of leaves with overlap at least  $1 - \epsilon$ , call them  $(B_i^\epsilon)$ , then the ranked weights  $(G_{N,\beta}(B_i^\epsilon))$  converge in law to the ranked weights of  $PD(\frac{\beta_c}{\beta}, 0)$ . This follows from an approximation argument combined with Talagrand's identities (see, e.g., [20, Theorem 6.3.5] or [28, Section 15.4]).

## References

- [1] Addario-Berry, L. and Reed, B. A. (2008). *Ballot theorems for random walks with finite variance*, <http://www.problab.ca/louigi/papers/ballot.pdf>.
- [2] Addario-Berry, L. and Reed, B. (2009). Minima in branching random walks. *Ann. Probab.* **37**, 3, 1044–1079. MR-2537549
- [3] Arguin, L.-P., *Extrema of log-correlated random variables: Principles and Examples*, arXiv:math.PR/1601.00582 (2016).
- [4] Arguin, L.-P., Bovier, A., and Kistler, N. (2013). The extremal process of branching Brownian motion. *Probab. Theory Related Fields* **157**, 3-4, 535–574. MR-3129797
- [5] Arguin, L.-P. and Zindy, O. (2014). Poisson-Dirichlet statistics for the extremes of a log-correlated Gaussian field. *Ann. Appl. Probab.* **24**, 4, 1446–1481. MR-3211001
- [6] Arguin, L.-P. and Zindy, O. (2015). Poisson-Dirichlet statistics for the extremes of the two-dimensional discrete Gaussian free field. *Electron. J. Probab.* **20**, no. 59, 19. MR-3354619
- [7] Austin, T. (2015). Exchangeable random measures. *Ann. Inst. Henri Poincaré Probab. Stat.* **51**, 3, 842–861. MR-3365963
- [8] Barral, J., Rhodes, R., and Vargas, V. (2012). Limiting laws of supercritical branching random walks. *C. R. Math. Acad. Sci. Paris* **350**, 9-10, 535–538. MR-2929063
- [9] Bolthausen, E. and Sznitman, A.-S. (2002). *Ten lectures on random media*. DMV Seminar, Vol. **32**. Birkhäuser Verlag, Basel. MR-1890289
- [10] Bovier, A., *Statistical mechanics of disordered systems*, Cambridge, 2012.
- [11] Bovier, A. and Kurkova, I. (2004). Derrida's generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.* **40**, 4, 439–480. MR-2070334
- [12] Bovier, A. and Kurkova, I. (2004). Derrida's generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.* **40**, 4, 481–495. MR-2070335
- [13] Bramson, M. D., Ding, J., and Zeitouni, O., Convergence in law of the maximum of nonlattice branching random walk, *Ann. Inst. H. Poincaré Probab. Statist.* (to appear). arXiv:math.PR/1404.3423
- [14] Bramson, M. D. (1978). Maximal displacement of branching Brownian motion. *Comm. Pure Appl. Math.* **31**, 5, 531–581. MR-0494541
- [15] Chauvin, B. and Rouault, A. (1997). Boltzmann-Gibbs weights in the branching random walk. In *Classical and modern branching processes (Minneapolis, MN, 1994)*. IMA Vol. Math. Appl., Vol. **84**. Springer, New York, 41–50. MR-1601693
- [16] Contucci, P., Mingione, E., and Starr, S. (2013). Factorization properties in  $d$ -dimensional spin glasses. Rigorous results and some perspectives. *J. Stat. Phys.* **151**, 5, 809–829. MR-3055377
- [17] Derrida, B. and Spohn, H. (1988). Polymers on disordered trees, spin glasses, and traveling waves. *J. Statist. Phys.* **51**, 5-6, 817–840. New directions in statistical mechanics (Santa Barbara, CA, 1987). MR-0971033
- [18] Ding, J. and Zeitouni, O. (2014). Extreme values for two-dimensional discrete Gaussian free field. *Ann. Probab.* **42**, 4, 1480–1515. MR-3262484
- [19] Ghirlanda, S. and Guerra, F. (1998). General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A* **31**, 46, 9149–9155. MR-1662161
- [20] Jagannath, A. (2014). *Approximate Ultrametricity for Random Measures and Applications to Spin Glasses*, arXiv:math.PR/1412.7076.

## On the overlap distribution of Branching Random Walks

- [21] Madaule, T. (2015). Convergence in law for the branching random walk seen from its tip, *Journal of Theoretical Probability* 1–37 (English).
- [22] Mézard, M., Parisi, G., and Virasoro, M. A. (1987). *Spin glass theory and beyond*. World Scientific Lecture Notes in Physics, Vol. **9**. World Scientific Publishing Co., Inc., Teaneck, NJ. MR-1026102
- [23] Panchenko, D. (2010). On the Dovbysh-Sudakov representation result. *Electron. Commun. Probab.* **15**, 330–338. MR-2679002
- [24] Panchenko, D. (2013). *The Sherrington-Kirkpatrick model*. Springer Monographs in Mathematics. Springer, New York. MR-3052333
- [25] Panchenko, D. (2014). Structure of 1-RSB asymptotic Gibbs measures in the diluted  $p$ -spin models. *J. Stat. Phys.* **155**, 1, 1–22. MR-3180967
- [26] Pitman, J. and Yor, M. (1997). The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25**, 2, 855–900. MR-1434129
- [27] Subag, E. and Zeitouni, O. (2015). Freezing and decorated Poisson point processes. *Comm. Math. Phys.* **337**, 1, 55–92. MR-3324155
- [28] Talagrand, M. (2011). *Mean field models for spin glasses. Volume II*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], Vol. **55**. Springer, Heidelberg. Advanced replica-symmetry and low temperature. MR-3024566