Boltzmann-Gibbs weights in the Branching Random Walk

B. CHAUVIN, A. ROUAULT

Université de Versailles-Saint-Quentin

Abstract

Considering a branching random walk as a tree model for many physical disordered systems, the a.s. convergence of the free energy is proved under minimal assumption (finite mean) on the partition function. The overlap of two nodes in the tree is their last common ancestor (or the common part of their branches). Under a "k log k-type" assumption, the overlap of two nodes of height n , picked up with Boltzmann-Gibbs weights is proved to have an explicit limit distribution. This extends a result of Joffe and simplify a proof of Derrida and Spohn.

1980 *AMS Classification* (1991 *Revision*) . Primary : 60J80, 60K35 . Secondary : 60F10, 82C41 .

Key words : Branching random walk, Boltzmann-Gibbs weights, disordered trees.

1. Introduction

The motivation for this paper comes both from statistical physics and branching processes. The physical reference is the paper by Derrida and Spohn ([DS]) where appears a tree model for disordered systems and also many well-known probabilistic quantities. On the other side, recent results for the branching random walk (Joffe, ([J2])) bring in the *overlap* of two nodes picked up in the n-th generation of a Galton-Watson tree. The overlap has a physical meaning too and is studied by physicists ([DS], [MPV]).

That is why, starting from a few physical motivations, we consider a branching random walk and try to work towards two objectives . The first one is to improve (or simply prove) physicists' results or conjectures (theorem 2) and the second one is to weaken the assumptions , mainly replacing the binary (or m-ary) trees used in spin glasses models by random trees satisfying $k \log k$ type assumptions or even first moment assumptions (theorem 1).

The paper is organized as follows : after preliminaries setting a branching random walk as a physical model, the asymptotics of the free energy is precised, under minimal assumptions. Then we turn out to the overlap in Section 3 where we prove a convergence result at high temperature, using a Boltzmann-Gibbs measure on the boundary of the tree.

2. Preliminaries : the tree model

Disordered systems, such as spin glasses and directed polymers in a random medium are considered ([MPV], [EA], [D2]). In the first family of models, each point i of a lattice has a spin $u(i)$ which is $+1$ or -1 valued so that for a N points lattice, a spin configuration u is an element of $\{-1,+1\}^N$. The dependencies between the points are modelled by J_{ij}, the intensity of the coupling between i and j and the energy of a spin configuration is

$$E_u = \sum_{i,j} J_{ij}\ u(i)\ u(j).$$

The disorder is introduced taking the J_{ij} as random variables. In the SK model introduced by Sherrington and Kirkpatrick (1975) the J_{ij} are i.i.d. gaussian random variables . In this model, the energies E_u are then sums of i.i.d. gaussian variables.

In a similar manner, in the case of directed polymers, i.i.d. energies are assigned to each bond (i,j) of a lattice and each path u has an energy E_u given by the sum of the energies of visited bonds.

The behaviour of the system is described by the Gibbs measure on the configurations

$$\rho_N(u) = \frac{1}{Z_N(\beta)} \exp(-\beta E_u)$$

where β is the inverse temperature and $Z_N(\beta)$ is a normalisation constant. Z_N is called the partition function and its asymptotic behaviour, as $N \to +\infty$ gives the phase transitions of the system.

Nevertheless, the SK model is hard to deal with, because of very general dependencies between the energies E_u and E_v of two configurations u and v . So Derrida ([D1]) introduced the REM (random energy model) which is easier since all the energies $((E_u,\ u \in$ space of configurations) are independent. Although the REM is useful (for intuition, comparisons and first results see [D1]), it clearly contains too much independence. Derrida and Spohn in [DS] proposed the regular binary tree as an intermediate way. They gave results on the partition function and on the overlap (the common part) of configurations.

Actually, we can extend the whole problem to random trees such as Galton-Watson trees. So, the complete model (tree + energies) will be described by a branching random walk.

The branching random walk has been defined in the papers of Biggins ([B1],[B2],[B3]). Start with a single initial ancestor at the origin. The offspring point process Z gives at once the number of children of the ancestor and their positions in $\mathbb{R}$. Each of these has chidren with positions relative to their parent given by independent copies of Z, and so on.

We assume that the offspring population $Z(\mathbb{R})$ is a.s. different from 0 and that the offspring mean $EZ(\mathbb{R}) > 1$. The underlying Galton-Watson process is then supercritical without extinction. Call z_n the n-th generation.

To emphasize the correspondence with physical models, let us look at branching random walk as Galton-Watson tree (in the terminology of Neveu ([N1])) where branches are marked by additive $\mathbb{R}$-valued jumps (and time is discrete). Recall that a tree is a subset ω of the set

$$U = \{\emptyset\} \cup \bigcup_{n \in \mathbb{N}^*} (\mathbb{N}^*)^n$$

of finite sequences of integers ($\mathbb{N} = \{0, 1, ...\}$ and $\mathbb{N}^* = \{1, 2...\}$), such that

. $\emptyset \in \omega$,
. $uv \in \omega \Rightarrow u \in \omega$ (the sequences are denoted by concatenation),
. $u \in \omega \ \Rightarrow$ there is an integer $N^u(\omega)$ such that $(uj \in \omega, j \in \mathbb{N}^*) \ \Leftrightarrow \ (1 \leq j \leq N^u(\omega))$

($N^u(\omega)$ denotes the number of children of u) The length of a sequence u in U is denoted by $|u|$. So $z_n = \{u \in \omega : \ \ |u| = n\}$.

According to this model, the previous spin configurations are now the nodes u of the tree (or the branch coming from the ancestor to node u), i.i.d. energies are now jumps along edges of the tree (these jumps are not identically distributed in general). The energy of a configuration u is its position X_u. That gives the partition function:

$$Z_n(\beta) = \sum_{|u|=n} \exp(-\beta X_u).$$

In the infinite temperature case $\beta = 0$, the branching random walk reduces to the Galton-Watson process, and $Z_n(0) = Z_n$ is the number of the individuals in the n-th generation. Notice that large deviations results on the branching random walk can be translated on the ground state energy (which is the minimal energy of all the configurations) ; indeed this is the leftmost position of the particles in the n-th generation, denoted "probabilistically" by L_n. An important probabilistic literature is devoted to $Z_n(\beta)$ and L_n. To define the main tool, let $\lambda = EZ$ be the intensity measure of Z and

$$m(\beta) = \int_{\mathbb{R}} e^{-\beta x} \lambda(dx)$$

be the Laplace transform of λ. We assume all along this paper that

$$(H_0) \qquad m(\beta) \ \ is \ \ finite \ \ for \ \ every \ \ \beta \in \mathbb{R}.$$

The logLaplace transform is denoted by

$$l(\beta) = \log m(\beta).$$

It is classical to introduce

$$W_n(\beta) = \frac{Z_n(\beta)}{m(\beta)^n} = \sum_{|u|=n} \exp(-\beta X_u - nl(\beta)),$$

which is a martingale with respect to the filtration $(\mathcal{F}_n, n > 0)$ where $\mathcal{F}_n$ is the $\sigma-$algebra generated by variables N^u and X_u indexed by nodes u with length less or equal to n .

Let us recall here the only results useful for our purpose. They can be found in Biggins ([B1], [B2], [B3]).

There exists two critical constants $\widetilde{\beta}_c$ and β_c such that

$$(1.1) \qquad \begin{cases} \beta \, l'(\beta) - l(\beta) > 0 \quad , \quad \beta < \widetilde{\beta}_c \text{ or } \beta > \beta_c \\ \beta \, l'(\beta) - l(\beta) < 0 \quad , \quad \widetilde{\beta}_c < \beta < \beta_c \end{cases}$$

Noticing that an additive shift on jumps yields a multiplicative shift on $Z_n(\beta)$ and does not affect its behaviour, we may assume without loss of generality that the displacements are centered so that $l'(0) = 0$ and $\widetilde{\beta}_c < 0 < \beta_c$.

In the following, since β is an inverse temperature, some results are expressed for the high temperature regime, $0 \leq \beta < \beta_c$, but actually they hold for $\widetilde{\beta}_c < \beta < \beta_c$. Those for the low temperature regime $\beta > \beta_c$ hold also for $\beta < \widetilde{\beta}_c$.

The main theorem ruling the behaviour of the martingale is the following.

Theorem A. *(Biggins (1977), [B1])*
(i) for $0 \le \beta < \beta_c$, *if*

$$(H_1) \qquad E(Z_1(\beta)\log^+ Z_1(\beta)) < \infty,$$

then a.s. and in L^1,

$$W_n(\beta) \longrightarrow W(\beta), \quad n \to +\infty$$

(ii) for $\beta \ge \beta_c$, a.s.

$$W_n(\beta) \longrightarrow 0, \quad n \to +\infty.$$

The ground state energy is first analyzed with the

Theorem B. *(Biggins (1977), [B2])*

$$\frac{L_n}{n} \stackrel{n\to+\infty}{\longrightarrow} -l'(\beta_c) \ \ a.s. \tag{1.2}$$

Further results can be found in Bramson ([Br1], [Br2]) and Dekking ([Dek]).

Let us come back to the asymptotic behaviour of $Z_n(\beta)$ when $n \to \infty$, and more precisely of the free energy per volume unit:

$$F_n(\beta) := -\frac{1}{n\beta}\log Z_n(\beta).$$

Theorem 1. *Under the first moment assumption* (H_0) *then,*

$$F_n(\beta) \stackrel{n\to\infty}{\longrightarrow} -\frac{l(\beta)}{\beta} \ \ a.s \ \ if \ \ \beta < \beta_c \tag{1.3}$$

$$F_n(\beta) \stackrel{n\to\infty}{\longrightarrow} -\frac{l(\beta_c)}{\beta_c} = -l'(\beta_c) \ \ a.s \ \ if \ \ \beta \ge \beta_c \tag{1.4}$$

Remark : Notice that the high temperature result is immediate for the m-ary (also called Cayley) trees used by the physicists, since in this case (H_1) is obviously satisfy and then (1.3) comes from theorem A, taking logarithms.

Proof : First, let us see how (1. 4) follows from (1. 3). Notice that $\beta \mapsto F_n(\beta)$ is increasing ($[BPP]$) so that for $\beta > \beta_c$ and ε small enough,

$$\liminf_n F_n(\beta) \ge \liminf_n F_n(\beta_c - \varepsilon) = -\frac{1}{\beta_c - \varepsilon} l(\beta_c - \varepsilon)$$

and by continuity,

$$\liminf_n F_n(\beta) \ge -\frac{l(\beta_c)}{\beta_c}.$$

Moreover, in view of (1. 2),

$$F_n(\beta) \le -\frac{1}{n\beta}\log e^{-\beta L_n} = \frac{L_n}{n} \longrightarrow -\frac{l(\beta_c)}{\beta_c} \qquad a.s.,$$

and (1. 4) is proved as soon as (1. 3) holds.

The a.s. lower bound in (1. 3) follows from

$$P(F_n(\beta) < -\frac{l(\beta)}{\beta} - \delta) = P(W_n(\beta) > e^{n\delta\beta}) \le e^{-n\delta\beta}$$

and the Borel Cantelli lemma.

For the a.s. upper bound, we use a standard truncation argument (Kingman [K], Biggins). Fix N and erase all the subtrees arising from nodes u such that $Z_1(\beta) \circ \Theta^u > N$ (where Θ^u denotes the shifted tree where u is the ancestor). Denote this modification by an index N. This provides a branching random walk satisfying $E(Z_1^N(\beta))^2 < \infty$. The martingale

$$W_n^N(\beta) := m_N(\beta)^{-n} Z_n^N(\beta)$$

is uniformly integrable and converges $a.s.$ to $W^N(\beta)$ with

$$P(W^N(\beta) > 0) = P(S_N)$$

where S_N is the survival set of the modified process.
For N fixed, $a.s.$ on S_N,

$$F_n^N(\beta) \overset{n\to\infty}{\longrightarrow} -\frac{l^N(\beta)}{\beta}$$

and since for any n and N,

$$F_n(\beta) \leq F_n^N(\beta)$$

we conclude that a.s. on S_N,

$$\limsup_n \; F_n(\beta) \leq \; -\frac{l^N(\beta)}{\beta}.$$

The extinction probability s_N is the smallest solution of $g_N(s) = s$, where

$$g_N(s) = E(s^{Z_1(0)}; Z_1(\beta) \leq N) + P(Z_1(\beta) > N).$$

Since

$$0 \leq g_N(s) - g(s) = E(1 - s^{Z_1(0)} \; ; \; Z_1(\beta) > N) \leq P(Z_1(\beta) > N),$$

g_N converges to g uniformly as $N \to +\infty$ and

$$s_N - g(s_N) = g_N(s_N) - g(s_N) \to 0.$$

Since s_N is decreasing we get that s_N goes to zero. Now it is easy to see that a.s. 1_{S_N} tends to 1 and that (by dominated convergence) :

$$l_N(\beta) = \log E(Z_1^N(\beta)) = \log E(Z_1(\beta); Z_1(\beta) \leq N) \longrightarrow l(\beta),$$

so we may conclude that a.s. :

$$\limsup_n F_n(\beta) \leq -\frac{l(\beta)}{\beta}$$

and this ends the proof of theorem 1.

3. Overlaps

Let us focus in this section on the overlap of two nodes of a tree. The boundary $\partial\omega$ of a tree ω is the set of infinite sequences $\mathbf{i}$ in

$$I = (\mathbb{N}^*)^{\mathbb{N}^*}$$

such that $\mathbf{i}|n \in \omega, \forall n \in \mathbb{N}^*$ (where $\mathbf{i}|n$ denotes the beginning of $\mathbf{i}$ with length n, i.e.

$$\mathbf{i} = i_1 i_2 ... i_n ... \ \Rightarrow \ \mathbf{i}|n = i_1 i_2 ... i_n \quad).$$

For $u \in \omega$ let

$$B(u) = \{\mathbf{j} \in \partial\omega : \mathbf{j}||u| = u\}$$

The balls $B(u), u \in \omega$ define a metrizable topology on $\partial\omega$ ([Li], [LPP]).

Definition. *For two nodes u and v of a tree ω the* <u>*last common ancestor*</u> *of u and v is denoted by $u \wedge v$. The* <u>*overlap*</u> *of u and v is its generation number $|u \wedge v|$. The same notations are available for $\mathbf{i}$ and $\mathbf{j}$ in $\partial\omega$.*

In a Galton-Watson tree, if u and v are two nodes picked up uniformly in the n-th generation, then the overlap depends on n and the question arises about its asymptotics. In the seventies Joffe ([J1] see also [JM] and [O'B]) has shown the convergence of the empirical distribution of the overlap under a L^1 assumption. Related results can be found in [Bu].

The challenge is now to see what happens in the spatial case for the branching random walk : the overlap is defined in the same way but its asymptotics will be understood via the *Boltzmann-Gibbs measure* which will be defined as below.

For $u \in \omega$ let Θ^u be the (sub)tree shifted at node u . Variables with an upper index u will denote variables shifted by Θ^u. The branching property yields (under H_1):

$$m(\beta)^k \ W(\beta) = \sum_{|u|=k} e^{-\beta X_u} \ W^u(\beta), \tag{3.1}$$

allowing the following definition.

Definition. *The Boltzmann-Gibbs (B-G) measure μ_β is the unique (random) measure defined on the boundary $\partial\omega$ of a tree ω by its value on the balls B_u*

$$\mu_\beta(B_u) = \frac{W^u(\beta)}{W(\beta)} \ e^{-\beta X_u - |u| l(\beta)}. \tag{3.2}$$

In the Galton-Watson case ($\beta = 0$) the B-G measure reduces to the *uniform measure* used by Lyons and others ([L], [LPP], [LR]).

Now let us try to see the B-G measure as a limit, just like in the Galton-Watson case the uniform measure on the boundary of the tree is the limit of the uniform measure on the n-th generation.

For every $\mathbf{i}$ in the boundary $\partial\omega$ let

$$\mu_\beta^n(B_{\mathbf{i}|n}) = \frac{1}{Z_n(\beta)} \ e^{-\beta X_{\mathbf{i}|n}}$$

$$\mu_\beta^n(B_{\mathbf{i}|k}) = \frac{1}{Z_n(\beta)} \ e^{-\beta X_{\mathbf{i}|k}} \ Z_{n-k}(\beta) \circ \Theta^{\mathbf{i}|k} \quad for \quad k < n$$

$$\mu_\beta^n(B_{\mathbf{i}|k}) = 0 \ , \qquad \mathbf{i}|k \neq (\mathbf{i}|n)11...1$$

$$= \mu_\beta^n(B_{\mathbf{i}|n}), \quad \mathbf{i}|k = (\mathbf{i}|n)11...1 \quad for \ k > n$$

This implies in particular:

$$\mu_\beta^n \ _{|\mathcal{F}_n} = \frac{1}{Z_n(\beta)} \sum_{|u|=n} e^{-\beta X_u} \ \delta_u$$

The interpretation of the (random) probability distribution μ_β^n is easy: Pick up an individual u in the n-th generation with Boltzmann-Gibbs weights and extend its branch downstream according to a prescribed rule: $u11111...$.

Theorem 2. *For $\beta < \beta_c$, under assumption (H_1), for a.e. ω,*

$$\mu_\beta^n \Longrightarrow \mu_\beta, \qquad \textit{when } n \to +\infty. \tag{3.3}$$

The proof is straightforward from theorem A and formula (3.1).

Moreover Liu and Rouault ([LR] theorem 6) proved that for $\beta < \beta_c$ and under (H_1), μ_β is non atomic. As an immediate consequence :

$$\mathbf{i} \wedge \mathbf{j} < \infty \ , \qquad \mu_\beta \otimes \mu_\beta \quad a.s., \tag{3.4}$$

and also,

Corollary 3. *Under assumption (H_1) and for $\beta < \beta_c$,*

$$\mu_\beta^n \otimes \mu_\beta^n \ (\mathbf{i} \wedge \mathbf{j} = a) \ \overset{n\to\infty}{\longrightarrow} \ \mu_\beta \otimes \mu_\beta \ (\mathbf{i} \wedge \mathbf{j} = a)$$

in other words, the last common ancestor of two individuals picked up in the n-th generation with B-G weights has an explicit limit distribution when $n \to \infty$:

$$\begin{aligned} \mu_\beta \otimes \mu_\beta \ (\mathbf{i} \wedge \mathbf{j} = a) &= \sum_{b \neq c \in z_1 \circ \Theta^a} \mu_\beta(B_b) \ \mu_\beta(B_c) \\ &= \Big(\mu_\beta(B_a)\Big)^2 - \sum_{b \in z_1 \circ \Theta^a} \Big(\mu_\beta(B_b)\Big)^2. \end{aligned}$$

Finally the following remarks mainly give relations between the literature and the above results.

Remarks

1. The convergence of the overlap in the Galton-Watson case ($\beta = 0$) holds without any "$k \ \log k$" assumption. The same strong result for the spatial high remperature case is a conjecture. A first step could be to extend the Seneta-Heyde theorem to the spatial case.

2. Instead of picking up two individuals in the n-th generation with B-G weights, it is possible to pick up three, four, ... individuals and to extend the above results.

3. In their paper, Derrida and Spohn ([DS]) show at high temperature that the empirical distribution of the relative overlap $\frac{|u \wedge v|}{n}$ (for u and v picked up in the n-th generation with B-G weights) converges to δ_0 as $n \to \infty$. For us it is actually a straightforward consequence of Corollary 3. Their proof is intricate but the method is quite interesting : they consider the continuous case (the branching random walk is replaced by a branching brownian motion) and let

$$u(t,x) = E\big(e^{-e^{-\beta x} \ Z_t(\beta)}\big)$$

which is a solution of the so-called K-P-P equation

$$\left\{ \begin{array}{lcl} \frac{\partial u}{\partial t} &=& \frac{1}{2}\frac{\partial^2 u}{\partial x^2} \quad + \quad u^2 - u \\ \\ u(0,x) &=& e^{-e^{-\beta x}}. \end{array} \right. \tag{E}$$

The asymptotic behaviour of the partition function $Z_t(\beta)$ and of the relative overlap are deduced from the asymptotics of $u(t,x)$ for large t and large x.

At low temperature $\beta > \beta_c$, they don't have a quite rigorous result but the conjecture can be expressed :

$$\mu_\beta^n \otimes \mu_\beta^n \left(\frac{|\mathbf{i} \wedge \mathbf{j}|}{n} \in A \right) \ \overset{n\to\infty}{\longrightarrow} \ \mathcal{B}(A),$$

where $\mathcal{B}$ is a mixing of Bernoulli distributions.

Acknowledgements

The authors would like to thank John Biggins and Bernard Derrida for valuable discussions. They are also indebted to the organisers of the workshop "Classical and Modern Branching Processes" at the IMA for the final form of this paper.

References

[AH] Asmussen, S. and Hering, H. Branching processes. Birkhauser, (1983).

[B1] Biggins, J. D. Martingale convergence in the branching random walk. *J. Appl. Prob.*, **14**, 25 - 37, (1977).

[B2] Biggins J. D. Chernoff theorem in the branching random walk. *J. Appl. Prob.*, **14**, 630 - 636, (1977).

[B3] Biggins, J. D. Growth rates in the branching random walk. *Z. Wahr. und Verw. Gebeite*, **48**, 17 - 34, (1979).

[Br1] Bramson M. Minimal displacement of branching random walk. *Zeit. Wahr.* **45**, 89-108 (1978).

[Br2] Bramson M. The convergence of solutions of the Kolmogorov nonlinear diffusion equation to travelling waves. *Memoirs Am. Math. Soc.*, **44**, 1-190 (1983).

[BPP] Buffet, E. Patrick, A. and Pulé, J. V. Directed polymers on trees : a martingale approach. *J. Phys. A. Math. Gen.* 1823-1834, (1993).

[Bu] Buhler, W. J. Generations and Degree of Relationship in Supercritical Markov Branching Processes. *Z. Wahr. und Verw. Gebeite*, **18**, 141 - 152, (1971).

[C] Chauvin, B. Arbres et processus de branchements. Thèse de Troisième Cycle Univ. Paris VI, (1985).

[CR] Chauvin, B. and Rouault, A. KPP equation and supercritical branching brownian motion in subcritical areas. Application to spatial trees. *Probab. Th. Rel. Fields*, **80**, 299 - 314, (1988).

[Dek] Dekking F.M. On the shape of the wavefront of branching random walk. *Preprint* (1994).

[D1] Derrida, B. Random-energy model : an exactly solvable model of disordered systems. *Phys. Rev.* **B24** 2613-1626 (1981).

[D2] Derrida, B. Directed polymers in a random medium. *Physica* **A163**, 71-84 (1990).

[DS] Derrida, B. and Spohn, H. Polymers on Disordered Trees, Spin Glasses and Travelling waves. *J. Stat. Phys.*, **51**, 817 - 840, (1988).

[EA] Edwards, S.F., Anderson, P.W. : Theory of spin glasses. *J. Phys.* **F5**, 965-974 (1975).

[J1] Joffe, A. Remarks on the structure of trees with applications to supercritical Galton-Watson processes. *In Advances in Prob.*, **5**, ed. A. Joffe and P. Ney, Dekker, New-York , 263 - 268, (1978).

[J2] Joffe, A. A new martingale in branching random walk. *Annals of Appl. Probab.*, **3**, 1145-1150, (1993).

[JM] Joffe, A. and Moncayo, A. Random variables, Trees and Branching Ramdom Walks. *Adv. in Math.*, **10**, 401 - 416, (1973).

[K] Kingman, J. F. C. The first birth problem for an age-dependent branching process. *Annals of Probab.*, **3**, 790 - 801, (1975).

[Li] Liu Q. Flows in networks and Hausdorff measures associated with applications to fractal sets in Euclidian space. Preprint Université Paris VI, (1993).

[LR] Liu Q. and Rouault A. On two measures defined on the boundary of a branching tree. *IMA Symposium on Classical and Modern Branching Processes.* (This volume)

[L] Lyons R. Probability and trees. *Preprint* (1994).

[LPP] Lyons R., Pemantle R., Peres Y. Ergodic theory on Galton-Watson trees, I and II. *Ergodic Theory Dynamical Systems*, to appear (1994).

[MPV] Mézard, M., Parisi, G. and Virasoro, M. Spin glass theory and beyond. World Scientific. Singapore. New Jersey, (1987).

[N1] Neveu, J. Arbres et processus de Galton-Watson. *Ann. IHP.*, **22**, 199 - 208, (1986).

[N2] Neveu, J. Multiplicative martingales for spatial branching processes. *Seminar. Stoch. Proc. Princeton*, 233 - 243, Birkhauser (London), (1987).

[0'B] O'Brien, G. L. A limit theorem for simple maxima and heavy branches in Galton-Watson trees. *J. Appl. Prob.*, **17**, 539 - 545, (1980).