Derrida's Generalized Random Energy models 4: Continuous state branching and coalescents.

Anton Bovier[1,2]
Weierstraß–Institut
für Angewandte Analysis und Stochastik
Mohrenstrasse 39, D-10117 Berlin, Germany

Irina Kurkova[3]
Laboratoire de Probabilités et Modèles Aléatoires
Université Paris 6
4, place Jussieu, B.C. 188
75252 Paris, Cedex 5, France

Abstract: In this paper we conclude our analysis of Derrida's Generalized Random Energy Models (GREM) by identifying the thermodynamic limit with a one-parameter family of probability measures related to a continuous state branching process introduced by Neveu. Using a construction introduced by Bertoin and Le Gall in terms of a coherent family of subordinators related to Neveu's branching process, we show how the Gibbs geometry of the limiting Gibbs measure is given in terms of the genealogy of this process via a deterministic time-change. This construction is fully universal in that all different models (characterized by the covariance of the underlying Gaussian process) differ only through that time change, which in turn is expressed in terms of Parisi's overlap distribution. The proof uses strongly the Ghirlanda-Guerra identities that impose the structure of Neveu's process as the only possible asymptotic random mechanism.

Keywords: Gaussian processes, generalized random energy model, continuous state branching process, subordinators, coalescent processes, genealogy, Ghirlanda-Guerra identities.

AMS Subject Classification: 82B44, 60G70,60K35

---

[1]e-mail: bovier@wias-berlin.de
[2]Research supported in part by the DFG in the Dutch-German Bilateral Research Group "Mathematics of Random Spatial Models from Physics and Biology".
[3]e-mail: kourkova@ccr.jussieu.fr

1. Introduction.

In a series of papers [BK1,BK2,BK3] we have recently taken up the analysis of a class of mean field spin glass models introduced by Derrida and Gardner in the 1980's[D,DG1,DG2]. In purely mathematical terms, these models can be described as follows. Consider the N dimensional hypercube $\mathcal{S}_N \equiv \{-1,1\}^N$ endowed with the (normalized) ultrametric $1-d_N$, where $d_N(\sigma,\tau) \equiv N^{-1}(\min(i : \sigma_i \neq \tau_i) - 1)$. Define a normal Gaussian processes X_σ indexed by $\mathcal{S}_N$ with covariance function

$$\mathbb{E} X_\sigma X_\tau = A(d_N(\sigma,\tau)) \tag{1.1}$$

for some nondecreasing function $A : [0,1] \to [0,1]$. The principal object of interest is the analysis of the asymptotic behaviour of the Gibbs measures

$$\mu_{\beta,N}(\sigma) \equiv \frac{e^{\beta\sqrt{N}X_\sigma}}{Z_{\beta,N}} \tag{1.2}$$

where the partition function $Z_{\beta,N}$ assures that $\mu_{\beta,N}$ is a probability measure.

Let us note as an aside that the particular formulation of the problem above is related to its history in the context of spin glasses. The problem can be posed, however, in a mathematically completely equivalent way in terms of families of Gaussian processes on the unit interval endowed with a covariance depending on a rescaled dy-adic distance. We will reformulate this problem in this context in Section 2. The corresponding processes then induce in a natural way a (Gibbs) probability measure in the unit interval. We will see that it is very natural to interpret our analysis of the Gibbs measure as an analysis of the fine structure of the limiting singular measure on the unit interval. Note that this point of view had been put forward already in [B], but was then somehow abandoned in [BK2,BK3].

Let us briefly dwell on the history of this problem. The model was introduced and analysed by Derrida and Gardner [D,DG1,DG2] in the case when A is a step function with finitely many steps (the corresponding models are called GREMs, or Derrida's GREMs) in the sense that the limit of the free energy

$$F_{\beta,N} \equiv -\frac{1}{\beta N} \ln Z_{\beta,N} \tag{1.3}$$

and some further thermodynamics functions were computed. The computation of the free energy was later done rigorously in [CCP]. Derrida and Gardner then also considered limits of their results as the number of steps tended to infinity, and interpreted this results as corresponding to continuous functions A [DG1]. These results were then also compared to those of the more commonly studied (and more difficult) class of Sherrington-Kirkpatrick models (which essentially differs from the class studied here in that the covariance is a function of the Hamming distance rather than our hierarchical distance).

While there were very few further rigorous results on these models (but see [GMP]), Ruelle in a seminal paper of 1988 [Ru] introduced a new class of models based on *Poisson cascades* (to which we will henceforth refer to as "Ruelle's GREM") which he apparently understood to be the appropriate asymptotic models to describe the limiting Gibbs measures of Derrida's GREMs. Ruelle noted a number of remarkable features of these models, and in particular observed that it was possible to construct limits as the number steps went to infinity in terms of projective limits. Surprisingly, his paper at no point contains a precise hint on how his models are to be related to the original spin glass models of Derrida.

Shortly after that, Neveu [Ne] noted a connection between Ruelle's models and continuous state branching processes. This paper also outlined a proof of the convergence of the rescaled partition function of the REM and GREM to a functional of the Poisson process, respectively Poisson cascades of Ruelle. Unfortunately, these observations are only contained in an internal report that was never published and that contains these ideas only in a somewhat embryonic form. Following a much later paper by Bolthausen and Sznitman [BoS], where it was explained how the results of replica theory of spin glasses can be interpreted in terms of a *coalescent process* (now known as the *Bolthausen-Sznitman coalescent*), Bertoin and Le Gall [BeLG] finally gave a precise and complete form of the relation between continuous state branching processes, the Ruelle's GREM, and the Bolthausen-Sznitman coalescent.

Around the time when these fascinating results appeared, we began to investigate more closely the link to the original spin glass models with Ruelle's models. In the REM, this connection was then made in a paper with M. Löwe [BKL] which was elaborated on in the lecture notes by one of us [B] (see also [T3,T4]). These results were extended to the GREMs in the papers [BK1,BK2], using essentially elementary methods. We observed, however, that the use of the so-called Ghirlanda-Guerra identities [GG] allowed a different approach that circumvented parts of these explicit computations (this fact was first observed in the REM by Talagrand [T3] who also exploited these identities heavily in his work on the p-spin SK models [T1,T2,T3,T4]). In fact, these identities impose structural constraints on any limit point that allow to prove convergence of the Gibbs measure (in a suitable sense) only on the basis of the convergence of the free energy, and that, moreover, allow to characterize the limit. These observations allowed us in [BK3] to extend our convergence results to the general class of models defined above.

In the present paper we want to conclude this investigation by linking our result up to the continuous state branching model, i.e. by identifying the limit proven to exist in [BK3] explicitly in terms of Neveu's branching process. This requires in fact little more than combining our results from [BK3] with those of Bertoin and Le Gall [BeLG], but we feel that the emerging complete picture is well worth to be put in evidence.

Let us insist that the main purpose of this and our preceding papers is to show that meaningful *infinite volume limits* exist in highly disordered mean field models, contrary to what is sometimes claimed. Quite on the contrary, there exist as we will see *universal* limiting random objects that serve as good approximations of the "large but finite" systems, in the best spirit of statistical mechanics. The fact that these objects turn out to be random, and that convergence tends to be in the sense of probability distributions is certainly unfamiliar to the traditionally trained mathematical statistical physicist, while this will hardly come as a surprise to probabilists or statisticians. Let us mention that the importance of ditributional limits of random measures in the context of spin glasses was strongly advocated in a series of papers by Newman and Stein, see e.g. [NS1,NS2,NS3]. When saying that limits are good approximations, care has to be taken of the topology used when constructing limits. There are indeed many pitfalls possible, and great care must be taken in order to get meaningful results. Rather unsurprisingly, the ingenious analysis of this problem introduced in the context of the replica method [MPV] is largely equivalent to formalism used in [BK3].

Let us recall the central problem one is faced with when analysing mean field spin glasses. What we want to do is to describe the geometric structure of a random probability measure on a set $\mathcal{S}_N$. One expects that this measure will concentrate (at low temperatures) on a relatively very small subset with rather complicated structure. Since due to randomness and symmetries there are no external references, we need a way to describe the structural geometric properties of such measures in an intrinsic, reference-free way. On the other hand, we need to allow sufficient compactness for limits to exist.

To resolve this problem, we introduced in [BK3] was what we called the *empirical distance distribution function*, i.e. the random measure

$$\mathcal{K}_{\beta,N} \equiv \sum_{\sigma \in \mathcal{S}_N} \mu_{\beta,N}(\sigma)\delta_{m_\sigma(\cdot)} \tag{1.4}$$

where

$$m_\sigma(t) \equiv \mu_{\beta,N}\left(\sigma' : d_N(\sigma,\sigma') > t\right) \tag{1.5}$$

This object describes the probability of a mass distribution around a randomly (according to the Gibbs measure) drawn point on $\mathcal{S}_N$.

A key object is the mean first moment of this random measure,

$$\int \mathcal{K}_{\beta,N}(dm)m(t) \equiv 1 - f_{\beta,N}(t) \tag{1.6}$$

which is nothing but probability that two configurations, σ, σ', drawn independently form the Gibbs sample satisfy $d_N(\sigma,\sigma') > t$. The function

$$f_{\beta,N}(t) \equiv \mu_{\beta,N}^{\otimes 2}\left(d_N(\sigma,\sigma') \leq t\right) \tag{1.7}$$

is now the analog of Parisi's order parameter[4]. In [BK3] we proved that

$$f_{\beta,N}(t) \to \mathbb{E}f_\beta = \min\left(\beta^{-1}\sqrt{2\ln 2}/\sqrt{\bar{a}(t)},\, 1\right) \tag{1.8}$$

where $\bar{a}$ is the right-derivative of the convex hull of the function A. Convergence in (1.8) holds both in mean and almost surely. We also showed that

$$\mathcal{K}_{\beta,N} \stackrel{\mathcal{D}}{\to} \mathcal{K}_\beta. \tag{1.9}$$

The limit is uniquely determined by Ghirlanda-Guerra relations, which give recursive formulas to compute all moments of $\mathcal{K}_\beta$ starting from the function f_β.

In fact, while the random measures $\mathcal{K}_{\beta,N}$ may look somewhat unfamiliar, their moments are closely linked and even equivalent to the more conventional n-replica distance distribution $\mathbb{Q}_{\beta,N}^{(n)}$. These are measures on the space $[0,1]^{n(n-1)/2}$

$$\mathbb{Q}_{\beta,N}^{(n)}(\mathcal{A}) \equiv \mathbb{E}\mu_{\beta,N}^{\otimes n}\left(\left(d(i,j)\right)_{1\leq i,j\leq N} \in \mathcal{A}\right) \tag{1.10}.$$

Note that these measures do of course give full measure to sets that respect the ultrametric triangle relations. In [BK3] we proved their convergence to a limiting distribution $\mathbb{Q}_\beta^{(n)}$. The Ghirlanda-Guerra identities (together with the fact that $1-d_N$ is an ultrametric distance) allow to compute $\mathbb{Q}_\beta^{n+1}$ in terms of $\mathbb{Q}_\beta^n$ recursively, while $\mathbb{Q}_\beta^2$ has distribution function $\mathbb{E}f_\beta(t)$.

[4] In the context of the SK models, this function is usually defined with d_N replaced by the "overlap parameter $R_N(\sigma,\sigma') \equiv N^{-1}\sum_i \sigma_i\sigma_i'$. In [BK2] we have shown that in the GREM, the choice of the distance used in the definition of $f_{\beta,N}$ does not affect the result in the limit $N\uparrow\infty$.

On the other hand, the full set of distributions $\mathbb{Q}_\beta^{(n)}$ determines the limiting random measures $\mathcal{K}_\beta$ through its moments.

It now remains to interpret these limiting objects in the context of Neveu's branching process. It will turn out that both $\mathcal{K}_\beta$ and $\mathbb{Q}_\beta^{(n)}$ have natural interpretations. The former will be interpreted in the language of the continuous state branching process, while the latter are naturally interpreted in the corresponding coalescent process on integer partitions.

The remainder of the paper is organised as follows. In Section 2 we describe a canonical construction of the genealogy associated to a flow of probability measures μ^t on $[0,1]$ in a general setting. In this process we introduce the empirical distribution $\mathcal{K}_t$ on the functions $m_x(t,\cdot)$ which describe the dependence of the family size of the individual x as a function of the degree of relatedness. We also define the genealogical distance between any two points on $[0,1]$ as the last time they had an ancestor in common. We are mostly interested in the case when the measures in the flow are random. Consequently, we define the distance on integers i,j as the distance between independent uniformly distributed random variables U_i, U_j on $[0,1]$ and consider the partitions of integers in blocks whose distance is most s. The family of these partitions as a process of s form a coalescent process. It turns out that $\mathcal{K}_t$ is completely determined by this coalescent, as we give explicit expressions for all its moments in terms of probabilities of random partitions.

In Section 3 we show how this construction works for a flow of the Gibbs measures $\mu_{\beta,N}$ on $\mathcal{S}_N$ (identified with $[0,1]$ via the canonical map (2.1)) of the CREM with an arbitrary function A. Namely, $\mathcal{K}_{\beta,N}$ of (1.4) is precisely $\mathcal{K}_1$ defined in Section 1 for this flow. We also explain how this construction can be reformulated for the Gaussian process on the unit interval.

In Section 4 we formulate our main theorem. It identifies the limit as $N\uparrow\infty$ of $K_{\beta,N}$ in terms of the flow of measures corresponding to Neveu's branching process: $X(t(y),x)/X(t(y),1)$, where $X(t,x)$ is the size of the population of this process at time t, provided that at the initial moment it was x. Here $t(y)$ is an appropriate time change defined only by $\mathbb{E}f_\beta(y)$ of (1.9). In fact, conceptually we do slightly more: we show that there is a flow of probability measures $\tilde{\mu}^t_{\beta,N}$ constructed via an embedding of Gibbs measures $\mu^t_{\beta,N}$ that converges to a limiting flow of measures $\Theta^t_{f_\beta}$ constructed from a time changed Neveu branching process, in the sense that the genealogies of the flow converges.

In Sections 5 and 6 we prove this theorem. Since $\mathcal{K}_t$ is determined by its moments, or equivalently, by the genealogical distances of integers for the corresponding coalescent, as was established at the end of Section 2, its suffices to show that the n-replica distance distribution functions (1.10) of our spin glass model converge to the genealogical distance distribution function of the Bolthausen-Sznitman coalescent (which corresponds to $\mathcal{K}_1$) under an appropriate time change. One way (short but indirect) to prove this is indicated in Section 5 and relies on the connection between Neveu's branching process and Ruelle's probability cascades established in [BeLG].

The second way (more direct) is given in Section 6: it consists in showing that the Bolthausen-Sznitman coalescent satisfies Ghirlanda-Guerra identities. For that purpose we use the Chinese restaurant process of J. Pitman [P].

We hope that the results presented in this class of models elucidate in a mathematically comprehensible context the fundamental and universal rôle played by Neveu's continuous state branching process as a universal random mechanism governing the extremal processes for a wide class of stochastic processes. If one accepts the common believe of theoretical physicists, its rôle goes well beyond the class of models we discuss here. Even on a slightly less speculative level, Neveu's process will emerge in any model for which the Ghirlanda-Guerra relations hold in their strong form, which means in particular that this will be the case if not for the actual SK model, then at least for models where weak additional fields have been added to the Hamiltonians (see [GG,Le,T2]). On the other hand, we also hope that these examples help to explain to a mathematical audience what physicist describe when they talk about "continuous replica symmetry breaking", and how such a phenomena can actually arise.

Acknowledgements: We are grateful for numerous enlightening discussions with Jean Bertoin on continuous state branching, subordinators, and coalescents. A.B. also acknowledges inspiring discussions with Anton Wakolbinger. I.K. thanks Ph. Marchal for helpful discussions. Part of this work was done while A.B. was resident at the Erwin-Schroedinger Institute for Mathematical Physics (ESI) in Vienna in the programme "Mathematical population genetics and statistical physics". He thanks the organisers, Ellen Baake, Michael Baake, and Reinhard Bürger for the kind invitation and the ESI for financial support. I.K. thanks the Weierstrass Institute for Applied Analysis and Stochastics, Berlin, for its hospitality and financial support. Women hen ganxie de Shi Zhan jiaoshuo yaoqing hen hao zhongguo fanguan.

2. Genealogy of a flow of probability measures

In [B] one of us proposed to describe the infinite volume limit of Gibbs measure in the Random Energy Model by considering its image on the unit interval through the canonical map $r_N : \mathcal{S}_N \to [0,1]$ defined as

$$r_N(\sigma) \equiv \sum_{i=1}^{N} 2^{-i}(1+\sigma_i)/2. \tag{2.1}$$

It was shown that the phase transition in the REM manifests itself by the fact that the resulting image measure converges to Lebesgue measure in the high temperature phase ($\beta \le \sqrt{2\ln 2}$) and towards a dense pure point measure in the low-temperature phase ($\beta > \sqrt{2\ln 2}$). While in the REM this appeared to give a rather nice picture, at first glance it seems to be difficult to encode the far more complex structure of the Gibbs measures of the GREM and CREM in such a simple embedding. Our purpose would be to identify a measure on $[0,1]$ that represents the limiting Gibbs measure. It will be instructive to explain how a naive approach to do so fails. On the hypercube we are interested in the masses of sets $b_\sigma(t) \equiv \{\sigma' : d_N(\sigma,\sigma') > t\}$. If we map such sets on the unit interval via r_N, we obtain an interval of length $2^{-[tN]}$. In fact there is no difficulty to express e.g. $\mathcal{K}_{\beta,N}$ for N fixed in terms of quantities defined with respect to the image measure on the hypercube. However, the construction involves masses of intervals of exponentially small size (in N). So what should one do in the limit when N is infinite? We cannot analyse the structure by looking at intervals of the size $2^{-t\infty}$.

What is needed is clearly a construction that does not refer explicitly to masses of intervals of exponentially small size while still revealing the fine structure of the measure at such a scale. In this section we show that a canonical construction exists when we consider a family of probability measures on $[0,1]$.

Let $\{\mu^t\}_{t\in\mathbb{R}^+}$ be a family of probability measures on $[0,1]$. Denote by Θ_t their (right-continuous) probability distribution functions $\Theta_t(x) = \int_0^x \mu^t(dx)$. Let us note that we will mostly be interested in cases when the measures μ^t are getting more and more irregular as t increases. We will consider Θ_t as a map $[0,1] \to [0,1]$, so that $\{\Theta_t\}_{t\in\mathbb{R}^+}$ represents a flow of maps on the unit interval. Define for $t' < t$

$$S^{(t',t)}(x) \equiv \Theta_t\left(\Theta_{t'}^{-1}(x)\right) \tag{2.2}$$

where the inverse of a right-continuous, non-decreasing function Θ is defined as

$$\Theta^{-1}(x) = \inf\{y \mid \Theta(y) \ge x\}. \tag{2.3}$$

We will need some elementary properties of the inverse function. We say that Θ increases at x, if for any $\epsilon > 0$, $\Theta(x) > \Theta(x - \epsilon)$.

Lemma 2.1: *Let Θ be a non-decreasing function. Then for any point x at which Θ is increasing,*

$$\Theta^{-1}(\Theta(x)) = x \tag{2.4}$$

Proof: We have

$$\Theta^{-1}(\Theta(x)) = \inf\{y : \Theta(y) \geq \Theta(x)\} \tag{2.5}$$

Since Θ is non-decreasing, for any $y \geq x$, $\Theta(y) \geq \Theta(x)$ and thus $\Theta^{-1}(\Theta(x)) \leq x$. Assume that $\Theta^{-1}(\Theta(x)) = y < x$. This implies that for some $y < x$, $\Theta(y) = \Theta(x)$, contradicting the assumption that x is a point of increase. $\diamondsuit$

Lemma 2.2: *For a given family of measures μ^t, let $\mathcal{I}_t$ denote of the set of points of increase of the function Θ_t. Assume that for $t'' < t' < t$, $\mathcal{I}_{t''} \subset \mathcal{I}_{t'}$. Then*

$$S^{(t'',t)}(x) = S^{(t',t)} \circ S^{(t'',t')}(x) \tag{2.6}$$

for any $x \in [0, 1]$.

Proof: By definition,

$$S^{(t',t)} \circ S^{(t'',t')}(x) = \Theta_t(\Theta_{t'}^{-1}(\Theta_{t'}(\Theta_{t''}^{-1}(x)))) \tag{2.7}$$

Note first that since $y = \Theta_{t''}^{-1}(x)$ is the smallest value for which $\Theta(y) \geq x$, for any $y' < y$ it must be true that $\Theta_{t''}(y') < \Theta_{t''}(y)$. Thus $\Theta_{t''}^{-1}(x) \in \mathcal{I}_{t''}$.

But if y is a point of increase of $\Theta_{t''}$, by assumption y is also a point of increase of $\Theta_{t'}$, hence

$$S^{(t',t)} \circ S^{(t'',t')}(x) = \Theta_t(y) = S^{(t'',t)}(x) \tag{2.8}$$

which proves the lemma. $\diamondsuit$

Remark: We see that the construction of the $S^{(t',t)}$ is best suited in situations where the distribution functions Θ_t are everywhere increasing. We will however encounter a more delicate situation where Θ^t are step functions with $\mathcal{I}_{t'} \subset \mathcal{I}_t$ for all $0 \leq t' \leq t < \infty$.

Assumption: In the remainder of this section we will assume that Θ_t satisfies $\mathcal{I}_{t'} \subset \mathcal{I}_t$ for all $0 \leq t' \leq t < \infty$.

It will be natural to think of Θ_t and $S^{(t',t)}$ as mapping subsets of $[0,1]$ to subsets of $[0,1]$ i.e. by abuse of notation if $(a,b)=I\subset[0,1]$, we will write $S^{(t',t)}(I)\equiv\left(S^{(t',t)}(a),S^{(t',t)}(b)\right)$, etc. We will think of $S^{(t',t)}(I)$ as the "offspring" at time t of a population I at time t'. For this to make sense we need of course conditions (2.6) to hold at all times. We now want to associate a genealogy to this flow. Ideally, we want to introduce the notion of the set $m_x(t,t')$ of points z at time t that have at time $t'<t$ the last time an ancestor in common with x. To allow some more flexibility it will be useful to have a softer version of this notion, where the ancestors of x and z are only required to be "close". Define now

$$m^\epsilon_x(t,t')\equiv S^{(t',t)}\left(V_\epsilon\left(\left(S^{(t',t)}\right)^{-1}(x)\right)\right) \tag{2.9}$$

where

$$V_\epsilon(y)\equiv(y-\epsilon,y+\epsilon) \tag{2.10}$$

$m^\epsilon_x(t,t')$ is the offspring at time t of a small neighbourhood of the ancestor of x at time t'. Thus, roughly, $|m^\epsilon_x(t,\cdot)|$ describes the dependence of the family size of the individual x as a function of the degree of relatedness. We will eventually tend $\epsilon\downarrow 0$. In cases when all functions $S^{(t',t)}$ are continuous, this of course produces a trivial answer, i.e. $\lim_{\epsilon\downarrow 0}m^\epsilon_x(t,t')=\{x\}$; if on the other hand $S^{(t',t)}$ has jumps at the position of the ancestor of x for certain values of t', $\lim_{\epsilon\downarrow 0}m^\epsilon_x(t,t')$ will be a non-trivial sequence of intervals.

Finally, define the associated empirical distribution of the functions $m^\epsilon_x(t,\cdot)$

$$\mathcal{K}^\epsilon_t=\int_0^1 dx\,\delta_{|m^\epsilon_x(t,\cdot)|} \tag{2.11}$$

and let

$$\mathcal{K}_t=\lim_{\epsilon\downarrow 0}\mathcal{K}^\epsilon_t. \tag{2.12}$$

Now we define the (ϵ)-genealogical "distance" of two points $x,y\in[0,1]$ with respect to μ^t as

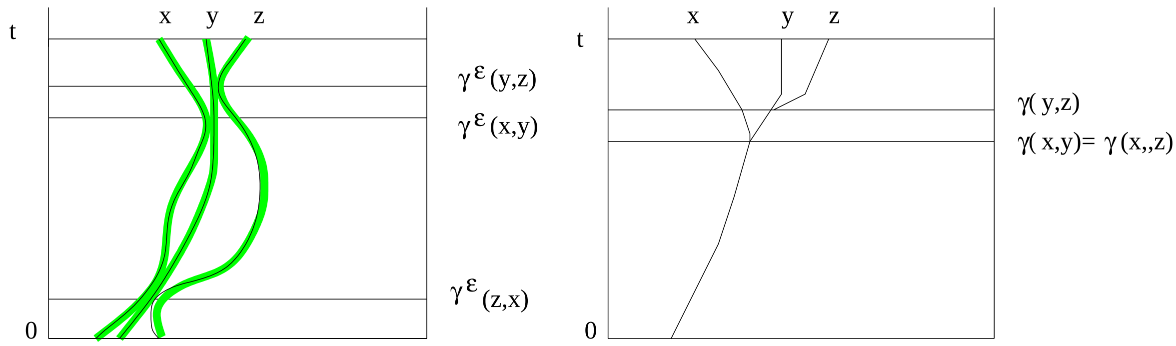
$$\gamma^\epsilon_t(x,y)\equiv\sup\left(t':y\in m^\epsilon_x(t,t')\right). \tag{2.13}$$

It would be nice if γ^ϵ_t defined an ultrametric. Unfortunately, this is in general not the case. The reason is simply that if the ancestors of x and y were at a distance ϵ for the last time at time t, they could have been farther apart at some later time. Thus the ancestor of z may get to a distance ϵ of the ancestor y at a time bigger than t, but may not be close to the ancestor of x after some time strictly less than t. However, this is not possible if we use the strict definition of genealogical distance as

$$\gamma_t \equiv \lim_{\epsilon \downarrow \infty} \gamma_t^\epsilon \tag{2.14}$$

In fact, we have:

Lemma 2.3: *If the hypothesis of Lemma 2.2 hold for all times $t'' \le t' \le t$, then $1 - \gamma_t$ defines an ultrametric distance on the unit interval.*



Ancestrol lines of three points and their neighbohoods. Note that the ε genealogies are not ultrametric

Ancestral lines in a flow with jumps. Here the genealogy is ultrametric

Proof: Let us note first off all that

$$y \in \lim_{\epsilon \downarrow 0} m_x^\epsilon(t, t'') \quad \forall t'' < t' = \gamma_t(x, y). \tag{2.15}$$

In fact it suffices to see that

$$\lim_{\epsilon \downarrow 0} S^{(t',t)}(V_\epsilon(S^{(t',t)})^{-1}(x)) \subset \lim_{\epsilon \downarrow 0} S^{(t'',t)}(V_\epsilon(S^{(t'',t)})^{-1}(x)) \tag{2.16}$$

which is by compatibility (2.6) equivalent to

$$\lim_{\epsilon \downarrow 0} V_\epsilon(S^{(t',t)})^{-1}(x)) \subset \lim_{\epsilon \downarrow 0} S^{(t'',t')}(V_\epsilon(S^{(t'',t)})^{-1}(x)) \tag{2.17}$$

This last equality again holds by (2.6). It is indeed trivial if $S^{(t'',t')}$ is continuous at $(S^{(t'',t)})^{-1}(x)$.

It follows from (2.15) that for any $x, y, z \in [0,1]$ if $\gamma_t(x,y) \neq \gamma_t(x,z)$ then $\gamma_t(y,z) = \min\{\gamma_t(x,z), \gamma_t(x,y)\}$. In fact, let e.g. $\gamma_t(x,z) > \gamma_t(x,y)$. Then $z \in \lim_{\epsilon \downarrow 0} m_x^\epsilon(t, \gamma_t(x,y))$ and then $\gamma_t(y,z) \ge \gamma_t(x,y)$. From the other point of view, if $\gamma_t(y,z) > \gamma_t(x,y)$, then either

$\gamma_t(x,z) \geq \gamma_t(y,z) > \gamma_t(x,y)$ or $\gamma_t(y,z) > \gamma_t(x,z) > \gamma(x,y)$. In the first case by (2.15) $x \in \lim_{\epsilon\downarrow 0} m_z^\epsilon(t, \gamma_t(y,z))$ and then $\gamma_t(x,y) \geq \gamma_t(y,z)$ which is impossible. In the second $y \in \lim_{\epsilon\downarrow 0} m_z^\epsilon(t, \gamma_t(x,z))$ from where $\gamma_t(x,y) \geq \gamma_t(x,z)$, which is again impossible. Thus $\gamma_t(y,z) = \gamma_t(x,y)$.

Note also that if $\gamma_t(x,y) = \gamma_t(x,z)$, then $\gamma_t(y,z) \geq \gamma_t(x,y) = \gamma_t(x,z)$. These observations imply that $1 - \gamma_t$ is an ultrametric distance on $[0,1]$. $\diamondsuit$

Again, γ_t is rather trivial if $S^{(t',t)}$ are all continuous, for then

$$\lim_{\epsilon\downarrow 0} \gamma_t^\epsilon(x,y) = \begin{cases} t, & \text{if} x = y \\ -\infty, & \text{if} x \neq y \end{cases} \tag{2.18}$$

and in a strict sense nobody has any relatives. On the other hand, in the discontinuous case, rather large families exist, and the ultrametric structure of the interval can be very rich.

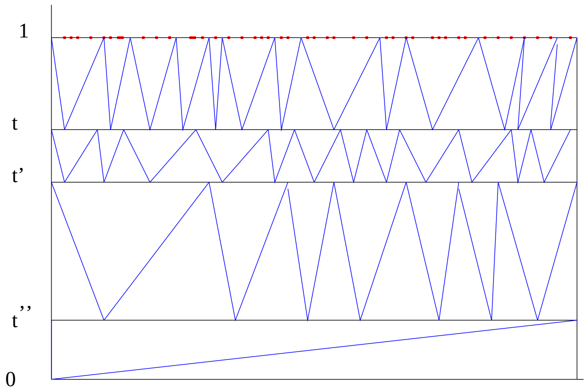
Note that we will usually be interested in cases where the family of measures μ^t is *random*. In that case $\mathcal{K}_t$ is a random probability measure. We will now describe a useful way of characterizing this random measure.

Having defined a distance $1-\gamma_t$ on $[0,1]$, we can define in a very natural way the analogous distance on the integers. To do this, consider a family of i.i.d. random variables $\{U_i\}_{i\in\mathbb{N}}$ distributed according to the uniform law on $[0,1]$. Given such a family, we set

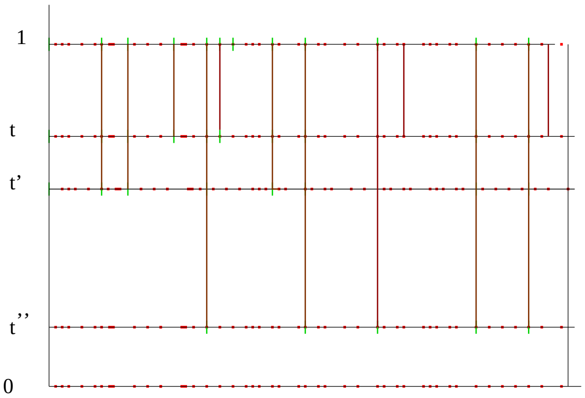
$$\rho_t(i,j) = \gamma_t(U_i, U_j) \tag{2.19}$$

Due to the ultrametric property of the γ_t and the independence of the U_i, for fixed t, the sets $B_i(t') \equiv \{j : \rho_t(i,j) \geq t'\}$ form an exchangeable random partition of the integers.

Moreover, the family of these partitions as a function of $t - t'$ is a stochastic process on the space of integer partitions with the property that for any $t' > t''$, the partition $B_i(t'')$ is a coarsening of the partition $B_i(t')$. Such a process is called a *coalescent process* (see e.g. [Be1,Be2,BeLG,,BePi,BeYo,BoS,P,Pi1,PPY]).



Genealogy of a flow of measures



Associated partions of random i.i.d. variables (dots).

The key observation for our purposes is the possibility to express $m_x(t,t')$ and its moments in terms of this coalescent [Be1]. Namely, it is plain that

$$\lim_{n\uparrow\infty} n^{-1}\sum_{j=1}^{n} \mathbb{1}_{j\in B_k(t')} = m_{U_i}(t,t'). \quad \text{a.s.} \tag{2.20}$$

for any i such that $i\in B_k(t')$. This implies, for instance, as shown in [Be1] that

$$\mathbb{E}\int dx m_x(t,t') = \mathbb{P}[2\in B_1(t')] \tag{2.21}$$

and more generally that

$$\mathbb{E}\int dx m_x^k(t,t') = \mathbb{P}[2,3,\dots,k+1\in B_1(t')] \tag{2.22}$$

Here the expectation $\mathbb{E}$ is with respect to the randomness of the family of measures μ^t, and $\mathbb{P}$ is the law with respect to the random genealogy (depending both on the random measures and the i.i.d. r.v. U_i). We will need slightly more general expressions, namely a family of moments that determine the law of the random probability measure $\mathcal{K}_t$. These can be written as follows. Let take any positive integer p, collection of positive real numbers $0<t_1,\dots,t_p\le t$, positive integers ℓ, and non-negative integers $k_{11},\dots,k_{1p},k_{21},\dots,k_{2p},\dots,k_{\ell 1},\dots,k_{\ell p}$. Then we need

$$M(p,\underline{t},\underline{\underline{k}}) \equiv \mathbb{E}\left(\int dx m_x^{k_{11}}(t_1)\dots m_x^{k_{1p}}(t_p)\right)\dots\left(\int dx m_x^{k_{\ell 1}}(t_1)\dots m_x^{k_{\ell p}}(t_p)\right) \tag{2.23}$$

By (2.20), we have that

$$\begin{aligned}&\int dx m_x^{k_{11}}(t_1)\dots m_x^{k_{1p}}(t_p)\\ &= \lim_{n\uparrow\infty} n^{-1-k_{11}-\dots-k_{1p}}\sum_{i=1}^{n}\sum_{j_1^1,\dots,j_{k_{11}}^1,\dots,j_1^p,\dots,j_{k_{1p}}^p}^{n} \mathbb{1}_{j_1^1,\dots,j_{k_{11}}^1\in B_{k(i,t_1)}(t_1)}\cdots\mathbb{1}_{j_1^p,\dots,j_{k_{1p}}^p\in B_{k(i,t_p)}(t_p)}\end{aligned} \tag{2.24}$$

Let us note first that in these expressions contributions from terms where two indices are equal can be neglected. Second, since

$$B_{k(i,t_p)}(t_p) \subset B_{k(i,t_{p-1})}(t_{p-1}) \subset \cdots \subset B_{k(i,t_1)}(t_1) \tag{2.25}$$

the summand in (2.24) is the same as

$$\mathbb{1}_{j_1^1,\dots,j_{k_{11}}^1,\dots,j_1^p,\dots,j_{k_{1p}}^p \in B_{k(i,t_1)}(t_1)} \cdots \mathbb{1}_{j_1^p,\dots,j_{k_{1p}}^p \in B_{k(i,t_p)}(t_p)} \tag{2.26}$$

Then

$$\begin{aligned} M(p,\underline{t},\underline{\underline{k}}) &= \lim_{n\uparrow\infty} n^{-\ell-k_{11}-\dots-k_{\ell p}} \mathbb{E} \sum_{i_1,\dots,i_\ell} \sum_{j_1^{1,1},\dots,j_{k_{11}}^{1,1},\dots,j_1^{p,1},\dots,j_{k_{1p}}^{p,1}} \cdots \sum_{j_1^{1,\ell},\dots,j_{k_{\ell 1}}^{1,\ell},\dots,j_1^{p,\ell},\dots,j_{k_{\ell p}}^{p,\ell}} \\ &\mathbb{1}_{j_1^{1,1},\dots,j_{k_{11}}^{1,1},\dots,j_1^{p,1},\dots,j_{k_{1p}}^{p,1} \in B_{k(i_1,t_1)}(t_1)} \cdots\cdots \mathbb{1}_{j_1^{1,\ell},\dots,j_{k_{\ell 1}}^{1,\ell},\dots,j_1^{p,\ell},\dots,j_{k_{\ell p}}^{p,\ell} \in B_{k(i_\ell,t_1)}(t_1)} \\ &\dots \mathbb{1}_{j_1^{p,1},\dots,j_{k_{1p}}^{p,1} \in B_{k(i_1,t_p)}(t_p)} \cdots\cdots \mathbb{1}_{j_1^{p,\ell},\dots,j_{k_{\ell p}}^{p,\ell} \in B_{k(i_\ell,t_p)}(t_p)} \\ &= \mathbb{P}\left[J_{11} \in B_1(t_1),\dots,J_{1p} \in B_1(t_p),\dots J_{\ell 1} \in B_\ell(t_1),\dots,J_{\ell p} \in B_\ell(t_p)\right] \end{aligned} \tag{2.27}$$

where

(i) $J_{11},\dots,J_{\ell p}$ is a partition of $\{1,\dots,k_{11}+\dots+k_{\ell p}\}$,

(ii) For all j,i, $J_{ji} \supset J_{ji+1}$, and

(iii) $|J_{ji}| = k_{j,i} + k_{j,i+1} + \dots + k_{j,p}$.

By exchangeability, the choice of the partition and the subsets is irrelevant. Note that we were forced in the last line to use a somewhat incongruent notations for the sets B: we enumerate $B_i(t_1)$ in their oder of appearance, and let $B_i(t_j)$ be the first block contained in $B_i(t_{i-1})$ on the level t_i.

The probabilities (2.27) can be expressed alternatively in the form

$$\mathbb{P}\left(\rho_t(1,2) \le t_1,\dots,\rho_t(n-1,n) \le t_{n(n-1)/2}\right). \tag{2.28}$$

Thus the random probability measure $\mathcal{K}_t$ is completely determined by the probabilities (2.27) or (2.28) of the corresponding coalescent through the family of its moments.

3. Finite N setting.

We will now show that for finite N we can use the general construction from the preceding section to relate the geometric description of the Gibbs measure on $\mathcal{S}_N$ to the genealogical description of a family of embedded measures on $[0,1]$.

Recall that our basic objects are

(i) $\sigma \in \mathcal{S}_N \equiv \{-1,1\}^N$ equipped with $d_N(\sigma,\tau) \equiv N^{-1}(\min(i : \sigma_i \neq \tau_i) - 1)$. Due to this notion, we can think of $\mathcal{S}_N$ as being the leaves of a binary tree $\mathcal{T}_N$, obtained by setting $\sigma(t) \equiv (\sigma_{i_1}, \sigma_{i_2}, \dots, \sigma_{i_{[tN]}})$. It will be convenient to be able to represent σ, for given $0 < t_1 < t_2 < \dots < t_n < 1$ in the form

$$\sigma = (\sigma(t_1), \sigma(t_2, t_1), \dots, \sigma(t_n, 1)) \tag{3.1}$$

where

$$\sigma(t_i, t_{i+1}) \in \{-1,1\}^{[N(t_{i+1}-t_i)]}. \tag{3.2}$$

(ii) For non-decreasing function $A : [0,1] \to [0,1]$, we have a Gaussian process X_σ with mean zero and covariance

$$\mathbb{E} X_\sigma X_\tau = A(d_N(\sigma,\tau)). \tag{3.3}$$

We extend this to a Gaussian process indexed by the tree $\mathcal{T}_N$ by setting

$$X_{\sigma(t)} = X_\sigma(t) \tag{3.4}$$

where $X_\sigma(t)$ was defined in [BK3] as the Gaussian process on $\mathcal{S}_N$ with mean zero and covariance

$$\mathbb{E} X_\sigma(t) X_\tau(t') = A(t \wedge t' \wedge d_N(\sigma,\tau)) \tag{3.5}$$

Note that the process $X_\sigma(t)$ has independent increments (in t).

(iii) Finally, we have the Gibbs measure $\mu^t_{\beta,N}$ defined on $\mathcal{S}_N$ by

$$\mu^t_{\beta,N}(\sigma) \equiv \frac{e^{\beta\sqrt{N} X_\sigma(t)}}{Z^t_{\beta,N}}. \tag{3.6}$$

Note that $\mu^t_{\beta,N}$ induces a natural measure on $\mathcal{T}_N$ by setting

$$\mu^t_{\beta,N}(\sigma(s)) \equiv \mu^t_{\beta,N}(\sigma' : \sigma'(s) = \sigma(s)) \tag{3.7}$$

We can now map the measures $\mu^t_{\beta,N}$ to the unit interval through the map r_N defined in (2.1) via

$$\tilde{\mu}^t_{\beta,N} \equiv \sum_{\sigma \in \mathcal{S}_N} \delta_{r_{[tN]}(\sigma)} \mu^t_{\beta,N}(\sigma) \tag{3.8}$$

We will apply the construction of Section 2 to this family of measures. Let us denote by

$$\widetilde{\Theta}^t_{\beta,N}(x) = \int_0^x \widetilde{\mu}^t_{\beta,N}(dy) \tag{3.9}$$

Note that by construction the functions $\widetilde{\Theta}^t_{\beta,N}$ increase on the set $r_{[tN]}(\mathcal{S}_N)$ and that $r_{[t'N]}(\mathcal{S}_N) \subset r_{[tN]}(\mathcal{S}_N)$, for all $0 < t' < t < \infty$, thus the assumptions of Lemma 2.2 are satisfied for all values of t, t', t''. It is also useful to realise that

$$\widetilde{\Theta}^t_{\beta,N}(r_{[tN]}(\sigma)) = \sum_{\sigma' : r_{[tN]}(\sigma') \leq r_{[tN]}(\sigma)} \mu^t_{\beta,N}(\sigma) \tag{3.10}$$

We set, for $t' \leq t$,

$$\widetilde{S}^{(t',t)}_{\beta,N}(x) = \widetilde{\Theta}^t_{\beta,N}\left(\left(\widetilde{\Theta}^{t'}_{\beta,N}\right)^{-1}(x)\right) \tag{3.11}$$

In explicit notation, this can be expressed as

$$\widetilde{S}^{(t',t)}_{\beta,N}(x) = \sum_\sigma \mu^t_{\beta,N}(\sigma) \mathbb{1}_{\widetilde{\Theta}^{t'}_N(r_{[tN]}(\sigma))<x} \tag{3.12}$$

Note that $\widetilde{S}^{(t',t)}_{\beta,N}$ are non-decreasing functions $[0,1] \to [0,1]$. Note also that $\widetilde{S}^{(t',t)}_{\beta,N}(x)$ is piecewise constant and jumps at the points $\widetilde{\Theta}^{t'}_{\beta,N}(r_{[Nt']}(\sigma(t')))$, $\sigma(t') \in \mathcal{S}_{[t'N]}$; the values of the increments at the point $\widetilde{\Theta}^{t'}_{\beta,N}(r_{[Nt']}(\sigma(t')))$ is $\mu^t_{\beta,N}(\sigma(t'))$. Of course most of these increments will tend to zero while some will be of order one, as N tends to infinity.

It also follows now directly from Lemma 2.2 that for any $t'' < t' < t$, $\widetilde{S}^{(t',t)}_{\beta,N}$ satisfies compatibility relations for "bridges" in the sense Bertoin-Le Gall [BeLG]:

Lemma 3.1: *For all* $0 \leq t'' < t' < t \leq 1$,

$$\widetilde{S}^{(t'',t)}_{\beta,N}(x) = \widetilde{S}^{(t',t)}_{\beta,N} \circ \widetilde{S}^{(t'',t')}_{\beta,N}(x) \tag{3.13}$$

where $\circ$ *denotes composition.*

We may define as in (2.9) $m_x(t,t') = \lim_{\epsilon\downarrow 0} m^\epsilon_x(t,t')$. In this case

$$m_x(t,t') \equiv \inf\left(\widetilde{S}^{(t',t)}_{\beta,N}(y) : \widetilde{S}^{(t',t)}_{\beta,N}(y) \geq x\right) - \sup\left(\widetilde{S}^{(t',t)}_{\beta,N}(y) : \widetilde{S}^{(t',t)}_{\beta,N}(y) < x\right) \tag{3.14}$$

which is nothing but the hight of the jump of the function $\widetilde{S}^{(t',t)}_{\beta,N}$ which comprises the point x.

¿From the remark following (3.11) we know that such a jump occurs at points y of the form $y = \tilde{\Theta}^{t'}_{\beta,N}(r_{[t'N]}(\sigma))$ and will be by an amount $\mu^t_{\beta,N}(\sigma(t'))$ and so

$$m_{\tilde{\Theta}^{t'}_{\beta,N}(r_N(\sigma))}(t,t') = m_\sigma(t',1) \tag{3.15}$$

where m_σ is defined in (1.5). This observation allows us to express the measure $\mathcal{K}_{\beta,N}$ defined in (1.4) describing the geometry of the Gibbs measures on the hypercube in terms of the genealogical distance distribution functions $\mathcal{K}^\epsilon_t$ defined in Section 2, with the family of measures ρ^t given by the measures $\widetilde{\mu}^t_{\beta,N}$. Namely:

Proposition 3.2: *Let $\mathcal{K}^\epsilon_t$ be as defined in (2.11) for the family of measures $\tilde{\mu}^t_{\beta,N}$, and let $\mathcal{K}_{\beta,N}$ be defined in (1.4). Then*

$$\lim_{\epsilon\downarrow 0}\mathcal{K}^\epsilon_1 = \mathcal{K}_{\beta,N} \tag{3.16}$$

In particular,

$$f_{\beta,N}(t') = 1 - \lim_{\epsilon\downarrow 0}\mathbb{E}\int_0^1 dx|m^\epsilon_x(1,t')| \tag{3.17}$$

is expressed in terms of the size biased average gap size in the function $\tilde{S}^{(t,1)}_{\beta,N}$.

Proposition 3.2 is obtained in a purely algebraic manner and simply shows that the general formalism introduced in Section 2 allows, in the case of the finite volume Gibbs measures, to express the measures $\mathcal{K}_{\beta,N}$ in terms of the genealogy of the embedded family of measures $\tilde{\mu}^t_{\beta,N}$. This opens the way to express the thermodynamic limit of $\mathcal{K}_{\beta,N}$ in terms of the genealogy of a suitable family of measures on $[0,1]$.

An alternative setting.

Before turning to this question, let us note that the above setting of Gaussian processes on $\{-1,1\}^N$ can be completely reformulated in terms of Gaussian processes on the unit interval, respectively unit square. To see this, let $\tilde{d}_1$ denote the standard dyadic valuation on $[0,1]$, i.e. if $x = \sum_{i\in\{0,1\}} x_i 2^{-i}$, then $\tilde{d}_1(x,x') = \inf(i : x_i \neq x'_i) - 1$. Set $\tilde{d}_N(x) \equiv N^{-1}\tilde{d}_N(x)$. Then let $X_N(t,x)$ be the centered Gaussian process on $\mathbb{R}_+ \times [0,1]$ such that

$$\mathbb{E}X_N(t,x)X_N(t',x') = A\left(t\wedge t'\wedge \tilde{d}_N(x,x')\right) \tag{3.18}$$

where A is as in (ii). Observe that, for fixed t, $X_N(t,x) = X_N(t,x')$ whenever $\widetilde{d}_N(x,x') \geq t$, i.e. $X_N(t,x)$ is piecewise constant as a function of x on blocks of size $2^{-[Nt]}$.

While it is plain that the process X_N does not converge to a sensible limit, it is a sensible question to ask for the extremal properties of the process $X_N(t,x)$ as $N \uparrow \infty$. To this end one may introduce the measures

$$\hat{\mu}^t_{\beta,N}(dx) \equiv \frac{e^{\beta\sqrt{N}X_N(t,x)}dx}{\int_0^1 dy e^{\beta\sqrt{N}X_N(t,y)}} \tag{3.19}$$

While these measures are absolutely continuous with respect to Lebesgue measure, they may converge to singular measures as N tends to infinity.

We can naturally apply the construction of Section 2 to the family of measures $\hat{\mu}^t_{\beta,N}$ and define in particular the corresponding measures $\widehat{\mathcal{K}}^\epsilon_{\beta,N,t}$, for $\epsilon > 0$. Of course, for N fixed, the limits as ϵ tends to zero will be trivial since the corresponding functions $S^{t',t}$ are continuous. But keeping $\epsilon > 0$ fixed, letting N tend to infinity, and taking then ϵ to zero, we get a non-trivial answer that coincides with the one we will obtain from the construction described in the first part of this section.

4. Continuous state branching, bridges, and all that.

We will now give a brief summary of the constructions in [BeLG] related to continuous state branching processes. The basic object here is a continuous state branching process characterized by its Laplace functional $u_t(\lambda)$. The process started in $a \geq 0$ will be denoted $X(\cdot,a)$. This can be extended to a genuine two parameter process using the fundamental branching property that state that if $X'(\cdot,b)$ and $X(\cdot,a)$ are independent copies, then $X(\cdot,a+b)$ has the same law as $X'(\cdot,b)+X(\cdot,a)$. Thus they construct their process by demanding that for any $a,b \geq 0$, $X(\cdot,a+b)-X(\cdot,a)$ is independent of the processes $X(\cdot,c)$, for all $c \leq a$, and its law is the same as that of $X(\cdot,b)$. The right continuous version of $X(t,\cdot)$ is then a subordinator. The Markov property of the branching process leads to the following extended construction, due to Bertoin and Le Gall [BeLG]:

Proposition 4.1: *There exists a process $\bar{S}^{(s,t)}(a), 0 \leq s \leq t, a \geq 0$, such that*

(i) For any $0 \leq s \leq t$, $\bar{S}^{(s,t)}$ is a subordinator with Laplace exponent $u_{t-s}(\lambda)$.

(ii) For any integer $p \geq 2$ and $0 \leq t_1 \leq t_2 \leq \cdots \leq t_p$, the subordinators $\bar{S}^{(t_1,t_2)}, \bar{S}^{(t_2,t_3)}, \ldots, \bar{S}^{(t_{p-1},t_p)}$ are independent and

$$\bar{S}^{(t_1,t_p)}(a) = \bar{S}^{(t_1,t_2)} \circ \bar{S}^{(t_2,t_3)} \circ \cdots \circ \bar{S}^{(t_{p-1},t_p)}(a), \forall_{a\geq 0}, \quad a.s. \tag{4.1}$$

(iii) The processes $\bar{S}^{(0,t)}(a)$ and $X(t,a)$ have the same finite dimensional marginals.

It is plain that the $\bar{S}^{(t',t)}$ constructed in [BeLG] are closely related the the $S^{(t',t)}$ of Section 2. In fact, we can first associate to the increasing process $X(t,a)$ the probability distribution function

$$\Theta^t(x) \equiv \frac{X(t,x)}{X(t,1)} \tag{4.2}$$

for $x \in [0,1]$. Then the normalized versions of the process $\bar{S}^{(s,t)}(a)$

$$S^{(s,t)}(x) \equiv \frac{1}{X(t,1)} \bar{S}^{(s,t)}(X(s,1)x) \tag{4.3}$$

can be represented as

$$S^{(s,t)}(a) = \Theta^t\left((\Theta^s)^{-1}(x)\right) \tag{4.4}$$

as in Section 2. Note that the hypothesis of Lemma 2.2 are always satisfied in this case with probability one, since the subordinators $\bar{S}^{(s,t)}$ increase is dense. The associated genealogy of this family of probability measures is then directly equivalent to the genealogy of the underlying branching process.

We denote by $\hat{\mathcal{K}}_t$ empirical genealogical process associated to this measure.

Now let us consider the above construction for the special case of Neveu's branching process. This is the process with Laplace functional $\Psi(u) = u \ln u$.

Bertoin and Le Gall [BeLG] showed that the coalescent process on the integers induced by Neveu's process (as explained in Section 2) $\hat{\mathcal{K}}_t$ coincides with the coalescent process constructed by Bolthausen and Sznitman [BoS]. They also proved the following remarkable result connecting the collection of subordinators to Ruelle's model. Let us state this result for our convenience. Take the parameters $0 < x_1 < \cdots < x_p < 1$ and $0 < t_1 < \cdots < t_p$ linked by the identities

$$t_k = \ln x_{k+1} - \ln x_1 \tag{4.5}$$

for $k = 0, \dots, p-1$, and $t_p = -\ln x_1$. Then the law of the family of jumps of the normalised subordinators $S^{(t_k,t_p)}$, for $k = 0, \dots, p-1$, is the same as the law of Ruelle's probability cascades with parameters x_i, $i = 1, \dots, p$.

Now consider a GREM with finitely many hierarchies and parameters such that the points $y_0 = 0$ and y_i, $i = 1, \dots, p$ are the extremal points of the convex hull of A. Let us remind that $\lim_{N\to\infty} \mathbb{E} f_{\beta,N}(y) = \mathbb{E} f_\beta(y)$ can be computed by (1.8) for any $y \in [0,1]$. Now set

$$\mathbb{E} f_\beta(y_i) = x_i \tag{4.6}$$

where all of the $x_i < 1$. In Theorem 1.5 of [BK2] we proved that the point process $\sum_\sigma \delta_{\mu_{\beta,N}(\sigma':d_N(\sigma,\sigma')>y_1),\dots,\mu_{\beta,N}(\sigma':d_N(\sigma,\sigma')>y_p)}$ in $[0,1]^p$ converge to Ruelle's probability cascades with parameters x_i, $i=1,\dots,p$.

Combining these two results yields

Proposition 4.2:_Let $\mu_{\beta,N}$ be the Gibbs measure associated to a GREM with finitely many hierarchies satisfying (4.6) at the extremal points y_i, $i=1,\dots,k$ of the convex hull of the function A. Then the family of functions $\widetilde{S}_{\beta,N}^{(y_k,y_p)}$, $k=0,1,\dots,p-1$ defined according to (3.11) with respect to $\mu_{\beta,N}$ converges in law, and the limit has the same distribution as the family of normalized subordinators (4.3) or (4.4) $S^{(t_k,t_p)}$, $k=0,1,\dots,p-1$ in the sense that the joint distribution of their jumps has the same law, provided t_k is chosen according to (4.6), (4.5)._

We can actually prove a more general result. From the preceding theorem we expect that Neveu's process will provide the universal limit for all of our models; the dependence on the particular model (i.e. the function A) and on the temperature must come from a rescaling of time. Set $x(y)\equiv \mathbb{E}f_\beta(y)$, where $\mathbb{E}f_\beta(y)$ is defined by the function A through (1.8). Set also $T=-\ln x(0)$, and for $y>0$

$$t(y)=T+\ln x(y). \tag{4.7}$$

Define the family of distribution functions

$$\Theta^y_{f_\beta}(x)\equiv\frac{X(t(y),x)}{X(t(y),1)} \tag{4.8}$$

where X is Neveu's process. Construct the genealogical process $\hat{\mathcal{K}}_y^{f_\beta}=\lim_{\epsilon\downarrow 0}\hat{\mathcal{K}}_y^{f_\beta,\epsilon}$ associated to this family of measures. Then

Theorem 4.3: _Consider a Derrida model with general function A such that A does not touch its convex hull $\bar{A}$ in the interior of any interval where $\bar{A}$ is linear. Then_

$$\mathcal{K}_{\beta,N}\xrightarrow{\mathcal{D}}\hat{\mathcal{K}}_1^{f_\beta} \tag{4.9}$$

Theorem 4.3 is the main result of this paper. The basis of the proof are the Ghirlanda-Guerra identities. In fact we will give two proofs. In Section 5 we will use a somewhat indirect argument that use Proposition 4.2 to show in a somewhat contorted way that the same Ghirlanda-Guerra identities hold for the limit of $\mathcal{K}_{\beta,N}$ and for $\mathcal{K}_1^{f_\beta}$. In Section 6 we give a direct proof of the Ghirlanda-Guerra identities for the Bolthausen-Snitzman coalescent.

5. Coalescence and Ghirlanda-Guerra identities.

In this section we will prove Theorem 4.3. As it was remarked in Section 2 $\mathcal{K}_t$ associated with a flow of measures is completely determined by its moments (2.23) which can be expressed via genealogical distance distributions of the corresponding coalescent. So, we will prove that the moments of $\mathcal{K}_{\beta,N}$, which are the n-replica distance distributions in our spin glass model (1.10) converge to the genealogical distance distributions on the integers constructed from the family of measures $\Theta^y_{f_\beta}$ based on the time-changed Neveu branching process.

It will be done by showing that the Ghirlanda-Guerra relations, which we have established in [BK3] to hold for the family of limiting measures and which determine them completely, are satisfied for the proposed limit.

In addition, it gives us the connection between the n-replica distance distribution function of the CREM with the genealogical distance distribution function of the Bolthausen-Sznitman coalescent.

Theorem 5.1: *Under the same assumptions as in Theorem 4.3, for any $n \in \mathbb{N}$,*

$$\begin{aligned}&\lim_{N\uparrow\infty} \mathbb{E}\mu^{y,\otimes n}_{\beta,N}\left(d_N(\sigma^1,\sigma^2) \le y_1, ,\dots, d_N(\sigma^{n-1},\sigma^n) \le y_{n(n-1)/2}\right)\\ &= \mathbb{P}\left(\rho_1(1,2) \le t(y_1),\dots,\rho_1(n-1,n) \le t(y_{n(n-1)/2})\right)\end{aligned} \tag{5.1}$$

where $t(y)$ is defined in Theorem 4.3. The distance ρ_1 is the distance on integers for the Bolthausen-Sznitman coalescent, induced through (2.19) by the genealogical distance γ_1 of the flow of measures $X(t,x)/X(t,1)$ of Neveu's branching process.

Proof: The proof of this theorem, and in fact the entire identifications of the limiting processes with objects constructed from Neveu's branching process relies on the Ghirlanda-Guerra identities [GG] that were derived for the models considered here in [BK3]. We restate this result in a slightly modified form. Let us remind that the family of measures $\mathbb{Q}^n_{\beta,N}$ is determined on the space $[0,1]^{n(n-1)/2}$ as $\mathbb{E}\mu^{\otimes n}_{\mathbb{N},\beta}(\underline{d_N} \in \cdot)$ where $\underline{d_N}$ denotes the vector of replica distances $d_N(\sigma^k,\sigma^l)$, $1 \le k < l \le n$. Denote by $\mathcal{B}_k$ the vector of the first $k(k-1)/2$ coordinates.

Theorem 5.2: [BK3] *The family of measures $\mathbb{Q}^{(n)}_{\beta,N}$ converge to limiting measures $\mathbb{Q}^{(n)}_\beta$ for all finite n, as $N \uparrow \infty$. Moreover, these measures are uniquely determined by the distance* *distribution functions* f_β*. They satisfy, for any* $y \in [0,1]$*,* $n \in \mathbb{N}$ *and* $k \leq n$*,*

$$\mathbb{Q}_\beta^{(n+1)}\left(d(k,n+1) \leq y|\mathcal{B}_n\right) = \frac{1}{n} f_\beta(y) + \frac{1}{n}\sum_{l \neq k}^{n} \mathbb{Q}_\beta^{(n)}\left(d(k,l) \leq y|\mathcal{B}_n\right) \tag{5.2}$$

Let us recall that due to the ultrametric property of d_N, these identities determine the measures $\mathbb{Q}_\beta^{(n)}$ uniquely. Thus, we must show that the right-hand side of (5.1) satisfies, for $t < 1$,

$$\mathbb{P}\left(\rho_1(k,n+1) \leq t \mid \mathcal{B}_n\right) = \frac{1}{n} e^{t-1} + \frac{1}{n}\sum_{l \leq n, l \neq k} \mathbb{P}\left(\rho_1(k,l) \leq t \mid \mathcal{B}_n\right) \tag{5.3}$$

There are *two* ways to verify that (5.3) holds for the Bolthausen-Sznitman coalescent.

The first one is to observe that relation (5.3) involves only the marginals of the coalescent at a finite set of times t_i. By Theorem 5 of Bertoin-Le Gall [BeLG], these can be expressed in terms of Ruelle's cascades modulo the appropriate time change. Thus, by Theorem 1.5 of [BK2] these probabilities can be expressed as limits of a suitably constructed GREM (with finitely many hierarchies) for which the Ghirlanda-Guerra relations do hold by Proposition 1.8 of [BK2]. Thus (5.3) is satisfied. $\diamondsuit$

The second way is to verify directly that Ghirlanda-Guerra relations (5.3) hold for the Bolthausen-Sznitman coalescent. This is the subject of the next Section 5.

6. Ghirlanda-Guerra identities and Chinese restaurant processes

The Ghirlanda-Guerra identities appear naturally if one considers exchangeable random partitions Π on $\mathbb{N}$, introduced by J. Pitman under the name of Chinese restaurant processes. For each parameter $0 < x < 1$ this partition can be constructed as follows. Let Π_n denote thet restriction of Π to the first n positive integers. Then, conditionally given $\Pi_n = \{A_1, \ldots, A_k\}$ for any particular partition of $\{1, 2, \ldots, n\}$ into k subsets (tables) A_i of sizes n_i, $i = 1, \ldots, k$, the partition Π_{n+1} is an extension of Π_n such that the number $n+1$ (new customer) is attached to the class (table) A_i with probability $(n_i - x)/n$, and forms a new class (sits at a new table) with probability kx/n. Let us denote by $p(n_1, \ldots, n_k)$ the probability of partitions Π with Π_n a particular partition of k classes of sizes $n_1, \ldots, n_k$ respectively. Then

$$p(n_1 \cup 1, n_2, \ldots, n_k) = \frac{1-x}{n} p(n_1, \ldots, n_k) \tag{6.1}$$

Let $q(n_1,\dots,n_k)$ be the probability of all coarser partitions than that, i.e.

$$\begin{aligned} q(n_1,\dots,n_k) = p(n_1,n_2,\dots,n_k) &+ \sum_{i<j} p(n_1,\dots,n_{i-1},n_i\cup n_j,n_{i+1},\dots,n_{j-1},n_{j+1},\dots,n_k) \\ &+ \sum_{i<j<r} p(n_1,\dots,n_i\cup n_j\cup n_r,\dots,n_k) + \cdots + p(n_1\cup n_2\cup\cdots\cup n_k) \end{aligned} \tag{6.2}$$

Then by (6.1)

$$\begin{aligned} q(n_1\cup 1,n_2,\dots,n_k) = &\frac{(1-x)}{n} q(n_1,n_2,\dots,n_k) \\ &+ \frac{1}{n}\sum_{i=2}^{k} n_i q(n_1\cup n_i,n_2,\dots,n_{i-1},n_{i+1},\dots,n_k) \\ &+ \frac{1}{n}(n_1-1) q(n_1,n_2,\dots,n_k) \end{aligned} \tag{6.3}$$

These formulas have been noticed by Pitman [P], Ch.10.6, and these are precisely Ghirlanda-Guerra identities for Ruelle's REM.

One way to see this, is to combine the results of [Pi1] and [PPY]. It should be said first that Π is a partially exchangeable random partition in the sense of [Pi1]. Then, given the sequence of its a.s. limiting relative frequencies of classes P_i *in order of appearance*, the conditional distribution of Π given the whole sequence (P_i) is as follows: for each n conditionally given P_i and $\Pi_n = \{A_1,\dots,A_k\}$ where A_i are in order of appearance, Π_{n+1} is an extension of Π_n in which $n+1$ attaches to class A_i with probability P_i, $1\le i\le k$ and forms a new class with probability $1-\sum_{i=1}^k P_i$. In other words

$$p(n_1,...,n_k) = \mathbb{E}\Big[\prod_{i=1}^{k} P_i^{n_i-1}\prod_{i=1}^{k-1}\Big(1-\sum_{j=1}^{i} P_j\Big)\Big] \tag{6.4}$$

In the case of the Chinese restaurant process

$$p(n_1,...,n_k) = \frac{x\times 2x\times\cdots\times kx}{n!}\prod_{i=1}^{k}(1-x)(2-x)\cdots(n_i-x) \tag{6.5}$$

The function $p(n_1,...,n_k)$ being symmetric, Π is an exchangeable random partition according to [Pi1]. Furthermore, again due to [Pi1], computing the moments from (6.4) and (6.5) one checks that the limiting relative frequencies in order of appearance in the Chinese restaurant process are $P_i = (1-W_1)(1-W_2)\cdots(1-W_{i-1})W_i$ with W_i independent beta $(1-x, ix)$. From the other point of view it has been shown in [PPY] that if $T=\sum_i \Delta_i$ has a stable distribution with index x, with $\Delta_1 > \Delta_2 > \dots$ being points of the Poisson point process on $(0,\infty)$, then the sequence $\Delta_{(i)}/T$ in size-biased order (this means, that given the whole sequence Δ_i/T, and U_i independent random variables uniform on $[0,1]$, then $U_1 \in \Delta_{(1)}/T$, $U_{\min\{j: U_j \notin \Delta_{(1)}/T\}} \in \Delta_{(2)}/T$ etc) has the same distribution of products of independent beta random variables. Thus the limiting frequencies of the Chinese restaurant process Π ranked by size are distributed as Δ_i/T, i.e. they have Poisson-Dirichlet distribution with parameter x. Then the partition Π of the Chinese restaurant process obeys the classical Kingman's construction: given the sequence of normalised jumps of the stable subordinator (Δ_i/T) with index x and given U_i independent uniform random variables on [0, 1], Π is distributed as a partition of blocks of indices of U_i belonging to the same intervals $\Delta_i/T \in [0,1]$.

We denote by $(x,0)$-partitions distributed as a Chinese restaurant process with parameter x.

To generalise this construction to the Bolthausen-Sznitman coalescent, let us introduce an operation of coagulation, see [Pi2]: for a partition $\pi = (A_1, A_2, \dots,)$ and $\Pi = (B_1, B_2, \dots)$, the Π-coagulation of π consists of blocks of the form $\bigcup_{j \in B_i} A_j$. Then by [BS] *the Markov kernels $(e^{-t},0)$-coagulation, $t \geq 0$, on partitions of $\mathbb{N}$ form a semi-group. The Markov process*

$$\mathbb{P}^{\pi}(\Pi(t+) \in \cdot) = (e^{t-T},0) - \textit{coagulation of } \pi$$

is distributed as the Bolthausen-Sznitman coalescent. It starts from a partition of singletons at time T and finishes by a partition of one block $\mathbb{N}$ at time 0. (The semi-group property can be also seen from the fact that the limiting frequencies of $(e^{-t},0)$-partitions are distributed as normalised jumps of stable subordinators and from their matching condition (4.1).) Then the marginals of $\Pi(t)$ at times $0 = t_0 < t_1 < \dots < t_{p-1} < t_p = T$ can be constructed as follows. Let $x_i = e^{t_{i-1}-t_p}$, $0 < x_1 < x_2 < \dots < x_p < 1$. Then $\Pi(t_{p-1}+)$ is distributed as the Chinese restaurant process with parameter x_p. Next, we define the partition $\Pi(t_{p-2}+)$ as the Chinese restaurant process on the classes of partition $\Pi(t_{p-1}+)$ with parameter $x_{p-1}/x_p = e^{t_{p-2}-t_{p-1}}$: this means that given already the classes $A_1^{p-1}, \dots, A_k^{p-1}$ obtained from $A_1^p, \dots, A_l^p$, where A_i^{p-1} consists of l_i blocks of Π^p, $i = 1, \dots, k$, $l_1 + \dots + l_k = l$, the block A_{l+1}^p joins A_i^{p-1} with probability $(l_i^{p-1} - x_{p-1}/x_p)/l$ and forms a new class with probability $kx_{p-1}/(x_p l)$. One iterates this procedure with parameters $x_{p-2}/x_{p-1}, \dots, x_1/x_2$ to construct the partitions $\Pi(t_{p-3}+), \dots, \Pi(t_0+)$. Then by the semi-group property $\Pi(t_i+)$ is distributed as a Chinese restaurant process with parameter $x_{i+1} = e^{t_i - t_p}$ for all $i = 0, 1, \dots, p-1$. It follows from all said above and from (6.3) that for all $t < t_p$

$$\mathbb{P}\left(\rho_{t_p}(k,n+1) > t \mid \mathcal{B}_n\right) = \frac{1-e^{t-t_p}}{n} + \frac{1}{n}\sum_{l\leq n, l\neq k} \mathbb{P}\left(\rho_{t_p}(k,l) > t \mid \mathcal{B}_n\right) \tag{6.6}$$

which is equivalent to (5.3).

Recently Ph. Marchal found another beautiful way to identify the limiting frequencies of the Chinese restaurant process with Ruelle's REM (or, equivalently, the range of the stable subordinator) and also the iterated Chinese restaurant process with the Bolthausen-Sznitman coalescent, see [M].

References

[AC] M. Aizenmann and P. Contucci, On the stability of the quenched state in mean field spin-glass models, J. Statist. Phys. **92**, 765-783 (1998).

[Be1] J. Bertoin, Homogeneous fragmentation processes, Probab. Theory Related Fields **121** , 301–318 (2001).

[Be2] J. Bertoin, Self-similar fragmentations, Ann. Inst. H. Poincaré (Probab. Statist.) **38**, 319–340 (2002).

[BeLG] J. Bertoin and J.F. Le Gall, The Bolthausen-Snitzman coalescent and the genealogy of continuous state branching processes, Probab. Theor. Tel. Fields **117**, 249-266 (2000).

[BePi] J. Bertoin and J. Pitman, Two coalescents derived from the ranges of stable subordinators, Electron. J. Probab. **5**, no. 7 (2000).

[BeYo] J. Bertoin and M. Yor, On subordinators, self-similar Markov processes and some factorizations of the exponential variable, Electron. Comm. Probab. **6**, 95–106 (2001)

[BoS] E. Bolthausen and A.-S. Snitzman, On Ruelle's probability cascades and an abstract cavity method, Commun. Math. Phys. **107**, 247-276 (1998).

[BK1] A. Bovier and I. Kurkova, Derrida's Generalized random Energy models 1. Poisson cascades and the extremal process, preprint 727, U. Paris 6 (2002).

[BK2] A. Bovier and I. Kurkova, Derrida's Generalized Random Energy models 2. Gibbs measures and probability cascades, preprint 728, U. Paris 6 (2002).

[BK3] A. Bovier and I. Kurkova, Derrida's Generalized Random Energy models 3. Models with continuous hierarchies, preprint 729, U. Paris 6 (2002).

[BKL] A. Bovier, I. Kurkova, and M. Löwe, Fluctuations of the free energy in the REM and the p-spin SK-models, Ann. Probab. **30**, 605-651 (2002).

[B] A. Bovier, Statistical mechanics of disordered systems, MaPhySto Lecture Notes 10, Aarhus, 2001.

[CCP] D. Capocaccia, M. Cassandro, and P. Picco, On the existence of thermodynamics for the Generalised Random Energy Model. J. Stat. Physics **46**, 493–505 (1987).

[D] B. Derrida, A generalization of the random energy model that includes correlations between the energies, J. Physique. Lett. **46**, 401–407 (1985).

[DG1] B. Derrida and E. Gardner, Solution of the generalized random energy model, J. Phys. C **19**, 2253–2274 (1986).

[DG2] B. Derrida and E. Gardner, Magnetic properties and function $q(x)$ of the generalised random energy model, J. Phys. C **19**, 5783-5798 (1986)

[GG] S. Ghirlanda, F. Guerra. General properties of the overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity, J. Phys. A **31**, 9144-9155 (1998).

[Gu] F. Guerra, Broken replica symmetry bounds in the mean field spin glass model, preprint 2002.

[GT] F. Guerra and F. Toninelli, The thermodynamics limit in mean field spin glass models, Commun. Math. Phys.**230**, 71-79 (2002).

[GMP] A. Galvez, S. Martinez, and P. Picco, Fluctuations in Derrida's random energy and generalized random enery models, J. Stat. Phys. **54**, 515-529 (1989).

[Le] M. Ledoux. On the distribution of overlaps in the Sherrington-Kirkpatrick spin glass model. J. Statist. Phys. 100 (2000), 871-892.

[M] Ph. Marchal Regenerative sets, random partitions and the Bolthausen-Sznitman coalescent. Preprint.

[MPV] M. Mézard, G. Parisi, M.A. Virasoro, Spin-glass theory and beyond. World Scientific, Singapore (1988).

[Ne] J. Neveu, A continuous state branching process in relation with the GREM model of spin glass theory, Rapport interne no 267, École Polytechnique (1992).

[NS1] Ch.M. Newman and D.L. Stein, Thermodynamic chaos and the structure of short range spin glasses, in "Mathematical aspects of spin glasses and neural networks", A. Bovier and P. Picco (Eds.), Progress in Probability, Birkhäuser, Boston (1997).

[NS2] C.M. Newman, D.L. Stein, Metastate approach to thermodynamic chaos. Phys. Rev. E **55**, 5194-5211 (1997).

[NS3] Ch.M. Newman and D.L. Stein, Ultrademocracy of ground states in the infinite ranged spin glass, preprint cond-mat/0301202, 2003.

[Ru] D. Ruelle, A mathematical reformulation of Derrida's REM and GREM. Commun. Math. Phys. **108**, 225-239 (1987).

[P] J. Pitman, Combinatorial stochastic processes, Lecture Notes of the 2002 École des Probabilités de St. Flour; to appear as Springer Lecture Notes in Mathematics.

[Pi1] J. Pitman, Exchangeable and partially exchangeable random partitions, Probab. Theor. Relat. Fields **102**, 145–158 (1995).

[PPY] M. Perman, J. Pitman, M. Yor, Size-biased sampling of Poisson point processes and excursions, Probab. Theor. Relat. Fields **92**, 21–39 (1992).

[SK] D. Sherrington and S. Kirkpatrick, Solvable model of a spin glass, Phys. Rev. Lett. **35**, 1792-1796 (1972).

[T1] M. Talagrand, Rigorous low temperature results for mean field p-spin interaction models, Probab. Theor. Rel. Fields. **117**, 303-360 (2000).

[T2] M. Talagrand, Self organization in the low-temperature region of a spin glass model, preprint 2002.

[T3] M. Talagrand, Mean field models for spin glasses: a first course, Lecture Notes of the 2000 École des Probabilités de St. Flour; to appear as Springer Lecture Notes in Mathematics.

[T4] M. Talagrand, Spin Glasses: a Challenge to Mathematicians, Ergebnisse der Mathematik und ihrer Grenzgebiete Vol. 46, Springer, Berlin-Heidelberg-New York, 2003.