Derrida's Generalized Random Energy models 1: Poisson cascades and extremal processes

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: This is the first of a series of three papers in which we present a full rigorous analysis of a class of spin glass models introduces by Derrida under the name of Generalized Random Energy Models (GREM). They are based on Gaussian random processes on the hypercube $\{-1,1\}^N$ with a hierarchical correlation structure. In this first paper we prove some result concerning the convergence of the extremal process to Poisson cascade processes which will form the basis of the analysis of the Gibbs measures to be presented in the sequel.

Keywords: Gaussian processes, extreme values, order statistics,generalized random energy model, spin glasses, Poisson cascades

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

[1]e-mail: bovier@wias-berlin.de
[2]Research supported in part by the DFG in the Concentration program "Interacting stochastic systems of high complexity".
[3]e-mail: kourkova@ccr.jussieu.fr

1. Introduction.

In 1980, Derrida proposed two simplified models for mean field spin glasses: the *random energy model* (REM)[D1,D2], and the *generalised random energy model*(GREM)[D3,DG1,DG2,DG3]. The former consisted of modelling the random energy landscape as simply i.i.d. Gaussian random variables on the set of spin configurations, $\{-1,1\}^N$. This model can be seen formerly as the limit of the so-called p-spin SK-models [SK], when p tends to infinity [D1]. In spite of its simplicity, this model has proven to be a rather instructive toy model, and has received considerable attention in the mathematical community [DG3,DW,Ei,OP,GMP,Ru,BKL,KP] [Bo]. Of course, in many respects this model is mathematically almost trivial, and physically quite unrealistic, as all the dependence structure that is present in more realistic models like the SK model, is absent. The GREM was introduced in view of *keeping* dependence, while simplifying it to a *hierarchical* structure to still yield a mathematically more tractable model. In fact, the GREM can be seen as a class of models that is obtained by equipping the hypercube $\{-1,1\}^N$ with a tree structure and an associated ultra-metric distance, and then considering standardized Gaussian random fields on the hypercube whose correlation function depends only on this distance. We will call these models *"Derrida's models"* in contrast to the *"Sherrington/Kirkpatrick (SK) models"* where the covariance depends on the Hamming distance, respectively the overlap $R_N(\sigma,\sigma') = N^{-1}\sum_{i=1}^N \sigma_i\sigma_j$. Let us mention in passing that the study of such processes may represent an independent interest in fields other than physics. E.g., in mathematical finance, such processes may represent a very reasonable model for the distribution of assets of a financial portfolio.

In [DG1], B. Derrida and E. Gardner presented a solution of the model with finitely many hierarchies in the sense that they computed the free energy in the thermodynamic limit. A rigorous derivation of this solution (in a somewhat more elegant form) was later obtained by Cappocaccia et al. [CCP]. Derrida and Gardner also considered the limit of their formulae when the number of hierarchies tends to infinity. They argued that for suitable choices of the covariance function, this limits yield approximations for the standard p-spin SK models, even though, as they point out, the quality of the approximations is not spectacular.

In this paper and two companion papers [BK2,BK3] we will undertake a comprehensive study of this class of models. In the present paper we consider only the case when the covariance depends on the ultra-metric valuation d_N via a step function with finitely many steps. This corresponds to the classical GREM of Derrida.

The GREM is mathematically much less understood than the simpler REM. As mentioned above, Cappocaccia et al. [CCP] derived the exact formula for the mean free energy. Galves et al. [GMP] construct a certain point process for the energies the case of the two-level tree. Most of the intuitive insight into the model stems from a paper by Ruelle [Ru] who in fact formulates a new model based on Poisson point processes that are suggested to represent the asymptotics of the fluctuations of the partition function and the Gibbs measures of Derrida's models in the thermodynamic limit. But this connection was established rigorously even in the simplest case of the REM only very recently [BKL].

An important ingredient in the analysis of the REM is the theory of convergence to Poisson processes of the extreme value statistics of (i.i.d.) random variables that is, of course, very well known (see e.g. [LLR,Re]). In the GREM, analogous results will be needed in the correlated case. Here, standard results on extremal processes are much rarer. There are criteria under which the extremal processes have the same limits as in the independent case. However, here we will be mostly interested in situations when this is not the case. The results of this first paper will be general Poisson convergence results for the extremes of hierarchically correlated Gaussian random processes, where the limits will in general be described as cascades of Poisson processes.

1.2. Definition of the models. We will consider Gaussian processes X_σ indexed by the hypercube $\mathcal{S}_N \equiv \{-1,1\}^N$. Let us equip the hypercube $\mathcal{S}_N$ with the natural ultra-metric valuation

$$d_N(\sigma,\sigma') \equiv \frac{1}{N}\left(\min\left(i|\sigma_i \neq \sigma_i'\right) - 1\right) \tag{1.1}$$

(Note that $\exp(-d_N(\sigma,\sigma'))$ is a ultra-metric on $\mathcal{S}_N$). We will consider processes whose covariance is given by a (non-decreasing) function of d_N, i.e. we assume that $\mathbb{E}X_\sigma = 0$, and

$$\operatorname{cov}\left(X_\sigma, X_{\sigma'}\right) = \mathbb{E}X_\sigma X_{\sigma'} = A(d_N(\sigma,\sigma')) \tag{1.2}$$

where $A(x)$ is a probability distribution function on $[0,1]$. Note that the models of the SK-family can be constructed in an analogous way, except that $d_N(\sigma,\sigma')$ is replaced by the overlap $R_N(\sigma,\sigma') \equiv N^{-1}\sum_{i=1}^N \sigma_i\sigma_i'$.

The proper GREM with *finitely* many hierarchies corresponds to the special case when A is the distribution function of a measure that is supported on a finite number, n, of points $x_1,\dots,x_n \in [0,1]$. In that case we denote the mass of the atoms x_i by a_i, and we set

$$\ln\alpha_i = (x_i - x_{i-1})\ln 2, \quad i = 1,\dots,n \tag{1.3}$$

where $x_0 \equiv 0$. Of course $\sum_{i=1}^n a_i = 1$, and $\prod_{i=1}^n \alpha_i = 2$. The hypercube $\mathcal{S}_N$ can then be considered as a n-fold product $\mathcal{S}_N = \prod_{i=1}^n \mathcal{S}_{N\ln\alpha_i/\ln 2}$. We will write $\sigma = \sigma_1\sigma_2\dots\sigma_n$ where

$\sigma_i \in \mathcal{S}_{N \ln \alpha_i / \ln 2}$. Usually we will assume that $x_1 > 0$, $x_n = 1$, and all $a_i > 0$, although at some later stage we will have to look also at what happens when $x_1 = 0$ and $a_n = 0$.

Then the Gaussian process X_σ can be represented in the following way.

Take $\alpha_1^N + (\alpha_1\alpha_2)^N + \cdots + (\alpha_1\alpha_2\cdots\alpha_n)^N$ independent standard Gaussian random variables enumerated as follows: The first α_1^N of them X_{σ_1} are enumerated by α_1^N configurations of spins $\sigma_1 \in \{-1,+1\}^{N \ln \alpha_1 / \ln 2}$, the next $(\alpha_1\alpha_2)^N$ of them $X_{\sigma_1\sigma_2}$ — by configurations of spins $\sigma_1\sigma_2 \in \{-1,+1\}^{N \ln(\alpha_1\alpha_2)/\ln 2}$, etc ... , finally the last $(\alpha_1\alpha_2\cdots\alpha_n)^N = 2^N$ of them $X_{\sigma_1\sigma_2,\dots,\sigma_n}$ are enumerated by configurations $\sigma_1\sigma_2\dots\sigma_n \in \{-1,+1\}^N$.

Then the process X_σ has the same distribution as

$$X_\sigma \equiv \sqrt{a_1} X_{\sigma_1} + \sqrt{a_2} X_{\sigma_1\sigma_2} + \cdots + \sqrt{a_n} X_{\sigma_1\sigma_2\dots\sigma_n}, \quad \text{if } \sigma = \sigma_1\sigma_2\dots\sigma_n, \tag{1.4}$$

The explicit construction in terms of (1.4) will be quite useful.

1.3. Convergence of associated point processes

We introduce the function $u_{\ln\alpha,N}(x)$, $x \in \mathbb{R}$, such that $P(X > u_{\ln\alpha,N}(x)) \sim \alpha^{-N} e^{-x}$, as $N \uparrow \infty$, where X is a standard Gaussian random variable. It can be written explicitly as (see e.g. [LLR])

$$u_{\ln\alpha,N}(x) = \sqrt{2\ln\alpha N} + \frac{x}{\sqrt{2\ln\alpha N}} - \frac{\ln N + \ln\ln\alpha + \ln 4\pi}{2\sqrt{2\ln\alpha N}}. \tag{1.5}$$

Theorem 1.1: *Let $n \in \mathbb{N}$, $n \geq 1$, $0 < a_i < 1$, $\alpha_i > 1$, $i = 1, 2, \dots, n$. Set $\bar{\alpha} \equiv \prod_{i=1}^n \alpha_i$ and assume that $\sum_{i=1}^n a_i = 1$. The point process*

$$\sum_{\sigma = \sigma_1\dots\sigma_n \in \{-1,+1\}^{\ln\bar{\alpha}/\ln 2}} \delta_{u^{-1}_{\ln\bar{\alpha},N}(\sqrt{a_1}X_{\sigma_1} + \sqrt{a_2}X_{\sigma_1\sigma_2} + \cdots + \sqrt{a_n}X_{\sigma_1\sigma_2\dots\sigma_n})}$$

converges weakly to the Poisson point process $\mathcal{P}$ on $\mathbb{R}$ with intensity measure $Ke^{-x}dx$, $K \in \mathbb{R}$ if and only if the following inequalities hold true:

$$a_i + a_{i+1} + \cdots + a_n \geq \ln(\alpha_i\alpha_{i+1}\cdots\alpha_n)/\ln\bar{\alpha} \quad \textit{for all } i = 2, 3, \dots, n. \tag{1.6}$$

Furthermore, if all inequalities in (1.6) are strict, then the constant $K = 1$. If some of them *are equalities, then K is given as*

$$
\begin{aligned}
K = \mathbb{P}\Big(\bigcap_{\substack{k: a_k+\cdots+a_n \\ =\ln(\alpha_k\cdots\alpha_n)/\ln\bar\alpha}} \Big(& Z_{k-1}\sqrt{\frac{a_{k-1}}{(a_{k-1}+\cdots+a_n)(a_k+\cdots+a_n)}} \\
& + Z_{k-2}\sqrt{\frac{a_{k-2}}{(a_{k-2}+\cdots+a_n)(a_{k-1}+\cdots+a_n)}} \\
& + \cdots \\
& + Z_1\sqrt{\frac{a_1}{(a_1+\cdots+a_n)(a_2+\cdots+a_n)}} < 0\Big)\Big)
\end{aligned}
$$

where $Z_1,\ldots,Z_{n-1}$ are $n-1$ independent standard Gaussian random variables.

Remark: The sharp conditions on the parameters a_i and α_i in this theorem are *strictly weaker* than the sufficient conditions one obtains e.g. from Slepian's Gaussian comparison lemma (see e.g. [LLR], Theorem 4.2.1).

Theorem 1.2: *Let $\alpha_i \geq 1$, $i=1,2,\ldots,k$, $\bar\alpha \equiv \prod_{i=1}^k \alpha_i$. Let $Y_{\sigma_1}, Y_{\sigma_1\sigma_2},\ldots,Y_{\sigma_1\ldots\sigma_k}$ be $\alpha_1^N+\cdots+(\alpha_1\cdots\alpha_k)^N$ identically distributed random variables enumerated as in the definition of the GREM. Assume that $1+\alpha_1^N+\cdots+(\alpha_1\cdots\alpha_{k-1})^N$ vectors $(Y_{\sigma_1})_{\sigma_1\in\{-1,1\}^{\ln\alpha_1/\ln\bar\alpha}}$, $(Y_{\sigma_1\sigma_2})_{\sigma_2\in\{-1,1\}^{\ln\alpha_2/\ln\bar\alpha}}$ $\forall\sigma_1\in\{-1,1\}^{\ln\alpha_1/\ln\bar\alpha},\ldots,(Y_{\sigma_1\sigma_2\ldots\sigma_k})_{\sigma_k\in\{-1,1\}^{\ln\alpha_k/\ln\bar\alpha}}$ $\forall\,\sigma_1\ldots\sigma_{k-1}\in\{-1,1\}^{\ln(\alpha_1\cdots\alpha_{k-1})/\ln\bar\alpha}$ are independent. Let also $v_{N,1}(x),\ldots,v_{N,k}(x)$ be functions on $\mathbb{R}$ such that the following point processes*

$$
\begin{aligned}
&\sum_{\sigma_1}\delta_{v_{N,1}(Y_{\sigma_1})} \to \mathcal{P}_1 \\
&\sum_{\sigma_2}\delta_{v_{N,2}(Y_{\sigma_1\sigma_2})} \to \mathcal{P}_2 \quad \forall\sigma_1 \\
&\cdots \\
&\sum_{\sigma_k}\delta_{v_{N,k}(Y_{\sigma_1\sigma_2\ldots\sigma_k})} \to \mathcal{P}_k \quad \forall\sigma_1\ldots\sigma_{k-1}
\end{aligned}
\tag{1.7}
$$

converge weakly to the Poisson point processes $\mathcal{P}_1,\ldots,\mathcal{P}_k$ on $\mathbb{R}$ with intensity measures $K_1e^{-x}dx,\ldots,K_ke^{-x}dx$ with some constants $K_1,\ldots,K_k$ respectively. Then the following point processes on $\mathbb{R}^k$

$$
\mathcal{P}_N^{(k)} \equiv \sum_{\sigma_1}\delta_{v_{N,1}(Y_{\sigma_1})}\sum_{\sigma_2}\delta_{v_{N,2}(Y_{\sigma_1\sigma_2})}\cdots\sum_{\sigma_k}\delta_{v_{N,k}(Y_{\sigma_1\sigma_2\ldots\sigma_k})} \to \mathcal{P}^{(k)}
$$

converge weakly to point process $\mathcal{P}^{(k)}$ on $\mathbb{R}^k$, which is characterised by the following generating functions:

$$
\begin{aligned}
F_{\Delta_1\times\cdots\times\Delta_k}(z) &\equiv \mathbb{E}z^{\sum_{x_1}\mathbb{1}_{\{x_1\in\Delta_1\}}\cdots\sum_{x_k}\mathbb{1}_{\{x_k\in\Delta_k\}}} \\
&= f_{1,\Delta_1}(f_{2,\Delta_2}(f_{3,\Delta_3}\cdots(f_{k-1,\Delta_{k-1}}(f_{k,\Delta_k}(z)))\cdots)), \qquad |z|<1
\end{aligned}
\tag{1.8}
$$

where $f_{i,\Delta_i}(z) = e^{K_i(z-1)(e^{-a_i}-e^{-b_i})}$, $\Delta_i = (a_i, b_i]$ *with* $a_i, b_i \in \mathbb{R}$ *or* $b_i = \infty$, $i = 1, 2, \dots, k$.

Moreover, the following independence properties of the counting random variables of the process $\mathcal{P}^{(k)}$, $\sum_{x_1} \mathbb{I}_{\{x_1 \in \Delta_1^j\}} \cdots \sum_{x_k} \mathbb{I}_{\{x_k \in \Delta_k^j\}}$, *corresponding to the intervals* $\Delta_1^j \times \cdots \times \Delta_k^j$, $\Delta_i^j = [a_i^j, b_i^j)$, $j = 1, 2, \dots, r$, $r > 1$, $i = 1, \dots, k$ *hold true:*

(i) If the first components of these intervals are disjoint, i.e. $a_1^1 \le b_1^1 \le a_1^2 \le b_1^2 \le \cdots a_1^r \le b_1^r$, *then these r.v. are independent.*

(ii) If the first $l-1$ *components of these intervals coincide and the* l*th components are disjoint, i.e.* $\Delta_i^1 = \cdots = \Delta_i^r$ *for* $i = 1, \dots, l-1$ *and* $a_l^1 \le b_l^1 \le a_l^2 \le b_l^2 \le \cdots a_l^k \le b_l^r$, *then these r.v. are conditionally independent under condition that* $\sum_{x_1} \mathbb{I}_{\{x_1 \in \Delta_1\}} \cdots \sum_{x_{l-1}} \mathbb{I}_{\{x_{l-1} \in \Delta_{l-1}\}}$ *is fixed.*

Remark: This theorem is a generalisation of Theorem 3 in [GMP]: we do not specify the distribution of the r.v. $Y_{\sigma_1 \dots \sigma_i}$ in the assumptions, and do not impose their independence (but only of their vectors).

Remark: The process $\mathcal{P}^{(k)}$ is called a *Poisson cascade* with k levels. It is best understood in terms of the following iterative construction. If $k = 1$, it is just a Poisson point process on $\mathbb{R}$ with intensity measure $K_1 e^{-x} dx$. To construct $\mathcal{P}^{(2)}$ on $\mathbb{R}^2$, we place the process $\mathcal{P}^{(1)}$ for $k = 1$ on the axis of the first coordinate and through each of its points draw a straight line parallel to the axis of the second coordinate. Then we put on each of these lines independently a Poisson point process with intensity measure $K_2 e^{-x} dx$. These points on $\mathbb{R}^2$ form the process $\mathcal{P}^{(2)}$ This procedure is now simply iterated k times.

Theorems 1.1 and 1.2 combined give a first important result that establish what different point processes may be constructed in the GREM.

Theorem 1.3: *Let* $\alpha_i \ge 1$, $0 < a_i < 1$, $i = 1, 2, \dots, n$, $\prod_{i=1}^n \alpha_i = 2$, $\sum_{i=1}^n a_i = 1$. *Let* $J_1, J_2, \dots, J_m \in \mathbb{N}$ *be the indices such that* $0 = J_0 < J_1 < J_2 < \cdots < J_m = n$. *We denote by* $\bar{a}_l \equiv \sum_{i=J_{l-1}+1}^{J_l} a_i$, $\bar{\alpha}_l \equiv \prod_{i=J_{l-1}+1}^{J_l} \alpha_i$, $l = 1, 2, \dots, m$, *and introduce standard Gaussian random variables*

$$\bar{X}^{\sigma_1 \dots \sigma_{J_{l-1}}}_{\sigma_{J_{l-1}+1}\sigma_{J_{l-1}+2}\cdots\sigma_{J_l}} \equiv \Big(\sqrt{a_{J_{l-1}+1}} X_{\sigma_{J_1} \dots \sigma_{J_{l-1}} \sigma_{J_{l-1}+1}} + \sqrt{a_{J_{l-1}+2}} X_{\sigma_{J_1} \dots \sigma_{J_{l-1}} \sigma_{J_{l-1}+1} \sigma_{J_{l-1}+2}} + \cdots + \sqrt{a_{J_l}} X_{\sigma_{J_1} \dots \sigma_{J_{l-1}} \sigma_{J_{l-1}+1} \dots \sigma_{J_l}} \Big) / \sqrt{\bar{a}_l}. \tag{1.9}$$

Assume that the partition $J_1, J_2, \dots, J_m$ *satisfies the following condition: for all* $l = 1, 2, \dots, m$

and all k such that $J_{l-1}+2 \leq k \leq J_l$

$$(a_k + a_{k+1} \cdots + a_{J_l-1} + a_{J_l})/\bar{a}_l \geq \ln(\alpha_k \alpha_{k+1} \cdots \alpha_{J_l-1}\alpha_{J_l})/\ln(\bar{\alpha}_l). \tag{1.10}$$

Then the point process

$$\mathcal{P}_N^{(m)} \equiv \sum_{\sigma_1 \ldots \sigma_{J_1}} \delta_{u^{-1}_{\ln \bar{\alpha}_1, N}(\bar{X}_{\sigma_1 \cdots \sigma_{J_1}})} \sum_{\sigma_{J_1+1} \ldots \sigma_{J_2}} \delta_{u^{-1}_{\ln \bar{\alpha}_2, N}(\bar{X}^{\sigma_1 \cdots \sigma_{J_1}}_{\sigma_{J_1+1} \cdots \sigma_{J_2}})} \cdots \sum_{\sigma_{J_{m-1}+1} \ldots \sigma_{J_m}} \delta_{u^{-1}_{\ln \bar{\alpha}_m, N}(\bar{X}^{\sigma_1 \cdots \sigma_{J_{m-1}}}_{\sigma_{J_{m-1}+1} \cdots \sigma_{J_m}})}$$

converges weakly to the process $\mathcal{P}^{(m)}$ on $\mathbb{R}^m$ defined in Theorem 1.2 with constants $K_1, \ldots, K_m$. The constant

$$K_l = 1 \tag{1.11}$$

if all $J_l - J_{l-1} - 1$ inequalities in (1.10) for $k = J_{l-1}+2, \ldots, J_l$ are strict. Otherwise

$$\begin{aligned} K_l = \mathbb{P}\Big(\bigcap_{\substack{k: J_{l-1}+1 \leq k \leq J_l, \\ (a_k + \cdots + a_{J_l})/\bar{a}_{J_l} = \ln(\alpha_k \cdots \alpha_{J_l})/\ln \bar{\alpha}_l}} & \Big(Z_{k-1} \sqrt{\frac{a_{k-1}}{(a_{k-1} + \cdots + a_{J_l})(a_k + \cdots + a_{J_l})}} \\ & + Z_{k-2} \sqrt{\frac{a_{k-2}}{(a_{k-2} + \cdots + a_{J_l})(a_{k-1} + \cdots + a_{J_l})}} \\ & + \cdots \\ & + Z_{J_{l-1}+1} \sqrt{\frac{a_{J_{l-1}+1}}{(a_{J_{l-1}+1} + \cdots + a_{J_l})(a_{J_{l-1}+2} + \cdots + a_{J_l})}} < 0 \Big) \Big) \end{aligned} \tag{1.12}$$

where $Z_{J_{l-1}+1}, \ldots, Z_{J_l}$ are independent standard Gaussian r.v.

Remark: Given a_i, α_i, $i = 1, \ldots, n$, the partition of indices $J_1, \ldots, J_m$ satisfying the condition (1.10) is generally not unique. Therefore this corollary yields a *family* of convergent point processes associated with the GREM. In the sequel we will frequently drop the upper index and write $\mathcal{P} \equiv \mathcal{P}^{(m)}$, in particular when the value of m can be read off from the arguments.

1.4. The extremal process.

We are now also in the position to formulate a result on the extreme order statistics of the random variables X_σ. It is clear that the extremal process should be constructed from one of the cascade processes that were constructed above. It remains to find out which one. The answer is of course simple: it is the one that provides the largest values, but still exists. It is easy to see how this will have to be done: one has simply to try to group as many intermediate hierarchies together to a single point process. In other words, among the possible choices for the integers $J_1,\dots,J_m$ allowed in Theorem 1.3 we must choose those that have maximal spacing. Algorithmically, this is achieved by setting $J_0 \equiv 0$, and

$$J_l \equiv \min\{J > J_{l-1} : A_{J_{l-1}+1,J} > A_{J+1,k}\ \forall k \geq J+1\}. \tag{1.13}$$

where we have put $A_{j,k} \equiv \sum_{i=j}^{k} a_i/(2\ln(\prod_{i=j}^{k} \alpha_i))$.

Of course it will be essential that

Proposition 1.4: *The sequence $J_1,\dots,J_m$ defined by (1.13) verifies the conditions (1.10) for all k such that $J_{l-1}+2 \leq k \leq J_l$ and all $l = 1,2,\dots,m$.*

Set now $\bar{a}_l \equiv \sum_{i=J_{l-1}+1}^{J_l} a_i$, $\bar{\alpha}_l \equiv \prod_{i=J_{l-1}+1}^{J_l} \alpha_i$, $\gamma_l \equiv \sqrt{\bar{a}_l}/\sqrt{2\ln\bar{\alpha}_l}$, $l = 1,2,\dots,m$. Clearly by (1.13) $\gamma_1 > \gamma_2 > \dots > \gamma_m$. Define the function $U_{J,N}$ by

$$U_{J,N}(x) \equiv \sum_{l=1}^{m} \left(\sqrt{2N\bar{a}_l\ln\bar{\alpha}_l} - N^{-1/2}\gamma_l(\ln(N(\ln\bar{\alpha}_l)) + \ln 4\pi)/2\right) + N^{-1/2}x \tag{1.14}$$

and the point process

$$\mathcal{E}_N \equiv \sum_{\sigma\in\{-1,1\}^N} \delta_{U_{J,N}^{-1}(X_\sigma)}. \tag{1.15}$$

Then the following holds true:

Theorem 1.5: *(i) The point process $\mathcal{E}_N$ converges weakly, as $N \uparrow \infty$, to the point process on $\mathbb{R}$*

$$\mathcal{E} \equiv \int_{\mathbb{R}^m} \mathcal{P}^{(m)}(dx_1,\dots,dx_m)\delta_{\sum_{l=1}^{m}\gamma_l x_l} \tag{1.16}$$

where $\mathcal{P}^{(m)}$ is the Poisson cascade introduced in Theorem 1.3 corresponding to the partition $J_1,\dots,J_m$ given by (1.13).

(ii) $\mathcal{E}$ exists provided by the fact that $\gamma_1 > \dots > \gamma_m$. It is the cluster point process on $\mathbb{R}$ containing an a.s. finite number of points in any interval $[a,\infty)$, $a \in \mathbb{R}$. The probability that there exists at least one point of $\mathcal{E}$ in the interval $[a,\infty)$ is decreasing exponentially as $a \to +\infty$.

(iii) Furthermore, we have $\max_\sigma(X_\sigma/\sqrt{N}) \to \sqrt{\bar{a}_1 2\ln\bar{\alpha}_1} + \dots + \sqrt{\bar{a}_m 2\ln\bar{\alpha}_m}$ a.s. and also $\mathbb{E}(\max_\sigma X_\sigma/\sqrt{N}) \to \sqrt{2\bar{a}_1\ln\bar{\alpha}_1} + \dots + \sqrt{2\bar{a}_m\ln\bar{\alpha}_m}$.

Remark: Note that if some of the inequalities in the condition (1.10) hold with equality, there are two (or several) of the point processes that to leading order give the same contribution to the extremes. The degeneracy is only weakly lifted in favour of the process with the longest increments in the J_i through the $\ln N$ term in the functions $u_{\ln\alpha,N}$ defined in (1.5). If one considers what happens here as a function of the parameters, one sees that the process with fewer levels 'dies out', and the process with extra levels takes over. While the values of the extremes are in leading order the same at the coexistence point, as the inequality gets strictly violated, these values now begin to drop substantially.

To our knowledge this result is new, and one of the rather few explicit results on the convergence of the extremal process in the case when correlations are strong enough to have an effect on the extremal process. Of course our result shows the expected effect of clustering of extremes in this case. Note that the case $m = 1$ arises when we are in the standard situation of convergence to the simple Poisson process.

In the remainder of this paper we will present the proofs of the theorems stated above.

Acknowledgements: A.B. thanks Francis Comets and the Université Paris 7 for their hospitality during the final period of the writing of this paper.

2. Proofs of point process convergence results.

In this section we will proof all the results related to the convergence of the various point processes, including the extremal process the partition function and Ruelle's processes $\mathcal{W}_{\beta,N}$. Throughout the rest of the paper we will use the relation $A(N) \sim B(N)$ to indicate that $\lim_{N\uparrow\infty} \frac{A(N)}{B(N)} = 1$.

The technical key to Theorem 1.1 is the following Lemma:

Lemma 2.1: *Let $n \in \mathbb{N}$, $n \geq 1$, $0 < a_i < 1$, $\alpha_i > 1$, $i = 1, 2, \dots, n$, and $\bar{\alpha} \equiv \prod_{i=1}^n \alpha_i$. Assume that $\sum_{i=1}^n a_i = 1$. If all inequalities (1.6) hold true, then for all $x \in \mathbb{R}$*

$$\mathbb{P}\Big(\forall\sigma \in \{-1,+1\}^{\ln\bar{\alpha}/\ln 2} : \sqrt{a_1}X_{\sigma_1} + \sqrt{a_2}X_{\sigma_1\sigma_2} + \dots + \sqrt{a_n}X_{\sigma_1\sigma_2\dots\sigma_n} < u_{\ln\bar{\alpha},N}(x)\Big) \sim e^{-Ke^{-x}}, \tag{2.1}$$

as $N \uparrow \infty$, where K is defined in Theorem 1.1. If at least one of inequalities (1.6) fails, then

$$\mathbb{P}\Big(\forall\sigma \in \{-1,+1\}^{\ln\bar{\alpha}/\ln 2} : \sqrt{a_1}X_{\sigma_1} + \sqrt{a_2}X_{\sigma_1\sigma_2} + \dots + \sqrt{a_n}X_{\sigma_1\sigma_2\dots\sigma_n} < u_{\ln\bar{\alpha},N}(x)\Big) \sim 1. \tag{2.2}$$

Proof. If $n = 1$, then all $\bar{\alpha}^N$ random variables in the event of (2.1) are independent. In this case the statement of the lemma is a classical result of extreme values theory, see e.g. Theorem 5.7.2 in [LLR].

Let now $n \geq 2$. We define

$$P(i,y) \equiv \mathbb{P}\Big(\exists \sigma_{i+1} \in \{-1,+1\}^{\ln \alpha_i/\ln 2}, \dots, \sigma_n \in \{-1,+1\}^{\ln \alpha_n/\ln 2} : \frac{\sqrt{a_i}X + \sqrt{a_{i+1}}X_{\sigma_{i+1}} + \sqrt{a_{i+2}}X_{\sigma_{i+1}\sigma_{i+2}} + \cdots + \sqrt{a_n}X_{\sigma_{i+1}\dots\sigma_n}}{\sqrt{a_i + \cdots + a_n}} > y\Big) \tag{2.3}$$

where $X, X_{\sigma_{i+1}}, \dots, X_{\sigma_{i+1}\dots\sigma_n}$ are all i.i.d. standard normal r.v.'s. We will estimate these probabilities by induction downwards from $i = n-1$, using the identity

$$P(i,y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Big(1 - \Big[1 - P\Big(i+1, \frac{\sqrt{a_i + \cdots + a_n}y - \sqrt{a_i}t}{\sqrt{a_{i+1} + \cdots + a_n}}\Big)\Big]^{\alpha_{i+1}^N}\Big) e^{-t^2/2}\, \mathrm{d}t. \tag{2.4}$$

First we prove the assertion of the lemma under assumption that all inequalities in (1.6) are strict. In this case we will prove inductively for $i = n-1, n-2, \dots, 1$ that

$$P(i,y) \sim \frac{(\alpha_{i+1} \cdots \alpha_n)^N e^{-y^2/2}}{\sqrt{2\pi}y}, \tag{2.5}$$

as $N \uparrow \infty$, uniformly for $y \in [\sqrt{2(\ln \bar{\alpha} - \delta_i)N(a_i + \cdots + a_n)}, \infty)$ with some $\delta_i > 0$.

Note that $u_{\ln \bar{\alpha},N}(x) \in [\sqrt{2(\ln \bar{\alpha} - \delta_1)N}, \infty)$ for any $x \in \mathbb{R}$, and Then inserting into (2.5) $i = 1$ and $y = u_{\ln \bar{\alpha},N}(x)$ implies that

$$P(1, u_{\ln \bar{\alpha},N}(x)) \sim (\alpha_2 \cdots \alpha_n)^N e^{-u^2_{\ln \bar{\alpha},N}(x)/2} / \sqrt{2\pi} u_{\ln \bar{\alpha},N}(x) \sim \alpha_1^{-N} e^{-x}.$$

Since the probability in the l.h.s. of (2.1) equals $(1 - P(1, u_{\ln \bar{\alpha},N}(x)))^{\alpha_1^N} \sim e^{\alpha_1^N \ln(1 - \alpha_1^{-N} e^{-x})} \sim e^{-e^{-x}}$, this proves the lemma in this case.

Let us now turn to the proof of (2.5). Note that $P(i,y) \leq (\alpha_{i+1} \dots \alpha_N)^N P(X > y)$, where X is a standard Gaussian random variable, and so the upper bound for (2.5) is immediate by Proposition 3.1.

We will establish the lower bound for (2.5) by induction. Let first $i = n-1$. The identity (2.4) reads

$$P(n-1,y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Big(1 - \Big[1 - \mathbb{P}\Big(X > \frac{\sqrt{a_{n-1} + a_n}y - \sqrt{a_{n-1}}t}{\sqrt{a_n}}\Big)\Big]^{\alpha_n^N}\Big) e^{-t^2/2}\, \mathrm{d}t. \tag{2.6}$$

Let us fix $\delta'_{n-1} > 0$ such that $\ln \alpha_n - (\ln \bar{\alpha} - \delta'_{n-1}) a_n < 0$, which is possible only due to the strict inequality in (1.6) for $i = n$. For all y, t satisfying

$$\frac{\sqrt{a_{n-1} + a_n}y - \sqrt{a_{n-1}}t}{\sqrt{a_n}} \geq \sqrt{2(\ln \bar{\alpha} - \delta'_{n-1})N a_n} \tag{2.7}$$

by Proposition 3.1 we have

$$\alpha_n^N \mathbb{P}\Big(X > \frac{\sqrt{a_{n-1}+a_n}y - \sqrt{a_{n-1}}t}{\sqrt{a_n}}\Big) \leq \alpha_n^N e^{(\ln\bar{\alpha}-\delta'_{n-1})Na_n} \to 0, \quad N\uparrow\infty.$$

Then, using the elementary inequalities

$$-x-x^2 \leq \ln(1-x) \leq -x, \quad 1+x\leq e^x \leq 1+x+x^2 \quad \text{for } |x|<1/2 \tag{2.8}$$

and again Proposition 3.1, it is easy to deduce the asymptotics of the integrand in (2.6):

$$1-\Big[1-\mathbb{P}\Big(X>\frac{\sqrt{a_{n-1}+a_n}y-\sqrt{a_{n-1}}t}{\sqrt{a_n}}\Big)\Big]^{\alpha_n^N}$$

$$\sim \frac{\alpha_n^N\sqrt{a_n}}{\sqrt{2\pi}(\sqrt{a_{n-1}+a_n}y-\sqrt{a_{n-1}}t)}e^{-(\sqrt{a_{n-1}+a_n}y-\sqrt{a_{n-1}}t)^2/(2a_n)}, \quad \text{as } N\uparrow\infty,$$

uniformly for all y,t satisfying (2.7). Observe that condition (2.7) is equivalent to $t\leq t_{cr}(y)\equiv(\sqrt{a_{n-1}+a_n}y-\sqrt{2(\ln\bar{\alpha}-\delta'_{n-1})Na_n})/\sqrt{a_{n-1}}$. Thus in view of (2.6), $P(n-1,y)$ obeys the following asymptotic lower bound

$$P(n-1,y)\geq\frac{\alpha_n^N}{2\pi}\int_{-\infty}^{t_{cr}(y)}\frac{\sqrt{a_n}}{\sqrt{2\pi}(\sqrt{a_{n-1}+a_n}y-\sqrt{a_{n-1}}t)}e^{-(\sqrt{a_{n-1}+a_n}y-\sqrt{a_{n-1}}t)^2/(2a_n)-t^2/2}\,\mathrm{d}t,$$

as $N\uparrow\infty$, uniformly in $y\in\mathbb{R}$. After the change of variables $s=(\sqrt{a_{n-1}+a_n}t-\sqrt{a_{n-1}}y)/\sqrt{a_n}$ the last integral becomes

$$\frac{\alpha_n^N e^{-y^2/2}}{\sqrt{2\pi}y}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{s_{cr}(y)}\frac{e^{-s^2/2}}{1-\frac{\sqrt{a_{n-1}}s}{\sqrt{a_n}y}}\,\mathrm{d}s$$

where $s_{cr}(y)=\big(y-\sqrt{2(\ln\bar{\alpha}-\delta'_{n-1})N(a_{n-1}+a_n)}\big)\sqrt{a_n}/\sqrt{a_{n-1}}$. It yields the necessary asymptotic lower bound for (2.5) as $N\uparrow\infty$ uniform for $y\in[\sqrt{2(\ln\bar{\alpha}-\delta_{n-1})N(a_{n-1}+a_n)},\infty)$ if we set $0<\delta_{n-1}<\delta'_{n-1}$.

Assuming that (2.1) is proved for $i=n-1,n-2,\dots,l+1$, we will prove it for $i=l$. Let us fix $\delta'_l>0$ such that $\delta'_l\leq\delta_{l+1}$ and $\ln(\alpha_{l+1}\cdots\alpha_n)-(\ln\bar{\alpha}-\delta'_l)(a_{l+1}+\cdots+a_n)<0$ due to the strict inequality in (1.6) for $i=l+1$. Then for all y and t verifying the condition

$$\frac{\sqrt{a_l+\cdots+a_n}y-\sqrt{a_l}t}{\sqrt{a_{l+1}+\cdots+a_n}}\geq\sqrt{2(\ln\bar{\alpha}-\delta'_l)N(a_{l+1}+\cdots+a_n)} \tag{2.9}$$

we have by the induction hypothesis (2.1) for $i=l+1$

$$\begin{aligned}\alpha_{l+1}^N P\Big(l+1,\frac{\sqrt{a_l+\cdots+a_n}y-\sqrt{a_l}t}{\sqrt{a_{l+1}+\cdots+a_n}}\Big) &\leq (\alpha_{l+1}\cdots\alpha_n)^N e^{-(\sqrt{a_l+\cdots+a_n}y-\sqrt{a_l}t)^2/2(a_{l+1}+\cdots+a_n)}\\ &\leq(\alpha_{l+1}\cdots\alpha_n)^N e^{-(\ln\bar{\alpha}-\delta'_l)N(a_{l+1}+\cdots+a_n)},\end{aligned}$$

which tends to zero, as $N \uparrow \infty$. Furthermore, again by the induction hypothesis for $i = l+1$ and elementary exponential inequalities (2.8)

$$1 - \left[1 - P\Big(l+1, \frac{\sqrt{a_l + \cdots + a_n}y - \sqrt{a_l}t}{\sqrt{a_{l+1} + \cdots + a_n}}\Big)\right]^{\alpha_{l+1}^N}$$
$$\sim \frac{(\alpha_{l+1}\cdots\alpha_n)^N \sqrt{a_{l+1} + \cdots + a_n}}{\sqrt{2\pi}(\sqrt{a_l + \cdots + a_n}y - \sqrt{a_l}t)} e^{-(\sqrt{a_l+\cdots+a_n}y - \sqrt{a_l}t)^2/2(a_{l+1}+\cdots+a_n)}$$

uniformly for all y, t satisfying (2.9), that is for

$$t \leq t_{cr(y)} \equiv (\sqrt{a_l + \cdots + a_n}y - \sqrt{2(\ln\bar{\alpha} - \delta_l')N}(a_{l+1} + \cdots + a_n))/\sqrt{a_l}.$$

In view of this estimate the identity (2.4) for $i = l$ gives the following asymptotic lower bound for $P(l, y)$ as $N \uparrow \infty$ uniform for $y \in \mathbb{R}$:

$$\begin{aligned} P(l,y) &\geq \frac{(\alpha_{l+1}\cdots\alpha_n)^N\sqrt{a_{l+1}+\cdots+a_n}}{2\pi}\int_{-\infty}^{t_{cr}(y)} \frac{e^{-(\sqrt{a_l+\cdots+a_n}y-\sqrt{a_l}t)^2/2(a_{l+1}+\cdots+a_n)-t^2/2}}{\sqrt{a_l+\cdots+a_n}y-\sqrt{a_l}t}\,\mathrm{d}t \\ &= \frac{(\alpha_{l+1}\cdots\alpha_n)^N e^{-y^2/2}}{\sqrt{2\pi}y}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{s_{cr}(y)} \frac{e^{-s^2/2}}{1-\frac{\sqrt{a_l}s}{\sqrt{a_{l+1}+\cdots+a_n y}}}\,\mathrm{d}s, \end{aligned}$$

where

$$s_{cr}(y) \equiv \big(y - \sqrt{(\ln\bar{\alpha} - \delta_l')N(a_l + \cdots + a_n)}\big)\sqrt{a_{l+1} + \cdots + a_n}/\sqrt{a_l}. \tag{2.10}$$

This bound has the asymptotics stated in (2.5), uniformly in $y \in [\sqrt{(\ln\bar{\alpha} - \delta_l)N(a_l + \cdots + a_n)}, \infty)$ if only $\delta_l > 0$ is chosen to be $\delta_l < \delta_l'$.

The analysis is more delicate whenever some of the inequalities (1.6) degenerate into equalities. Assume that $k = \max\{i > 1, a_i + \cdots + a_n = \ln(\alpha_i\cdots\alpha_n)/\ln\bar{\alpha}\}$. Then for $i = n-1, n-2, \ldots, k$ (2.5) remains valid in the respective domains of y. For $i = k-1$ we will prove the following two assertions ($\mathbf{a}_{k-1}$) and ($\mathbf{b}_{k-1}$).

($\mathbf{a}_{k-1}$) For any $C \in \mathbb{R}$ we have

$$\begin{aligned} P(k-1,y) &\sim \frac{(\alpha_k\cdots\alpha_n)^N e^{-y^2/2}}{\sqrt{2\pi}y} \\ &\times \mathbb{P}\Big(Z_{k-1} > \big[y - \sqrt{2(\ln\bar{\alpha})(a_{k-1}+\cdots+a_n)N}\big]\sqrt{\frac{a_k+\cdots+a_n}{a_{k-1}}}\Big) \quad \text{as} \quad N\uparrow\infty \end{aligned} \tag{2.11}$$

uniformly in $y \in [\sqrt{2\ln\bar{\alpha}(a_{k-1} + \cdots + a_n)N} + C, \infty)$.

($\mathbf{b}_{k-1}$) For any $\epsilon > 0$, there exist a negative constant $\tilde{C}_{k-1}(\epsilon) \in \mathbb{R}$ of large enough absolute value, and $N(\epsilon) \in \mathbb{N}$, such that, for all $N \geq N(\epsilon)$ and all $y \in \left(0\,,\, \sqrt{2\ln\bar{\alpha}(a_{k-1}+\cdots+a_n)N} + \tilde{C}_{k-1}(\epsilon)\right]$

$$P(k-1,y) \leq \frac{\epsilon(\alpha_k\cdots\alpha_n)^N e^{-y^2/2}}{\sqrt{2\pi}y}. \tag{2.12}$$

The proof is based again on (2.4) for $i = k-1$. First of all, let us remark that the asymptotics of $P(k,y)$ is determined by (2.5) uniformly in the interval of $y \in [u_{\ln(\alpha_k\cdots\alpha_n),N}(x), \infty)$ for any $x \in \mathbb{R}$. This is ensured by the equality (1.1) for $i = k$, since consequently

$$u_{\ln(\alpha_k\ldots\alpha_n),N}(x) = \sqrt{2(\ln\bar{\alpha})N(a_k+\cdots+a_n)} + o(1) \quad N\uparrow\infty. \tag{2.13}$$

Therefore for any $x \in \mathbb{R}$

$$\begin{aligned}
&\alpha_k^N P\Big(k, \frac{\sqrt{a_{k-1}+\cdots+a_n}y - \sqrt{a_{k-1}}t}{\sqrt{a_k+\cdots+a_n}}\Big) \\
&\sim \frac{(\alpha_k\cdots\alpha_n)^N\sqrt{a_k+\cdots+a_n}}{\sqrt{2\pi}(\sqrt{a_{k-1}+\cdots+a_n}y - \sqrt{a_{k-1}}t)} 7e^{-(\sqrt{a_{k-1}+\cdots+a_n}y-\sqrt{a_{k-1}}t)^2/2(a_k+\cdots+a_n)} \\
&\leq \frac{(\alpha_k\cdots\alpha_n)^N e^{-u^2_{\ln(\alpha_k\cdots\alpha_n),N}(x)/2}}{\sqrt{2\pi}u_{\ln(\alpha_k\cdots\alpha_n),N}(x)} \sim e^{-x}, \quad N\uparrow\infty
\end{aligned}$$

uniformly for y, t such that

$$\frac{\sqrt{a_{k-1}+\cdots+a_n}y - \sqrt{a_{k-1}}t}{\sqrt{a_k+\cdots+a_n}} \geq u_{\ln(\alpha_k\cdots\alpha_n),N}(x). \tag{2.14}$$

Then, given an arbitrary $\epsilon > 0$, using the inequalities (2.8), one can choose a sufficiently large $x_0 = x_0(\epsilon) \in \mathbb{R}$ and $N(\epsilon, x_0)$ to bound the integrand in (2.4) for $i = k-1$ as follows:

$$\begin{aligned}
&(1-\epsilon)\frac{(\alpha_k\cdots\alpha_n)^N\sqrt{a_k+\cdots+a_n}}{\sqrt{2\pi}(\sqrt{a_{k-1}+\cdots+a_n}y - \sqrt{a_{k-1}}t)} e^{-(\sqrt{a_{k-1}+\cdots+a_n}y-\sqrt{a_{k-1}}t)^2/2(a_k+\cdots+a_n)} \\
&\leq 1 - \Big[1 - P\Big(k, \frac{\sqrt{a_{k-1}+\cdots+a_n}y - \sqrt{a_{k-1}}t}{\sqrt{a_k+\cdots+a_n}}\Big)\Big]^{\alpha_k^N} \\
&\leq (1+\epsilon)\frac{(\alpha_k\cdots\alpha_n)^N\sqrt{a_k+\cdots+a_n}}{\sqrt{2\pi}(\sqrt{a_{k-1}+\cdots+a_n}y - \sqrt{a_{k-1}}t)} e^{-(\sqrt{a_{k-1}+\cdots+a_n}y-\sqrt{a_{k-1}}t)^2/2(a_k+\cdots+a_n)}
\end{aligned} \tag{2.15}$$

for all $N \geq N(\epsilon, x_0)$ and all y, t under condition (2.14), that is for t up to $t_{cr}(y) \equiv (y\sqrt{a_{k-1}+\cdots+a_n} - u_{\ln(\alpha_k\cdots\alpha_n),N}(x_0)\sqrt{a_k+\cdots+a_n})/\sqrt{a_{k-1}}$. After making this choice, let us decompose the integral (2.4) for $i = k-1$ into two terms I_1 and I_2: in the first one the integration is from $-\infty$ to $t_{cr}(y)$ and in the second from $t_{cr}(y)$ to $+\infty$. In view of (2.15), for all $N \geq N(\epsilon, x_0)$, I_1 is bounded from below and above by

$$\frac{1}{2\pi}\int_{-\infty}^{t_{cr}(y)} \frac{(\alpha_k \cdots \alpha_n)^N \sqrt{a_k + \cdots + a_n}}{(\sqrt{a_{k-1} + \cdots + a_n}y - \sqrt{a_{k-1}}t)} e^{-(\sqrt{a_{k-1}+\cdots+a_n}y - \sqrt{a_{k-1}}t)^2/2(a_k+\cdots+a_n) - t^2/2}\,\mathrm{d}t$$
$$= \frac{(\alpha_k \cdots \alpha_n)^N e^{-y^2/2}}{2\pi y}\int_{-\infty}^{\left(y\sqrt{a_k+\cdots+a_n} - u_{\ln(\alpha_k\cdots\alpha_n),N}(x_0)\sqrt{a_{k-1}+\cdots+a_n}\right)/\sqrt{a_{k-1}}} \frac{e^{-s^2/2}}{1 - \frac{\sqrt{a_{k-1}}s}{\sqrt{a_k+\cdots+a_n}y}}\,\mathrm{d}s$$

times the factors $(1-\epsilon)$ and $(1+\epsilon)$ respectively. Combined with the observation (2.13), this integral yields the asymptotics stated in $(\mathbf{a}_{k-1})$. Moreover for any $\epsilon > 0$ one can find $\tilde{C}_{k-1}(\epsilon)$ and $N(\epsilon)$ such that for all $y \in (0, \sqrt{2(\ln\bar{\alpha})N(a_{k-1} + \cdots + a_n)} + \tilde{C}_{k-1}(\epsilon)]$ and $N \geq N(\epsilon)$ I_1 obeys the bound of $(\mathbf{b}_{k-1})$.

The term I_2 brings no significant contribution to the asymptotics of $(\mathbf{a}_{k-1})$ as it verifies the bound (2.12) uniformly on the whole interval $y \in (0,\infty)$. To show this, we simply bound its integrand from above by $e^{-t^2/2}$ and consider separately two intervals of y: $\Delta_1 = \left(0, \sqrt{\frac{2(a_k+\cdots+a_n)\ln(\alpha_k\cdots\alpha_n)+\tilde{\delta}_{k-1}}{a_{k-1}+\cdots+a_n}N}\right]$ and $\Delta_2 = \left(\sqrt{\frac{2(a_k+\cdots+a_n)\ln(\alpha_k\cdots\alpha_n)+\tilde{\delta}_{k-1}}{a_{k-1}+\cdots+a_n}N}, \infty\right)$ where $\tilde{\delta}_{k-1}$ is any fixed number $0 < \tilde{\delta}_{k-1} < a_{k-1}$. The estimate of I_2 in Δ_1 is trivial: for any $\epsilon > 0$ one can find $N(\epsilon)$ such that for all $N \geq N(\epsilon)$ and all $y \in \Delta_1$

$$\begin{aligned} I_2 &\leq \frac{\epsilon \ln(\alpha_k \cdots \alpha_n)^N e^{-N((a_k+\cdots+a_n)\ln(\alpha_k\cdots\alpha_n)+\tilde{\delta}_{k-1})/(a_{k-1}+\cdots+a_n)}}{\sqrt{2\pi\frac{2(a_k+\cdots+a_n)\ln(\alpha_k\cdots\alpha_n)+\tilde{\delta}_{k-1}}{a_{k-1}+\cdots+a_n}N}} \\ &\leq \frac{\epsilon\ln(\alpha_k\cdots\alpha_n)^N e^{-y^2/2}}{\sqrt{2\pi}y}. \end{aligned} \tag{2.16}$$

Since $\tilde{\delta}_{k-1} > 0$, then for all $y \in \Delta_2$ $\tau_{cr}(y) > 0$. Here Proposition 3.1 applies:

$$I_2 \leq \frac{1}{\sqrt{2\pi}t_{cr}(y)} e^{-t_{cr}(y)^2/2} \leq \frac{(\alpha_k\cdots\alpha_n)^N C(\tilde{\delta}_{k-1})}{\sqrt{2\pi}y} e^{-y^2/2} e^{f(y)/2} \tag{2.17}$$

where

$$\begin{aligned} f(y) = \big(&-y^2(a_k + \cdots + a_n) - u^2_{\ln(\alpha_k\cdots\alpha_n),N}(x_0)(a_k + \cdots + a_n) \\ &+ 2u_{\ln(\alpha_k\cdots\alpha_n),N}(x_0)y\sqrt{a_{k-1} + \cdots + a_n}\sqrt{a_k + \cdots + a_n} - 2a_{k-1}N\ln(\alpha_k\cdots\alpha_n)\big)/a_{k-1}. \end{aligned}$$

The function $f(y)$ reaches its maximum at

$$y = u_{\ln(\alpha_k\cdots\alpha_n),N}(x_0)\sqrt{a_{k-1} + \cdots + a_n}/\sqrt{a_k + \cdots + a_n}$$

where it equals

$$u^2_{\ln(\alpha_k\cdots\alpha_n),N}(x_0) - 2N\ln(\alpha_k\cdots\alpha_n) = -\ln N - \ln\ln(\alpha_k\cdots\alpha_n) - \ln 4\pi + 2x_0 + o(1).$$

Hence, $e^{f(y)/2}$ being of the order at most $N^{-1/2}$, I_2 does not contribute to the asymptotics of $(\mathbf{a}_{k-1})$ and satisfies the bound of $(\mathbf{b}_{k-1})$ for all $y \in \Delta_2$ as well when N is bigger some $N(\epsilon)$. This concludes the proof of $(\mathbf{a}_{k-1})$ and $(\mathbf{b}_{k-1})$.

To continue the proof of the lemma, let us assume that the inequalities (1.6) are strict for $i = k-1, k-2, \ldots, l-1$ and $(\mathbf{a}_i^1)$, $(\mathbf{b}_i)$ (see below) hold true for $i = k-2, \ldots, l$ respectively. Under these assumptions we show that

- if the inequality (1.6) is also strict for $i = l$, then $(\mathbf{a}_{l-1}^1)$ and $(\mathbf{b}_{l-1})$ remain true;
- if (1.6) turns into the equality for $i = l$, then $(\mathbf{a}_{l-1}^2)$ and $(\mathbf{b}_{l-1})$ take place.

The assertion $(\mathbf{b}_i)$ is the same as $(\mathbf{b}_{k-1})$ above with $k-1$ replaced by i. The assertions $(\mathbf{a}_i^1)$, $(\mathbf{a}_i^2)$ are the following:

$(\mathbf{a}_i^1)$ For any $C \in \mathbb{R}$ we have

$$P(i, y) \sim \frac{(\alpha_{i+1} \cdots \alpha_n)^N e^{-y^2/2}}{\sqrt{2\pi} y} \mathbb{P}(A_i^k(y))$$

uniformly in the interval of $y \in [\sqrt{2 \ln \bar{\alpha}(a_i + \cdots + a_n) N} + C, \infty)$.

$(\mathbf{a}_i^2)$ For any $C \in \mathbb{R}$ we have

$$\begin{aligned} P(i, y) \sim & \frac{(\alpha_{i+1} \cdots \alpha_n)^N e^{-y^2/2}}{\sqrt{2\pi} y} \\ & \times \mathbb{P}\Big(A_i^k(y) \bigcap \Big[Z_i > \big(y - \sqrt{2 \ln \bar{\alpha}(a_i + \cdots + a_n) N}\big) \sqrt{\frac{a_{i+1} + \cdots + a_n}{a_i}}\Big]\Big), \end{aligned}$$

uniformly in the interval of $y \in [\sqrt{2 \ln \bar{\alpha}(a_i + \cdots + a_n) N} + C, \infty)$. The events $A_i^k(y)$ are defined as:

$$\begin{aligned} A_i^k(y) \equiv \Bigg(Z_{k-1} > \Bigg[& y \sqrt{\frac{a_{k-1} + \cdots + a_n}{a_i + \cdots + a_n}} - Z_i \sqrt{\frac{a_i(a_{k-1} + \cdots + a_n)}{(a_i + \cdots + a_n)(a_{i+1} + \cdots a_n)}} \\ & - Z_{i+1} \sqrt{\frac{a_{i+1}(a_{k-1} + \cdots + a_n)}{(a_{i+1} + \cdots + a_n)(a_{i+2} + \cdots + a_n)}} - \cdots \\ & - Z_{k-3} \sqrt{\frac{a_{k-3}(a_{k-1} + \cdots + a_n)}{(a_{k-3} + \cdots + a_n)(a_{k-2} + \cdots + a_n)}} \\ & - Z_{k-2} \sqrt{\frac{a_{k-2}}{a_{k-2} + \cdots + a_n}} - \sqrt{2 (\ln \bar{\alpha})(a_{k-1} + \cdots + a_n) N} \Bigg] \sqrt{\frac{a_k + \cdots + a_n}{a_{k-1}}} \Bigg). \end{aligned} \tag{2.18}$$

By virtue of $(\mathbf{a}_l^1)$, $(\mathbf{b}_l)$, and the inequalities (2.8), for any given $\epsilon > 0$ there exist a sufficiently large $x_0 = x_0(\epsilon) \in \mathbb{R}$, and $N(x_0, \epsilon)$, such that, for all $N \geq N(x_0, \epsilon)$, all $y \in \mathbb{R}$, and all $t \leq t_{cr}(y) \equiv (y\sqrt{a_{l-1} + \cdots + a_n} - u_{\ln(\alpha_l \cdots \alpha_n),N}(x_0)\sqrt{a_l + \cdots + a_n})/\sqrt{a_{l-1}}$, the integrand in (1.6) for $i = l-1$ can be estimated from above and below by $\alpha_l^N P(l, \cdot)e^{-t^2/2}((1 \pm \epsilon)$ times. Having x_0 fixed, we again split the integral (1.6) into two terms, I_1 and I_2, where the integration is from $-\infty$ to $t_{cr}(y)$ and from $t_{cr}(y)$ to $+\infty$, respectively. By the same arguments as in (2.17), the term I_2 does not contribute to $(\mathbf{a}_{l-1}^1)$ or $(\mathbf{a}_{l-1}^2)$, and satisfies the bound of condition $(\mathbf{b}_{l-1})$.

To treat the term I_1 we distinguish the two cases when the inequality (1.6) is strict for $i = l$, and when equality holds, respectively.

In the second case, the assumed equality entails

$$u_{\ln(\alpha_l \cdots \alpha_n),N}(x_0) = \sqrt{2 \ln \bar{\alpha}(a_l + \cdots + a_n)N} + o(1). \tag{2.19}$$

This means that the probability $P(l, \cdot)$ entering the integrand of I_1 has the asymptotics claimed in $(\mathbf{a}_l^1)$ uniformly in the interval of integration $t \leq t_{cr}(y)$ and all y. Then for any $\epsilon > 0$, one can find $N(x_0, \epsilon)$ such that, for all $N \geq N(x_0, \epsilon)$, the term I_1 is bounded from above and below by

$$\begin{aligned} &\frac{(\alpha_l \cdots \alpha_n)^N \sqrt{a_l + \cdots + a_n}}{2\pi} \int_{-\infty}^{t_{cr}(y)} \frac{e^{-(\sqrt{a_{l-1}+\cdots+a_n}y - \sqrt{a_{l-1}}t)^2/2(a_l+\cdots+a_n) - t^2/2}}{\sqrt{a_{l-1} + \cdots + a_n}y - \sqrt{a_{l-1}}t} \\ &\qquad \times \mathbb{P}\Big(A_l^k\Big(\frac{y\sqrt{a_{l-1} + \cdots + a_n} - t\sqrt{a_{l-1}}}{\sqrt{a_l + \cdots + a_n}}\Big)\Big)\, \mathrm{d}t (1 \pm \epsilon)^2 \end{aligned} \tag{2.20}$$

$$\begin{aligned} =&\frac{(\alpha_l \cdots \alpha_n)^N e^{-y^2/2}}{2\pi y} \int_{-\infty}^{s_{cr}(y)} \frac{e^{-s^2/2}}{1 - \frac{\sqrt{a_{l-1}}s}{\sqrt{a_l+\cdots+a_n}y}} \\ &\times \mathbb{P}\Big(A_l^k\Big(y\sqrt{\frac{a_l + \cdots + a_n}{a_{l-1} + \cdots + a_n}} - s\sqrt{\frac{a_{l-1}}{a_{l-1} + \cdots + a_n}}\Big)\Big)\, \mathrm{d}s (1 \pm \epsilon)^2 \\ =&\frac{(\alpha_l \cdots \alpha_n)^N e^{-y^2/2}}{2\pi y} \int_{-\infty}^{s_{cr}(y)} \frac{e^{-s^2/2}}{1 - \frac{\sqrt{a_{l-1}}s}{\sqrt{a_l+\cdots+a_n}y}} \mathbb{P}(A_{l-1}^k(y, s))\, \mathrm{d}s (1 \pm \epsilon)^2, \end{aligned} \tag{2.21}$$

where $s_{cr}(y) = (y\sqrt{a_l + \cdots + a_n} - u_{\ln(\alpha_l \cdots \alpha_n),N}(x_0)\sqrt{a_{l-1} + \cdots + a_n})/\sqrt{a_{l-1}}$ and $A_{l-1}^k(y, s)$ is the event $A_{l-1}^k(y)$ with Z_{l-1} replaced by the variable of integration s. Finally, we substitute (2.19) into the upper limit of integration $s_{cr(y)}$ and the assertions $(\mathbf{a}_{l-1}^2)$ and $(\mathbf{b}_{l-1})$ follow.

Let us now turn to the case of the inequality (1.6) strict for $i = l$. Given $\epsilon > 0$, one can choose a negative constant $\tilde{C}_l(\epsilon)$ of sufficiently large modulus such that $(\mathbf{a}_l^1)$ is satisfied and that

$$P(Z < \tilde{C}_l(\epsilon)\sqrt{a_{l-1}+\cdots+a_n}/\sqrt{a_{l-1}}) < \epsilon, \quad \text{where } Z \in \mathcal{N}(0,1). \tag{2.22}$$

Next, let us decompose $I_1 \equiv I_1' + I_2''$, s.t. in I_1' the integration is from $-\infty$ to $\tilde{t}_{cr}(y)$ and in I_2 – from $\tilde{t}_{cr}(y)$ to $t_{cr}(y)$, where

$$\tilde{t}_{cr}(y) \equiv \big(y\sqrt{a_{l-1}+\cdots+a_n} - \sqrt{a_l+\cdots+a_n}[\sqrt{2(\ln\bar{\alpha})N(a_l+\cdots+a_n)} + \tilde{C}_l(\epsilon)]\big)/\sqrt{a_{l-1}}.$$

The choice of $\tilde{C}_l(\epsilon)$ enables us to apply to the integrand of I_2'' the estimate $(\mathbf{b}_l)$. It is easy to see, changing variables as in (2.20)), that I_2'' is bounded by $\epsilon(\alpha_l\cdots\alpha_n)^N e^{-y^2/2}/\sqrt{2\pi}y$, for all N large enough, and all y in the intervals of $(\mathbf{a}^1_{l-1})$ and $(\mathbf{b}_{l-1})$.

To consider I_1', we are entitled to estimate its integrand by $(\mathbf{a}^1_l)$ uniformly in the interval of integration. Then for all N large enough this term is bounded from above and below by the integral (2.20) with $t_{cr}(y)$ replaced by $\tilde{t}_{cr}(y)$. Consequently, it equals (2.21) with $s_{cr}(y)$ replaced by

$$\tilde{s}_{cr}(y) = \big(y\sqrt{a_l+\cdots+a_n} - [\sqrt{2(\ln\bar{\alpha})N(a_l+\cdots+a_n)} + \tilde{C}_l(\epsilon)]\sqrt{a_{l-1}+\cdots+a_n}\big)/\sqrt{a_{l-1}}.$$

Combined with (2.22), this integral yields the asymptotics $(\mathbf{a}^1_{l-1})$ and the bound $(\mathbf{b}_{l-1})$.

Following this recurrence procedure up to $i=1$, we obtain for a particular $y = u_{\ln\bar{\alpha},N}(x)$

$$P(1, u_{\ln\bar{\alpha},N}(x)) \sim K(\alpha_2\cdots\alpha_n)^N e^{-u^2_{\ln\bar{\alpha},N}(x)/2}/\sqrt{2\pi}u_{\ln\bar{\alpha},N}(x) \sim \alpha_1^{-N}Ke^{-x},$$

where K is as defined in Theorem 1.1. Thus the probability in the l.h.s. of (2.1) equals

$$(1 - P(1, u_{\ln\bar{\alpha},N}(x)))^{\alpha_1^N} \sim e^{\alpha_1^N \ln(1-K\alpha_1^{-N}e^{-x})} \sim e^{-Ke^{-x}}. \tag{2.23}$$

It remains to establish the assertion (2.2) if for some k, $2 \le k \le n$, the inequality (1.6) is violated. For this purpose we prove by induction for $i = k-1, k-2, \ldots, 1$ that for any $\epsilon > 0$ and $\bar{\delta}_i > \bar{\delta}_{i+1} > 0$ there exists a number $N(\epsilon, \bar{\delta}_i)$ such that

$$\mathbb{P}(i,y) \le \frac{\epsilon\ln(\alpha_k\cdots\alpha_n)^N e^{-y^2/2}}{\sqrt{2\pi}y} \tag{2.24}$$

for all $N \ge N(\epsilon,\bar{\delta}_i)$ and all $y \in \Big(0, \sqrt{\frac{2(a_i+\cdots+a_n)\ln(\alpha_k\cdots\alpha_n)-2\bar{\delta}_i}{a_k+\cdots+a_n}N}\Big]$. The arguments of the proof are very similar to those above. Then starting from $\bar{\delta}_{k-1}$ such that $(\mathbf{a}_k + \cdots a_n)\ln\bar{\alpha} < \ln(\alpha_k\cdots\alpha_n) - \bar{\delta}_{k-1}$ and ending with $\bar{\delta}_1$ sufficiently close to $\bar{\delta}_{k-1}$, we make $y = u_{\ln\bar{\alpha},N}(x)$

belong to the interval of y where (2.24) with $i=1$ holds true. Thus, $P(1,u_{\ln\bar\alpha,N}(x))=o(\alpha_1^{-N}e^{-x})$ and the assertion (2.2) follows as in (2.23). This completes the proof of the lemma. $\diamondsuit$

Proof of Theorem 1.1. The necessity of (1.6) is immediate from the assertion (2.2) of Lemma 2.1. The sufficiency follows from Kallenberg's theorem [Ka] (see also [LLR]) on the weak convergence of a point process $\mathcal{P}_N$ to the Poisson process $\mathcal{P}$. Applying his theorem in our situation, weak convergence holds whenever

(i) For all intervals $(a,b]\in\mathbb{R}$

$$\mathbb{E}\sum_{\sigma}\mathbb{I}_{\{u^{-1}_{\ln\bar\alpha,N}(X_\sigma)\in(a,b]\}}\to Ke^{-a}-Ke^{-b},\quad \text{as}\quad N\uparrow\infty.$$

(ii) For all finite unions of half-open disjoint intervals $\Delta=\Delta_1\cup\dots\cup\Delta_k$, $\Delta_i=(a_i,b_i]$, $a_1<b_1<a_2<b_2<\dots<a_k<b_k$

$$\mathbb{P}\{\forall\sigma: u^{-1}_{\ln\bar\alpha,N}(X_\sigma)\not\in\Delta\}\to\sum_{i=1}^{k}e^{-K(e^{-a_i}-e^{-b_i})}.$$

The first statement is obvious: the l.h.s. of (i) equals

$$\sum_\sigma\mathbb{P}(u_{\ln\bar\alpha,N}(a)<X_\sigma\le u_{\ln\bar\alpha,N}(b))=2^N[\mathbb{P}(X_\sigma>u_{\ln\bar\alpha,N}(a))-\mathbb{P}(X_\sigma\ge u_{\ln\bar\alpha,N}(b))]$$
$$\sim Ke^{-a}-Ke^{-b}.$$

The statement (ii) for $\Delta=(x,\infty)$ has been proven in Lemma 2.1. The proof proceeds along the same lines in the general case of (ii). We just outline formal differences. Instead of $P(i,y)$, we analyse by recurrence $P(i,y_1,z_1,y_2,z_2,\dots y_k,z_k)$ which is the probability that no one of the corresponding random variables belongs to the union $\bigcup_{i=1}^k(y_i,z_i]$. In the case of only strict inequalities in (1.6) the assertion (2.5) turns into

$$P(i,y_1,z_1,\dots,y_k,z_k)\sim\frac{(\alpha_{i+1}\cdots\alpha_n)^N}{\sqrt{2\pi}}\sum_{i=1}^k\frac{e^{-y_i^2}}{y_i}-\frac{e^{-z_i^2}}{z_i}$$

uniformly for all $y_1\in[\sqrt{2(\ln\bar\alpha-\delta_i)N(a_i+\dots+a_n)},\infty)$ and all $z_k>y_k>\dots>z_1>y_1$. If some equalities in (1.6) are present, then we have the assertions like $(\mathbf{a}_i^1)$ and $(\mathbf{a}_i^2)$ with e^{-y^2}/y replaced by $\sum_{i=1}^k e^{-y_i^2}/y_i-e^{-z_i^2}/z_i$, and $P(A_i^k(y))$ by $P(A_i^k(y_1))$. In the assertion $(\mathbf{b}_i)$ we insert $e^{y_1^2/2}/y_1$ instead of $e^{y^2/2}/y$. All of $(\mathbf{a}_i^1)$, $(\mathbf{a}_i^2)$, $(\mathbf{b}_i)$ hold uniformly in the corresponding intervals of parameters where y_1 substitutes y and where $z_k>y_k>\dots>z_1>y_1$.

The proofs of all these assertions are completely analogous to those in Lemma 2.1, note only that $t_{cr}(y)$ and $\tilde{t}_{cr}(y)$ become $t_{cr}(y_1)$ and $\tilde{t}_{cr}(y_1)$ respectively. Finally, considering a particular case $i=1$, $y_l = u_{\ln \bar{\alpha},N}(a_l)$, $z_l = u_{\ln \bar{\alpha},N}(b_l)$ for $l=1,2,\dots,k$ as in (2.23) leads to (ii). $\diamondsuit$

Proof of Theorem 1.2. We proceed by induction. In the case $k=1$, $\mathcal{P}^{(1)} = \mathcal{P}_1$ is the Poisson point process on $\mathbb{R}$ and the statement of the theorem is already contained in the assumptions.

Assume that (1.8) is proved for $k-1$. Taking into account the identical distribution of the r.v. $Y_{\sigma_1\cdots\sigma_i}$, $i=1,2,\dots,k$, and the independence imposed on their vectors we may write:

$$
\begin{aligned}
F^N_{\Delta_1\times\cdots\times\Delta_k}(z) &\equiv \mathbb{E} z^{\sum_{\sigma_1}\mathbb{1}_{\{v_{N,1}(Y_{\sigma_1})\in\Delta_1\}}\cdots\sum_{\sigma_k}\mathbb{1}_{\{v_{N,k}(Y_{\sigma_1\ldots\sigma_k})\in\Delta_k\}}}\\
&= \mathbb{E}\Big(\mathbb{E}\Big(z^{\sum_{\sigma_k}\mathbb{1}_{\{v_{N,k}(Y_{\sigma_1\ldots\sigma_k}\})\in\Delta_n}}\;\Big|\;\sum_{\sigma_1}\mathbb{1}_{\{v_{N,1}(Y_{\sigma_1})\in\Delta_1\}}\cdots\sum_{\sigma_k}\mathbb{1}_{\{v_{N,k-1}(Y_{\sigma_1\ldots\sigma_{k-1}})\in\Delta_{k-1}\}}\Big)\Big)\\
&= \mathbb{E}\Big(\mathbb{E} z^{\sum_{\sigma_k}\mathbb{1}_{\{v_{N,k}(Y_{\sigma_1\ldots\sigma_k})\in\Delta_k\}}}\Big)^{\sum_{\sigma_1}\mathbb{1}_{\{v_{N,1}(Y_{\sigma_1})\in\Delta_1\}}\cdots\sum_{\sigma_{k-1}}\mathbb{1}_{\{v_{N,k-1}(Y_{\sigma_1\ldots\sigma_{k-1}})\in\Delta_{k-1}\}}}.
\end{aligned}
$$

Clearly, $\mathbb{E} z^{\sum_{\sigma_k}\mathbb{1}_{\{v_{N,k}(Y_{\sigma_1\ldots\sigma_k})\in\Delta_k\}}} \to f_k(z)$ as $N\uparrow\infty$. It will be convenient to introduce

$$
\zeta_N(z) = \mathbb{E} z^{\sum_{\sigma_k}\mathbb{1}_{\{v_{N,k}(Y_{\sigma_1\ldots\sigma_k})\in\Delta_k\}}}/f_k(z) - 1
$$

(which tends to zero as $N\uparrow\infty$) and

$$
\eta_{N,k-1} \equiv \sum_{\sigma_1}\mathbb{1}_{\{v_{N,1}(Y_{\sigma_1})\in\Delta_1\}}\cdots\sum_{\sigma_{k-1}}\mathbb{1}_{\{v_{N,k-1}(Y_{\sigma_1\ldots\sigma_{k-1}})\in\Delta_{k-1}\}}.
$$

With this notations we have

$$
F^N_{\Delta_1\times\cdots\times\Delta_k}(z) = \mathbb{E}\big(f_k(z)\big)^{\eta_{N,k-1}} + \mathbb{E}\Big[f_k(z)^{\eta_{N,k-1}}\Big(\big(1+\zeta_N(z)\big)^{\eta_{N,k-1}} - 1\Big)\Big] \equiv I_1^N + I_2^N. \quad (2.25)
$$

The term I_2^N in this representation converges to zero as $N\uparrow\infty$. In fact, due to the convergence (1.7) and $a_1 > -\infty$, for any $\epsilon>0$ one can find a sufficiently large $C(\epsilon)>0$ and $N(\epsilon)$ such that

$$
P(\eta_{N,k-1}\ge C(\epsilon)) \le P\Big(\sum_{\sigma_1}\mathbb{1}_{v_{N,1}(Y_{\sigma_1})\in\Delta_1} \ge C(\epsilon)\Big) < \epsilon
$$

for all $N\ge N(\epsilon)$. Then

$$
|I_2^N| \le 2\epsilon + \mathbb{E}\Big(|f_k(z)|^{\eta_{N,k-1}}\Big|\big(1+\zeta_N(z)\big)^{\eta_{N,k-1}} - 1\Big|\mathbb{1}_{\eta_{N,k-1}<C(\epsilon)}\Big)
$$

for all $N \geq N(\epsilon)$. The second term in (2.25) tends to zero, as $N \uparrow \infty$, since $\zeta_N(Z) \to 0$. It remains to analyse I_1^N. But this is the generating function of the model with $k-1$ levels at the point $f_k(z)$ and $|f_k(z)| < 1$. Thus by the induction hypothesis

$$\lim_{N\uparrow\infty} F^N_{\Delta_1\times\cdots\times\Delta_k}(z) = \lim_{N\uparrow\infty} I_1^N = f_1(f_2(\cdots(f_k(z))\cdots))$$

and (1.8) is proved.

Let us establish the independence property (i). We take $r=2$ for shortness, the case of $r>2$ is analogous. Again in view of the identical distribution of the r.v. $Y_{\sigma_1\cdots\sigma_i}$, and of the independence of their vectors, we get:

$$\begin{aligned}
&F_{\Delta_1^1\times\cdots\times\Delta_k^1\cup\Delta_1^2\times\cdots\times\Delta_k^2}(z) = \lim_{N\uparrow\infty} F^N_{\Delta_1^1\times\cdots\times\Delta_k^1\cup\Delta_1^2\times\cdots\times\Delta_k^2}(z)\\
&= \lim_{N\uparrow\infty} \mathbb{E}\Big[(\mathbb{E}z^{\sum_{\sigma_2}\mathbb{1}_{\{v_{N,2}(Y_{\sigma_2})\in\Delta_2^1\}}\cdots\sum_{\sigma_n}\mathbb{1}_{\{v_{N,k}(Y_{\sigma_1\ldots\sigma_k})\in\Delta_k^1\}}})^{\sum_{\sigma_1}\mathbb{1}_{\{v_{N,1}(Y_{\sigma_1})\in\Delta_1^1\}}}\\
&\qquad\times(\mathbb{E}z^{\sum_{\sigma_2}\mathbb{1}_{\{v_{N,2}(Y_{\sigma_2})\in\Delta_2^2\}}\cdots\sum_{\sigma_k}\mathbb{1}_{\{v_{N,k}(Y_{\sigma_1\ldots\sigma_k})\in\Delta_k^2\}}})^{\sum_{\sigma_1}\mathbb{1}_{\{v_{N,1}(Y_{\sigma_1})\in\Delta_1^2\}}}\Big] \qquad (2.26)\\
&= \lim_{N\uparrow\infty} \mathbb{E}\Big[(f_{2,\Delta_2^1}(f_{3,\Delta_3^1}(\cdots(f_{k,\Delta_k^1}(z))\cdots)))^{\sum_{\sigma_1}\mathbb{1}_{\{v_{N,1}(Y_{\sigma_1})\in\Delta_1^1\}}}\\
&\qquad\times(f_{2,\Delta_2^2}(f_{3,\Delta_3^2}(\cdots(f_{k,\Delta_k^2}(z))\cdots)))^{\sum_{\sigma_1}\mathbb{1}_{\{v_{N,1}(Y_{\sigma_1})\in\Delta_1^2\}}}\Big]
\end{aligned}$$

where in the last line we applied (1.8) with $n-1$ levels together with the arguments similar to (2.25). Recall that the counting variables $\sum_{x_1}\mathbb{1}_{x_1\in\Delta_1}$ of the Poisson point process $\mathcal{P}_1$ corresponding to non-intersecting segments Δ_1^1 and Δ_1^2 are independent. Therefore (2.26) equals the product of the generating functions $f_{1,\Delta_1^1}(\cdots(f_{n,\Delta_n^1}(z))\cdots)$ and $f_{1,\Delta_1^2}(\cdots(f_{k,\Delta_k^2}(z))\cdots)$ finishing the proof of the property (i). The proof of (ii) is analogous. $\diamondsuit$

Proof of Proposition 1.4 Suppose that for some l and some k, $J_{l-1}+2 \leq k \leq J_l$, the inequality (1.10) is not true. Then $A_{J_{l-1}+1,k-1} > A_{k,J_l}$. In view of the definition of J_l either for some $j_0 > J_l$ $A_{J_{l-1}+1,k-1} \leq A_{k,j_0}$ or there exists p, $k \leq p < J_l$, such that $A_{J_{l-1}+1,k-1} \leq A_{k,p}$ and for all $j > p$ $A_{J_{l-1}+1,k-1} > A_{k,j}$. Under the former condition $A_{J_{l-1}+1,k-1} < A_{J_l+1,j_0}$ and $A_{k,J_l} < A_{J_l+1,j_0}$, whence $A_{J_{l-1}+1,J_l} < A_{J_l+1,j_0}$ contradicting the definition of J_l. Under the latter condition $A_{k,p} > A_{p+1,j}$, $A_{J_{l-1}+1,k-1} > A_{p+1,j}$, and thus $A_{J_{l-1}+1,p} > A_{p+1,j}$ for all $j > p$. This implies $J_l \leq p$ contradicting $p < J_l$. $\diamondsuit$

Proof of Theorem 1.5. To prove Theorem 1.5 we need two auxiliary results:

Proposition 2.2: *Assume that the numbers $\gamma_1,\ldots,\gamma_m$ are such that $\gamma_1 > \gamma_2 > \cdots >$*

$\gamma_m > 1$. *Then the limit*

$$\lim_{x\to-\infty}\int_{\substack{(x_1,\dots,x_m)\in\mathbb{R}^m,\\ \exists i,1\le i\le m:\gamma_1x_1+\cdots+\gamma_ix_i>(\gamma_1+\cdots+\gamma_i)x}} e^{\gamma_1x_1+\cdots+\beta\gamma_mx_m}\mathcal{P}^{(m)}(dx_1\dots dx_m) \tag{2.27}$$

is finite a.s.

Proposition 2.3: *Assume that the numbers* $\gamma_1,\dots,\gamma_m$ *are such that* $\gamma_1>\gamma_2>\cdots>\gamma_m>1$. *Then for any* $a\in R$

$$\mathbb{P}(\exists\quad \textit{infinitely many}\quad (x_1,\dots,x_m)\in\mathcal{P}^{(m)}:\gamma_1x_1+\cdots x_m\gamma_m>a)=0 \tag{2.28}$$

In fact we will prove statements (i) and (ii) of Theorem 1.5 together with (2.27) and (2.28) by induction.

For $m=1$ (i) and (ii) are immediate from Theorem 1.3. To see (2.27) and (2.28) note that $\mathcal{P}^{(1)}$ has a finite number of points in $[x,\infty)$ a.s. for any $x\in\mathbb{R}$. Then the integral $\int_x^\infty e^{c_1x_1}\mathcal{P}^{(1)}(dx)$ is finite a.s. while $\mathbb{E}\int_{-\infty}^x e^{c_1x_1}\mathcal{P}^{(1)}(dx)=e^{(c_1-1)x}/(c_1-1)\to 0$ as $x\to-\infty$.

Assume that (i), (ii) (2.27) and (2.28) are true up to $m-1$. First of all let us show the existence of $\mathcal{E}$. We may consider $\mathcal{E}$ as the cluster process (see [DV] Chapter 8) with the process N_c of the centers of clusters $\gamma_1x_1+\cdots+\gamma_{m-1}x_{m-1}$ at the projection points of $\mathcal{P}^{(m)}$ on the plane of the first $m-1$ coordinates and the independent Poisson point processes of the intensity $e^{-x}dx$ corresponding to each cluster. Then by Proposition 8.2.II in [DV] it suffices to show that for any bounded $A\in\mathbb{R}$ $\int_R\mathbb{P}(A\mid\gamma_1x_1+\cdots+\gamma_{m-1}x_{m-1})N_c(dx)$ is finite a.s. where $\mathbb{P}(A\mid\gamma_1x_1+\dots\gamma_{m-1}x_{m-1})$ is the probability that A contains no point of the cluster with the center $\gamma_1x_1+\ldots+\gamma_{m-1}x_{m-1}$. It suffices to prove this for $A=[a,\infty)$. A simple calculation gives that

$$\begin{aligned}\mathbb{P}(A\mid\gamma_1x_1+\ldots+\gamma_{m-1}x_{m-1}) &= 1-\exp\{-e^{-(a-\gamma_1x_1-\cdots-\gamma_{m-1}x_{m-1})/\gamma_m}\}\\ &\le e^{(-a+\gamma_1x_1+\cdots+\gamma_{m-1}x_{m-1})/\gamma_m}\end{aligned} \tag{2.29}$$

and

$$\int_R e^{(\gamma_1x_1+\cdots+\gamma_{m-1}x_{m-1})/\gamma_m}N_c(dx)=\int_{R^{m-1}}e^{(\gamma_1x_1+\cdots+\gamma_{m-1}x_{m-1})/\gamma_m}\mathcal{P}^{(m-1)}(dx_1,\dots dx_{m-1})<\infty\quad\text{a.s.} \tag{2.30}$$

by (2.27) for $m-1$, since $\gamma_1/\gamma_m > \cdots > \gamma_{m-1}/\gamma_m > 1$. Then $\mathcal{E}$ exists and has an a.s. finite number of points in any bounded subset of $\mathcal{R}$. Furthermore, let $y = a/(\gamma_1 + \cdots + \gamma_m)$. Then

$$P_j \equiv \mathbb{P}\Big(\exists (x_1, \ldots, x_m) \in \mathcal{P}^{(m)} : \forall i = 1, 2, \ldots, j-1 \ \gamma_1 x_1 + \cdots + \gamma_i x_i \le (\gamma_1 + \cdots + \gamma_i) y,$$
$$\gamma_1 x_1 + \cdots + \gamma_j x_j > (\gamma_1 + \cdots + \gamma_j) y\Big)$$
$$\le \int_{\substack{(x_1, \ldots, x_k) \in \mathbb{R}^k : \\ \forall i=1,2,\ldots,j-1 \ \gamma_1 x_1 + \cdots + \gamma_i x_i \le (\gamma_1 + \cdots + \gamma_i) y, \\ \gamma_1 x_1 + \cdots + \gamma_j x_j > (\gamma_1 + \cdots + \gamma_j) y}} e^{-x_1 - \cdots - x_j} \, dx_1 \cdots dx_j = e^{-jy} \tag{2.31}$$

leading to the exponential bound $\mathbb{P}\Big(\exists x \in \mathcal{E} : x > a\Big) \le \sum_{j=1}^m P_j \le \sum_{j=1}^m e^{-ja/(\gamma_1 + \cdots + \gamma_j)}$. It follows by Borel-Cantelli lemma that $\mathcal{E}$ has an a.s. finite number of points in any interval $[a, \infty)$ and (ii) is proven.

Now (2.28) is almost immediate. For any $\epsilon > 0$ one can find a large $A(\epsilon, a)$ such that the probability (2.28) is not bigger than

$$\begin{aligned} &\mathbb{P}(\exists \quad \text{inf. many} \quad (x_1, \ldots, x_m) \in \mathcal{P}^{(m)} \gamma_1 x_1 + \cdots x_m \gamma_m > a \\ &\forall i = 1, \ldots, m-1, |\gamma_1 x_1 + \cdots + \gamma_i x_i| < (\gamma_1 + \cdots + \gamma_i) A(\epsilon, a)) + \epsilon. \end{aligned} \tag{2.32}$$

This follows from the exponential bound analogous to (2.31) and the property (ii) applied to the processes $\mathcal{E}$ constructed from $\mathcal{P}^{(1)}, \ldots, \mathcal{P}^{(m-1)}$. The probability (2.32) does not exceed

$$\sum_{j=1}^{m-1} \mathbb{P}(\exists \quad \text{inf. many} \quad (x_{m-j+1}, \ldots, x_m) \in \mathcal{P}^{(j)} :$$
$$\gamma_{m-j+1} x_{m-j+1} + \cdots x_m \gamma_m > a - (\gamma_1 + \cdots + \gamma_j) A(a, \epsilon)) + \epsilon$$

where the terms in the sum equal zero by induction assumption. Since $\epsilon > 0$ is arbitrary, (2.28) follows.

To show (2.27) we consider

$$\begin{aligned} &\int_{\substack{(x_1, \ldots, x_m) \in \mathbb{R}^m, \\ \exists i, 1 \le i \le m : c_1 x_1 + \cdots + c_i x_i > (c_1 + \cdots + c_i) y}} e^{c_1 x_1 + \cdots + c_m x_m} \mathcal{P}^{(m)}(dx_1 \ldots dx_m) \\ &= \sum_{k=1}^m \int_{\substack{(x_1, \ldots, x_m) \in \mathbb{R}^m, \\ \forall i=1,\ldots,k-1 : c_1 x_1 + \cdots + c_i x_i \le (c_1 + \cdots + c_i) y \\ c_1 x_1 + \cdots + c_k x_k > (c_1 + \cdots + c_k) y}} e^{c_1 x_1 + \cdots + c_m x_m} \mathcal{P}^{(m)}(dx_1 \ldots dx_m) \end{aligned} \tag{2.33}$$

where each term is determined by induction. In fact, for the kth term, given any realisation of $\mathcal{P}^{(m)}$ in $\mathbb{R}^m$, take its projection on the plane of the first k coordinates. By (2.28) for $k < m$ it contains a finite number of points $(x_1, \ldots, x_k)$ in the domain of $\mathbb{R}^k$ where $c_1 x_1 + \cdots c_k x_k >$

$(c_1+\cdots+c_k)y$. Whenever the first k coordinates of a point of $\mathcal{P}^{(m)}$ in $\mathbb{R}^m$ are fixed, the remaining $m-k$ coordinates are distributed as $\mathcal{P}^{(m-k)}$ in $\mathbb{R}^{m-k}$. Then the integral of the function $e^{c_{k+1}x_{k+1}+\cdots+c_m x_m}$ over these coordinates is defined by induction and is finite a.s. Thus the kth term of (2.33) is the sum of an a.s. finite number of terms and each of them is a.s. finite. Finally we put

$$\begin{aligned}&\int_{\mathbb{R}^m} e^{c_1x_1+\cdots+c_mx_m}\mathcal{P}^{(m)}(dx_1\dots dx_m)\\ &\equiv \lim_{x\to-\infty}\int_{\substack{(x_1,\dots,x_m)\in\mathbb{R}^m,\\ \exists i,1\le i\le m: c_1x_1+\cdots+c_ix_i>(c_1+\cdots+c_i)y}} e^{c_1x_1+\cdots+c_mc_m}\mathcal{P}^{(m)}(dx_1\dots dx_m).\end{aligned}$$

This limit is finite a.s. as it verifies Cauchy's criterion. Denote by

$$I(y)=\int_{\substack{(x_1,\dots,x_m)\in\mathbb{R}^m,\\ \exists i,1\le i\le m: c_1x_1+\cdots+c_ix_i>(c_1+\cdots+c_i)y}} e^{c_1x_1+\cdots+c_mx_m}\mathcal{P}^{(m)}(dx_1\dots dx_m).$$

Then for any $\epsilon>0$ and any $y,y'\in\mathbb{R}$, $y'<y$ under assumptions $c_1>\cdots>c_m>1$

$$\begin{aligned}&P(|I(y)-I(y')|>\epsilon)\le\epsilon^{-2}\mathbb{E}(I(y')-I(y))\\ &\le\epsilon^{-2}\int_{\substack{(x_1,\dots,x_m)\in\mathbb{R}^m:\\ \forall i=1,2,\dots,m\;\; c_1x_1+\cdots+c_ix_i\le(c_1+\cdots+c_i)x}} e^{(c_1-1)x_1+\cdots+(c_m-1)x_m}dx_1\cdots dx_m\\ &=\epsilon^{-2}e^{(c_1-1)y+\cdots+(c_m-1)y}\to 0,\qquad y\to-\infty\end{aligned}$$

and Borel-Cantelli lemma yields the result.

To start with (i), let us remark that

$$\begin{aligned}&U^{-1}_{J,N}(X_\sigma)\\ &=\gamma_1u^{-1}_{\ln\bar\alpha_1,N}(\bar X_{\sigma_1,\dots\sigma_{J_1}})+\gamma_2u^{-1}_{\ln\bar\alpha_2,N}(\bar X^{\sigma_1,\dots\sigma_{J_1}}_{\sigma_{J_1+1}+\cdots+\sigma_{J_2}})+\cdots+\gamma_1u^{-1}_{\ln\bar\alpha_1,N}(\bar X^{\sigma_1,\dots\sigma_{J_{m-1}}}_{\sigma_{J_{m-1}+1}+\cdots+\sigma_{J_m}}).\end{aligned}$$

The convergence (i) is equivalent to the convergence of the following probability generating functionals (p.g.fl.s): for any continuous function $h(x)$, $0<h(x)\le 1$, with $1-h$ vanishing outside a bounded set $\mathbb{E}\prod_{x\in\mathbb{E}_N}h(x)\to\mathbb{E}\prod_{x\in E}h(x)$, see Proposition 9.1.VII in [DV]. Let us write

$$\begin{aligned}&\mathbb{E}\prod_{x\in\mathbb{E}_N}h(x)-\mathbb{E}\prod_{x\in E}h(x)\\ &=\mathbb{E}\prod_{\substack{x_1^N,\dots,x_{m-1}^N\\ \exists x_m^N:\gamma_1x_1^N+\cdots\gamma_mx_m^N\in A}}\mathbb{E}\Big[\Big(\prod_{x_m^N}h(\gamma_1x_1^N+\cdots+\gamma_{m-1}x_{m-1}^N+\gamma_mx_m^N)\\ &\qquad-\prod_{x_m}h(\gamma_1x_1^N+\cdots+\gamma_{m-1}x_{m-1}^N+\gamma_mx_m)\Big)\;\Big|\;(x_1^N,\dots,x_{m-1}^N)\Big] \qquad (2.34)\\ &+\mathbb{E}\prod_{x_1^N,\dots,x_{m-1}^N}\prod_{x_m}h(\gamma_1x_1^N+\cdots+\gamma_{m-1}x_{m-1}^N+\gamma_mx_m)\\ &\qquad-\mathbb{E}\prod_{x_1,\dots,x_{m-1}}\prod_{x_m}h(\gamma_1x_1+\cdots+\gamma_{m-1}x_{m-1}+\gamma_mx_m)\end{aligned}$$

where we denote by $(x_1^N,\dots,x_m^N)$, $(x_1,\dots,x_m)$ the points of the processes $\mathcal{P}_N^{(m)}$ and $\mathcal{P}^{(m)}$ in $\mathbb{R}^m$ respectively and by A the support of $1-h$. For any fixed $x_1^N,\dots,x_{m-1}^N$ the term in square brackets converges to zero by Theorem 1.3 and p.g.fl.s. convergence criterion applied to $\mathcal{P}_N^{(1)}$ and $\mathcal{P}^{(1)}$ in $\mathbb{R}$. The statement (i) for $m-1$ and the fact that $\mathcal{E}$ has an a.s. finite number of points in any bounded set imply that for any $\epsilon>0$ one can find $K(\epsilon)$ such that the number of terms $\mathbb{E}[\cdot]$ in the product over $\{x_1^N,\dots,x_{m-1}^N:\ \exists x_m^N, \gamma_1 x_1^N+\cdots \gamma_m x_m^N\in A\}$ is smaller than $K(\epsilon)$ with probability at least $1-\epsilon$ for all N large enough. Taking into account also that $0<h(x)<1$, we derive that the first term in (2.34) converges to zero. The second term converges to zero by the p.g.fl.s. convergence criterion applied to $\mathcal{E}_N$ and $\mathcal{E}$ constructed from the first $m-1$ coordinates on the function

$$\tilde{h}(x)=\prod_{y\in \operatorname{supp}\mathcal{P}^{(1)}} h(x+\gamma_m y).$$

Finally we turn to (iii) of Theorem 1.5. First of all, consider the case of Theorem 1.1, i.e. $m=1$. By Lemma 2.1 we have $\mathbb{P}(\max(X_\sigma/\sqrt{N})<u_{\ln\bar{\alpha},N}(x)/\sqrt{N})\to e^{-e^{-x}}$ as $N\to\infty$ where $e^{-e^{-x}}\to 1$ if $x\to+\infty$ and $e^{-e^{-x}}\to 0$ if $x\to-\infty$. It follows that $\max(X_\sigma/\sqrt{N})\to\sqrt{2\ln\bar{\alpha}}$ in probability. Let $m>1$. Take *any* partition J satisfying (1.10) and $\bar{a}_1,\dots,\bar{a}_m,\bar{\alpha}_1,\dots,\bar{\alpha}_m$ subjected to it. Then from the tree's structure of X_σ in the GREM (1.4) and the previous result for $m=1$ applied to each cell of the partition J $\liminf_{N\to\infty}\max(X_\sigma/\sqrt{N})\geq\sqrt{2\bar{a}_1\ln\bar{\alpha}_1}+\cdots+\sqrt{2\bar{a}_m\ln\bar{\alpha}_m}$ in probability. The maximal value $\max_J\sqrt{2\bar{a}_1\ln\bar{\alpha}_1}+\cdots+\sqrt{2\bar{a}_m\ln\bar{\alpha}_m}$ is attained on the partition (1.13). And it gives the true value of the maximum of X_σ: by (i) and (ii) of Theorem 1.5 for some $h>0$ $\mathbb{P}(\max(X_\sigma/\sqrt{N})>(\sqrt{\bar{a}_1}u_{\ln\bar{\alpha}_1,N}(x)+\cdots+\sqrt{\bar{a}_m}u_{\ln\bar{\alpha}_m,N}(x))/\sqrt{N})\to\mathbb{P}(\exists y\in\mathcal{E}: y>x)\leq e^{-hx}$. Thus $\max(X_\sigma/\sqrt{N})\to\sqrt{2\bar{a}_1\ln\bar{\alpha}_1}+\cdots+\sqrt{2\bar{a}_m\ln\bar{\alpha}_m}$ in probability. To deduce from this the convergence of the expectation, note that the function $\max(X_\sigma)$ is the Lipschitz function of $\alpha_1^N+\cdots+(\alpha_1\cdots\alpha_n)^N$ independent Gaussian r.v. $X_{\sigma_1},\dots,X_{\sigma_1\dots\sigma_n}$ with the Lipschitz constant equal to 1 and apply the concentration of measure inequality (3.1) to it with $\epsilon=\delta\sqrt{N}$. Once the convergence of the expectation proven, the inequality (3.1) applied for the second time shows that $\max(X_\sigma/\sqrt{N})$ converges in probability exponentially fast. Then Borel-Cantelli lemma proves the a.s. convergence. $\diamondsuit$

Appendix

Proposition 3.1: *For any $x > 0$*

$$\left(\frac{1}{x} - \frac{1}{x^3}\right) e^{-x^2/2} \leq \int_x^\infty e^{-t^2/2} dt \leq \frac{1}{x} e^{-x^2/2}.$$

(see [Fe], Lemma 2, page 175).

Proposition 3.2: Let $f(x_1, \dots x_M)$ be a function of M variables with Lipschitz constant L. and $X_1, \dots, X_M$ independent standard Gaussian r.v.

$$\mathbb{P}(|f(X_1, \dots, X_M) - \mathbb{E}f(X_1, \dots, X_M)| > \epsilon) \leq \exp(-\epsilon^2/(2L)). \tag{3.1}$$

(see e.g. [LT], page 21).

References

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

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

[BK3] A. Bovier and I. Kurkova, Derrida's Generalized Random Energy models 3. Models with continuous hierarchies. in preparation (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** (2002).

[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).

[D1] B. Derrida, Random energy model: limit of a family of disordered models, Phys. Rev. Letts. **45**, 79-82 (1980).

[D2] B. Derrida, Random energy model: An exactly solvable model of disordered systems, Phys. Rev. B **24**, 2613-2626 (1981)

[D3] 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)

[DG3] E. Gardner and B. Derrida, The probability distribution of the partition function of the random energy model, J. Phys. A**22**, 1975-1981 (1989)

[DV] D.J Daley and D. Vere-Jones, An introduction to the theory of point processes, Speinger, New York, 1988.

[DW] T.C. Dorlas and J.R. Wedagedera, Large deviations and the random energy model. Int. J. Mod. Phys. **B 15**, 1–15 (2001).

[Ei] TH. Eisele, On a third order phase transition, Commun. Math. Phys. **90**, 125-159 (1983).

[Fe] W. Feller, An introduction to probability theory and its applications. Vol. I, John Wiley & Sons, New York-London-Sydney, 1966.

[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).

[Ka] O. Kallenberg, Random measures, 4th edition, Akademie Verlag, Berlin, 1986.

[KP] M.F. Kratz and P. Picco, On a representation of Gibbs measure for R.E.M., preprint Samos 151 (2002:

[LLR] M. R. Leadbetter, G. Lindgren, H. Rootzén, Extremes and Related Properties of Random Sequences and Processes, Springer, Berlin-Heidelberg-New York, 1983.

[LT] M. Ledoux and L. Talagrand, Probability in Banach space. Spinger, Berlin-Heidelberg-New York, 1991.

[OP] E. Olivieri and P. Picco, On the existence of thermodynamics for the random energy model, Commun. Math. Phys. **96**, 125-144 (1991).

[Re] S.I. Resnick, Extreme values, regular variation , and point processes, Springer, New York, 1987.

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

[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, Concentration of measure and isoperimetric inequalities in product space, Publ. Math. I.H.E.S., **81**, 73-205 (1995).