HAL open science

Poisson–Dirichlet statistics for the extremes of a log-correlated Gaussian field

Louis-Pierre Arguin, Olivier Zindy

▶ **To cite this version:**

Louis-Pierre Arguin, Olivier Zindy. Poisson–Dirichlet statistics for the extremes of a log-correlated Gaussian field. 2012. hal-00680873v1

HAL Id: hal-00680873

https://hal.science/hal-00680873v1

Preprint submitted on 20 Mar 2012 (v1), last revised 20 Jul 2017 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

POISSON-DIRICHLET STATISTICS FOR THE EXTREMES OF A LOG-CORRELATED GAUSSIAN FIELD

LOUIS-PIERRE ARGUIN AND OLIVIER ZINDY

Abstract. We study the statistics of the extremes of a discrete Gaussian field with logarithmic correlations at the level of the Gibbs measure. The model is defined on the periodic interval $[0, 1]$. It is based on a model introduced by Bacry and Muzy [3], and is similar to the logarithmic Random Energy Model studied by Carpentier and Le Doussal [14] and more recently by Fyodorov and Bouchaud [23]. At low temperature, it is shown that the normalized covariance of two points sampled from the Gibbs measure is either 0 or 1. This is used to prove that the joint distribution of the Gibbs weights converges in a suitable sense to that of a Poisson-Dirichlet variable. In particular, this proves a conjecture of Carpentier and Le Doussal that the statistics of the extremes of the log-correlated field behave as those of i.i.d. Gaussian variables and of branching Brownian motion at the level of the Gibbs measure. The proof is based on the computation of the free energy of a perturbation of the model, where a scale-dependent variance is introduced, and on general tools of spin glass theory.

1. Introduction

This paper studies the statistics of the extremes of a Gaussian field whose correlations decays logarithmically with the distance. The model is related to the process introduced by Bacry and Muzy [3], and similar to the logarithmic Random Energy Model or *log-REM* studied by Carpentier and Le Doussal [14], and Fyodorov and Bouchaud [23]. Another important log-correlated model is the two-dimensional discrete Gaussian free field. The model studied here has the advantages of having a graphical representation of the correlations and a continuous scale parameter, cf. Section 1.1, which might make the ideas of the proof more transparent. The method developed here is expected to hold for the two-dimensional discrete Gaussian free field.

The statistics of the extremes of log-correlated Gaussian fields are expected to resemble those of i.i.d. Gaussian variables or *Random Energy Model* (REM) and to a finer level, those of branching Brownian motion. In fact, log-correlated fields are conjectured to be the critical case where correlations start to affect the statistics of the extremes. The reader is referred to the works of Carpentier and Le Doussal [14]; Fyodorov and Bouchaud [23]; and Fyodorov, Le Doussal and Rosso [24] for physical motivations of this fact. The analysis for general log-correlated Gaussian field is complicated by the fact that, unlike branching Brownian motion, the correlations do not necessarily exhibit a tree structure.

The approach of this paper is in the spirit of the seminal work of Derrida and Spohn [18] who studied the extremes of branching Brownian motion using the Gibbs

2000 *Mathematics Subject Classification.* primary 60G15, 60F05; secondary 82B44, 60G70, 82B26.
Key words and phrases. Log-correlated Gaussian fields, Gibbs measure, Poisson-Dirichlet variable, tree approximation, spin glasses.

measure. Even though correlations are not tree-like for general log-correlated models, such fields can often be decomposed as a sum of independent fields acting on different scales. The main results of the paper are Theorem 1.4 on the correlations of the extremes and Theorem 1.5 on the statistics of the Gibbs weights. The results show that, in effect, the statistics of the extremes of the log-correlated field are the same as those of branching Brownian motion at the level of the Gibbs measure, as conjectured by Carpentier and Le Doussal [14]. The proof of the first theorem is based on an adaptation of a technique of Bovier and Kurkova [10, 11] originally developed for hierarchical Gaussian fields such as branching Brownian motion. For this purpose, we need to introduce a family of log-correlated Gaussian models where the variance of the fields in the scale-decomposition depends on the scale. The free energy of the perturbed models is computed using ideas of Daviaud [16]. The second theorem on the Poisson-Dirichlet statistics of the Gibbs weights is proved using the first theorem on correlations and general spin glass theory results. The approach is robust, cf. Theorem 2.5, and could be of independent interest to prove Poisson-Dirichlet statistics for the extremes of other Gaussian fields.

1.1. **A log-correlated Gaussian field.** Following [3], we consider the half-infinite cylinder

$$\mathcal{C}^+ := \{(x,y)\,;\, x \in [0,1]_\sim\,,\, y \in \mathbb{R}_+^*\},$$

where $[0,1]_\sim$ stands for the unit interval where the two endpoints are identified. We write $\|x-x'\| := \min\{|x-x'|, 1-|x-x'|\}$ for the distance on $[0,1]_\sim$.

The following measure is put on $\mathcal{C}^+$:

$$\theta(\mathrm{d}x,\mathrm{d}y) := y^{-2}\,\mathrm{d}x\,\mathrm{d}y.$$

Note that θ is invariant under homogeneous scaling $(x,y) \mapsto (\lambda x, \lambda y)$. For $\sigma > 0$, *the variance parameter*, there exists a random measure μ on $\mathcal{C}^+$ that satisfies:

i) for any measurable set A in $\mathcal{B}(\mathcal{C}^+)$, the random variable $\mu(A)$ is a centered Gaussian with variance $\sigma^2\,\theta(A)$.
ii) for every sequence of disjoint sets $(A_n)_n$ in $\mathcal{B}(\mathcal{C}^+)$, the Borel σ-algebra associated with $\mathcal{C}^+$, the random variables $(\mu(A_n))_n$ are independent and

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n), \qquad \text{a.s.}$$

Let Ω be the probability space on which μ is defined and let $\mathbb{P}$ be the law of μ. Ω is endowed with the σ-algebras $\mathcal{F}_u$ generated by the random variables $\mu(A)$, for all the sets A at a distance greater than u from the x-axis. The reader is referred to [3] for the existence of the probability space $(\Omega, (\mathcal{F}_u)_u, \mathbb{P})$.

The subsets needed for the definition of the Gaussian field are the cone-like subsets $A_u(x)$ of $\mathcal{C}^+$,

$$A_u(x) := \{(s,y) \in \mathcal{C}^+ \,:\, y \geq u\,,\, -f(y)/2 \leq s - x \leq f(y)/2\},$$

where $f(y) = y$ for $y \in (0,1/2)$ and $f(y) = 1/2$ otherwise. See Figure 1 for a depiction of the subsets. Observe that, by construction, if $\|x-x'\| = \ell > u$, then $A_u(x)$ and $A_u(x')$ intersect exactly above the line $y = \ell$.

The Gaussian process $\omega_u = \big(\omega_u(x), x \in [0,1]_\sim\big)$ is defined using the random measure μ,

$$\omega_u(x) := \mu(A_u(x)), \qquad x \in [0,1]_\sim. \tag{1.1}$$

By the properties i) and ii) of μ listed above, the covariance between $\omega_u(x)$ and $\omega_u(x')$ is given by the integral over θ of the intersection of $A_u(x)$ and $A_u(x')$:

$$\mathbb{E}[\omega_u(x)\omega_u(x')] = \int_{A_u(x)\cap A_u(x')} \theta(\mathrm{d}s, \mathrm{d}y). \tag{1.2}$$

The paper focuses on a discrete version of ω_u. Let $N \in \mathbb{N}$ and take $\varepsilon = 1/N$. Define the set

$$\mathcal{X}_N = \mathcal{X}_\varepsilon := \left\{0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{i}{N}, \dots, \frac{N-1}{N}\right\}.$$

The notation $\mathcal{X}_N$ and $\mathcal{X}_\varepsilon$ will be used equally depending on the context. For a given N, the log-correlated Gaussian field is the collection of Gaussian centered random variables $\omega_\varepsilon(x)$ for $x \in \mathcal{X}_N$:

$$X = (X_x, x \in \mathcal{X}_N) = (\omega_\varepsilon(x), x \in \mathcal{X}_N). \tag{1.3}$$

A compelling feature of this construction is that a *scale decomposition* for X is easily obtained from property ii) above. Indeed, it suffices to write the variable X_x as a sum of independent Gaussian fields corresponding to disjoint horizontal strips of $\mathcal{C}^+$. The y-axis then plays the role of the scale.

The covariances of the field are computed from (1.2) by straightforward integration (see also Figure 1).

Lemma 1.1. *For any* $0 < \varepsilon = 1/N < 1/2$,

$$\begin{aligned} \mathbb{E}[X_x^2] &= \sigma^2(\log N + 1 - \log 2), \qquad x \in \mathcal{X}_N, \\ \mathbb{E}[X_x X_{x'}] &= \sigma^2(\log(1/\|x - x'\|) - \log 2), \qquad x \neq x' \in \mathcal{X}_N. \end{aligned}$$

Similar constructions of log-correlated Gaussian fields using a random measure on cone-like subsets are also possible in two dimensions, see e.g. [28].

1.2. **Main results.** Without loss of generality, the results of this section are stated for the variance parameter $\sigma = 1$. The points where the field is unusually high, *the extremes* or *the high points*, can be studied using a minor adaptation of the arguments of Daviaud for the two-dimensional discrete Gaussian free field [16]. We denote by $|\mathcal{A}|$ the cardinality of a finite set $\mathcal{A}$.

Theorem 1.2. *Let*

$$\mathcal{H}_N(\gamma) := \left\{x \in \mathcal{X}_N : X_x \geq \sqrt{2}\gamma \log N\right\}$$

be the set of γ*-high points. Then for any* $0 < \gamma < 1$,

$$\lim_{N\to\infty} \frac{\log|\mathcal{H}_N(\gamma)|}{\log N} = 1 - \gamma^2, \qquad \text{in probability.}$$

Moreover, for all $\rho > 0$ *there exists a constant* $c = c(\rho) > 0$ *such that*

$$\mathbb{P}\left(|\mathcal{H}_N(\gamma)| \leq N^{(1-\gamma^2)-\rho}\right) \leq \exp\{-c(\log N)^2\},$$

for N *large enough.*

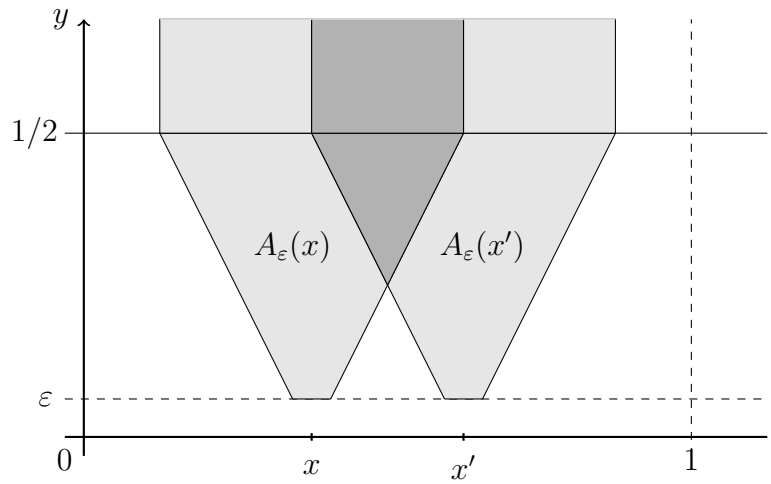

FIGURE 1. The two subsets $A_\varepsilon(x)$ and $A_\varepsilon(x')$ for $\varepsilon = 1/N$. The variance of the variables is given by the integral over $\theta(\mathrm{d}t, \mathrm{d}y) = y^{-2}\mathrm{d}t\ \mathrm{d}y$ of the lighter grey area above $\varepsilon = 1/N$, and the covariance by the integral over the intersection of the subsets, the darker grey region.

The technique of Daviaud is based on a tree approximation introduced by Bolthausen, Deuschel and Giacomin [5] for the discrete two-dimensional Gaussian free field. There, the technique is used to obtain the first order of the maximum. The same argument applies here. Theorem 1.2 and simple Gaussian estimates yield

$$\lim_{N\to\infty} \frac{\max_{x\in\mathcal{X}_N} X_x}{\log N} = \sqrt{2}, \qquad \text{in probability.} \tag{1.4}$$

The important feature of Theorem 1.2 and Equation (1.4) is that they are identical to the results for N i.i.d. Gaussian variables of variance $\log N$. In other words, the above observables of the high points are not affected by the correlations of the field. The i.i.d. case is called the *Random Energy Model* (REM) in the spin glass literature.

The starting point of the paper is to understand to which extent i.i.d. statistics is a good approximation for more refined observables of the extremes of log-correlated Gaussian fields. To this end, we turn to tools of statistical physics which allow for a good control of the correlations.

First, consider the *partition function* $Z_N(\beta)$ of the model (β stands for the inverse-temperature):

$$Z_N(\beta) := \sum_{x\in\mathcal{X}_N} \exp\{\beta X_x\}, \qquad \forall\ \beta > 0,$$

and the *free energy*

$$f_N(\beta) := \frac{1}{\log N} \log Z_N(\beta), \qquad \forall\ \beta > 0.$$

Theorem 1.2 is used to compute the free energy of the model.

Corollary 1.3. *Let $\beta_c := \sqrt{2}$. Then, for all $\beta > 0$*

$$f(\beta) := \lim_{N\to\infty} f_N(\beta) = \begin{cases} 1 + \frac{\beta^2}{2}, & \text{if } \beta < \beta_c, \\ \sqrt{2}\beta, & \text{if } \beta \geq \beta_c, \end{cases} \qquad \textit{a.s. and in } L^1.$$

The free energy is the same as for the REM with variance $\log N$. In particular, the model undergoes *freezing* above β_c in the sense that the quantity $f(\beta)/\beta$ is constant.

More importantly, consider the *normalized Gibbs weights* or *Gibbs measure*

$$G_{\beta,N}(x) := \frac{\mathrm{e}^{\beta X_x}}{Z_N(\beta)}, \qquad x \in \mathcal{X}_N.$$

By design, the Gibbs measure concentrates on the high points of the Gaussian field. The first main result of the paper is to achieve a control of the correlations at the level of the Gibbs measure. Precisely, with spin glasses in mind, we consider the normalized covariance or *overlap*

$$q(x,y) = q^{(N)}(x,y) := -\frac{\log \|y-x\|}{\log N}, \qquad x,y \in \mathcal{X}_N. \tag{1.5}$$

Clearly, $\|x-y\| = \varepsilon^{q(x,y)}$ and $0 \le q(x,y) \le 1$. Moreover, the overlap $q(x,y)$ is equal to the normalized correlations $\mathbb{E}[X_x X_y]/\mathbb{E}[X_x^2]$ plus a term that goes to zero as N goes to infinity.

A fundamental object, that records the correlations of high points, is the *distribution function of the overlap* sampled from the Gibbs measure. Namely, denote by $G_{\beta,N}^{\times 2}$ the product measure on $\mathcal{X}_N \times \mathcal{X}_N$. Let (x_1, x_2) be two *replicas* sampled from $G_{\beta,N}^{\times 2}$. Write for simplicity q_{12} for $q(x_1,x_2)$. The averaged distribution function of the overlap is:

$$x_\beta^{(N)}(q) := \mathbb{E}\left[G_{\beta,N}^{\times 2}\left\{q_{12} \le q\right\}\right], \qquad 0 \le q \le 1. \tag{1.6}$$

More generally, the product measure on s *replicas* $(x_1, ..., x_s) \in \mathcal{X}_N^s$ sampled from the Gibbs measure will be denoted by $G_{\beta,N}^{\times s}$. Let $F : [0,1]^{\frac{s(s-1)}{2}}$ be a continuous function on the overlaps of s replicas, that is a function that depends smoothly on $q_{ll'} := q(x_l, x_{l'})$, $l \neq l'$, for $(x_1, ..., x_s) \in \mathcal{X}_N^s$. We will write $\mathbb{E}G_{\beta,N}^{\times s}\big(F(q_{ll'})\big)$ for the averaged expectation of F when $(x_1, ..., x_s)$ is sampled from $G_{\beta,N}^{\times s}$.

The first result is the analogue of results of Derrida and Spohn for the Gibbs measure of branching Brownian motion (see Equation (6.19) in [18]), of Chauvin and Rouault on branching random walks [15] and of Bovier and Kurkova on Derrida's *Generalized Random Energy Models* (GREM) [17], [10].

Theorem 1.4. *For $\beta > \beta_c$,*

$$\lim_{N\to\infty} x_\beta^{(N)}(q) = \lim_{N\to\infty} \mathbb{E}\left[G_{\beta,N}^{\times 2}\left\{q_{12} \le q\right\}\right] = \begin{cases} \frac{\beta_c}{\beta}, & \text{for } 0 \le q < 1, \\ 1, & \text{for } q = 1. \end{cases}$$

In other words, the theorem states that for large N, the only possible normalized correlations between high points are 0 or 1. This had been conjectured for this type of Gaussian field by Carpentier and Le Doussal, see page 16 in [14].

Similarly to [11], the control of the correlations is achieved by introducing a perturbed version of the model, cf. Section 2.1. In the present case, the proof is more intricate since the structure of correlations of the Gaussian field for finite N is not tree-like or *ultrametric* as in the cases of branching Brownian motion and GREM's. For example, for branching Brownian motion, $q(x,y)$ corresponds to the branching time of the common ancestor of two particles at time t, x and y, divided by t. Because of the branching structure,

$$\text{the inequality } \; q(x,y) \ge \min\{q(x,z), q(y,z)\} \; \text{ is satisfied for all } x,y,z. \tag{1.7}$$

(The terminology *ultrametric* comes from the fact that the distance induced by the form $q(\cdot,\cdot)$ is ultrametric when (1.7) holds.) The *Parisi Ultrametricity Conjecture* in the spin-glass literature states that, even though tree-like correlations might not be present for finite N, ultrametric correlations are recovered in the limit $N \to \infty$ for a large class of Gaussian fields at the level of the Gibbs measure, that is:

$$\lim_{N\to\infty} \mathbb{E}\left[G_{\beta,N}^{\times 3}\left\{q_{12} \geq \min\{q_{13}, q_{23}\}\right\}\right] = 1. \tag{1.8}$$

It is not hard to see that Theorem 1.4 implies the ultrametricity conjecture for the Gaussian field considered, since the overlaps can only take value 0 or 1. (In the language of spin glasses, the field is said to admit a *one-step replica symmetry breaking* at low temperature.)

The second main result is to describe the entire joint distribution of overlaps sampled from the Gibbs measure. For the purpose of the statement, we recall the definition of a Poisson-Dirichlet variable. For $0 < \alpha < 1$, let $\eta = (\eta_i, i \in \mathbb{N})$ be the atoms of a Poisson random measure on $(0,\infty)$ of intensity measure $s^{-\alpha-1}\, ds$. A *Poisson-Dirichlet variable* ξ of parameter α is a probability measure on the space of decreasing weights $\boldsymbol{s} = (s_1, s_2, \dots)$ with $1 \geq s_1 \geq s_2 \geq \dots \geq 0$ and $\sum_i s_i \leq 1$ which has the same law as

$$\xi \overset{law}{=} \left(\frac{\eta_i}{\sum_j \eta_j}, i \in \mathbb{N}\right)_{\downarrow},$$

where $\downarrow$ stands for the decreasing rearrangement.

Theorem 1.5. *Let $\beta > \beta_c$ and $\xi = (\xi_k, k \in \mathbb{N})$ be a Poisson-Dirichlet variable of parameter β_c/β. Denote by E the expectation with respect to ξ. For any continuous function $F : [0,1]^{\frac{s(s-1)}{2}} \to \mathbb{R}$ of the overlaps of s replicas:*

$$\lim_{N\to\infty} \mathbb{E}\left[G_{\beta,N}^{\times s}\left(F(q_{ll'})\right)\right] = E\left[\sum_{k_1\in\mathbb{N},\dots,k_s\in\mathbb{N}} \xi_{k_1}\dots\xi_{k_s}\ F(\delta_{k_l k_{l'}})\right].$$

Essentially, the theorem shows that the Gibbs weights of the high points converge in law to a Poisson-Dirichlet variable. However, it is important to stress that, as in the case of branching Brownian motion (and unlike the REM), it is not the collection $(G_{\beta,N}(x), x \in \mathcal{X}_N)_{\downarrow}$ *per se* that converges to a Poisson-Dirichlet variable. This is because the continuity of the function F naturally identify points x, y for which $q(x,y)$ tends to 1 in the limit $N \to \infty$. Rather, the result shows that the Poisson-Dirichlet weights are formed by the sum of the Gibbs weights of high points that are arbitrarily close to each other.

1.3. **Relation to Previous Results.** Bolthausen and Kistler have studied a family of models called *generalized GREM's* for which the correlations are not ultrametric [7, 8] for finite N. By construction, the overlaps of these models can only take a finite number of values (uniformly in N, the number of variables). They compute the free energies and the Gibbs measure and prove the Parisi ultrametricity conjecture for these. Bovier and Kurkova [10, 11] have obtained the distribution of the Gibbs measure for Gaussian fields, called the CREM's, where the values of the overlaps are not *a priori* restricted. Their analysis is restricted to models with ultrametric correlations and include the case of branching Brownian motion.

The works of Bolthausen, Deuschel and Zeitouni [6], Bramson and Zeitouni [12] and Ding [19] establish the tightness of the recentered maximum of the two-dimensional discrete Gaussian free field. We expect that their method can be applied to the Gaussian field we consider.

We note that Fang and Zeitouni [22] have studied a branching random walk model where the variance of the motion is time-dependent. This model is related to the simpler GREM model of spin glasses and to the CREM of Bovier and Kurkova. The family of log-correlated Gaussian fields introduced in Section 2.2 is akin to these hierarchical models, where the scale parameter replaces the time parameter.

2. Outline of the Proof

2.1. **A family of perturbed models.** In this section, we define a family of Gaussian fields for which the variance parameter σ is scale-dependent. It can be seen as the GREM analogue for the non-hierarchical Gaussian field considered here. We restrict ourselves to the case where σ takes three values, which is the one needed for the proof of Theorem 1.4. However, the construction and the results can hold for any finite number of values.

Fix $\varepsilon = 1/N$. We introduce a scale (or time) parameter t by defining for any $t \in [0,1]$,

$$X_x(t) := \omega_{\varepsilon^t}(x), \qquad x \in \mathcal{X}_\varepsilon.$$

Observe that for any fixed x, the process $(X_x(t))_{0\le t\le 1}$ has independent increments and is a martingale for the filtration $(\mathcal{F}_{\varepsilon^t}, t \ge 0)$:

$$\mathbb{E}[X_x(t) \,|\, \mathcal{F}_{\varepsilon^s}] = X_x(s), \qquad \text{for } t > s.$$

This is a consequence of the defining property ii) of the random measure μ.

The parameters of the family of perturbed models are $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, 1)$, where $0 < \alpha_1 < \alpha_2 < 1$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ with $\sigma_i > 0$, $i = 1,2,3$. For the sake of clarity and to avoid repetitive trivial corrections, it is assumed throughout the paper that N^{α_1}, $N^{\alpha_2-\alpha_1}$, and $N^{1-\alpha_2}$ are integers. The Gaussian field $Y^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(t) = (Y_x^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(t), x \in \mathcal{X}_\varepsilon)$ is defined from the field X as follows

(2.1)

$$Y_x^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(t) = \begin{cases} \sigma_1 X_x(t), & \text{if } 0 < t \le \alpha_1, \\ \sigma_1 X_x(\alpha_1) + \sigma_2 \left(X_x(t) - X_x(\alpha_1)\right), & \text{if } \alpha_1 < t \le \alpha_2, \\ \sigma_1 X_x(\alpha_1) + \sigma_2 \left(X_x(\alpha_2) - X_x(\alpha_1)\right) + \sigma_3 \left(X_x(t) - X_x(\alpha_2)\right), & \text{if } \alpha_2 < t \le 1. \end{cases}$$

The construction is depicted in Figure 2. We write $Y^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}$ for the field $(Y_x^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(1), x \in \mathcal{X}_\varepsilon)$. The dependence on $\boldsymbol{\sigma}$ and $\boldsymbol{\alpha}$ will sometimes be dropped in the notation of Y for simplicity.

Consider the partition function $Z_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta)$ of the perturbed model

$$Z_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) := \sum_{x\in\mathcal{X}_N} \exp\{\beta\, Y_x\}, \tag{2.2}$$

and the free energy,

$$f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) := \frac{1}{\log N} \log Z_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta), \qquad \forall\, \beta > 0.$$

The log number of high points can be computed for the Gaussian field Y using Daviaud's technique recursively. The free energy is then obtained by doing an explicit

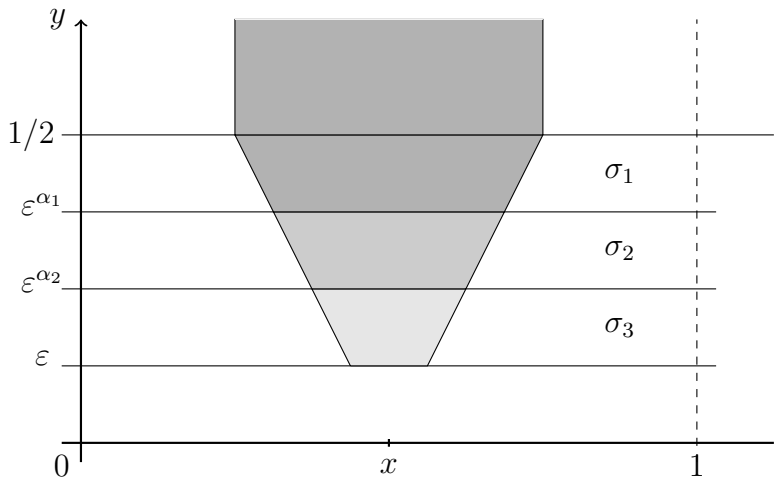

FIGURE 2. The cone associated with the process $Y_x(\cdot)$.

sum on these high points. This is the object of Section 3 and Section 4. We only write the results for the two cases needed for Theorem 1.4 as will be explained in Section 2.2. The result is better expressed in terms of the free energy of the REM with N i.i.d. Gaussian variables of variance $\sigma^2 \log N$:

$$f(\beta;\sigma^2) := \begin{cases} 1+\frac{\beta^2\sigma^2}{2}, & \text{if } \beta \le \beta_c(\sigma^2) := \frac{\sqrt{2}}{\sigma}, \\ \sqrt{2}\sigma\beta, & \text{if } \beta \ge \beta_c(\sigma^2). \end{cases}$$

Corollary 1.3 from the next result with the choice $\sigma_1 = \sigma_2 = \sigma_3$.

Proposition 2.1. *Let $V_{12} := \sigma_1^2\alpha_1 + \sigma_2^2(\alpha_2-\alpha_1)$, and $V_{23} := \sigma_2^2(\alpha_2-\alpha_1)+\sigma_3^2(1-\alpha_2)$. Then:*

- **Case 1:** *if $\sigma_1 \le \sigma_2$ and $\frac{V_{12}}{\alpha_2} \ge \sigma_3^2$,*
$$\lim_{N\to\infty} f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) = \alpha_2 f(\beta;\frac{V_{12}}{\alpha_2}) + (1-\alpha_2) f(\beta;\sigma_3^2),$$
- **Case 2:** *if $\sigma_1 \ge \sigma_2$, $\sigma_2 \le \sigma_3$, and $\sigma_1^2 \ge \frac{V_{23}}{1-\alpha_1}$,*
$$\lim_{N\to\infty} f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) = \alpha_1 f(\beta;\sigma_1^2) + (1-\alpha_1) f(\beta;\frac{V_{23}}{1-\alpha_1}),$$

where the convergence holds almost surely and in L^1.

The expressions are identical to the free energy of a GREM with three levels where the parameters σ_i fails to satisfy monotonicity conditions and is reduced to a GREM with two effective levels. The conditions are more easily understood by defining a piecewise linear function of slopes σ_1^2, σ_2^2 and σ_3^2 on the intervals $[0,\alpha_1]$, $[\alpha_1,\alpha_2]$, $[\alpha_2,1]$ respectively. In the two cases above, this functions fails to be concave. However, it is easily verified that the effective parameters define the concave hull of the function. The reader is referred to [13] and [10] for more details on the concavity conditions. Moreover, in both cases, there are two critical values for β corresponding to the respective $\beta_c(\sigma^2)$ of the two effective parameters σ^2. In Case 1, the two critical β's are $\sqrt{2\alpha_2/V_{12}}$ and $\sqrt{2/\sigma_3^2}$, whereas they are $\sqrt{2/\sigma_1^2}$ and $\sqrt{2(1-\alpha_1)/V_{23}}$ in the Case 2.

2.2. **The Bovier-Kurkova technique.** The proof of Theorem 1.4 relies on determining the overlap distribution of the original model from the free energy of the perturbed ones. This approach has been used by Bovier and Kurkova in the case of the GREM-type models [10, 11].

For $u \in (-1,1)$ and $t, \delta \in (0,1)$ such that $t+\delta < 1$, consider the field $(Y_x, x \in \mathcal{X}_\varepsilon)$ defined in (2.1) with the choice of parameters $\boldsymbol{\sigma} = (1, (1+u), 1)$ and $\boldsymbol{\alpha} = (t, t+\delta, 1)$, see Figure 3. (Again, for the sake of clarity, it is assumed that N^t, N^δ and $N^{1-(t+\delta)}$ are integers.) The original Gaussian field (X_x) is recovered at $u = 0$. Note that if $u > 0$, the parameters correspond to the first case of Proposition 2.1 and if $u < 0$, to the second. The field Y can also be represented as follows:

$$Y_x = X_x + u\left(X_x(t+\delta) - X_x(t)\right), \qquad 1 \le i \le N. \tag{2.3}$$

The proof of the next lemma is a simple integration and is postponed to the Appendix, see Section 5.2.

Lemma 2.2. *Fix $0 < \varepsilon = 1/N < 1/2$, and $t, \delta \in (0,1)$ such that $t+\delta < 1$. Let $\tilde{X}_x := X_x(t+\delta) - X_x(t)$. Then, for $x \in \mathcal{X}_\varepsilon$*

$$\mathbb{E}[\tilde{X}_x^2] = \mathbb{E}[\tilde{X}_x X_x] = \delta \log N, \qquad x \in \mathcal{X}_\varepsilon,$$

and, for $x, x' \in \mathcal{X}_\varepsilon$,

$$\mathbb{E}[\tilde{X}_x X_{x'}] = \begin{cases} \delta \log N + O(1), & \text{if } t+\delta \le q(x,x') \le 1, \\ (q(x,x') - t)\log N + O(1), & \text{if } t < q(x,x') < t+\delta, \\ 0, & \text{if } 0 \le q(x,x') \le t, \end{cases} \tag{2.4}$$

where we recall that $\|x - x'\| = \varepsilon^{q(x,x')}$.

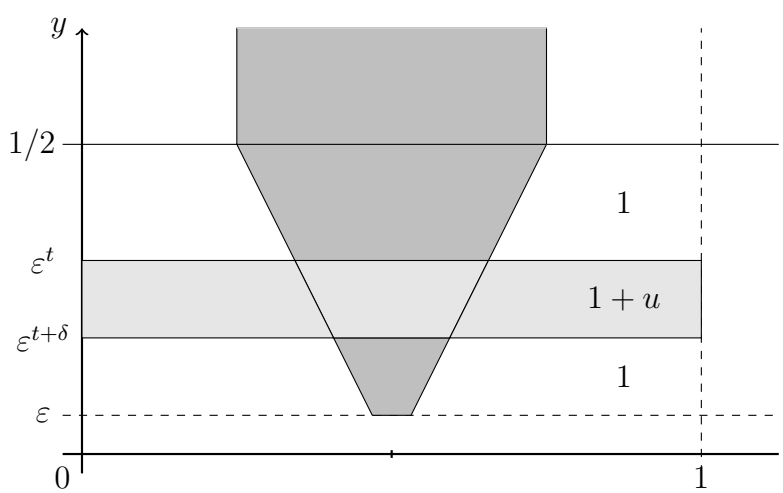

FIGURE 3. The perturbed model where the variance parameter is $(1+u)$ on the strip $[\varepsilon^{t+\delta}, \varepsilon^t]$ where $\varepsilon = 1/N$.

This result together with a Gaussian integration by part yield an important lemma.

Lemma 2.3. *For all $t, \delta \in (0,1)$, such that $t+\delta < 1$, we have*

$$\beta \int_t^{t+\delta} x_\beta^{(N)}(s)\mathrm{d}s + o_N(1) = \frac{1}{\log N}\mathbb{E}\left[\sum_{x \in \mathcal{X}_\varepsilon} G_{\beta,N}(x)(X_x(t+\delta) - X_x(t))\right],$$

where $o_N(1)$ stands for a term that goes to 0 as N goes to ∞.

Proof. Fix $\varepsilon = 1/N$, t and δ. Note that $(\tilde{X}_x; (X_{x'}, x' \in \mathcal{X}_\varepsilon))$ is a Gaussian vector of $N+1$ variables. Therefore, Gaussian integration by part (see Lemma 5.3) yields, for all $x \in \mathcal{X}_\varepsilon$,

$$\begin{aligned}\beta^{-1}\mathbb{E}\left[\frac{\tilde{X}_x \mathrm{e}^{\beta X_x}}{\sum_{x'\in\mathcal{X}_\varepsilon} \mathrm{e}^{\beta X_{x'}}}\right] &= -\sum_{x'\in\mathcal{X}_\varepsilon} \mathbb{E}\left[\tilde{X}_x X_{x'}\right] \mathbb{E}\left[\frac{\mathrm{e}^{\beta(X_x+X_{x'})}}{\left(\sum_{z\in\mathcal{X}_\varepsilon} \mathrm{e}^{\beta X_z}\right)^2}\right] \\ &\quad + \mathbb{E}\left[\tilde{X}_x X_x\right] \mathbb{E}\left[\frac{\mathrm{e}^{\beta X_x}}{\sum_{z\in\mathcal{X}_\varepsilon} \mathrm{e}^{\beta X_z}}\right].\end{aligned}$$

Lemma 2.2 and elementary manipulations imply

$$\begin{aligned}&(\beta \log N)^{-1}\mathbb{E}\left[\sum_{x\in\mathcal{X}_\varepsilon} \tilde{X}_x G_{\beta,N}(x)\right] \\ &= \sum_{x,x'\in\mathcal{X}_\varepsilon} \left(\int_t^{t+\delta} \mathbf{1}_{\{q(x,x')\le s\}} \mathrm{d}s\right) \mathbb{E}\left[G_{\beta,N}(x) G_{\beta,N}(x')\right] + O\left(\frac{1}{\log N}\right) \\ &= \int_t^{t+\delta} \mathbb{E}\left[G_{\beta,N}^{\times 2}\{q_{12}\le s\}\right] \mathrm{d}s + O\left(\frac{1}{\log N}\right),\end{aligned}$$

which concludes the proof of the lemma. □

Proof of Theorem 1.4. Fix $\beta > \beta_c = \sqrt{2}$. Write $Z_N^{(u,t,\delta)}(\beta)$ for the partition function (2.2) for the choices $\boldsymbol{\sigma} = (1, (1+u), 1)$ and $\boldsymbol{\alpha} = (t, t+\delta, 1)$. Direct differentiation and Equation (2.3) give

$$\frac{\mathrm{d}}{\mathrm{d}u}\left(\mathbb{E}\log Z_N^{(u,t,\delta)}(\beta)\right)_{u=0} = \beta\,\mathbb{E}\sum_{x\in\mathcal{X}_\varepsilon} (X_x(t+\delta) - X_x(t)) G_{\beta,N}(x),$$

which, together with Lemma 2.3, yields

$$\int_t^{t+\delta} x_\beta^{(N)}(s)\mathrm{d}s = \beta^{-2}(\log N)^{-1}\frac{\mathrm{d}}{\mathrm{d}u}\left(\mathbb{E}\log Z_N^{(u,t,\delta)}(\beta)\right)_{u=0} + o_N(1). \tag{2.5}$$

Observe that $\mathbb{E}f_N^{(u,t,\delta)}(\beta) = (\log N)^{-1}\mathbb{E}\log Z_N^{(u,t,\delta)}(\beta)$ is a convex function of u. Moreover, by Proposition 2.1, $\mathbb{E}f_N^{(u,t,\delta)}(\beta)$ converges. Write $f^{(u,t,\delta)}(\beta)$ for the limit. Recall that the expression for $f^{(u,t,\delta)}(\beta)$ depends on the sign of u and of course on β. Convexity in u implies that

$$\lim_{N\to\infty}\frac{\mathrm{d}}{\mathrm{d}u}\mathbb{E}f_N^{(u,t,\delta)}(\beta) = \frac{\mathrm{d}}{\mathrm{d}u}f^{(u,t,\delta)}(\beta) \text{ for any } u \text{ where } u \mapsto f^{(u,t,\delta)}(\beta) \text{ is differentiable.} \tag{2.6}$$

We show the function is differentiable at $u = 0$. The derivative can be computed by Proposition 2.1. For u small enough, β is larger than the two critical β's. Thus

$$\frac{\mathrm{d}}{\mathrm{d}u}f^{(u,t,\delta)}(\beta) = \begin{cases} \sqrt{2}\beta\frac{(1+u)(t+\delta)\delta}{\sqrt{(t+\delta)(t+(1+u)^2\delta)}}, & \text{if } u > 0, \\ \sqrt{2}\beta\frac{(1+u)(1-t)\delta}{\sqrt{(1-t)((1+u)^2\delta+1-(t+\delta))}}, & \text{if } u < 0. \end{cases}$$

From this, it is easily verified that $f^{(u,t,\delta)}(\beta)$ is differentiable at $u = 0$ and

$$\frac{\mathrm{d}}{\mathrm{d}u}\left(f^{(u,t,\delta)}(\beta)\right)_{u=0} = \sqrt{2}\beta\delta. \tag{2.7}$$

Equations (2.5), (2.6), and (2.7) together imply

$$\lim_{N\to\infty}\int_t^{t+\delta} x_\beta^{(N)}(s)\mathrm{d}s = \frac{\sqrt{2}}{\beta}\delta \quad , \text{ for all } t,\delta\in(0,1), \text{ with } t+\delta<1. \tag{2.8}$$

This shows weak convergence of the sequence $(x_\beta^{(N)})_N$ to the distribution function of the random variable taking values 0 and 1 with respective probability $\sqrt{2}/\beta$ and $1-\sqrt{2}/\beta$. Indeed, suppose the convergence does not hold. Since $(x_\beta^{(N)})_N$ is tight, there must exist a subsequence that converges weakly to a distribution function x_β where, for some $s_0\in(0,1)$, $x_\beta(s_0)>\sqrt{2}/\beta$ or $x_\beta(s_0)<\sqrt{2}/\beta$. But x_β is non-decreasing, so (2.8) must be violated for some $t>s_0$ or $t<s_0$ in the limit $N\to\infty$ for δ small enough. This concludes the proof of Theorem 1.4. □

2.3. **A spin-glass approach to Poisson-Dirichlet variables.** In this section, the link between Theorem 1.4 and Theorem 1.5 is explained. The technique, inspired from the study of spin glasses, is general and is of independent interest to prove convergence to Poisson-Dirichlet statistics.

The first step is to find a good space for the convergence of the random measure $G_{\beta,N}$. To this aim, note that the collection of functionals $\mathbb{E}G_{\beta,N}^{\times s}\big[F(q_{ll'})\big]$ over all $s\in\mathbb{N}$ and all continuous functions on the overlaps of s replicas determine the law of a $\mathbb{N}\times\mathbb{N}$ random matrix, say $R^{(N)}=(R_{ll'}^{(N)})_{ll'\in\mathbb{N}}$ through the identity:

$$\mathbb{E}G_{\beta,N}^{\times s}\big[F(q_{ll'})\big] = E\big[F(R_{ll'}^{(N)})\big].$$

$R_{ll'}^{(N)}$ is the overlap of the l-th and l'-th points sampled from $G_{\beta,N}$. We write E for the expectation of the law of $R^{(N)}$. $R^{(N)}$ is a covariance matrix almost surely and has only 1's on the diagonal. Moreover, since each point is sampled independently from the same measure, its law is *weakly exchangeable*, that is for any permutation π of a finite number of indices:

$$(R_{\pi(l)\pi(l')}^{(N)}) \overset{law}{=} (R_{ll'}^{(N)}).$$

It is not hard to see that the laws of random covariance matrices with 1's on the diagonal and with this above symmetry form a compact space under the weak topology induced by the convergence of expectation of the continuous functions F on s replicas, $s\in\mathbb{N}$. This space is called the space of *Random Overlap Structures* in [2]. In particular, there exists a subsequence $\{R^{(N_m)}\}$ that converges. Denote the limit random matrix by R. Since R is also weakly exchangeable, it is constructed by sampling from a random measure exactly as for $R^{(N)}$ by a representation of Dovbysh and Sudakov [20]. Precisely, there exists a random probability measure, say μ_β, on a Hilbert space $\mathcal{H}$, say $\ell^2(\mathbb{N})$, such that for any continuous function F on s replicas:

$$\lim_{m\to\infty}\mathbb{E}G_{\beta,N_m}^{\times s}\Big[F(q_{ll'})\Big] = E[F(R_{ll'})] = E\mu_\beta^{\times s}[F(v_l\cdot v_l')]. \tag{2.9}$$

In the above notation, s vectors of $\mathcal{H}$ are sampled independently from μ_β. The inner product between the l-th and l'-th copy is denoted by $v_l\cdot v_l'$. E is the expectation on the random measure μ_β. Note that, since $q(x,x)\le 1$, the random measure μ_β is supported on the unit ball.

The first consequence of Theorem 1.4 is that for any limit μ_β of a converging subsequence:

$$\mathbb{E}\left[\mu_\beta^{\times 2}\left\{v_1\cdot v_2\le q\right\}\right]=\lim_{N\to\infty}\mathbb{E}\left[G_{\beta,N}^{\times 2}\left\{q_{12}\le q\right\}\right]=\frac{\beta_c}{\beta}\,1_{[0,1)}(q)+1_{\{1\}}(q). \tag{2.10}$$

(The first equality is obtained by bounding $1_{[0,q]}(q_{ll'})$ by continuous functions on two replicas above and below and by applying (2.9).) Equation (2.10) implies that μ_β is an atomic measure.

Corollary 2.4. *If a subsequence of $(G_{\beta,N})$ converges weakly to μ_β in the sense of (2.9), then there exist random orthonormal vectors $(e_i; i\in\mathbb{N})\subset\mathcal{H}$, i.e. such that $e_i\cdot e_j=\delta_{ij}$; and random weights $\xi=(\xi_i; i\in\mathbb{N})_\downarrow$ with $\xi_i\ge 0$, $\sum_{i\in\mathbb{N}}\xi_i=1$ such that:*

$$\mu_\beta=\sum_{i\in\mathbb{N}}\xi_i\;\delta_{e_i},\qquad P-a.s.$$

Moreover, from (2.10), $E[\sum_{i\in\mathbb{N}}\xi_i^2]=1-\frac{\beta_c}{\beta}$.

Proof. μ_β is a random probability measure on the unit sphere of $\mathcal{H}$. Fix a realization of μ_β. Let B_ϵ be a ball in $\mathcal{H}$ of radius ϵ such that $\mu_\beta(B_\epsilon)>0$. Let $(v_l, l\in\mathbb{N})$ be iid vectors of $\mathcal{H}$ sampled from μ_β. There must be an infinite number of vectors of this sequence in B_ϵ by the Borel-Cantelli lemma 2. On the other hand by (2.10) the only possible values for $v_l\cdot v_{l'}$ is 0 or 1, P-a.s. By taking ϵ small enough, this shows that $v_l\cdot v_{l'}=1$ for every vector sampled from B_ϵ. Thus the v_l's sampled from B_ϵ are all equal showing that if $\mu_\beta(B_\epsilon)>0$, ϵ small enough, there exists a unique vector $e_0\in B_\epsilon$ such that $\mu\{e_0\}=\mu(B_\epsilon)$. Since this holds for any B_ϵ, we conclude there exists a countable (maybe finite) collection $\{e_i\}\subset\mathcal{H}$ such that $\mu_\beta\{e_i\}>0$. Moreover, $e_i\cdot e_j=0$ if $i\ne j$ since $v_l\cdot v_{l'}=0$ or 1, P-a.s. for the sequence of i.i.d. vectors sampled from μ_β. □

To finish the proof of Theorem 1.5, it remains to show that the random weights ξ are distributed like a Poisson-Dirichlet variable of parameter $\frac{\beta_c}{\beta}$. In fact, the parameter is already determined by Corollary 2.4, since for a Poisson-Dirichlet variable ξ' of parameter x, $E[\sum_k(\xi'_k)^2]=1-x$ holds, see e.g. Corollary 2.2 in [29]. This will also imply that for any converging sequence of $(G_{\beta,N})$ in the sense of (2.9), the limit is the same. In particular, it implies convergence of the whole sequence by compactness.

To prove the Poisson-Dirichlet statistics of the weights ξ, we use the following characterization theorem of the law, see [30] p. 22 for details. Define for all $m\in\mathbb{N}$ the joint moments of the weights

$$S(n_1,...,n_m)=E\sum_{k_1,...,k_m}\xi_{k_1}^{n_1}\dots\xi_{k_m}^{n_m},\qquad \text{for } n_1,...,n_m\ge 1. \tag{2.11}$$

The collection of $S(n_1,...,n_m)$, $m\in\mathbb{N}$, determines the law of a random mass-partition, that is a random variable on ordered sequences $1\ge r_1\ge r_2\ge\dots\ge 0$ with $\sum_{i\in\mathbb{N}}r_i\le 1$. If ξ is a Poisson-Dirichlet variable, it is shown in [30], Proposition 1.2.8, that the moments satisfy the recursion relations:

$$\begin{aligned}S(n_1+1,...,n_m)=&\frac{S(2)}{s}S(n_1,\dots,n_m)+\frac{n_1-1}{s}S(n_1,\dots,n_m)\\&+\sum_{2\le l\le m}\frac{n_l}{s}S(n_1+n_l,n_2,\dots,n_{l-1},n_{l+1},\dots,n_m),\end{aligned} \tag{2.12}$$

where $s = n_1 + ... + n_m$. It is not hard to verify that all moments $S(n_1, ..., n_m)$ (and thus the law of ξ) are determined by recursion from $S(2)$ and the identities (2.12).

It turns out that these identities are satisfied by ξ defined by Corollary 2.4.

Theorem 2.5. *Let ξ be a random mass-partition satisfying the assumptions of Corollary 2.4. The moments $S(n_1, ..., n_m)$ of ξ satisfy (2.12) for any $m \in \mathbb{N}$ and any $n_1, \dots, n_m \in \mathbb{N}$. In particular, ξ has the law of a Poisson-Dirichlet variable of parameter $1 - S(2)$.*

Proof. The identities are a general property of the Gibbs measure $(G_{\beta,N}(x)\,,\, x \in \mathcal{X}_N)$ of centered Gaussian fields known as the Ghirlanda-Guerra identities. They were introduced in [25]. It is shown in [27] that, for any β where the free energy $f(\beta)$ is differentiable, the following concentration holds:

$$\lim_{N\to\infty} \frac{1}{\log N} \mathbb{E} G_{\beta,N}\Big(\big| X_{x_1} - \mathbb{E} G_{\beta,N}(X_{x_1}) \big| \Big) = 0. \tag{2.13}$$

Note that by Corollary 1.3, differentiability holds at all β for the Gaussian field considered. Let F be a continuous function on the overlaps of s replicas. Observe that (2.13) and Cauchy-Schwartz inequality imply

$$\lim_{N} \frac{1}{\log N} \Big(\mathbb{E} G_{\beta,N}^{\times s}(X_{x_1} F(q_{ll'})) - \mathbb{E} G_{\beta,N}(X_{x_1}) \mathbb{E} G_{\beta,N}^{\times s}(F(q_{ll'})) \Big) = 0. \tag{2.14}$$

The two terms can be evaluated by Gaussian integrations by part, see Lemma 5.3,

$$\frac{1}{\beta \log N}\, \mathbb{E} G_{\beta,N}(X_{x_1}) = 1 - \mathbb{E} G_{\beta,N}^{\times 2}(q_{12}) + O\left(\frac{1}{\log N}\right), \tag{2.15}$$

and

$$\begin{aligned} &\frac{1}{\beta \log N}\, \mathbb{E} G_{\beta,N}^{\times s}(X_{x_1} F(q_{ll'})) \\ &= -s\, \mathbb{E} G_{\beta,N}^{\times s+1}(q_{1,s+1} F(q_{ll'})) + \sum_{1\le k\le s} \mathbb{E} G_{\beta,N}^{\times s}(q_{1k} F(q_{ll'})) + O\left(\frac{1}{\log N}\right). \end{aligned} \tag{2.16}$$

Finally recalling (2.14) and assembling (2.15)–(2.16) yields the Ghirlanda-Guerra identities (see Equation (16) in [25]):

$$\begin{aligned} &\mathbb{E} G_{\beta,N}^{\times s+1}\Big[q_{1,s+1}\ F(q_{ll'}) \Big] = \\ &\qquad \frac{1}{s} \mathbb{E} G_{\beta,N}^{\times 2}\Big[q_{12} \Big] \mathbb{E} G_{\beta,N}^{\times s}\Big[F(q_{ll'}) \Big] + \frac{1}{s} \sum_{k=2}^{s} \mathbb{E} G_{\beta,N}^{\times s}\Big[q_{1k}\ F(q_{ll'}) \Big] + o_N(1). \end{aligned} \tag{2.17}$$

(Note that the term for $k = 1$ cancels with the 1 since $q_{11} = 1 + o_N(1)$.)

In particular, for any converging subsequence of $(G_{\beta,N})_N$ in the sense of (2.9), one obtains by Corollary 2.4

$$\begin{aligned} &E\Big[\sum_{k_1,\dots,k_{s+1}} \xi_{k_1} \dots \xi_{k_{s+1}}\ \delta_{k_1 k_{s+1}} F(\delta_{k_l k_{l'}}) \Big] = \\ &\frac{1}{s} E\Big[\sum_k \xi_k^2 \Big] E\Big[\sum_{k_1,\dots,k_s} \xi_{k_1} \dots \xi_{k_s}\ F(\delta_{k_l k_{l'}}) \Big] + \frac{1}{s} \sum_{r=2}^{s} E\Big[\sum_{k_1,\dots,k_s} \xi_{k_1} \dots \xi_{k_s}\ \delta_{k_1 k_r}\ F(\delta_{k_l k_{l'}}) \Big]. \end{aligned} \tag{2.18}$$

To deduce (2.12) from (2.18), we follow ([30], pages 24, 25). The set $\{1, \dots, s\}$ can be decomposed into the disjoint union of sets $I_1, \dots, I_m$ with $|I_j| = n_j$ for all

$1 \le j \le m$. Consider the functions $(F_j)_{1\le j\le m}$ given by $F_j(\delta_{k_l k_{l'}}) := \prod_{k_l,k_{l'}\in I_j} \delta_{k_l k_{l'}}$ and define $F := \prod_{1\le j\le m} F_j$. Then, elementary manipulations imply (2.12). Note that the second term on the right side of (2.18) yields the last two terms of (2.12). □

3. High points of the perturbed models

In this section, the log-number of high points at a given level is computed for the perturbed models introduced in Section 2. The focus is on the two cases described in Theorem 2.1, though the technique applies to any perturbed model with a finite number of parameters. The free energies of the models are computed in Section 4.

Let $Y = (Y_x, x \in \mathcal{X}_\varepsilon)$ be the Gaussian field introduced in Section 2.1. Recall the notation and the two choices of parameters $(\boldsymbol{\sigma},\boldsymbol{\alpha})$ in Proposition 2.1:

$$\begin{aligned} &\textbf{Case 1:}\quad \sigma_1 \le \sigma_2, \qquad \frac{V_{12}}{\alpha_2} \ge \sigma_3^2; \\ &\textbf{Case 2:}\quad \sigma_1 \ge \sigma_2, \qquad \sigma_2 \le \sigma_3, \qquad \sigma_1^2 \ge \frac{V_{23}}{1-\alpha_1}. \end{aligned} \tag{3.1}$$

Define also as before $V_{12} := \sigma_1^2\alpha_1 + \sigma_2^2(\alpha_2-\alpha_1)$, $V_{23} := \sigma_2^2(\alpha_2-\alpha_1) + \sigma_3^2(1-\alpha_2)$, and $V_{123} := \sigma_1^2\alpha_1 + \sigma_2^2(\alpha_2-\alpha_1) + \sigma_3^2(1-\alpha_2)$.

Proposition 3.1.

$$\lim_{N\to\infty} \mathbb{P}\left(\max_{x\in\mathcal{X}_\varepsilon} Y_x \ge \sqrt{2}\gamma_{max}\log N\right) = 0,$$

where

$$\gamma_{max} = \gamma_{max}(\boldsymbol{\sigma},\boldsymbol{\alpha}) := \begin{cases} \sqrt{V_{12}\alpha_2} + \sigma_3(1-\alpha_2), & \textit{for Case 1;} \\ \sigma_1\alpha_1 + \sqrt{V_{23}(1-\alpha_1)}, & \textit{for Case 2.} \end{cases}$$

Proposition 3.2. *Let $\mathcal{H}_N^Y(\gamma) := \left\{x\in\mathcal{X}_\varepsilon : Y_x \ge \sqrt{2}\gamma\log N\right\}$ be the set of γ-high points. Then, for all $0<\gamma<\gamma_{max}$,*

$$\lim_{N\to\infty} \frac{\log|\mathcal{H}_N^Y(\gamma)|}{\log N} = \mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma), \qquad \textit{in probability,}$$

where in Case 1:

$$\mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma) := \begin{cases} 1-\frac{\gamma^2}{V_{123}}, & \textit{if } \gamma < V_{123}\sqrt{\frac{\alpha_2}{V_{12}}}, \\ (1-\alpha_2) + \frac{(\gamma-\sqrt{V_{12}\alpha_2})^2}{\sigma_3^2(1-\alpha_2)}, & \textit{if } \gamma \ge V_{123}\sqrt{\frac{\alpha_2}{V_{12}}}; \end{cases}$$

and in Case 2:

$$\mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma) := \begin{cases} 1-\frac{\gamma^2}{V_{123}}, & \textit{if } \gamma < \frac{V_{123}}{\sigma_1}, \\ (1-\alpha_1) + \frac{(\gamma-\sigma_1\alpha_1)^2}{V_{23}}, & \textit{if } \gamma \ge \frac{V_{123}}{\sigma_1}. \end{cases}$$

Moreover, for any $\mathcal{E} < \mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)$, there exists c such that

$$\mathbb{P}\left(|\mathcal{H}_N^Y(\gamma)| \le N^{\mathcal{E}}\right) \le \exp\{-c(\log N)^2\}.$$

The two propositions will be proved for Case 1, the reasoning for Case 2 being identical.

3.1. **Proof of Proposition 3.1.** The idea is to construct a Gaussian field with hierarchical correlations that dominates Y at the level of the covariances. The result will follow by comparison using Slepian's lemma. The same field will be used in the proof of the upper bound in Proposition 3.2.

Notice that if $\varepsilon^{\alpha_2} < \|x - x'\| \leq \varepsilon^{\alpha_1}$, the corresponding cone-like sets for Y_x and $Y_{x'}$ in $\mathcal{C}^+$ intersect between the lines $y = \varepsilon^{\alpha_2}$ and $y = \varepsilon^{\alpha_1}$. Therefore the covariance of the variables satisfies, writing $\ell := \|x - x'\|$,

$$\begin{aligned}\mathbb{E}[Y_x Y_{x'}] &= \sigma_2^2 \int_{\ell}^{\varepsilon^{\alpha_1}} \frac{y-\ell}{y^2}\mathrm{d}y + \sigma_1^2 \left(\int_{\varepsilon^{\alpha_1}}^{1/2} \frac{y-\ell}{y^2}\mathrm{d}y + \int_{1/2}^{\infty} \frac{1/2-\ell}{y^2}\mathrm{d}y \right) \\ &\geq \sigma_1^2 \left(\log \frac{1/2}{\varepsilon^{\alpha_1}} - 1 \right).\end{aligned}$$

By applying the same reasoning when $\varepsilon \leq \|x - x'\| \leq \varepsilon^{\alpha_2}$, one obtains the following lower bound for the covariance

$$\mathbb{E}[Y_x Y_{x'}] \geq \begin{cases} 0, & \text{if } \|x-x'\| > \varepsilon^{\alpha_1}, \\ \sigma_1^2 \left(\log \frac{1/2}{\varepsilon^{\alpha_1}} - 1 \right), & \text{if } \varepsilon^{\alpha_2} < \|x-x'\| \leq \varepsilon^{\alpha_1}, \\ \sigma_1^2 \left(\log \frac{1/2}{\varepsilon^{\alpha_1}} - 1 \right) + \sigma_2^2 \left(\log \frac{\varepsilon^{\alpha_1}}{\varepsilon^{\alpha_2}} - 1 \right), & \text{if } \varepsilon \leq \|x-x'\| \leq \varepsilon^{\alpha_2}. \end{cases} \tag{3.2}$$

Equation (3.2) is used to construct a Gaussian field $\tilde{Y}$. Define the map π

$$\begin{aligned}\pi : \mathcal{X}_\varepsilon \to &\quad \mathcal{X}_{\varepsilon^{\alpha_1}} \times \mathcal{X}_{\varepsilon^{\alpha_2}} \times \mathcal{X}_\varepsilon \\ x \to &\quad (\pi_1(x), \pi_2(x), x)\end{aligned}$$

where $\pi_1(x)$ is the unique $y \in \mathcal{X}_{\varepsilon^{\alpha_1}}$ such that $\|x - y\| \leq \frac{\varepsilon^{\alpha_1}}{2}$; $\pi_2(x)$ is the unique $y \in \mathcal{X}_{\varepsilon^{\alpha_2}}$ such that $\|x - y\| \leq \frac{\varepsilon^{\alpha_2}}{2}$. (If $\|x - y\| = \frac{\varepsilon^{\alpha_1}}{2}$, there are two possibilities for y. We take the right point). The pre-image of $y \in \mathcal{X}_{\varepsilon^{\alpha_1}}$ under π_1 are exactly the points in $\mathcal{X}_\varepsilon$ that are at a distance less than $\frac{\varepsilon^{\alpha_1}}{2}$ from y. One can think of $\pi_1(x)$ as the *ancestor* of x at the scale ε^{α_1} and $\pi_2(x)$ as the *ancestor* of x at the scale ε^{α_2}.

Consider the following Gaussian variables

$$\begin{aligned} &(g_x^{(1)}, x \in \mathcal{X}_{\varepsilon^{\alpha^1}}) \text{ i.i.d. Gaussians of variance } \sigma_1^2 \alpha_1 \log N - \sigma_1^2 \log 2 - \sigma_1^2, \\ &(g_x^{(2)}, x \in \mathcal{X}_{\varepsilon^{\alpha^2}}) \text{ i.i.d. Gaussians of variance } \sigma_2^2 (\alpha_2 - \alpha_1) \log N - \sigma_2^2, \\ &(g_x^{(3)}, x \in \mathcal{X}_{\varepsilon}) \text{ i.i.d. Gaussians of variance } \sigma_3^2 (1 - \alpha_2) \log N + 2\sigma_1^2 + \sigma_2^2. \end{aligned} \tag{3.3}$$

These three families are also taken independent. Then, the field $\tilde{Y}$ is defined, using the map π above and the Gaussian random variables $g_x^{(i)}$, by

$$\tilde{Y}_x = g^{(1)}_{\pi_1(x)} + g^{(2)}_{\pi_2(x)} + g^{(3)}_x. \tag{3.4}$$

This construction and Equation (3.2) directly imply the following comparison lemma.

Lemma 3.3.

$$\begin{aligned} \mathbb{E}[\tilde{Y}_x^2] &= \mathbb{E}[Y_x^2], \qquad \forall x \in \mathcal{X}_\varepsilon, \\ \mathbb{E}[\tilde{Y}_x \tilde{Y}_y] &\leq \mathbb{E}[Y_x Y_y], \qquad \forall\, x \neq y,\ x, y \in \mathcal{X}_\varepsilon. \end{aligned} \tag{3.5}$$

The following corollary is a straightforward consequence of the above lemma and Slepian's lemma, see Corollary 3.12 in [26].

Corollary 3.4. *For any $\lambda > 0$*

$$\mathbb{P}\left(\max_{x\in\mathcal{X}_\varepsilon} Y_x \geq \lambda\right) \leq \mathbb{P}\left(\max_{x\in\mathcal{X}_\varepsilon} \tilde{Y}_x \geq \lambda\right). \tag{3.6}$$

The Gaussian field $\tilde{Y}$ is almost identical to a GREM model with three levels with parameters $0 < \alpha_1 < \alpha_2 < 1$ and $\sigma_1, \sigma_2, \sigma_3$, see e.g. [17, 10]. In fact the only aspect different from an exact GREM are the terms of order one in the variances of the Gaussian random variables $g_x^{(i)}$'s. However, these do not affect the entropy of the high points. In Case 1, the field reduces to a two-level GREM with effective parameters $(V_{12}/\alpha_2, \sigma_3^2)$, $(\alpha_2, 1)$ whereas in Case 2, the effective parameters are $(\sigma_1^2, V_{23}/(1-\alpha_1))$, $(\alpha_1, 1)$. The proofs of Proposition 3.1 and of the upper bound of Proposition 3.2 is based on the following standard GREM result. A proof is given for completeness, but some details will be omitted. The reader is referred to Theorem 1.1 in [10] where a stronger result on the maximum is given and to [9], Lecture 9, for more details on the free energy and on the log-number of high points of a two-level GREM.

Lemma 3.5. *Let $\tilde{Y}$ be the Gaussian field constructed above. Then*

$$\mathbb{P}\left(\max_{x\in\mathcal{X}_\varepsilon} \tilde{Y}_x \geq \sqrt{2}\gamma_{max}\log N\right) \to 0, \qquad N \to \infty,$$

where γ_{max} is defined in Proposition 3.1. Moreover,

$$\lim_{N\to\infty} \frac{\log|\mathcal{H}_N^{\tilde{Y}}(\gamma)|}{\log N} = \mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma) \qquad \textit{in probability}, \tag{3.7}$$

where $\mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)$ is defined in Proposition 3.2.

Proof. We only prove the Case 1, the reasoning in the Case 2 being similar. Consider the field $(\tilde{Y}_x(\alpha_2), x \in \mathcal{X}_{\varepsilon^{\alpha_2}})$ where $\tilde{Y}_x(\alpha_2) := g^{(1)}_{\pi_1(x)} + g^{(2)}_{\pi_2(x)}$. Markov's inequality and a Gaussian estimate, see Lemma 5.1, yield

$$\mathbb{P}\left(\max_{x\in\mathcal{X}_{\varepsilon^{\alpha_2}}} \tilde{Y}_x(\alpha_2) \geq \sqrt{2}\sqrt{V_{12}\alpha_2}\log N\right) \to 0, \qquad N \to \infty. \tag{3.8}$$

Define

$$\mathcal{H}_N^{\tilde{Y}}(\gamma_2,\gamma_3) := \{x \in \mathcal{X}_\varepsilon : \tilde{Y}_x(\alpha_2) \geq \sqrt{2}\gamma_2\log N\ ,\ g_x^{(3)} \geq \sqrt{2}\gamma_3\log N\}.$$

Again, Markov's inequality together with a Gaussian estimate gives for $\gamma_2, \gamma_3 > 0$,

$$\mathbb{P}\left(|\mathcal{H}_N^{\tilde{Y}}(\gamma_2,\gamma_3)| \geq 1\right) \leq C\frac{\sqrt{V_{12}\ \sigma_3^2(1-\alpha_2)}}{\gamma_2\gamma_3\log N}\ N^{1-\frac{\gamma_2^2}{V_{12}}-\frac{\gamma_3^2}{\sigma_3^2(1-\alpha_2)}}.$$

Equation (3.8) implies that $|\mathcal{H}_N^{\tilde{Y}}(\gamma_2,\gamma_3)|$ is zero with probability tending to one if $\gamma_2 \geq \sqrt{V_{12}\alpha_2}$. Suppose $0 < \gamma_2 < \sqrt{V_{12}\alpha_2}$. Then, if $\gamma_2 + \gamma_3 \geq \gamma_{max}$, the second parameter γ_3 must be greater than $\sigma_3(1-\alpha_2)$. Therefore $\mathbb{P}(|\mathcal{H}_N^{\tilde{Y}}(\gamma_2,\gamma_3)| \geq 1)$ goes to 0, when N tends to infinity, in the case $\gamma_2 + \gamma_3 \geq \gamma_{max}$. This implies the first claim.

For the second claim, we note first that there is a self-averaging of the log-number of high points:

$$\lim_{N\to\infty} \frac{\log|\mathcal{H}_N^{\tilde{Y}}(\gamma_2,\gamma_3)|}{\log N} = \lim_{N\to\infty} \frac{\log\mathbb{E}|\mathcal{H}_N^{\tilde{Y}}(\gamma_2,\gamma_3)|}{\log N}, \qquad \text{in probability.}$$

This self-averaging holds under the two conditions on γ_2 and γ_3 imposed with high probability by the first part of the proof : $\gamma_2 < \sqrt{V_{12}\alpha_2}$ and $\gamma_2 + \gamma_3 < \gamma_{max}$. This is a

straightforward computation using the second moment method and is done in Lecture 9 in [9]. Note also that a Laplace-method argument yields

$$\lim_{N\to\infty} \frac{\log|\mathcal{H}_N^{\tilde{Y}}(\gamma)|}{\log N} = \lim_{N\to\infty} \frac{1}{\log N} \max_{\gamma_2:|\gamma_2|<\sqrt{V_{12}\alpha_2}} \log|\mathcal{H}_N^{\tilde{Y}}(\gamma_2,\gamma-\gamma_2)|, \qquad \text{in probability.}$$

It remains to notice that, by linearity of the expectation,

$$\frac{\log \mathbb{E}|\mathcal{H}_N^{\tilde{Y}}(\gamma_2,\gamma_3)|}{\log N} = 1 - \frac{\gamma_2^2}{V_{12}} - \frac{\gamma_3^2}{\sigma_3^2(1-\alpha_2)} + o_N(1). \tag{3.9}$$

For a given $\gamma = \gamma_2 + \gamma_3$, the expression on the right in (3.9) is maximized at

$$\gamma_2^* = \gamma\,\frac{V_{12}}{V_{123}}, \qquad \gamma_3^* = \gamma - \gamma_2^* = \gamma\,\frac{\sigma_3^2(1-\alpha_2)}{V_{123}}.$$

If $\gamma \le \gamma_{crit} = V_{123}\sqrt{\frac{\alpha_2}{V_{12}}}$, then γ_2^* and γ_3^* satisfy these conditions. Equation (3.9) evaluated at γ_2^* and γ_3^* equals $\mathcal{E}^{\boldsymbol{\sigma},\boldsymbol{\alpha}}(\gamma)$. If $\gamma > \gamma_{crit}$, then (3.9) is maximized for γ_2 tending to $\sqrt{V_{12}\alpha_2}$ and $\mathcal{E}^{\boldsymbol{\sigma},\boldsymbol{\alpha}}(\gamma)$ is again recovered. □

3.2. **Proof of Proposition 3.2.** Proposition 3.2 asserts that, for all $\rho > 0$,

$$\mathbb{P}\left(\left|\frac{\log|\mathcal{H}_N^{Y}(\gamma)|}{\log N} - \mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)\right| > \rho\right) \to 0, \qquad N\to\infty.$$

The proof is split in two parts, proving first that the upper bound $\mathbb{P}(|\mathcal{H}_N^{Y}(\gamma)| > N^{\mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)+\rho})$ converges to 0 by comparing to the field $\tilde{Y}$ constructed in the last section. Second, proving that the lower bound $\mathbb{P}(|\mathcal{H}_N^{Y}(\gamma)| < N^{\mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)-\rho})$ decays to zero following the argument of Daviaud [16].

3.2.1. *Proof of the upper bound in Proposition 3.2.* The first result is a comparison in the spirit of Corollary 3.4.

Corollary 3.6. *Let $\mathcal{H}_N^{Y}(\gamma) = \{x \in \mathcal{X}_\varepsilon : Y_x \ge \sqrt{2}\gamma\log N\}$ and similarly for $\tilde{Y}$. For any $M \in \mathbb{N}$,*

$$\mathbb{P}\left(|\mathcal{H}_N^{Y}(\gamma)| \ge M\right) \le \mathbb{P}\left(|\mathcal{H}_N^{\tilde{Y}}(\gamma)| \ge M\right). \tag{3.10}$$

Proof. The proof is a again a consequence of Lemma 3.3 and Slepian's lemma, see Corollary 3.12 in [26]. They imply that for any $\lambda_x \in \mathbb{R}$, $x \in \mathcal{X}_\varepsilon$,

$$\mathbb{P}\Big(Y_x \ge \lambda_x,\ x\in\mathcal{X}_\varepsilon\Big) \ge \mathbb{P}\Big(\tilde{Y}_x \ge \lambda_x,\ x\in\mathcal{X}_\varepsilon\Big). \tag{3.11}$$

The integer moments of $|\mathcal{H}_N(\gamma)|$ can be expressed as a linear combination of probabilities

$$\begin{aligned}\mathbb{E}[|\mathcal{H}_N^{Y}(\gamma)|^k] &= \sum_{x_1,\dots,x_k\in\mathcal{X}_\varepsilon} \mathbb{P}\left(Y_{x_1} \ge \sqrt{2}\gamma\log N,\dots,Y_{x_k} \ge \sqrt{2}\gamma\log N\right)\\ &\le \sum_{x_1,\dots,x_k\in\mathcal{X}_\varepsilon} \mathbb{P}\left(\tilde{Y}_{x_1} \ge \sqrt{2}\gamma\log N,\dots,\tilde{Y}_{x_k} \ge \sqrt{2}\gamma\log N\right) \quad = \mathbb{E}[|\mathcal{H}_N^{\tilde{Y}}(\gamma)|^k].\end{aligned}$$

The corollary follows from the inequality for the moments because the variables $|\mathcal{H}_N^{Y}(\gamma)|$ and $|\mathcal{H}_N^{\tilde{Y}}(\gamma)|$ are nonnegative and bounded by N. □

The proof of the upper bound in Proposition 3.2 can now be concluded. Let $\rho > 0$. Corollary 3.6 implies

$$\mathbb{P}\left(|\mathcal{H}_N^Y(\gamma)| \geq N^{\mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)+\rho}\right) \leq \mathbb{P}\left(|\mathcal{H}_N^{\tilde{Y}}(\gamma)| \geq N^{\mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)+\rho}\right).$$

On the other hand, the right side goes to zero by Lemma 3.5 since

$$\mathbb{P}\left(|\mathcal{H}_N^{\tilde{Y}}(\gamma)| \geq N^{\mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)+\rho}\right) \leq \mathbb{P}\left(\left|\frac{\log|\mathcal{H}_N^{\tilde{Y}}(\gamma)|}{\log N} - \mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)\right| \geq \rho\right).$$

□

3.2.2. *Proof of the lower bound in Proposition 3.2.* The proof of the lower bound is a finite recursive argument. Two lemmas are needed. The first is a generalization of the lower bound in Daviaud's theorem (see Theorem 1.2 or [16]).

Lemma 3.7. *Let $0 < \alpha_0 < \alpha \leq 1$. Suppose that the parameter σ is constant on the strip $[0,1]_\sim \times [\varepsilon^\alpha, \varepsilon^{\alpha_0}]$, and that the event*

$$\Xi_0 := \left\{\#\{x \in \mathcal{X}_{\varepsilon^{\alpha_0}} : Y_x(\alpha_0) \geq \sqrt{2}\gamma_0 \log N\} \geq N^{\mathcal{E}_0}\right\},$$

is such that

$$\mathbb{P}(\Xi_0^c) \leq \exp\{-c_0(\log N)^2\},$$

for some $\gamma_0 \geq 0$, $\mathcal{E}_0 > 0$ and $c_0 > 0$.

Let

$$\mathcal{E}(\gamma) := \mathcal{E}_0 + (\alpha - \alpha_0) - \frac{(\gamma - \gamma_0)^2}{\sigma^2(\alpha - \alpha_0)} > 0.$$

Then, for any γ such that $\mathcal{E}(\gamma) > 0$ and any $\mathcal{E} < \mathcal{E}(\gamma)$, there exists c such that

$$\mathbb{P}\left(\#\{x \in \mathcal{X}_{\varepsilon^\alpha} : Y_x(\alpha) \geq \sqrt{2}\gamma \log N\} \leq N^{\mathcal{E}}\right) \leq \exp\{-c(\log N)^2\}.$$

We stress that γ may be such that $\mathcal{E}(\gamma) < \mathcal{E}_0$. The second lemma, which follows, serves as the starting point of the recursion and is analogous to Lemma 8 in [5].

Lemma 3.8. *For any $0 < \alpha < \alpha_1$, there exists $\mathcal{E} = \mathcal{E}(\alpha)$ and $c = c(\alpha)$ such that*

$$\mathbb{P}\left(\#\{x \in \mathcal{X}_{\varepsilon^\alpha} : Y_x(\alpha) \geq 0\} \leq N^{\mathcal{E}}\right) \leq \exp\{-c(\log N)^2\}.$$

We first conclude the proof of the lower bound in Proposition 3.2 using the two above lemmas.

Proof of the lower bound of Proposition 3.2. Let γ such that $0 < \gamma < \gamma_{max}$. Choose $\mathcal{E}$ such that $\mathcal{E} < \mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\gamma)$. It will be shown that for some $c > 0$

$$\mathbb{P}\left(|\mathcal{H}_N^Y(\gamma)| \leq N^{\varepsilon}\right) \leq \exp\{-c(\log N)^2\}. \tag{3.12}$$

By Lemma 3.8, for $\alpha_0 < \alpha_1$ arbitrarily close to 0 and $\gamma_0 = 0$, there exists $\mathcal{E}_0 = \mathcal{E}_0(\alpha_0) > 0$ and $c_0 = c_0(\alpha_0) > 0$, such that

$$\mathbb{P}\left(\#\{x \in \mathcal{X}_{\varepsilon^{\alpha_0}} : Y_x(\alpha_0) \geq 0\} \leq N^{\mathcal{E}_0}\right) \leq \exp\{-c_0(\log N)^2\}. \tag{3.13}$$

Observe that we have $0 \leq \mathcal{E}_0 \leq \alpha_0$. Moreover, let

$$\mathcal{E}_1(\gamma_1) := \mathcal{E}_0 + (\alpha_1 - \alpha_0) - \frac{\gamma_1^2}{\sigma_1^2(\alpha_1 - \alpha_0)}. \tag{3.14}$$

Lemma 3.7 is applied from α_0 to α_1. For any γ_1 with $\mathcal{E}_1(\gamma_1) > 0$ and any $\mathcal{E}_1 < \mathcal{E}_1(\gamma_1)$, there exists $c_1 > 0$ such that

$$\mathbb{P}\left(\#\{x \in \mathcal{X}_{\varepsilon^{\alpha_1}} : Y_x(\alpha_1) \geq \sqrt{2}\gamma_1 \log N\} \leq N^{\mathcal{E}_1}\right) \leq \exp\{-c_1(\log N)^2\}.$$

Therefore, Lemma 3.7 can be applied from α_1 to α_2 for any γ_1 with $\mathcal{E}_1(\gamma_1) > 0$. Define similarly

$$\mathcal{E}_2(\gamma_1, \gamma_2) := \mathcal{E}_1(\gamma_1) + (\alpha_2 - \alpha_1) - \frac{(\gamma_2 - \gamma_1)^2}{\sigma_2^2(\alpha_2 - \alpha_1)}. \tag{3.15}$$

Then, for any γ_2 with $\mathcal{E}_2(\gamma_1, \gamma_2) > 0$, and $\mathcal{E}_2 < \mathcal{E}_2(\gamma_1, \gamma_2)$, there exists $c_2 > 0$ such that

$$\mathbb{P}\left(\#\{x \in \mathcal{X}_{\varepsilon^{\alpha_2}} : Y_x(\alpha_2) \geq \sqrt{2}\gamma_2 \log N\} \leq N^{\mathcal{E}_2}\right) \leq \exp\{-c_2(\log N)^2\}.$$

Finally, the lemma is applied from α_2 to 1 (where γ_1, γ_2 are such that $\mathcal{E}_1(\gamma_1) > 0$ and $\mathcal{E}_2(\gamma_1, \gamma_2) > 0$). Define

$$\mathcal{E}_3(\gamma_1, \gamma_2, \gamma_3) := \mathcal{E}_2(\gamma_1, \gamma_2) + (1 - \alpha_2) - \frac{(\gamma_3 - \gamma_2)^2}{\sigma_3^2(1 - \alpha_2)}. \tag{3.16}$$

Then for any γ_3 with $\mathcal{E}_3(\gamma_1, \gamma_2, \gamma_3) > 0$ and $\mathcal{E}_3 < \mathcal{E}_3(\gamma_1, \gamma_2, \gamma_3)$, there exists $c_3 > 0$ such that

$$\mathbb{P}\left(\#\{x \in \mathcal{X}_\varepsilon : Y_x \geq \sqrt{2}\gamma_3 \log N\} \leq N^{\mathcal{E}_3}\right) \leq \exp\{-c_3(\log N)^2\}. \tag{3.17}$$

Recalling that $0 \leq \mathcal{E}_0 \leq \alpha_0$, Equation (3.12) follows from (3.17) if it is proved that $\lim_{\alpha_0 \to 0} \mathcal{E}_3(\gamma_1, \gamma_2, \gamma) = \mathcal{E}^{(\boldsymbol{\sigma}, \boldsymbol{\alpha})}(\gamma)$ for an appropriate choice of γ_1 and γ_2 (in particular such that $\mathcal{E}_1(\gamma_1) > 0$ and $\mathcal{E}_2(\gamma_1, \gamma_2) > 0$). It is easily verified that, for a given γ, the quantity $\mathcal{E}_3(\gamma_1, \gamma_2, \gamma)$ is maximized at

$$\gamma_1^* = \gamma \frac{\sigma_1^2(\alpha_1 - \alpha_0)}{V_{123} - \sigma_1^2\alpha_0}, \qquad \gamma_2^* = \gamma \frac{V_{12} - \sigma_1^2\alpha_0}{V_{123} - \sigma_1^2\alpha_0}.$$

Plugging these back in (3.14) and (3.15) shows that $\mathcal{E}_1(\gamma_1^*) > 0$ and $\mathcal{E}_2(\gamma_1^*, \gamma_2^*) > 0$ provided that

$$\gamma < \frac{V_{123}}{\sigma_1} \qquad \text{and} \qquad \gamma < V_{123}\sqrt{\frac{\alpha_2}{V_{12}}} =: \gamma_{crit},$$

with α_0 small enough (depending on γ). Note that the second condition on γ implies the first since $\sigma_1 \leq \sigma_2$. Furthermore, since

$$\mathcal{E}_3(\gamma_1^*, \gamma_2^*, \gamma) = \mathcal{E}_0 + (1 - \alpha_0) - \frac{\gamma^2}{V_{123} - \sigma_1^2\alpha_0},$$

we obtain $\lim_{\alpha_0 \to 0} \mathcal{E}_3(\gamma_1^*, \gamma_2^*, \gamma) = \mathcal{E}^{(\boldsymbol{\sigma}, \boldsymbol{\alpha})}(\gamma)$, which concludes the proof in the case $0 < \gamma < \gamma_{crit}$.

If $\gamma_{crit} \leq \gamma < \gamma_{max}$, the condition $\mathcal{E}_2(\gamma_1^*, \gamma_2^*) > 0$ will be violated as α_0 goes to zero (note however that $\mathcal{E}_1(\gamma_1^*)$ remains positive). In this case, for $\nu > 0$, pick $\gamma_2^{**} = \sqrt{V_{12}\alpha_2} - \nu$ such that $\mathcal{E}_2(\gamma_1^*, \gamma_2^{**}) > 0$. The first term in γ_2^{**} corresponds to γ_2^* evaluated at γ_{crit} for $\alpha_0 = 0$. In particular, $\lim_{\alpha_0 \to 0, \nu \to 0} \mathcal{E}_2(\gamma_1^*, \gamma_2^{**}) = 0$. From (3.16), this shows that

$$\lim_{\alpha_0 \to 0, \nu \to 0} \mathcal{E}_3(\gamma_1^*, \gamma_2^{**}, \gamma) = (1 - \alpha_2) - \frac{(\gamma - \sqrt{V_{12}\alpha_2})^2}{\sigma_3^2(1 - \alpha_2)} = \mathcal{E}^{(\boldsymbol{\sigma}, \boldsymbol{\alpha})}(\gamma).$$

Note that $\mathcal{E}^{(\boldsymbol{\sigma}, \boldsymbol{\alpha})}(\gamma)$ is strictly positive if and only if $\gamma < \sqrt{V_{12}\alpha_2} + \sigma_3(1 - \alpha_2) = \gamma_{max}$. This concludes the proof of (3.12). □

Proof of Lemma 3.7. Let γ such that $\mathcal{E}(\gamma) > 0$ and $\mathcal{E}$ such that $0 < \mathcal{E} < \mathcal{E}(\gamma)$. Pick $\overline{\gamma} > \gamma$ such that

$$\mathcal{E}(\overline{\gamma}) > \mathcal{E} > 0. \tag{3.18}$$

Since $\overline{\gamma} > \gamma$, there exists $\varsigma \in (0,1)$ such that

$$\overline{\gamma}(1-\varsigma) \geq \gamma. \tag{3.19}$$

For $K \in \mathbb{N}$ (which will be fixed later), we set

$$\begin{aligned} \eta_\ell &:= \alpha_0 + \frac{\ell-1}{K}(\alpha-\alpha_0), \qquad 1 \leq \ell \leq K+1, \\ \lambda_\ell &:= \gamma_0 + \frac{\ell-1}{K}(\overline{\gamma}-\gamma_0)(1-\varsigma), \qquad 1 \leq \ell \leq K+1. \end{aligned}$$

Observe that the η_ℓ's and the λ_ℓ's satisfy $\eta_1 = \alpha_0 < \eta_2 < \cdots < \eta_K < \eta_{K+1} = \alpha$, and $\lambda_1 = \gamma_0 < \lambda_2 < \cdots < \lambda_K < \lambda_{K+1} = (1-\varsigma)\overline{\gamma} + \varsigma\gamma_0$. Consider the sets $\mathcal{A}_\ell$ given by:

$$\mathcal{A}_\ell := \left\{ \underline{x}^{(\ell)} = (x_1, \ldots, x_\ell) \,:\, x_i \in \mathcal{X}_{2\varepsilon^{\eta_i}} \,,\, \forall\, 1 \leq i \leq \ell \text{ and } \|x_{i+1} - x_i\| \leq \varepsilon^{\eta_i}/2 \right\},$$

for $1 \leq \ell \leq K+1$. Note that only half of the x_i's in $\mathcal{X}_{\varepsilon^{\eta_i}}$'s are considered. Also, to each x_i we consider the points x_{i+1} in $\mathcal{X}_{2\varepsilon^{\eta_{i+1}}}$ that are close to x_i. By analogy with a branching process, these points can be thought of as the *children* of x_i. The reason for these two choices is that the cones corresponding to the variables $Y_{x_{i+1}}(\eta_{i+1})$ and $Y_{x'_{i+1}}(\eta_{i+1})$ do not intersect below the line $y = \varepsilon^{\eta_i}$ if $x_i \neq x'_i$, see Figure 4.

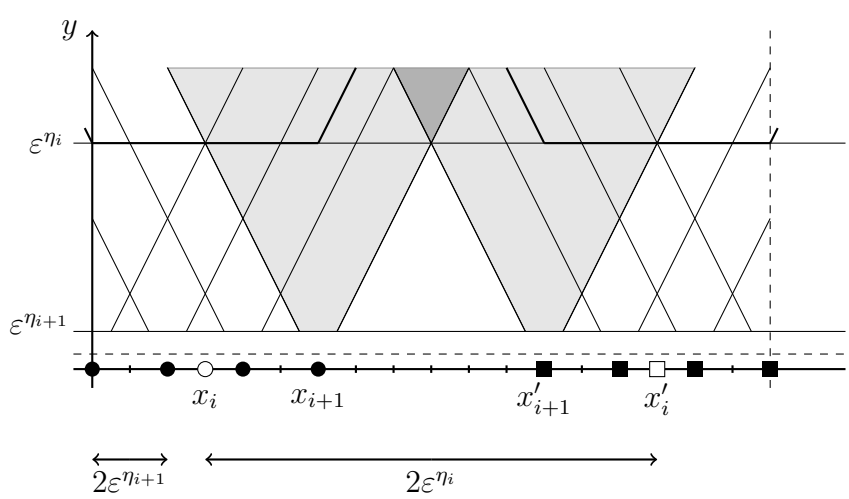

FIGURE 4. Approximation by a tree-like structure. The black circles symbolize the children of the white circle, while the black squares symbolize the children of the white square.

Now consider, the sets of high points of $\mathcal{A}_\ell$:

$$A_\ell := \left\{ \underline{x}^{(\ell)} \in \mathcal{A}_\ell \,:\, Y_{x_i}(\eta_i) \geq \sqrt{2}\lambda_i \log N \,,\, \forall\, 1 \leq i \leq \ell \right\}, \qquad 1 \leq \ell \leq K+1,$$

and

$$B_\ell := \{ \# A_\ell \geq n_\ell \}, \qquad 1 \leq \ell \leq K+1,$$

where

$$n_\ell := N^{\mathcal{E}_0 + \frac{\ell-1}{K}\left((\alpha-\alpha_0) - \frac{(\overline{\gamma}-\gamma_0)^2}{\sigma^2(\alpha-\alpha_0)}\right)}, \qquad 1 \leq \ell \leq K+1, \tag{3.20}$$

such that $N^{\mathcal{E}_0} = n_1$ and $n_{K+1} = N^{\mathcal{E}(\overline{\gamma})}$. Furthermore, with these definitions and the choice of $\overline{\gamma}$ in (3.19) and (3.18), we have for large N

$$\begin{aligned} B_{K+1} &= \{\#A_{K+1} > n_{K+1}\} \\ &\subset \Big\{\#\{x \in \mathcal{X}_{\varepsilon^\alpha} : Y_x(\alpha) \geq \sqrt{2}((1-\varsigma)\overline{\gamma} + \varsigma\gamma)\log N\} > N^{\mathcal{E}(\overline{\gamma})}\Big\} \\ &\subset \Big\{\#\{x \in \mathcal{X}_{\varepsilon^\alpha} : Y_x(\alpha) \geq \sqrt{2}\gamma \log N\} > N^{\mathcal{E}}\Big\}. \end{aligned}$$

It is thus sufficient to find a bound for $\mathbb{P}(B_{K+1}^c)$ to prove the lemma. For events C_ℓ to be defined in (3.23), we use the elementary bound $\mathbb{P}(B_{K+1}^c) \leq \mathbb{P}(B_{K+1}^c \cap B_K \cap C_K^c) + \mathbb{P}(C_K) + \mathbb{P}(B_K^c)$ which applied recursively gives

$$\mathbb{P}(B_{K+1}^c) \leq \sum_{\ell=2}^{K+1} \big(\mathbb{P}(B_\ell^c \cap B_{\ell-1} \cap C_{\ell-1}^c) + \mathbb{P}(C_{\ell-1})\big) + \mathbb{P}(B_1^c). \tag{3.21}$$

The last term has the correct bound by assumption. It remains to bound the ones appearing in the sum.

On the event B_ℓ, there exist at least n_ℓ high ℓ-branches $\underline{x}^{(\ell)} = (x_1, \ldots, x_\ell)$, these are branches that satisfy $Y_{x_i}(\eta_i) \geq \sqrt{2}\lambda_i \log N$ for $1 \leq i \leq \ell$. Select the first n_ℓ such ℓ-branches and denote them by $\underline{x}_j^{(\ell)} = (x_{j,1}, \ldots, x_{j,\ell})$, for all $1 \leq j \leq n_\ell$. Consider the set $\mathcal{A}_{j,\ell}$, the children of $x_{j,\ell}$ at level $\eta_{\ell+1}$: $\mathcal{A}_{j,\ell} := \{x \in \mathcal{X}_{2\varepsilon^{\eta_{\ell+1}}} : \|x - x_{j,\ell}\| \leq \varepsilon^{\eta_\ell}/2\}$. It holds

$$\begin{aligned} B_\ell \cap B_{\ell+1}^c \quad &\subset \quad B_\ell \cap \left\{ \sum_{j=1}^{n_\ell} \sum_{x \in \mathcal{A}_{j,\ell}} \mathbf{1}_{\left\{Y_x(\eta_{\ell+1}) - Y_{x_{j,\ell}}(\eta_\ell) \geq \sqrt{2}\frac{(\overline{\gamma}-\gamma_0)(1-\varsigma)}{K}\log N\right\}} \leq n_{\ell+1} \right\} \\ &\subset \quad B_\ell \cap \left\{ \sum_{j=1}^{n_\ell} \zeta_j \leq \frac{2n_{\ell+1}}{N^{(\alpha-\alpha_0)/K}} \right\}, \end{aligned}$$

where

$$\zeta_j := \frac{1}{|\mathcal{A}_{j,\ell}|} \sum_{x \in \mathcal{A}_{j,\ell}} \mathbf{1}_{\left\{Y_x(\eta_{\ell+1}) - Y_{x_{j,\ell}}(\eta_\ell) \geq \sqrt{2}\frac{(\overline{\gamma}-\gamma_0)(1-\varsigma)}{K}\log N\right\}}, \tag{3.22}$$

and $|\mathcal{A}_{j,\ell}| = N^{(\alpha-\alpha_0)/K}/2$. A crucial point is that $Y_{x_{j,\ell}}(\eta_\ell)$ is not equal to $Y_x(\eta_\ell)$ since $x \neq x_{j,\ell}$ in general. However, it turns out that their value must be very close since the variance of the difference is essentially a constant due to the logarithmic correlations. Precisely, let

$$C_\ell := \bigcup_{\underline{x}^{(\ell)} \in \mathcal{A}_\ell} \bigcup_{\substack{x \in \mathcal{X}_{2\varepsilon^{\eta_{\ell+1}}} : \\ \|x - x_\ell\| \leq \varepsilon^{\eta_\ell}/2}} \left\{ \big|Y_{\underline{x}^\ell}(\eta_\ell) - Y_x(\eta_\ell)\big| \geq \sqrt{2}\nu \frac{(\overline{\gamma}-\gamma_0)(1-\varsigma)}{K} \log N \right\}, \tag{3.23}$$

for $\nu > 0$ which is fixed and will be chosen small later. By Lemma 5.4 of the Appendix, $\mathrm{Var}(Y_x(\eta_\ell) - Y_{x'}(\eta_\ell)) \leq \max\{\sigma_1^2, \sigma_2^2, \sigma_3^2\} < \infty$, for every $1 \leq \ell \leq K$, and any $x \in \mathcal{X}_{2\varepsilon^{\eta_\ell}}$, $x' \in \mathcal{X}_{2\varepsilon^{\eta_{\ell+1}}}$ such that $\|x' - x\| \leq \varepsilon^{\eta_\ell}/2$. Therefore, a standard Gaussian estimate, see Lemma 5.1, together with the union-bound give

$$\mathbb{P}(C_\ell) \leq \exp\{-d(\log N)^2\}, \tag{3.24}$$

for all $1 \leq \ell \leq K$ and some $d > 0$.

It remains to bound the first term appearing in the sum of (3.21). On C_ℓ^c, $Y_{x_{j,\ell}}(\eta_\ell)$ can be replaced by $Y_x(\eta_\ell)$ in (3.22), making a small error that depends on ν. Namely, one has $\zeta_j \geq \tilde{\zeta}_j$, where

$$\tilde{\zeta}_j := \frac{1}{|\mathcal{A}_{j,\ell}|} \sum_{x \in \mathcal{A}_{j,\ell}} \mathbf{1}_{\left\{Y_x(\eta_{\ell+1}) - Y_x(\eta_\ell) \geq \sqrt{2}(1+\nu)\frac{(\overline{\gamma}-\gamma_0)(1-\varsigma)}{K}\log N\right\}}.$$

Note that conditionally on $\mathcal{F}_{\varepsilon^{\eta_\ell}}$, the $\tilde{\zeta}_j$'s are i.i.d. Moreover, since the $\tilde{\zeta}_j$'s are independent of $\mathcal{F}_{\varepsilon^{\eta_\ell}}$, they are also independent of each other. Lemma 5.2 of the Appendix guarantees that the sum of the $\tilde{\zeta}_j$ cannot be too low. Observe that

$$\mathbb{E}\left[\tilde{\zeta}_j\right] = \mathbb{P}\left(z \geq \sqrt{2}(1+\nu)\frac{(\overline{\gamma}-\gamma_0)(1-\varsigma)}{K}\log N\right),$$

where z is a centered Gaussian with variance $\sigma^2 \log\left(\frac{\varepsilon^{\eta_\ell}}{\varepsilon^{\eta_{\ell+1}}}\right) = \sigma^2 \frac{(\alpha-\alpha_0)}{K}\log N$. By a Gaussian estimate, Lemma 5.1,

$$\mathbb{E}\left[\tilde{\zeta}_j\right] \geq \exp\left\{-\frac{1}{K}\frac{(1+2\nu)^2(\overline{\gamma}-\gamma_0)^2(1-\varsigma)^2}{\sigma^2(\alpha-\alpha_0)}\log N\right\},$$

where $(1+\nu)$ has been replaced by $(1+2\nu)$ to absorb the $1/\sqrt{\log N}$ term in front of the exponential. Consequently, using elementary manipulations,

$$\begin{aligned} B_{\ell+1}^c \cap B_\ell \cap C_\ell^c &\subset \left\{\sum_{j=1}^{n_\ell}\left(\tilde{\zeta}_j - \mathbb{E}\left[\tilde{\zeta}_j\right]\right) \leq \frac{2n_{\ell+1}}{N^{(\alpha-\alpha_0)/K}} - n_\ell N^{-\frac{1}{K}\frac{(1+2\nu)^2(\overline{\gamma}-\gamma_0)^2(1-\varsigma)^2}{\sigma^2(\alpha-\alpha_0)}}\right\} \\ &\subset \left\{\left|\sum_{j=1}^{n_\ell}\left(\tilde{\zeta}_j - \mathbb{E}\left[\tilde{\zeta}_j\right]\right)\right| \geq \frac{1}{2}n_\ell N^{-\frac{1}{K}\frac{(1+2\nu)^2(\overline{\gamma}-\gamma_0)^2(1-\varsigma)^2}{\sigma^2(\alpha-\alpha_0)}}\right\}, \end{aligned}$$

provided

$$\frac{1}{K}\frac{(1+2\nu)^2(\overline{\gamma}-\gamma_0)^2(1-\varsigma)^2}{\sigma^2(\alpha-\alpha_0)} < \frac{1}{K}\frac{(\overline{\gamma}-\gamma_0)^2}{\sigma^2(\alpha-\alpha_0)},$$

that is

$$(1+2\nu)(1-\varsigma) < 1. \tag{3.25}$$

Fix ν small enough such that (3.25) is satisfied. Write for short

$$\mu := \frac{1}{K}\frac{(1+2\nu)^2(\overline{\gamma}-\gamma_0)^2(1-\varsigma)^2}{\sigma^2(\alpha-\alpha_0)}.$$

Then, taking $n = n_\ell$ and $t = n_\ell N^{-\mu}$ in Lemma 5.2, we get

$$\mathbb{P}(B_{\ell+1}^c \cap B_\ell \cap C_\ell^c) \leq 2\exp\left\{\frac{n_\ell^2 N^{-2\mu}}{2n_\ell + \frac{2}{3}n_\ell N^{-\mu}}\right\}.$$

By the form of n_ℓ in (3.20), K can be taken large enough so that $n_\ell N^{-2\mu} > N^\delta$ for some $\delta > 0$ and all $\ell = 1, ..., K+1$. This concludes the proof of the lemma. □

Proof of Lemma 3.8. Take $\alpha' < \alpha$ in such a way that $\mathcal{X}_{\varepsilon^{\alpha'}} \subset \mathcal{X}_{\varepsilon^\alpha}$. Consider the set

$$\Lambda := \{x \in \mathcal{X}_{\varepsilon^{\alpha'}} : Y_x(\alpha') \geq -\sigma_1(\alpha-\alpha')\log N\},$$

and the event

$$A = A_\delta := \left\{|\Lambda| \geq N^\delta\right\}, \qquad \delta > 0.$$

The parameters $\mathcal{E}$, δ and α' will be chosen later as a function of α. By splitting the probability on the event A,

$$\begin{aligned}
&\mathbb{P}\left(\#\{x \in \mathcal{X}_{\varepsilon^\alpha} : Y_x(\alpha) \geq 0\} \leq N^{\mathcal{E}}\right) \\
&\leq \mathbb{P}\left(\#\{x \in \mathcal{X}_{\varepsilon^\alpha} : Y_x(\alpha) \geq 0\} \leq N^{\mathcal{E}}; A\right) + \mathbb{P}(A^c) \\
&\leq \mathbb{E}\left[\mathbb{P}\left(\#\{x \in \Lambda : Y_x(\alpha) - Y_x(\alpha') \geq \sigma_1(\alpha-\alpha')\log N\} \leq N^{\mathcal{E}} \mid \mathcal{F}_{\varepsilon^{\alpha'}}\right); A\right] + \mathbb{P}(A^c),
\end{aligned}$$

where the second inequality is obtained by restricting to the set $\Lambda \subset \mathcal{X}_{\varepsilon^\alpha}$.

First we prove that the definition of A yields a super-exponential decay of the first term for $\mathcal{E}$ and δ depending on $\alpha - \alpha'$. The variables $Y_x(\alpha) - Y_x(\alpha')$, $x \in \mathcal{X}_{\varepsilon^{\alpha'}}$, are i.i.d. Gaussians of variance $\sigma_1^2(\alpha-\alpha')\log N$. Write for simplicity $(z_i, i = 1, ..., N^\delta)$ for i.i.d. Gaussians random variables with variance $\sigma_1^2(\alpha-\alpha')\log N$. A standard Gaussian estimate (see Lemma 5.1 of the Appendix) implies

$$\mathbb{P}\left(z_i \geq \sigma_1(\alpha-\alpha')\log N\right) \geq \frac{1}{2}\frac{\mathrm{e}^{-\frac{1}{2}(\alpha-\alpha')\log N}}{\sqrt{(\alpha-\alpha')\log N}} \geq \mathrm{e}^{-\frac{2}{3}(\alpha-\alpha')\log N}.$$

Therefore

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{P}\left(\#\{x \in \Lambda : Y_x(\alpha) - Y_x(\alpha') \geq \sigma_1(\alpha-\alpha')\log N\} \leq N^{\mathcal{E}} \mid \mathcal{F}_{\varepsilon^{\alpha'}}\right); A\right] \\
&\leq \mathbb{P}\left(\sum_{i=1}^{N^\delta}\left(\mathbf{1}_{\{z_i \geq \sigma_1(\alpha-\alpha')\log N\}} - \mathbb{P}\left(z_i \geq \sigma_1(\alpha-\alpha')\log N\right)\right) \leq N^{\mathcal{E}} - N^{\delta - \frac{2}{3}(\alpha-\alpha')}\right).
\end{aligned}$$

Lemma 5.2 in the Appendix gives a super-exponential decay of the above probability for the choice $\delta > \frac{4}{3}(\alpha-\alpha')$ and $\mathcal{E} - \delta + \frac{2}{3}(\alpha-\alpha') < 0$, for example $\delta = 2(\alpha-\alpha')$ and $\mathcal{E} = \alpha - \alpha'$.

It remains to show that $\mathbb{P}(A^c)$ has super-exponential decay. We have

$$\mathbb{P}(A^c) \leq P(A^c, \max_{x \in \mathcal{X}_{\varepsilon^{\alpha'}}} Y_x(\alpha') \leq (\log N)^2) + P(\max_{x \in \mathcal{X}_{\varepsilon^{\alpha'}}} Y_x(\alpha') > (\log N)^2).$$

The second term is easily shown to have the desired decay. We focus on the first. On the event $A^c \cap \{\max_{x \in \mathcal{X}_{\alpha'}} Y_x(\alpha') \leq (\log N)^2\}$,

$$\begin{aligned}
\frac{1}{|\mathcal{X}_{\varepsilon^{\alpha'}}|}\sum_{x \in \mathcal{X}_{\varepsilon^{\alpha'}}} \omega_{\alpha'}(x) &= \frac{1}{|\mathcal{X}_{\varepsilon^{\alpha'}}|}\sum_{x \in \Lambda} \omega_{\alpha'}(x) + \frac{1}{|\mathcal{X}_{\varepsilon^{\alpha'}}|}\sum_{x \in \Lambda^c} \omega_{\alpha'}(x) \\
&\leq \frac{|\Lambda|}{|\mathcal{X}_{\varepsilon^{\alpha'}}|}(\log N)^2 + \left(1 - \frac{|\Lambda|}{|\mathcal{X}_{\varepsilon^{\alpha'}}|}\right)(-\sigma_1(\alpha-\alpha')\log N).
\end{aligned} \tag{3.26}$$

Since $|\mathcal{X}_{\varepsilon^{\alpha'}}| = N^{\alpha'}$, it is easily checked that for $\delta = 2(\alpha-\alpha') < \alpha'$, the above is smaller than $-\frac{2}{3}\sigma_1(\alpha-\alpha')\log N$. Therefore we choose α' such that $\alpha < 3\alpha'/2$. Finally the left side of (3.26) is a Gaussian random variable, whose variance is of order 1. Therefore the probability that it is smaller than $-\frac{2}{3}\sigma_1(\alpha-\alpha')\log N$ is super-exponentially small. This completes the proof of the lemma. □

4. The free energy from the high points: proof of Proposition 2.1

In this section, we compute the free energy of the perturbed models introduced in Section 2.1. The free energy $f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta)$ is shown to converge in probability to the claimed expression. The L^1-convergence then follows from the fact that the variables

$(f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta))_{N\geq 1}$ are uniformly integrable. This is a consequence of Borell-TIS inequality. (Another more specific approach used by Capocaccia, Cassandro and Picco [13] for the GREM models could also have been applied here, see Section 3.1 in [13].) Indeed, we clearly have

$$\beta\frac{\max_{x\in\mathcal{X}_N} Y_x}{\log N} \leq f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) \leq 1+\beta\frac{\max_{x\in\mathcal{X}_N} Y_x}{\log N}.$$

Therefore, uniform integrability follows if it is proved that $\frac{1}{(\log N)^2}\mathbb{E}\big[\big(\max_{x\in\mathcal{X}_N} Y_x\big)^2\big]$ is uniformly bounded. It equals

$$\frac{1}{(\log N)^2}\mathbb{E}\big[\big(\max_{x\in\mathcal{X}_N} Y_x - \mathbb{E}[\max_{x\in\mathcal{X}_N} Y_x]\big)^2\big] + \frac{1}{(\log N)^2}\mathbb{E}[\max_{x\in\mathcal{X}_N} Y_x]^2.$$

The second term is uniformly bounded by comparing with i.i.d. centered Gaussian random variables of variance $V_{123}\log N$ and using Slepian's inequality (see e.g. [1], page 57). For the second term, we use Borell-TIS inequality (see [1], page 50)

$$\mathbb{P}\Big(\big|\max_{x\in\mathcal{X}_N} Y_x - \mathbb{E}\max_{x\in\mathcal{X}_N} Y_x\big| > r\Big) \leq 2e^{-\frac{r^2}{2V_{123}\log N}}, \qquad \forall r>0,$$

to get

$$\mathbb{E}\left[\left(\frac{\max_{x\in\mathcal{X}_N} Y_x - \mathbb{E}[\max_{x\in\mathcal{X}_N} Y_x]}{\log N}\right)^2\right] \leq 4\int_0^\infty re^{-\frac{r^2}{2V_{123}}\log N}\mathrm{d}r,$$

which goes to zero for $N\to\infty$. The almost-sure convergence is straightforward from the L^1-convergence and the almost-sure self-averaging property of the free energy:

$$\lim_{N\to\infty}|f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) - \mathbb{E}f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta)| = 0, \qquad \text{a.s.}$$

This is a standard consequence of concentration of measure (see [30], page 32) since the free energy is a Lipschitz function of i.i.d. Gaussian variables of Lipschitz constant smaller than $\beta/\sqrt{\log N}$. (Note that the Y_x's can be written as a linear combination of i.i.d. standard Gaussians with coefficients chosen to get the correct covariances.)

It remains to prove that the free energy $f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta)$ converges in probability to the claimed expression in Proposition 2.1. For fixed $\beta>0$ and $\nu>0$, we prove that

$$\lim_{N\to\infty}\mathbb{P}\left(f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) \leq f^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta)-\nu\right) = 0, \tag{4.1}$$

$$\lim_{N\to\infty}\mathbb{P}\left(f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) \geq f^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta)+\nu\right) = 0. \tag{4.2}$$

First, we introduce some notations and give a preliminary result. For simplicity, we will write $\mathcal{E}$ for $\mathcal{E}^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}$ throughout the proof. For any $M\in\mathbb{N}$, consider the partition of $[0,\gamma_{max}]$ into M intervals $[\gamma_{i-1},\gamma_i[$, where the γ_i's are given by

$$\gamma_i := \frac{i}{M}\gamma_{max}, \qquad i=0,1,\ldots,M.$$

Moreover for any $N\geq 2$, any $M\in\mathbb{N}$ and any $\delta>0$, define the random variables

$$K_{N,M}^+(i) := \#\left\{x\in\mathcal{X}_N : \frac{Y_x}{\sqrt{2}\log N}\in[\gamma_{i-1},\gamma_i[\right\}, \qquad 1\leq i\leq M,$$

$$K_{N,M}^-(i) := \#\left\{x\in\mathcal{X}_N : -\frac{Y_x}{\sqrt{2}\log N}\in[\gamma_{i-1},\gamma_i[\right\}, \qquad 1\leq i\leq M,$$

and the events

$$B^{\pm}_{N,M,\delta} := \bigcap_{i=1}^{M} \{N^{\mathcal{E}(\gamma_{i-1})-\delta} - N^{\mathcal{E}(\gamma_i)+\delta} \le K^{\pm}_{N,M}(i) \le N^{\mathcal{E}(\gamma_{i-1})+\delta} - N^{\mathcal{E}(\gamma_i)-\delta}\}$$
$$\bigcap \Big\{\#\{x \in \mathcal{X}_N \,:\, |Y_x| \ge \sqrt{2}\gamma_{max}\log N\} = 0\Big\}.$$

The next result is a straightforward consequence of Proposition 3.1, Proposition 3.2 and the fact that Gaussian random variables are symmetric.

Lemma 4.1. *For any $M \in \mathbb{N}$ and any $\delta > 0$, we have*

$$\lim_{N\to\infty} \mathbb{P}\left(B^{+}_{N,M,\delta} \cap B^{-}_{N,M,\delta}\right) = 1.$$

Define the continuous function

$$P_\beta(\gamma) := \mathcal{E}(\gamma) + \sqrt{2}\beta\gamma, \qquad \forall \gamma \in [0, \gamma_{max}].$$

Using the expression of $\mathcal{E}$ in Proposition 3.2 on the different intervals, it is easily checked by differentiation that

$$\sup_{\gamma\in[0,\gamma_{max}]} P_\beta(\gamma) = f^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta). \tag{4.3}$$

Furthermore, the continuity of $\gamma \mapsto P_\beta(\gamma)$ on $[0, \gamma_{max}]$ yields

$$\max_{0\le i\le M-1} P_\beta(\gamma_i) \longrightarrow \sup_{\gamma\in[0,\gamma_{max}]} P_\beta(\gamma) = f^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta), \qquad M \to \infty.$$

Fix $M \in \mathbb{N}$ large enough and $\delta > 0$ small enough, such that

$$\max_{0\le i\le M-1} P_\beta(\gamma_i) \ \ge\ f^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) - \frac{\nu}{3}, \tag{4.4}$$

$$\frac{\sqrt{2}\beta}{M} \ <\ \frac{\nu}{3}, \tag{4.5}$$

$$\delta \ <\ \min\left\{-\frac{1}{2}\min_{1\le i\le M}\{\mathcal{E}(\gamma_i) - \mathcal{E}(\gamma_{i-1})\}, \frac{\nu}{3}\right\}. \tag{4.6}$$

Note that for fixed M, $\min_{1\le i\le M}\{\mathcal{E}(\gamma_i) - \mathcal{E}(\gamma_{i-1})\} < 0$ since $\gamma \mapsto \mathcal{E}(\gamma)$ is a decreasing function on $[0, \gamma_{max}]$.

Proof of the lower bound (4.1). Observe that the partition function $Z_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta)$ associated with the perturbed model satisfies $Z_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) \ge \sum_{i=1}^{M} K^{+}_{N,M}(i) N^{\sqrt{2}\gamma_{i-1}\beta}$. Therefore on $B^{+}_{N,M,\delta}$ we get

$$Z_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) \ge \sum_{i=1}^{M} \left(1 - N^{\mathcal{E}(\gamma_i)-\mathcal{E}(\gamma_{i-1})+2\delta}\right) N^{P_\beta(\gamma_{i-1})-\delta}.$$

This yields on $B^{+}_{N,M,\delta}$

$$f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) \ge \frac{\log(1 - N^{\min_{1\le i\le M}\{\mathcal{E}(\gamma_i)-\mathcal{E}(\gamma_{i-1})\}+2\delta})}{\log N} + \max_{0\le i\le M-1} P_\beta(\gamma_i) - \delta.$$

Since for δ in (4.6)

$$\lim_{N\to\infty} (\log N)^{-1} \log(1 - N^{\min_{1\le i\le M}\{\mathcal{E}(\gamma_i)-\mathcal{E}(\gamma_{i-1})\}+2\delta}) = 0,$$

the choices of M, δ in (4.4) and (4.6) give that $f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) - f^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) > -\nu$ on $B^{+}_{N,M,\delta}$ for N large enough. Therefore, (4.1) is a consequence of Lemma 4.1.

Proof of the upper bound (4.2).

Observe first that the partition function $Z_N(\beta)$ satisfies on $B^+_{N,M,\delta} \cap B^-_{N,M,\delta}$

$$Z_N(\beta) \le \sum_{i=1}^{M} K^+_{N,M}(i) N^{\sqrt{2}\gamma_i\beta} + \sum_{i=1}^{M} K^-_{N,M}(i) N^{-\sqrt{2}\gamma_{i-1}\beta}.$$

Since on $B^+_{N,M,\delta} \cap B^-_{N,M,\delta}$ the random variables $K^+_{N,M}(i)$ and $K^-_{N,M}(i)$ are less than $N^{\mathcal{E}(\gamma_{i-1})+\delta}$ for all $1 \le i \le M$, this yields

$$Z_N(\beta) \le \sum_{i=1}^{M} N^{\mathcal{E}(\gamma_{i-1})+\delta} \left(N^{\sqrt{2}\gamma_i\beta} + N^{-\sqrt{2}\gamma_{i-1}\beta} \right) \le 2 \sum_{i=1}^{M} N^{\mathcal{E}(\gamma_{i-1})+\sqrt{2}\gamma_i\beta+\delta}.$$

Therefore, we get

$$f_N(\beta) \le \frac{\log(2M)}{\log N} + \sup_{\gamma \in [0,\gamma_{max}]} P_\beta(\gamma) + \frac{\sqrt{2}\beta}{M} + \delta,$$

on $B^+_{N,M,\delta} \cap B^-_{N,M,\delta}$. Recalling (4.3) and since $\lim_{N\to\infty} (\log N)^{-1} \log(2M) = 0$, the choices of M and δ in (4.5) and (4.6) imply that $f_N^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) - f^{(\boldsymbol{\sigma},\boldsymbol{\alpha})}(\beta) < \nu$ on $B^+_{N,M,\delta} \cap B^-_{N,M,\delta}$ for N large enough. Therefore (4.2) is a consequence of Lemma 4.1. □

5. Appendix

5.1. Gaussian estimates, large deviation result and integration by part.

Lemma 5.1 (see e.g. [21]). *Let X be a standard Gaussian random variable.*

(1) *For any $a \ge 0$,*
$$\mathbb{P}(|X| \ge a) \le \mathrm{e}^{-a^2/2}.$$
(2) *For any $a \ge 1$,*
$$\mathbb{P}(|X| \ge a) \ge \frac{\mathrm{e}^{-a^2/2}}{\sqrt{2\pi}a}.$$
(3) *Moreover we have the following approximation for a large*
$$\frac{(1-2a^{-2})}{\sqrt{2\pi}a}\mathrm{e}^{-a^2/2} \le \mathbb{P}(X \ge a) \le \frac{1}{\sqrt{2\pi}a}\mathrm{e}^{-a^2/2}.$$

Lemma 5.2 (see e.g. [4]). *Let $Z_1, \dots, Z_n$ be i.i.d. real valued random variables satisfying $\mathbb{E}[Z_i] = 0$, $\sigma^2 = \mathbb{E}[Z_i^2]$ and $\|Z_i\|_\infty \le 1$. Then for any $t > 0$,*

$$\mathbb{P}\left(\left| \sum_{i=1}^{n} Z_i \right| \ge t \right) \le 2 \exp\left\{ -\frac{t^2}{2n\sigma^2 + 2t/3} \right\}.$$

Lemma 5.3 (see e.g. Appendix of [30]). *Let $(X, Z_1, \dots, Z_d)$ be a centered Gaussian random vector. Then, for any C^1 function $F : \mathbb{R}^d \mapsto \mathbb{R}$, of moderate growth at infinity, we have*

$$\mathbb{E}\left[X F(Z_1, \dots, Z_d) \right] = \sum_{i=1}^{d} \mathbb{E}\left[X Z_i \right] \mathbb{E}\left[\frac{\partial F}{\partial z_i}(Z_1, \dots, Z_d) \right].$$

5.2. **Proof of Lemma 2.2.** Recall that $0 < \varepsilon = 1/N < 1/2$, and $t, \delta \in (0,1)$ is such that $t + \delta < 1$. Also by definition, $\|x' - x\| = \varepsilon^{q(x,x')}$.

It is clear that $\mathbb{E}[\tilde{X}_x X_x] = E[(\tilde{X}_x)^2]$, which is the variance of the centered Gaussian random variable $\mu(A_{\varepsilon^{t+\delta}}(x) \setminus A_{\varepsilon^t}(x))$. This variance can be computed and equals

$$\int_{\varepsilon^{t+\delta}}^{\varepsilon^t} y^{-1} \mathrm{d}y = [\log y]_{\varepsilon^{t+\delta}}^{\varepsilon^t} = \delta \log N.$$

For the covariance, observe that $\mathbb{E}[\tilde{X}_x X_{x'}]$ is equal to the variance of the random variable $\mu((A_{\varepsilon^{t+\delta}}(x) \setminus A_{\varepsilon^t}(x)) \cap A_\varepsilon(x'))$. If $\varepsilon \le \ell := \|x' - x\| \le \varepsilon^{t+\delta}$ (i.e. $t + \delta \le q(x,x') \le 1$), then the subsets $A_\varepsilon(x)$ and $A_\varepsilon(x')$ intersect below the line $y = \varepsilon(t+\delta)$ thus, the covariance is given by

$$\mathbb{E}[\tilde{X}_x X_{x'}] = \int_{\varepsilon^{t+\delta}}^{\varepsilon^t} \frac{y - \ell}{y^2} \mathrm{d}y = [\log y]_{\varepsilon^{t+\delta}}^{\varepsilon^t} + \ell \left[\frac{1}{y}\right]_{\varepsilon^{t+\delta}}^{\varepsilon^t} = \delta \log N + O(1).$$

If $\varepsilon^{t+\delta} < \ell = \|x' - x\| < \varepsilon^t$ (i.e. $t < q(x,x') \le t + \delta$), then the subsets intersect in between the lines $y = \varepsilon^{t+\delta}$ and $y = \varepsilon^t$, thus

$$\mathbb{E}[\tilde{X}_x X_{x'}] = \int_{\ell}^{\varepsilon^t} \frac{y - \ell}{y^2} \mathrm{d}y = [\log y]_{\ell}^{\varepsilon^t} + \ell \left[\frac{1}{y}\right]_{\ell}^{\varepsilon^t} = (q(x,x') - t) \log N + O(1).$$

Finally if $\ell = \|x' - x\| \ge \varepsilon^t$ (i.e. $0 \le q(x,x') \le t$), then the set $(A_{\varepsilon^{t+\delta}}(x) \setminus A_{\varepsilon^t}(x)) \cap A_\varepsilon(x')$ is empty and thus $\mathbb{E}[\tilde{X}_x X_{x'}] = 0$. □

5.3. **A key property of the perturbed models.** The following lemma is a key tool to approximate the Gaussian field we consider by a tree. Indeed the difference between the contribution to the Gaussian field at a certain scale for two points that are close can be explicitly computed by integrating parallelograms, see Figure 5 below, and is shown to be small.

Lemma 5.4. *Fix α, α_0 as in Lemma 3.7, u such that $\alpha_0 < u < \alpha$ and $\delta \in (0,1)$. Then for all $x, x' \in \mathcal{X}_\varepsilon$ such that $\|x - x'\| \le \delta \varepsilon^u$, we have*

$$\mathrm{Var}\,(Y_x(u) - Y_{x'}(u)) \le 2\,\overline{\sigma}^2\,\delta,$$

where $\overline{\sigma}$ denotes an upper bound for the σ_i's.

Proof. Writing $A := A_{\varepsilon^u}(x) \Delta\, A_{\varepsilon^u}(x')$, we have

$$\begin{aligned} \mathrm{Var}\,(Y_x(u) - Y_{x'}(u)) &\le \overline{\sigma}^2 \int_A y^{-2} \mathrm{d}s\mathrm{d}y = 2\,\overline{\sigma}^2 \|x - x'\| \int_{\varepsilon^u}^{\infty} y^{-2} \mathrm{d}y \\ &= 2\,\overline{\sigma}^2 \frac{\|x - x'\|}{\varepsilon^u} \le 2\,\overline{\sigma}^2 \delta, \end{aligned}$$

which concludes the proof of the lemma. □

Acknowledgements: The authors thank Yan Fyodorov, Nicola Kistler, Irina Kurkova, and Vincent Vargas for helpful discussions. O.Z. would like to thank the Courant Institute of Mathematical Science and the Université de Montréal for hospitality and financial support.

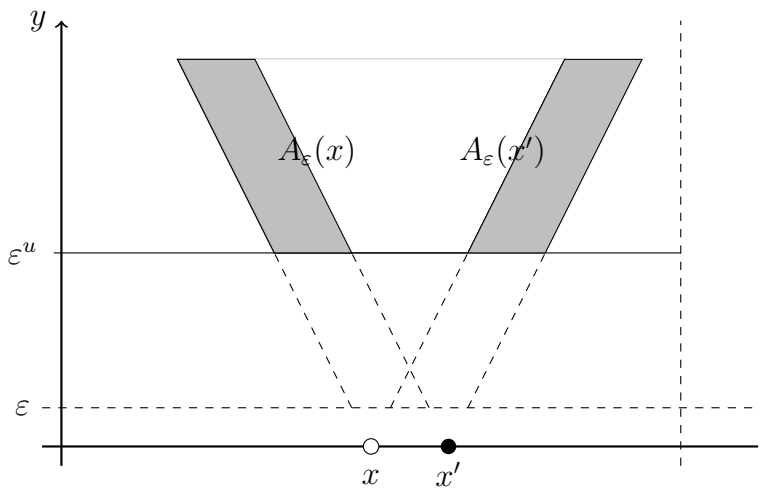

FIGURE 5. The error terms in the tree approximation correspond to the two grey parallelograms in Lemma 5.4.

REFERENCES

[1] Adler, R. J. and Taylor, J. E. (2007). *Random fields and geometry.* Springer Monographs in Mathematics.

[2] Arguin, L.-P. and Chatterjee, S. (2011). Random Overlap Structures: Properties and Applications to Spin Glasses. Preprint. `ArXiv:1011.1823`.

[3] Bacry, E. and Muzy, J.-F. (2003). Log-infinitely divisible multifractal processes. *Comm. Math. Phys.* **236**, 449–475.

[4] Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57**, 33–45.

[5] Bolthausen, E., Deuschel, J.-D. and Giacomin, G. (2001). Entropic repulsion and the maximum of the two-dimensional harmonic crystal. *Ann. Probab.* **29**, 1670–1692.

[6] Bolthausen, E., Deuschel, J.-D. and Zeitouni, O. (2011). Recursions and tightness for the maximum of the discrete, two dimensional Gaussian Free Field. *Elec. Comm. Probab.* **16**, 114–119.

[7] Bolthausen, E. and Kistler, N. (2006). On a nonhierarchical version of the Generalized Random Energy Model. *Ann. Appl. Prob.* **16**, 1–14.

[8] Bolthausen, E. and Kistler, N. (2009). On a nonhierarchical version of the Generalized Random Energy Model II. Ultrametricity. *Stoch. Proc. Appl.* **119**, 2357–2386.

[9] Bolthausen, E. and Sznitman, A.-S. (2002). *Ten Lectures on Random Media.* Birkhaüser

[10] Bovier, A. and Kurkova, I. (2004). Derrida's generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.* **40**, 439–480.

[11] Bovier, A. and Kurkova, I. (2004). Derrida's generalised random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.* **40**, 481–495.

[12] Bramson, M. and Zeitouni, O. (2011). Tightness of the recentered maximum of the two-dimensional discrete Gaussian Free Field. *To appear in Comm. Pure Appl. Math.*

[13] Capocaccia, D., Cassandro, M. and Picco, P. (1987). On the existence of thermodynamics for the generalized random energy model. *J. Statist. Phys.* **46**, 493–505.

[14] Carpentier, D. and Le Doussal, P. (2001). Glass transition for a particle in a random potential, front selection in nonlinear renormalization group, and entropic phenomena in Liouville and Sinh-Gordon models. *Phys. Rev. E* **63**, 026110.

[15] Chauvin, B. and Rouault, A. (1996). Boltzmann-Gibbs weights in the branching random walk. In *Classical and Modern Branching Processes*. K. B. Athreya and P. Jagers, eds. Lecture Notes in Maths. **84**, Springer, 41–50.

[16] Daviaud, O. (2006). Extremes of the discrete two-dimensional Gaussian free field. *Ann. Probab.* **34**, 962–986.

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

[18] Derrida, B. and Spohn, H. (1988). Polymers on disordered trees, spin glasses, and traveling waves. *J. Statist. Phys.* **51**, 817–840.

[19] Ding, J. (2011). Exponential and double exponential tails for maximum of two-dimensional discrete Gaussian free field. Preprint. `ArXiv:1105.5833`.

[20] Dovbysh, L. and Sudakov, V. (1982). Gram-de Finetti matrices. *J. Soviet. Math.* **24**, 3047–3054.

[21] Durrett, R. (2004). *Probability: Theory and Examples* (third edition). Duxbury, Belmont.

[22] Fang, M. and Zeitouni, O. (2011). Branching random walks in time inhomogeneous environments. Preprint. `ArXiv:1112.1113`.

[23] Fyodorov, Y. V. and Bouchaud, J.-P. (2008). Freezing and extreme value statistics in a Random Energy Model with logarithmically correlated potential. *Phys.A: Math. Theor.* **41**, 372001 (12pp).

[24] Fyodorov, Y. V., Le Doussal, P. and Rosso, A. (2009). Statistical mechanics of logarithmic REM: duality, freezing and extreme value statistics of $1/f$ noises generated by Gaussian free fields. *J. Stat. Mech.* P10005 (32pp).

[25] Ghirlanda, S. and Guerra, F. (1998). General Properties of overlap probability distributions in disordered spin systems. *J.Phys. A* **31**, 9149–9155.

[26] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*. Springer-Verlag.

[27] Panchenko, D. (2010). The Ghirlanda-Guerra identities for mixed p-spin model. *C.R. Acad. Sci. Paris* Ser. I **348**, 189–192.

[28] Robert, R. and Vargas, V. (2010). Gaussian multiplicative chaos revisited. *Ann. Probab.* **38**, 605–631.

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

[30] Talagrand, M. (2003). *Spin glasses: a challenge for mathematicians. Cavity and mean field models.* Springer-Verlag.

Département de Mathématiques et Statistique, Université de Montréal, Montréal, H3T 1J4, Canada

E-mail address: `arguinlp@dms.umontreal.ca`

Laboratoire de Probabilités et Modèles Aléatoires, CNRS UMR 7599, Université Paris 6, 4 place Jussieu, 75252 Paris Cedex 05, France

E-mail address: `olivier.zindy@upmc.fr`