



# Extremes of the two-dimensional Gaussian free field with scale-dependent variance

Louis-Pierre Arguin<sup>1</sup> and Frédéric Ouimet<sup>2</sup>

Baruch College and Graduate Center (CUNY), Department of Mathematics  
New York, NY 10010, USA.

*E-mail address:* [louis-pierre.arguin@baruch.cuny.edu](mailto:louis-pierre.arguin@baruch.cuny.edu)

*URL:* <https://arguin.commonscs.gc.cuny.edu>

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

*E-mail address:* [ouimetfr@dms.umontreal.ca](mailto:ouimetfr@dms.umontreal.ca)

*URL:* <https://sites.google.com/site/fouimet26>

**Abstract.** In this paper, we study a random field constructed from the two-dimensional Gaussian free field (GFF) by modifying the variance along the scales in the neighborhood of each point. The construction can be seen as a local martingale transform and is akin to the time-inhomogeneous branching random walk. In the case where the variance takes finitely many values, we compute the first order of the maximum and the log-number of high points. These quantities were obtained by [Bolthausen et al. \(2001\)](#) and [Daviaud \(2006\)](#) when the variance is constant on all scales. The proof relies on a truncated second moment method proposed by [Kistler \(2015\)](#), which streamlines the proof of the previous results. We also discuss possible extensions of the construction to the continuous GFF.

## 1. Introduction

1.1. *The model.* Let  $(W_k)_{k \geq 0}$  be a simple random walk starting at  $u \in \mathbb{Z}^2$  with law  $\mathcal{P}_u$ . For every finite box  $B \subseteq \mathbb{Z}^2$ , the Gaussian free field (GFF) on  $B$  is a centered

---

*Received by the editors September 9, 2015; accepted May 20, 2016.*

2010 *Mathematics Subject Classification.* 60G70, 82B44.

*Key words and phrases.* extreme value theory, Gaussian free field, branching random walk.

<sup>1</sup>L.-P. Arguin is supported by the NSF grant DMS 1513441, the PSC-CUNY Research Award 68784-00 46, a Eugene M. Lang Junior Faculty Research Fellowship, and partially by a NSERC Discovery grant and a FQRNT *Nouveaux chercheurs* grant.

<sup>2</sup>F. Ouimet is supported by a NSERC Doctoral Program Alexander Graham Bell scholarship and partially by a NSERC Master Program Alexander Graham Bell scholarship.

Gaussian field  $\phi \doteq \{\phi_v\}_{v \in B}$  with covariance matrix

$$G_B(u, v) \doteq \frac{\pi}{2} \cdot \mathcal{E}_u \left[ \sum_{k=0}^{\tau_{\partial B}-1} 1_{\{W_k=v\}} \right], \quad u, v \in B, \tag{1.1}$$

where  $\tau_{\partial B}$  is the first hitting time of  $(W_k)_{k \geq 0}$  on the *boundary* of  $B$ ,

$$\partial B \doteq \{v \in B \mid \exists z \notin B \text{ such that } \|v - z\|_2 = 1\},$$

and  $\|\cdot\|_2$  denotes the Euclidean distance in  $\mathbb{Z}^2$ . With this definition,  $B$  contains its boundary. We let  $B^\circ \doteq B \setminus \partial B$ . By convention, summations are zero when there are no indices, so  $\phi$  is identically zero on  $\partial B$ . This is the *Dirichlet boundary condition*. The constant  $\pi/2$  in (1.1) is a convenient normalization for the variance.

In this paper, we consider a family of Gaussian fields constructed from the GFF  $\{\phi_v\}_{v \in V_N}$  on the square box  $V_N \doteq \{0, 1, \dots, N\}^2$ . These Gaussian fields are the analogues, in the context of the GFF, of the time-inhomogeneous branching random walks studied in [Bovier and Kurkova \(2004\)](#); [Fang and Zeitouni \(2012a\)](#); [Bovier and Hartung \(2014\)](#); [Ouimet \(2015\)](#). We study the maxima and the number of high points of this family of Gaussian fields as  $N \rightarrow \infty$ .

The construction is very natural for any Gaussian field on a metric space and bears strong similarities with martingale transforms. It is based on the modification of the variance in neighborhoods around every point along different mesoscopic scales. More precisely, for  $\lambda \in (0, 1)$  and  $v = (v_1, v_2) \in V_N$ , consider the closed neighborhood  $[v]_\lambda$  in  $V_N$  consisting of the square box of width  $N^{1-\lambda}$  centered at  $v$  that has been cut off by the boundary of  $V_N$  :

$$[v]_\lambda \doteq \left( \left[ v_1 - \frac{1}{2}N^{1-\lambda}, v_1 + \frac{1}{2}N^{1-\lambda} \right] \times \left[ v_2 - \frac{1}{2}N^{1-\lambda}, v_2 + \frac{1}{2}N^{1-\lambda} \right] \right) \cap V_N .$$

By convention, we define  $[v]_0 \doteq V_N$  and  $[v]_1 \doteq \{v\}$ . We stress that square boxes are not essential to the construction; any neighborhood centered at  $v$  containing points at distance roughly  $N^{1-\lambda}$  would do. Let  $\mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c} \doteq \sigma(\{\phi_v, v \notin [v]_\lambda^c\})$  be the  $\sigma$ -algebra generated by the variables on the boundary of the box  $[v]_\lambda$  and those outside of it. Since the neighborhoods are shrinking with  $\lambda$ , for any  $v \in V_N$ , the collection  $\mathbb{F}_v \doteq \{\mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}\}_{\lambda \in [0,1]}$  is a filtration. In particular, if we let

$$\phi_v(\lambda) \doteq \mathbb{E} \left[ \phi_v \mid \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c} \right],$$

then

for every  $v \in V_N$ ,  $(\phi_v(\lambda))_{\lambda \in [0,1]}$  is a  $\mathbb{F}_v$ -martingale.

It is also a Gaussian field, therefore disjoint increments of the form  $\phi_v(\lambda') - \phi_v(\lambda)$  are independent. These observations motivate the definition of *scale-inhomogeneous Gaussian free field*, which can be seen as a martingale-transform of  $(\phi_v(\lambda))_{\lambda \in [0,1]}$  applied simultaneously for every  $v \in V_N$ .

Fix  $M \in \mathbb{N}$  and consider the parameters

$$\boldsymbol{\sigma} \doteq (\sigma_1, \sigma_2, \dots, \sigma_M) \in (0, \infty)^M, \tag{variance parameters}$$

$$\boldsymbol{\lambda} \doteq (\lambda_1, \lambda_2, \dots, \lambda_M) \in (0, 1]^M, \tag{scale parameters}$$

where  $0 \doteq \lambda_0 < \lambda_1 < \dots < \lambda_M \doteq 1$ . The parameters  $(\sigma, \lambda)$  can be encoded simultaneously in the left-continuous step function

$$\sigma(s) \doteq \sigma_1 1_{\{0\}}(s) + \sum_{i=1}^M \sigma_i 1_{(\lambda_{i-1}, \lambda_i]}(s), \quad s \in [0, 1].$$

We write  $\nabla_i$  for the difference operator with respect to the index  $i$ . When the index variable is obvious, we omit the subscript. For example,

$$\nabla \phi_v(\lambda_i) \doteq \phi_v(\lambda_i) - \phi_v(\lambda_{i-1}).$$

**Definition 1.1** (Scale-inhomogeneous Gaussian free field). Let  $\phi \doteq \{\phi_v\}_{v \in V_N}$  be the GFF on  $V_N$ . The  $(\sigma, \lambda)$ -GFF on  $V_N$  is a Gaussian field  $\psi \doteq \{\psi_v\}_{v \in V_N}$  defined by

$$\psi_v \doteq \sum_{i=1}^M \sigma_i \nabla \phi_v(\lambda_i) = \sum_{i=1}^M \sigma_i (\phi_v(\lambda_i) - \phi_v(\lambda_{i-1})). \tag{1.2}$$

Similarly to the GFF, we define

$$\psi_v(\lambda) \doteq \mathbb{E} \left[ \psi_v \mid \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c} \right].$$

The field with two variances ( $M = 2$ ) was presented in [Arguin and Zindy \(2015\)](#), where it was used to prove Poisson-Dirichlet statistics of the Gibbs measure in the homogeneous case ( $M = 1$ ).

1.2. *Main results.* The main results of this paper are the derivation of the first order of the maximum and the log-number of high points for the scale-inhomogeneous Gaussian free field of Definition 1.1. The methods of proof are general and directly applicable to time-inhomogeneous branching random walks and to other log-correlated Gaussian fields.

First, we need to introduce some notations. For any positive measurable function  $f : [0, 1] \rightarrow \mathbb{R}$ , define the integral operators

$$\mathcal{J}_f(s) \doteq \int_0^s f(r) dr \quad \text{and} \quad \mathcal{J}_f(s_1, s_2) \doteq \int_{s_1}^{s_2} f(r) dr.$$

It turns out that the first order of the maximum and the log-number of high points are controlled by the *concavification* of  $\mathcal{J}_{\sigma^2}(\cdot)$ . Let  $\hat{\mathcal{J}}_{\sigma^2}$  be the function whose graph is the concave hull of  $\mathcal{J}_{\sigma^2}$ . Its graph is an increasing and concave polygonal line, see Figure 1.1 for an example. There exists a unique non-increasing left-continuous step function  $s \mapsto \bar{\sigma}(s)$  such that

$$\hat{\mathcal{J}}_{\sigma^2}(s) = \mathcal{J}_{\bar{\sigma}^2}(s) = \int_0^s \bar{\sigma}^2(r) dr \quad \text{for all } s \in (0, 1].$$

The points on  $[0, 1]$  where  $\bar{\sigma}$  jumps will be denoted by

$$0 \doteq \lambda^0 < \lambda^1 < \dots < \lambda^m \doteq 1, \tag{1.3}$$

where  $m \leq M$ . To be consistent with previous notations, we set  $\bar{\sigma}_l \doteq \bar{\sigma}(\lambda^l)$ .

**Theorem 1.2** (First order of the maximum). *Let  $\{\psi_v\}_{v \in V_N}$  be the  $(\sigma, \lambda)$ -GFF on  $V_N$  of Definition 1.1, then*

$$\lim_{N \rightarrow \infty} \frac{\max_{v \in V_N} \psi_v}{\log N^2} = \mathcal{J}_{\bar{\sigma}^2/\bar{\sigma}}(1) \doteq \gamma^* \quad \text{in probability.}$$

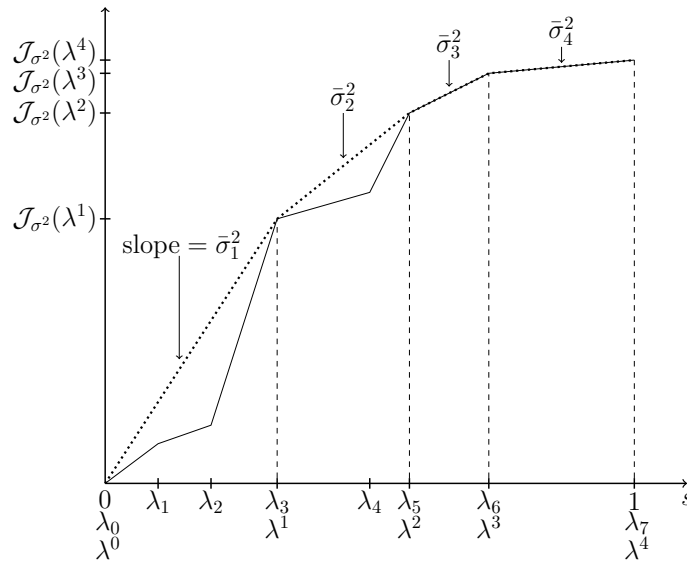


FIGURE 1.1. Example of  $\mathcal{J}_{\sigma^2}$  (closed line) and  $\hat{\mathcal{J}}_{\sigma^2}$  (dotted line) with 7 values for  $\sigma^2$ .

In the homogeneous case where  $M = 1$  and  $\sigma_1 = 1$ , the result reduces to  $\gamma^* = 1$ , as proved in Bolthausen et al. (2001), which corresponds to the first order of the maximum of  $N^2$  i.i.d. Gaussian variables of mean 0 and variance  $\log N$ . Note that the result of Theorem 1.2 can be written as follows :

$$\gamma^* = \mathcal{J}_{\sigma^2/\bar{\sigma}}(1) = \sum_{l=1}^m \int_{\lambda^{l-1}}^{\lambda^l} \frac{\sigma^2(s)}{\bar{\sigma}(s)} ds = \int_0^1 \bar{\sigma}(s) ds . \tag{1.4}$$

This is simply a weighted average of homogeneous cases on the intervals  $[\lambda^{l-1}, \lambda^l]$  with variance parameter  $\bar{\sigma}_l$ . We say that  $s \mapsto \bar{\sigma}^2(s)$  act as the *effective variance* of the field. We stress that  $\gamma^*$  is strictly smaller than  $\bar{\sigma}_1$  in cases where the concave hull is not a straight line. In particular, the upper bound on the level of the maximum cannot be proved by a simple union bound as in the homogeneous case.

The set of  $\gamma$ -high points of the field  $\psi$  is defined as

$$\mathcal{H}_N^\gamma \doteq \{v \in V_N \mid \psi_v \geq \gamma \log N^2\}, \text{ for all } 0 \leq \gamma < \gamma^* .$$

The number of high points will depend on *critical levels* defined by

$$\gamma^l \doteq \int_0^1 \frac{\sigma^2(s)}{\bar{\sigma}(s \wedge \lambda^l)} ds = \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^l) + \frac{\mathcal{J}_{\sigma^2}(\lambda^l, 1)}{\bar{\sigma}_l}, \quad 1 \leq l \leq m, \quad \gamma^0 \doteq 0. \tag{1.5}$$

**Theorem 1.3** (Log-number of high points or Entropy). *Let  $\{\psi_v\}_{v \in V_N}$  be the  $(\sigma, \lambda)$ -GFF on  $V_N$  of Definition 1.1 and let  $\gamma^{\lambda^{-1}} \leq \gamma < \gamma^l$  for some  $l \in \{1, \dots, m\}$ , then*

$$\lim_{N \rightarrow \infty} \frac{\log |\mathcal{H}_N^\gamma|}{\log N^2} = (1 - \lambda^{l-1}) - \frac{(\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))^2}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} \doteq \mathcal{E}_\gamma \quad \text{in probability.}$$

The homogeneous case where  $M = 1$  and  $\sigma_1 = 1$  was proved in Daviaud (2006). In that case, we have  $\mathcal{E}_\gamma = 1 - \gamma^2$  as for  $N^2$  i.i.d. Gaussian variables of mean 0 and variance  $\log N$ . The proofs of Theorems 1.2 and 1.3 are deferred to Section 3.

The method of proof is explained in Section 2. It is a refinement of the second moment method based on the control of the increments of high points at every scale. The method was used in Kistler (2015) to obtain a new proof of the first order of the maximum in the homogeneous case. Here we extend this method to the log-number of high points in all settings and to the first order of the maximum in the inhomogeneous setting. In the scale-dependent case, as opposed to the homogeneous case, it is necessary to truncate the first moment using the information at every scale  $\lambda^l$  to get the correct upper bound.

1.3. *Related works and conjectures.* The scale-inhomogeneous GFF is the equivalent of the time-inhomogeneous branching random walk (IBRW) where the variance of the random walk is a function of time. In particular, Theorems 1.2 and 1.3 can be proved for branching random walks using the same technique, see Section 2 of Ouimet (2014). In fact, much more precise information is known about the maxima of these models. In Bovier and Kurkova (2004), the authors introduce a continuous version of Derrida's *Generalized Random Energy Model* (GREM) Derrida (1985), which is akin to a time-inhomogeneous branching random walk, for which they obtain the first order of the maximum and the free energy. In particular, they noticed the concavification phenomenon for the first order. This observation also appears in Capocaccia et al. (1987) for the GREM. A model interpolating between the GREM and the branching random walk was introduced in Kistler and Schmidt (2015) where Poisson statistics of the extremes are proved. For Gaussian IBRWs with two values of the variance ( $M = 2$ ), the lower order corrections for the maximum and tightness of the law were proved in Fang and Zeitouni (2012a). In this case, convergence of the extremal processes and of the law of the recentered maximum have been shown in Bovier and Hartung (2014). This is also proved in the case where the integral of the variance remains strictly below its concave hull (for example, in the case of increasing variances), see Bovier and Hartung (2015). For strictly decreasing variances, the lower order corrections for IBRWs exhibit a slowdown of the order  $t^{1/3}$  as proved in Fang and Zeitouni (2012b); Maillard and Zeitouni (2016). Similar results for non-Gaussian IBRWs and more general variances are proved in Mallein (2015), though not at the level of convergence of the law. In Ouimet (2015), the second order of the maximum for the Gaussian IBRW with a finite number of variances is shown by generalizing the approach of Fang and Zeitouni (2012a) and the tightness follows from Fang (2012).

In general, we expect that the scale-inhomogeneous GFF with a finite number of variances behave as the time-inhomogeneous branching random walk with the same parameters for the lower order correction term of the maximum and for its law. For the homogeneous GFF, the convergence of the law of the recentered maximum was proved in Bramson et al. (2016). In Arguin and Zindy (2015), the scale-inhomogeneous GFF with two values of the variance was introduced to prove Poisson-Dirichlet statistics for the extremes of the homogeneous GFF. Actual Poisson statistics for local extremes was proved later in Biskup and Louidor (2016).

One interest of Definition 1.1 for the scale-inhomogeneous GFF is that it can be extended to a piecewise smooth variance function  $\sigma : [0, 1] \rightarrow [a, b]$  where  $a > 0$ . Consider the two-dimensional continuous Gaussian free field  $\phi = \{\phi_v\}_{v \in [0, 1]^2}$  on the unit square  $[0, 1]^2$ , see e.g. Sheffield (2007) for a definition. The field  $\phi$  cannot be defined as a random function. However, averages over sets make sense as random

variables. In particular, for every  $v \in [0, 1]^2$  and  $\lambda \in [0, 1]$ , one can define  $\phi_v^r(\lambda)$  as the average of the field over a circle of radius  $r^\lambda$  :

$$\phi_v^r(\lambda) \doteq \frac{1}{2\pi r^\lambda} \int_0^{2\pi} \phi_{v+r^\lambda e^{i\theta}} d\theta. \quad (1.6)$$

The parameter  $r$  plays the role of  $N^{-1}$  in the discrete setting. The continuous scale-inhomogeneous GFF for the variance function  $\lambda \mapsto \sigma(\lambda)$  can then be defined in terms of these averages :

$$\psi_v^r(1) \doteq \int_0^1 \sigma(\lambda) d\phi_v^r(\lambda), \quad v \in [0, 1]^2.$$

The stochastic integral makes sense because  $(\phi_v^r(\lambda))_{\lambda \in [0, 1]}$  is a Gaussian martingale. Following the definition in Duplantier and Sheffield (2011) (up to a factor 2), a point  $v \in [0, 1]^2$  is called  $\gamma$ -thick if

$$\lim_{r \rightarrow 0} \frac{\psi_v^r(1)}{\log(r^{-2})} \geq \gamma$$

where it is assumed that the continuous Green function on  $[0, 1]^2$  associated to  $\phi$  has been normalized as in (1.1). This is the notion analogous to  $\gamma$ -high points. It was shown in Hu et al. (2010) that the Hausdorff dimension of the set of  $\gamma$ -thick points is  $2(1 - \gamma^2)$  when  $\sigma \equiv 1$ . In view of Theorem 1.3, it is reasonable to conjecture that the Hausdorff dimension of the set of  $\gamma$ -thick points of  $\psi$  is

$$2 \left( (1 - \lambda_\star) - \frac{(\gamma - \mathcal{J}_{\sigma^2/\sigma}(\lambda_\star))^2}{\mathcal{J}_{\sigma^2}(\lambda_\star, 1)} \right), \quad (1.7)$$

where  $\lambda_\star \doteq \inf\{\lambda \in [0, 1] : \gamma \leq \int_0^1 \frac{\sigma^2(s)}{\sigma(s\wedge\lambda)} ds\}$ .

## 2. Outline of Proof

As stated before, the results of this paper are applicable to time-inhomogeneous branching random walks and, more generally, to any scale-dependent log-correlated Gaussian field. The proof relies on two main ingredients: an underlying approximate tree structure present in log-correlated models and an adaptation of the multiscale refinement of the second moment method introduced in Kistler (2015). In particular, the method requires understanding the increments of high points along every scale to prove tight upper and lower bounds. In Kistler (2015), this method was used to streamline the proof of Bolthausen et al. (2001) for the first order of the maximum of the homogeneous GFF. Here, we adapt the method to deal with scale-inhomogeneous fields and log-number of high points.

To see the tree structure, define the *branching scale between  $v$  and  $v'$*  in  $V_N$  :

$$\rho(v, v') \doteq \max\{\lambda \in [0, 1] : [v]_\lambda \cap [v']_\lambda \neq \emptyset\}. \quad (2.1)$$

This is the largest  $\lambda$  for which the two neighborhoods  $[v]_\lambda$  and  $[v']_\lambda$  intersect. We always have by definition that  $\|v - v'\|_2$  is of order  $N^{1-\rho(v, v')}$ . The branching scale plays the same role as the branching time (normalized to lie in  $[0, 1]$ ) in branching random walk. More precisely, let  $\{\phi_v\}_{v \in V_N}$  be a homogeneous GFF and consider the increments  $\phi_v(\lambda') - \phi_v(\lambda)$  and  $\phi_{v'}(\mu') - \phi_{v'}(\mu)$  for some choice of  $\lambda < \lambda'$  and  $\mu < \mu'$ . The Markov property of the Gaussian free field (see Section A) implies that for  $\lambda, \mu > \rho(v, v')$ ,

$$\phi_v(\lambda') - \phi_v(\lambda) \quad \text{is independent of} \quad \phi_{v'}(\mu') - \phi_{v'}(\mu),$$

because the neighborhoods  $[v]_\lambda$  and  $[v']_{\mu'}$  are disjoint, see Figure 2.2. This means that the increments after the branching scale are independent.

On the other hand, if  $\lambda < \rho$ , it can be shown using Green function estimates (see e.g. Lemma 12 in Bolthausen et al., 2001) that

$$\mathbb{V}(\phi_v(\lambda) - \phi_{v'}(\lambda)) = O(1) .$$

In other words, the values of  $\phi_v(\lambda)$  and  $\phi_{v'}(\lambda)$  must be close. This suggests that the increments before the branching scale are almost identical. In particular, without losing much information, we can restrict the field  $\{\phi_v(\lambda)\}_{v \in V_N}$  to a set  $R_\lambda \subseteq V_N$  containing  $\lfloor N^\lambda \rfloor^2$   $v$ 's with neighborhoods  $[v]_\lambda$  that can only touch at their boundary and are not cut off by  $\partial V_N$ . To remove any ambiguity, define  $R_\lambda$  in such a way that  $\max_{v \in V_N} \min_{z \in R_\lambda} \|v - z\|_2$  is minimum. We call  $R_\lambda$  the *set of representatives at scale  $\lambda$*  and define  $R_1 \doteq V_N$ . For instance, if  $N = 2^n$ ,  $\lambda \in [0, 1)$  and  $\lambda n \in \mathbb{N}$ , then divide  $V_N$  into a grid with  $N^{2\lambda}$  squares of side length  $N^{1-\lambda}$ , the center point of each square is a representative at scale  $\lambda$ .

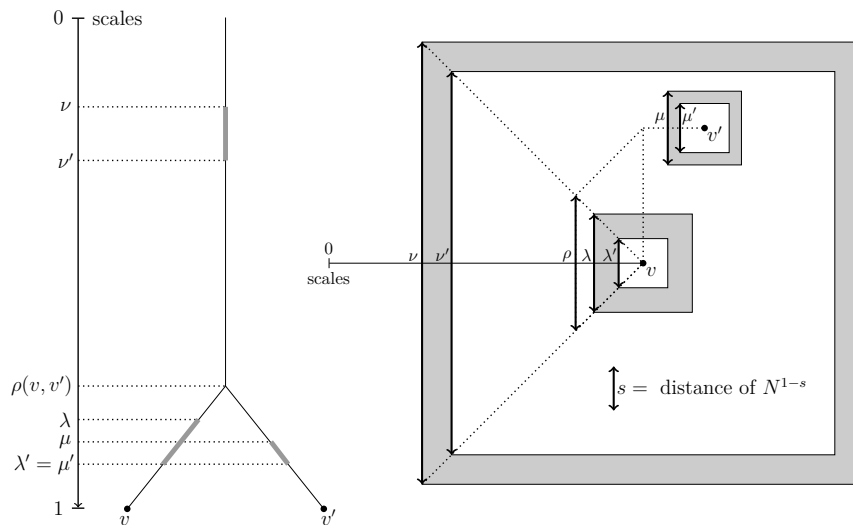


FIGURE 2.2. The branching structure of the GFF.

Of course, the branching structure here is not exact as in branching random walk. In particular, nothing precise can be said on the increments  $\phi_v(\lambda') - \phi_v(\lambda)$  and  $\phi_{v'}(\lambda') - \phi_{v'}(\lambda)$  in the case where  $\lambda < \rho < \lambda'$ . However, the contribution of such increments can be made negligible by considering a large number of increments, as we shall do. This branching structure holds also for the  $(\sigma, \lambda)$ -GFF, since it is defined in terms of the increments of  $\phi$ , see (1.2) and Lemma A.1.

For  $0 < \gamma < \gamma^*$ , the  $\gamma$ -high points are such that  $\psi_v \geq \gamma \log N^2$ . It is reasonable to expect that for these points, there exists a unique optimal path  $\lambda \mapsto L_N^\gamma(\lambda)$  such that  $\psi_v(\lambda) \geq L_N^\gamma(\lambda)$  at each scale  $\lambda$ . We write  $L_N^*$  for the corresponding optimal path in the case of the maximum level  $\gamma^*$ . It is the information on these paths along the scales that is crucial for the method to yield tight upper and lower bounds. We explain heuristically how to determine these optimal paths using first moments.

Consider the set of  $v$ 's for which the increments of the field  $\psi$  reach level  $\nabla\gamma_i$  between each scale  $\lambda_i$  :

$$\Lambda_{N,M} \doteq \{v \in V_N \mid \nabla\psi_v(\lambda_i) \geq \nabla\gamma_i \log N^2 \text{ for all } i \in \{1, 2, \dots, M\}\},$$

where  $\gamma_0 \doteq 0$ . By construction,  $|\Lambda_{N,M}|$  is a lower bound on the number of points in  $V_N$  reaching a height of  $\gamma_M \log N^2$ . We also consider the corresponding quantity at intermediate scales  $\lambda_k < \lambda_M$ . In this case, because of correlations, we can restrict ourselves to representatives at scale  $\lambda_k$  :

$$\Lambda_{N,k} \doteq \{v \in R_{\lambda_k} \mid \nabla\psi_v(\lambda_i) \geq \nabla\gamma_i \log N^2 \text{ for all } i \in \{1, 2, \dots, k\}\}.$$

There are  $O(N^{2\lambda_k})$  representatives at scale  $\lambda_k$  and the variance of the increments is  $\mathbb{V}(\nabla\psi_v(\lambda_i)) = \sigma_i^2 \nabla\lambda_i \log N + O(1)$  if we ignore the boundary effect. Therefore, using the independence between the increments and standard Gaussian estimates (see Lemma A.7, it will be used repeatedly) :

$$\mathbb{E}[|\Lambda_{N,k}|] \asymp N^{2\lambda_k} \prod_{i=1}^k \mathbb{P}(\nabla\psi_v(\lambda_i) \geq 2\nabla\gamma_i \log N) \asymp \frac{N^{2\lambda_k} N^{-2\sum_{i=1}^k \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i}}}{(\log N)^{k/2}},$$

where  $\asymp$  means that the ratio of the two sides lies in a compact interval bounded away from 0, for  $N$  large enough. In other words,

$$\lim_{N \rightarrow \infty} \frac{\log(\mathbb{E}[|\Lambda_{N,k}|])}{\log N^2} = \sum_{i=1}^k \left( \nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i} \right).$$

Since there should be representatives at each scale  $\lambda_k$  that ultimately yield a high value at scale  $\lambda_M$ , it is intuitive that the level of the maximum can be found by maximizing

$$\gamma_M = \sum_{i=1}^M \nabla\gamma_i \text{ under the constraints } \sum_{i=1}^k \left( \nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i} \right) \geq 0, \quad 1 \leq k \leq M.$$

This optimization problem can be solved using the Karush-Kuhn-Tucker theorem (see Lemma B.2). We write  $(\gamma_1^*, \gamma_2^*, \dots, \gamma_M^*)$  for the unique solution. We will make extensive use of the polygonal line  $L_N^*(\cdot)$  linking the points  $(0, 0)$ ,  $(\lambda_1, \gamma_1^* \log N^2)$ ,  $(\lambda_2, \gamma_2^* \log N^2)$ ,  $\dots$ ,  $(\lambda_M, \gamma_M^* \log N^2)$  to prove Theorem 1.2 and 1.3 :

$$L_N^*(s) \doteq \int_0^s \frac{\sigma^2(r)}{\bar{\sigma}(r)} dr \log N^2 = \mathcal{J}_{\sigma^2/\bar{\sigma}}(s) \log N^2, \quad s \in [0, 1]. \tag{2.2}$$

This is the *optimal path for the maximum*. Figure 2.3 shows an example of such a path. In particular, it is important to note that the optimal path coincides with its concave hull at each scale  $\lambda^l$ , namely

$$L_N^*(\lambda^l) = \hat{L}_N^*(\lambda^l) = \hat{\mathcal{J}}_{\sigma^2/\bar{\sigma}}(\lambda^l) \log N^2 = \mathcal{J}_{\bar{\sigma}}(\lambda^l) \log N^2, \quad 1 \leq l \leq m. \tag{2.3}$$

The same heuristic can be used to determine the optimal path  $L_N^\gamma(\cdot)$  for  $\gamma$ -high points,  $0 < \gamma < \gamma^*$ . Setting now  $\gamma_M = \gamma$ , we get

$$\lim_{N \rightarrow \infty} \frac{\log(\mathbb{E}[|\Lambda_{N,M}|])}{\log N^2} = \sum_{i=1}^{M-1} \left( \nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i} \right) + \left( \nabla\lambda_M - \frac{(\gamma - \gamma_{M-1})^2}{\sigma_M^2 \nabla\lambda_M} \right). \tag{2.4}$$



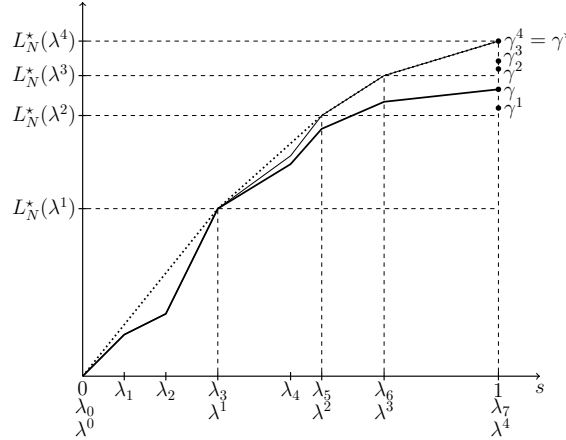


FIGURE 2.3. Example of  $L_N^\gamma$  (bold line),  $L_N^*$  (thin line) and its concavified version  $\hat{L}_N^*$  (dotted line), with 7 values for  $\sigma^2$  and  $\gamma^1 < \gamma < \gamma^2$ .

A lower bound for the log-number of  $\gamma$ -high points can be found by maximizing (2.4) with respect to  $\gamma_1, \gamma_2, \dots, \gamma_{M-1}$  and under the constraints

$$\sum_{i=1}^k \left( \nabla \lambda_i - \frac{(\nabla \gamma_i)^2}{\sigma_i^2 \nabla \lambda_i} \right) \geq 0, \quad 1 \leq k \leq M - 1. \tag{2.5}$$

The unique solution to this problem is found in Lemma B.3 using again the Karush-Kuhn-Tucker theorem. The form of the path will always depend on the critical levels defined in (1.5). Whenever  $\gamma^{l-1} \leq \gamma < \gamma^l$ , the optimal path for  $\gamma$ -high points is :

$$L_N^\gamma(s) \doteq \begin{cases} \mathcal{J}_{\sigma^2/\bar{\sigma}}(s) \log N^2, & 0 \leq s \leq \lambda^{l-1} \\ \left( \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}) + \frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, s)}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} (\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1})) \right) \log N^2, & \lambda^{l-1} \leq s \leq 1. \end{cases} \tag{2.6}$$

The path coincide on  $[0, \lambda^{l-1}]$  with the optimal path for the maximum. Also, note that  $L_N^\gamma$  is continuous and converges uniformly to  $L_N^*$  as  $\gamma \rightarrow \gamma^*$  (which yields that  $L_N^*$  is continuous as well).

### 3. Proofs of the main results

3.1. Preliminaries. For all  $\lambda \in [0, 1]$ , recall that  $\psi_v(\lambda) \doteq \mathbb{E}[\psi_v \mid \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^\zeta}]$ . By the Markov property of the GFF (see Lemma A.1), it is not hard to show that for any partition  $0 \doteq s_0 < s_1 < \dots < s_K \doteq 1$  of  $[0, 1]$  such that  $\{\lambda_i\}_{i=0}^M \subseteq \{s_j\}_{j=0}^K$ , we have for all  $1 \leq k \leq l \leq K$  :

$$\psi_v(s_l) - \psi_v(s_{k-1}) = \sum_{j=k}^l \sigma(s_j) \nabla \phi_v(s_j) .$$

In particular, the independence of the increments of  $\psi$  follows directly from the one for  $\phi$ . Moreover, using standard estimates on Green functions, Lemma A.2 shows that

$$-C_1(\delta) \leq \mathbb{V}(\psi_v(s_l) - \psi_v(s_{k-1})) - \mathcal{J}_{\sigma^2}(s_{k-1}, s_l) \log N \leq C_2 \tag{3.1}$$

for all  $v \in V_N^\delta$  and  $N$  large enough (depending on  $\delta$ ), where

$$V_N^\delta \doteq \{v \in V_N \mid \min_{z \in \partial V_N} \|v - z\|_2 \geq \delta N\}, \quad \delta \in (0, 1/2].$$

The set  $V_N^\delta$  contains the points that are at a distance at least  $\delta N$  from the boundary of  $V_N$ . Lemma A.3 proves that the upper bound in (3.1) holds on  $V_N$ , that is

$$\max_{v \in V_N} \mathbb{V}(\psi_v(s_l) - \psi_v(s_{k-1})) \leq \mathcal{J}_{\sigma^2}(s_{k-1}, s_l) \log N + C \tag{3.2}$$

for  $N$  large enough.

Throughout the proofs,  $c$  and  $C$  will denote positive constants whose value can change at different occurrences and might depend on the parameters  $(\sigma, \lambda)$ . For simplicity, equations in the proofs are implicitly stated to hold for  $N$  large enough where it is needed.

3.2. *First order of the maximum.* Theorem 1.2 is a direct consequence of Lemma 3.1, which proves that  $\gamma^* \log N^2$  is an upper bound on the first order of the maximum, and Lemma 3.3 which shows the corresponding lower bound.

**Lemma 3.1** (Upper bound on the first order of the maximum). *Let  $\{\psi_v\}_{v \in V_N}$  be the  $(\sigma, \lambda)$ -GFF on  $V_N$  of Definition 1.1 and  $\gamma^*$  as in Theorem 1.2. For all  $\varepsilon > 0$ , there exists a constant  $c = c(\varepsilon, \sigma, \lambda) > 0$  such that*

$$\mathbb{P} \left( \max_{v \in V_N} \psi_v \geq (\gamma^* + m\varepsilon) \log N^2 \right) \leq N^{-c} \tag{3.3}$$

for  $N$  large enough.

*Proof:* Recall the definition of the optimal path  $L_N^*$  from (2.2) and define

$$L_N^{*,z}(s) \doteq L_N^*(s) + z \log N^2, \quad s \in [0, 1].$$

Recall the definition of  $\lambda^j$  in (1.3) and the notation  $R_{\lambda^j}$  for the set of representatives at scale  $\lambda^j$ . Consider the set of representatives whose value reached just over the optimal level at  $\lambda^j$  :

$$\mathcal{H}_{N,j}^{*,\varepsilon} \doteq \left\{ v \in R_{\lambda^j} \mid \psi_v(\lambda^j) \geq L_N^{*,j\varepsilon}(\lambda^j) \right\}, \quad 1 \leq j \leq m.$$

The idea of the proof is to split the probability that at least one point in  $V_N$  reaches just over the optimal height by looking at the first scale  $\lambda^j$ ,  $1 \leq j \leq m$ , where the set  $\mathcal{H}_{N,j}^{*,\varepsilon}$  is not empty. This provides the appropriate constraints along the scales to get the correct upper bound. For  $0 < \eta_\varepsilon < \varepsilon/m$ , define

$$A_\varepsilon \doteq \left\{ |\psi_v(\lambda^j) - \psi_{v_{\lambda^j}}(\lambda^j)| \leq \eta_\varepsilon \log N^2 \text{ for all } j \in \{1, \dots, m\} \text{ and all } v \in V_N \right\}$$

where  $v_\lambda$  denotes any representative in  $R_\lambda$  that is closest to  $v$ . Here we introduced the event  $A_\varepsilon$  to approximate the branching structure of the field  $\psi$ . Since  $R_{\lambda^m} = V_N$  by definition and  $L_N^{*,m\varepsilon}(\lambda^m) = (\gamma^* + m\varepsilon) \log N^2$ , a union bound gives the following

upper bound on the probability in (3.3) :

$$\begin{aligned}
 \mathbb{P}\left(|\mathcal{H}_{N,m}^{*,\varepsilon}| \geq 1\right) &\leq \mathbb{P}(A_\varepsilon^c) + \sum_{l=1}^m \mathbb{P}\left(\left\{|\mathcal{H}_{N,1}^{*,\varepsilon}| = \dots = |\mathcal{H}_{N,l-1}^{*,\varepsilon}| = 0, |\mathcal{H}_{N,l}^{*,\varepsilon}| \geq 1\right\} \cap A_\varepsilon\right) \\
 &\leq \mathbb{P}(A_\varepsilon^c) + \sum_{l=1}^m \mathbb{P}\left(\left\{\begin{array}{l} \exists v \in R_{\lambda^l} \cap V_N^o \text{ s.t.} \\ \psi_v(\lambda^l) \geq L_N^{*,l\varepsilon}(\lambda^l) \text{ and} \\ \psi_{v_{\lambda^j}}(\lambda^j) < L_N^{*,j\varepsilon}(\lambda^j) \\ \text{for all } 1 \leq j \leq l-1 \end{array}\right\} \cap A_\varepsilon\right) \\
 &\leq C e^{-c(\eta_\varepsilon)(\log N)^2} \\
 &\quad + \sum_{l=1}^m N^{2\lambda^l} \max_{v \in R_{\lambda^l} \cap V_N^o} \mathbb{P}\left(\left\{\begin{array}{l} \psi_v(\lambda^l) \geq L_N^{*,l\varepsilon-\eta_\varepsilon}(\lambda^l) \text{ and} \\ \psi_v(\lambda^j) < L_N^{*,j(\varepsilon+\eta_\varepsilon)}(\lambda^j) \\ \text{for all } 1 \leq j \leq l-1 \end{array}\right\}\right) \quad (3.4)
 \end{aligned}$$

The bound on  $\mathbb{P}(A_\varepsilon^c)$  follows easily from a union bound (with  $m \cdot (N + 1)^2$  terms), Gaussian estimates (Lemma A.7) and the variance estimates of Lemma A.6.

It remains to consider the terms in the sum in (3.4). We look at the case  $l = 1$ . Since  $\max_{v \in V_N} \mathbb{V}(\psi_v(\lambda^1)) \leq \lambda^1 \bar{\sigma}_1^2 \log N + C$  from (3.2) and  $L_N^*(\lambda^1) = \lambda^1 \bar{\sigma}_1 \log N^2$ , a Gaussian estimate shows that

$$\begin{aligned}
 \mathbb{P}\left(\psi_v(\lambda^1) \geq L_N^{*,\varepsilon-\eta_\varepsilon}(\lambda^1)\right) &\leq \frac{\sqrt{\mathbb{V}(\psi_v(\lambda^1))}}{L_N^{*,\varepsilon-\eta_\varepsilon}(\lambda^1)} \exp\left(-\frac{(L_N^{*,\varepsilon-\eta_\varepsilon}(\lambda^1))^2}{2\mathbb{V}(\psi_v(\lambda^1))}\right) \\
 &\leq \frac{C}{\sqrt{\log N}} N^{-2\lambda^1} N^{-4\frac{(\varepsilon-\eta_\varepsilon)}{\sigma_1}}.
 \end{aligned}$$

After multiplying by  $N^{2\lambda^1}$ , we conclude that the  $l = 1$  term in (3.4) goes to 0 like  $N^{-c(\varepsilon)}$ . We now show a similar estimate for a fixed  $l \in \{2, \dots, m\}$ . To simplify the notation, denote  $(X_v^1, \dots, X_v^l) \doteq (\psi_v(\lambda^1), \dots, \psi_v(\lambda^l))$ . By conditioning on the value of the vector  $\mathbf{X} \doteq (X_v^1, \dots, X_v^{l-1})$ , the probability in (3.4) is equal to

$$\int_{-\infty}^{L_N^{*,1(\varepsilon+\eta_\varepsilon)}(\lambda^1)} \dots \int_{-\infty}^{L_N^{*,(l-1)(\varepsilon+\eta_\varepsilon)}(\lambda^{l-1})} \mathbb{P}\left(X_v^l \geq L_N^{*,l\varepsilon-\eta_\varepsilon}(\lambda^l) \mid \mathbf{X} = \mathbf{x}\right) f_v(\mathbf{x}) \, d\mathbf{x}$$

where  $f_v$  is the density function of  $\mathbf{X}$ . By independence of the increments, the last integral is equal to

$$\int_{-\infty}^{L_N^{*,1(\varepsilon+\eta_\varepsilon)}(\lambda^1)} \dots \int_{-\infty}^{L_N^{*,(l-1)(\varepsilon+\eta_\varepsilon)}(\lambda^{l-1})} \mathbb{P}\left(\nabla X_v^l \geq L_N^{*,l\varepsilon-\eta_\varepsilon}(\lambda^l) - x_{l-1}\right) f_v(\mathbf{x}) \, d\mathbf{x} \quad (3.5)$$

Since  $l\varepsilon - \eta_\varepsilon = (\varepsilon - l\eta_\varepsilon) + (l - 1)(\varepsilon + \eta_\varepsilon)$ , a Gaussian estimate and the bound  $\max_{v \in V_N} \mathbb{V}(\nabla X_v^l) \leq \bar{\sigma}_l^2 \nabla \lambda^l \log N + C$  from (3.2) give

$$\begin{aligned} & \mathbb{P}\left(\nabla X_v^l \geq L_N^{*,l\varepsilon-\eta_\varepsilon}(\lambda^l) - x_{l-1}\right) \\ & \leq \frac{\sqrt{\mathbb{V}(\nabla X_v^l)}}{L_N^{*,l\varepsilon-\eta_\varepsilon}(\lambda^l) - x_{l-1}} \exp\left(\frac{-(\nabla L_N^*(\lambda^l) + L_N^{*,l\varepsilon-\eta_\varepsilon}(\lambda^{l-1}) - x_{l-1})^2}{2\mathbb{V}(\nabla X_v^l)}\right) \\ & \leq \frac{C}{\sqrt{\log N}} N^{-2\nabla \lambda^l} \exp\left(-2\frac{(L_N^{*,l\varepsilon-\eta_\varepsilon}(\lambda^{l-1}) - x_{l-1})}{\bar{\sigma}_l}\right) \\ & = \frac{C}{\sqrt{\log N}} N^{-2\nabla \lambda^l} N^{-4\frac{\varepsilon-l\eta_\varepsilon}{\bar{\sigma}_l}} \exp\left(-2\frac{(L_N^{*(l-1)(\varepsilon+\eta_\varepsilon)}(\lambda^{l-1}) - x_{l-1})}{\bar{\sigma}_l}\right). \end{aligned} \tag{3.6}$$

To get the second inequality, we bounded the ratio using

$$L_N^{*,l\varepsilon-\eta_\varepsilon}(\lambda^l) - x_{l-1} \geq \nabla L_N^*(\lambda^l) = \bar{\sigma}_l \nabla \lambda^l \log N^2$$

from the integration limits of  $x_{l-1}$  in (3.5). It is convenient to do the change of variables  $Y_{v,j} \doteq (\varepsilon + \eta_\varepsilon) \log N^2 + \nabla L_N^*(\lambda^j) - \nabla X_v^j$  for all  $j \in \{1, \dots, l-1\}$ . Equation (3.5) is then bounded, using (3.6), by

$$\frac{CN^{-4\frac{\varepsilon-l\eta_\varepsilon}{\bar{\sigma}_l}}}{N^{2\lambda^l} \sqrt{\log N}} N^{2\lambda^{l-1}} \int_0^\infty \int_{-y_1}^\infty \dots \int_{-\sum_{j=1}^{l-2} y_j}^\infty \prod_{j=1}^{l-1} e^{-2\frac{y_j}{\bar{\sigma}_l}} e^{-\frac{(y_j - (\varepsilon + \eta_\varepsilon) \log N^2 - \nabla L_N^*(\lambda^j))^2}{2\mathbb{V}(Y_{v,j})}} \frac{1}{\sqrt{2\pi\mathbb{V}(Y_{v,j})}} dy. \tag{3.7}$$

After multiplying by  $N^{2\lambda^l}$ , the  $l$ -th term of the sum in (3.4) has the right decay if we show that the integral in (3.7) is bounded by  $\tilde{C}N^{-2\lambda^{l-1}}$ . From (3.2), we have

$$0 < \mathbb{V}(Y_{v,j}) \leq \bar{\sigma}_j^2 \nabla \lambda^j \log N + C$$

for all  $v \in V_N^0$ . If the variances were all equal to  $\bar{\sigma}_j^2 \nabla \lambda^j \log N + C$ , the argument would be simpler. Extra work is needed to take care of the boundary effect of the GFF. We gather the result into a lemma for later use in the proof of Lemma 3.4.  $\square$

**Lemma 3.2.** *Let  $2 \leq l \leq m$  and  $\mathbf{z} \doteq (z_j)_{j=1}^{l-1}$  be such that  $0 < z_j \leq \bar{\sigma}_j^2 \nabla \lambda^j \log N + C$ . For all  $\varepsilon > 0$ , consider the integral*

$$I_\varepsilon(\mathbf{z}) \doteq \int_0^\infty g_1(y_1) \dots \int_{-\sum_{j=1}^{l-3} y_j}^\infty g_{l-2}(y_{l-2}) \int_{-\sum_{j=1}^{l-2} y_j}^\infty e^{-2a_l \sum_{j=1}^{l-1} y_j} g_{l-1}(y_{l-1}) dy$$

where  $a_l > 1/\bar{\sigma}_{l-1}$  and

$$g_j(y) \doteq \frac{1}{\sqrt{2\pi z_j}} \exp\left(-\frac{1}{2z_j} \left((y - \varepsilon \log N^2) - \nabla L_N^*(\lambda^j)\right)^2\right), \quad 1 \leq j \leq l-1.$$

Then  $I_\varepsilon(\mathbf{z}) \leq \tilde{C}N^{-2\lambda^{l-1}}$ .

*Proof:* Let  $\beta_j \doteq \frac{\nabla L_N^*(\lambda^j)}{2z_j}$ ,  $1 \leq j \leq l-1$ . When  $a_l - \beta_{l-1} \geq 1/\sqrt{z_{l-1}}$ , the first integral with respect to  $y_{l-1}$  in  $I_\varepsilon(\mathbf{z})$  is equal to

$$\begin{aligned} & e^{-2a_l \sum_{j=1}^{l-2} y_j} \int_{-\sum_{j=1}^{l-2} y_j}^{\infty} e^{-2a_l y_{l-1}} \frac{1}{\sqrt{2\pi z_{l-1}}} e^{-\frac{1}{2z_{l-1}}((y_{l-1} - \varepsilon \log N^2) - \nabla L_N^*(\lambda^{l-1}))^2} dy_{l-1} \\ & \leq e^{-2a_l \sum_{j=1}^{l-2} y_j} \frac{1}{\sqrt{z_{l-1}}} e^{-\frac{1}{2z_{l-1}}(\nabla L_N^*(\lambda^{l-1}))^2} \int_{-\sum_{j=1}^{l-2} y_j}^{\infty} e^{-2(a_l - \beta_{l-1})y_{l-1}} dy_{l-1} \\ & = e^{-2a_l \sum_{j=1}^{l-2} y_j} \frac{1}{\sqrt{z_{l-1}}} e^{-\frac{1}{2z_{l-1}}(\nabla L_N^*(\lambda^{l-1}))^2} \frac{1}{2(a_l - \beta_{l-1})} e^{2(a_l - \beta_{l-1}) \sum_{j=1}^{l-2} y_j}. \end{aligned}$$

Since  $z_{l-1} \leq \bar{\sigma}_{l-1}^2 \nabla \lambda^{l-1} \log N + C$  and  $\nabla L_N^*(\lambda^{l-1}) = \bar{\sigma}_{l-1} \nabla \lambda^{l-1} \log N^2$ , the above is smaller than

$$C e^{-2\beta_{l-1} \sum_{j=1}^{l-2} y_j} N^{-2\nabla \lambda^{l-1}}. \quad (3.8)$$

When  $a_l - \beta_{l-1} < 1/\sqrt{z_{l-1}}$ , we have by completing the square :

$$\begin{aligned} & e^{-2a_l \sum_{j=1}^{l-2} y_j} \int_{-\sum_{j=1}^{l-2} y_j}^{\infty} \frac{e^{-2a_l y_{l-1}}}{\sqrt{2\pi z_{l-1}}} e^{-\frac{((y_{l-1} - \varepsilon \log N^2) - \nabla L_N^*(\lambda^{l-1}))^2}{2z_{l-1}}} dy_{l-1} \\ & \leq e^{-2a_l \sum_{j=1}^{l-2} y_j} \int_{-\sum_{j=1}^{l-2} y_j}^{\infty} \frac{e^{-2a_l(y_{l-1} - \varepsilon \log N^2)}}{\sqrt{2\pi z_{l-1}}} e^{-\frac{((y_{l-1} - \varepsilon \log N^2) - \nabla L_N^*(\lambda^{l-1}))^2}{2z_{l-1}}} dy_{l-1} \\ & = e^{-2a_l \sum_{j=1}^{l-2} y_j} e^{-2a_l \nabla L_N^*(\lambda^{l-1})} e^{2a_l^2 z_{l-1}} \\ & \quad \cdot \int_{-\sum_{j=1}^{l-2} y_j}^{\infty} \frac{1}{\sqrt{2\pi z_{l-1}}} e^{-\frac{((y_{l-1} - \varepsilon \log N^2) - (\nabla L_N^*(\lambda^{l-1}) - 2a_l z_{l-1}))^2}{2z_{l-1}}} dy_{l-1} \\ & \leq \exp\left(-2a_l \sum_{j=1}^{l-2} y_j - 2a_l \nabla L_N^*(\lambda^{l-1}) + 2a_l^2 z_{l-1}\right). \end{aligned}$$

In the regime  $a_l - \beta_{l-1} < 1/\sqrt{z_{l-1}}$ , note that  $2a_l^2 z_{l-1} < a_l \nabla L_N^*(\lambda^{l-1}) + 2a_l \sqrt{z_{l-1}}$ . Since  $z_{l-1} \leq \bar{\sigma}_{l-1}^2 \nabla \lambda^{l-1} \log N + C$ , the above is smaller than

$$\exp\left(-2a_l \sum_{j=1}^{l-2} y_j - a_l \nabla L_N^*(\lambda^{l-1}) + C\sqrt{\log N}\right).$$

By assumption,  $a_l > 1/\bar{\sigma}_{l-1}$ . Therefore, the above is smaller than

$$e^{-2a_l \sum_{j=1}^{l-2} y_j} N^{-2\nabla \lambda^{l-1}}. \quad (3.9)$$

The second integral with respect to  $y_{l-2}$  in  $I_\varepsilon(\mathbf{z})$  is evaluated similarly using (3.8) and (3.9) by taking

$$a_{l-1} \doteq \min\{a_l, \beta_{l-1}\}$$

and considering whether  $a_{l-1} - \beta_{l-2} \geq 1/\sqrt{z_{l-2}}$  or not. Note that  $a_{l-1} > 1/\bar{\sigma}_{l-2}$  holds since  $a_l > 1/\bar{\sigma}_{l-1} > 1/\bar{\sigma}_{l-2}$  (because the steps of  $\bar{\sigma}$  are decreasing in height) and  $\beta_{l-1} \geq 1/\bar{\sigma}_{l-1} - O((\log N)^{-1}) > 1/\bar{\sigma}_{l-2}$  from the bounds on  $z_{l-1}$ . This recursive reasoning shows that  $I_\varepsilon(\mathbf{z})$  is smaller than  $\tilde{C}N^{-2\lambda^{l-1}}$ .  $\square$

**Lemma 3.3** (Lower bound on the first order of the maximum). *Let  $\{\psi_v\}_{v \in V_N}$  be the  $(\sigma, \lambda)$ -GFF on  $V_N$  of Definition 1.1 and  $\gamma^*$  as in Theorem 1.2. For all  $0 < \varepsilon < 1$ ,*

there exists a constant  $c = c(\varepsilon, \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$  such that

$$\mathbb{P}\left(\max_{v \in V_N} \psi_v \leq (1 - \varepsilon)\gamma^* \log N^2\right) \leq N^{-c} \tag{3.10}$$

for  $N$  large enough.

Without loss of generality, we can assume that  $\lambda_i \in \mathbb{Q}$  for all  $i \in \{0, \dots, M\}$ . To see this, define  $\tilde{\lambda}_i \doteq \lambda_i + \eta_i$  where  $0 < \eta_i < \min_i \nabla \lambda_i$  and such that  $\tilde{\lambda}_i \in \mathbb{Q}$  for all  $i \in \{1, \dots, M - 1\}$ . Now, define a new scale-inhomogeneous Gaussian free field :

$$\tilde{\psi}_v \doteq \sum_{i=1}^M \sigma_i \nabla \phi_v(\tilde{\lambda}_i) = \psi_v + \sum_{i=1}^{M-1} (\sigma_i - \sigma_{i+1})(\phi_v(\tilde{\lambda}_i) - \phi_v(\lambda_i)) .$$

As a particular case of Lemma A.3, note that

$$\max_{v \in V_N} \mathbb{V}\left(\phi_v(\tilde{\lambda}_i) - \phi_v(\lambda_i)\right) \leq (\tilde{\lambda}_i - \lambda_i) \log N + C = \eta_i \log N + C .$$

If we can show Lemma 3.3 when the  $\lambda_i$ 's are rational numbers, then a union bound and a Gaussian estimate yield

$$\begin{aligned} \mathbb{P}\left(\max_{v \in V_N} \psi_v \leq (1 - 2\varepsilon)\gamma^* \log N^2\right) &\leq \mathbb{P}\left(\max_{v \in V_N} \tilde{\psi}_v \leq (1 - \varepsilon)\gamma^* \log N^2\right) \\ &\quad + \sum_{v \in V_N} \sum_{i=1}^{M-1} \mathbb{P}\left(|\sigma_i - \sigma_{i+1}| \left|\phi_v(\tilde{\lambda}_i) - \phi_v(\lambda_i)\right| \geq (\varepsilon/(M - 1))\gamma^* \log N^2\right) \\ &\leq N^{-c(\varepsilon, \boldsymbol{\sigma}, \boldsymbol{\lambda})} + N^2(M - 1) \exp\left(-\frac{((\varepsilon/(M - 1))\gamma^* \log N^2)^2}{2 \max_i |\sigma_i - \sigma_{i+1}|^2 (\eta_i \log N + C)}\right) . \end{aligned}$$

The second term can be made  $O(N^{-\tilde{c}(\varepsilon, \boldsymbol{\sigma}, \boldsymbol{\lambda})})$  where  $\tilde{c} > 0$  is arbitrarily large, by choosing the  $\eta_i$ 's small enough with respect to  $\varepsilon$ .

The proof of Lemma 3.3 is based on a coarse-graining of the scales introduced in Kistler (2015). Consider  $\alpha_k \doteq \frac{k}{K}$ ,  $0 \leq k \leq K$ . The parameter  $K \in \mathbb{N}$  will be chosen large enough depending on  $\varepsilon$  during the proof. By the argument above, we can assume that  $\lambda_i K \in \mathbb{N}_0$  for all  $i \in \{0, \dots, M\}$ , so that the  $\alpha_k$ 's form a finer partition of  $[0, 1]$  than the  $\lambda_i$ 's. The bounds in (3.1) imply that for all  $k \in \{1, \dots, K\}$  and for all  $v \in V_N^\delta$  :

$$|\mathbb{V}(\nabla \psi_v(\alpha_k)) - \sigma^2(\alpha_k) \nabla \alpha_k \log N| \leq C(\delta) . \tag{3.11}$$

The parameter  $\delta \in (0, 1/2)$  remains fixed to an arbitrary value in the rest of this section. For all  $0 < \varepsilon < 1$ , denote by  $L_{N, \varepsilon}^*$  the following sub-optimal path :

$$L_{N, \varepsilon}^*(s) = (1 - \varepsilon)L_N^*(s) = (1 - \varepsilon)\mathcal{J}_{\sigma^2/\bar{\sigma}}(s) \log N^2, \quad s \in [0, 1] .$$

The proof relies on the Paley-Zygmund inequality (see Lemma A.8) applied to a *modified number of exceedances*. In fact, we consider only points in  $V_N^\delta$  whose increments are almost optimal. Moreover, and crucially, we drop the first  $r$  increments. We will choose  $r$  during the proof. This allows more independence between the variables of the field, which is needed to find a tight lower bound using the Paley-Zygmund inequality. More precisely, define

$$\mathcal{N}_\varepsilon^* \doteq \sum_{v \in V_N^\delta} 1_{A_v} \quad \text{where} \quad A_v \doteq \{\nabla \psi_v(\alpha_j) \geq \nabla L_{N, \varepsilon}^*(\alpha_j) \quad \forall j \in \{r + 1, \dots, K\}\} .$$

For a fixed  $\varepsilon > 0$ , there is the following inequality for  $c = c(\varepsilon) > 0$  :

$$\mathbb{P} \left( \max_{v \in V_N} \psi_v \geq (1 - 3\varepsilon)\gamma^* \log N^2 \right) \geq \mathbb{P}(\mathcal{N}_\varepsilon^* \geq 1) - O(N^{-c}) . \tag{3.12}$$

Indeed, on the event  $\{\mathcal{N}_\varepsilon^* \geq 1\}$ , we have

$$\begin{aligned} \max_{v \in V_N^\delta} \psi_v - \psi_v(\alpha_r) &\geq (1 - \varepsilon)\mathcal{J}_{\sigma^2/\bar{\sigma}}(\alpha_r, 1) \log N^2 \\ &= (1 - \varepsilon)\gamma^* \log N^2 - (1 - \varepsilon)\mathcal{J}_{\sigma^2/\bar{\sigma}}(\alpha_r) \log N^2 \\ &\geq (1 - 2\varepsilon)\gamma^* \log N^2 \end{aligned}$$

where we take  $K$  large enough that  $(1 - \varepsilon)\mathcal{J}_{\sigma^2/\bar{\sigma}}(\alpha_r) < \varepsilon\gamma^*$ . Furthermore, the probability  $\mathbb{P}(\max_{v \in V_N^\delta} \psi_v - \psi_v(\alpha_r) \geq (1 - 2\varepsilon)\gamma^* \log N^2)$  is equal to

$$\begin{aligned} &\mathbb{P} \left( \max_{v \in V_N^\delta} \psi_v - \psi_v(\alpha_r) \geq (1 - 2\varepsilon)\gamma^* \log N^2, \min_{v \in V_N^\delta} \psi_v(\alpha_r) > -\varepsilon\gamma^* \log N^2 \right) \\ &+ \mathbb{P} \left( \max_{v \in V_N^\delta} \psi_v - \psi_v(\alpha_r) \geq (1 - 2\varepsilon)\gamma^* \log N^2, \min_{v \in V_N^\delta} \psi_v(\alpha_r) \leq -\varepsilon\gamma^* \log N^2 \right) . \end{aligned}$$

The distribution of  $\psi_v(\alpha_r)$  is symmetric, so the second term is smaller than

$$\mathbb{P} \left( \max_{v \in V_N^\delta} \psi_v(\alpha_r) \geq \varepsilon\gamma^* \log N^2 \right) \leq N^2 \exp \left( -\frac{(\varepsilon\gamma^*)^2 \log N^2}{\max_i \sigma_i^2 \alpha_r} \right) \tag{3.13}$$

where we used a union bound, a Gaussian estimate and (3.2) to get the inequality. This is  $O(N^{-c})$  by choosing  $K$  large enough for a fixed  $\varepsilon$  and  $r$ . On the other hand, the first term is smaller than  $\mathbb{P}(\max_{v \in V_N^\delta} \psi_v \geq (1 - 3\varepsilon)\gamma^* \log N^2)$ . Since  $V_N \supseteq V_N^\delta$ , this implies (3.12) as claimed.

*Proof of Lemma 3.3:* In view of (3.12), it suffices to show  $\mathbb{P}(\mathcal{N}_\varepsilon^* \geq 1) = 1 - O(N^{-c})$ . The Paley-Zygmund inequality implies

$$\mathbb{P}(\mathcal{N}_\varepsilon^* \geq 1) \geq \frac{(\mathbb{E}[\mathcal{N}_\varepsilon^*])^2}{\mathbb{E}[(\mathcal{N}_\varepsilon^*)^2]} .$$

We show

$$\mathbb{E}[(\mathcal{N}_\varepsilon^*)^2] \leq (1 + O(N^{-\frac{1}{2K}(1-(1-\varepsilon)^2)})) (\mathbb{E}[\mathcal{N}_\varepsilon^*])^2, \tag{3.14}$$

which proves the claim.

The first moment is easily evaluated by the independence of the increments :

$$\mathbb{E}[\mathcal{N}_\varepsilon^*] = \sum_{v \in V_N^\delta} \mathbb{P}(A_v) = \sum_{v \in V_N^\delta} \prod_{j=r+1}^K \mathbb{P}(\nabla\psi_v(\alpha_j) \geq \nabla L_{N,\varepsilon}^*(\alpha_j)) .$$

Using Gaussian estimates and the variance estimates in (3.11), the probabilities are for every  $j$  and  $v \in V_N^\delta$  :

$$\begin{aligned} p_{v,j} &\doteq \mathbb{P}(\nabla\psi_v(\alpha_j) \geq \nabla L_{N,\varepsilon}^*(\alpha_j)) \\ &\asymp \frac{1}{\sqrt{\log N}} \exp \left( -(1 - \varepsilon)^2 \frac{\sigma^2(\alpha_j)}{\bar{\sigma}^2(\alpha_j)} \nabla\alpha_j \log N^2 \right) . \end{aligned} \tag{3.15}$$

Write  $e_j$  for the exponential term on the right-hand side of (3.15). The first moment satisfies

$$\mathbb{E}[\mathcal{N}_\varepsilon^*] = \sum_{v \in V_N^\delta} \mathbb{P}(A_v) \geq \frac{c(\varepsilon, \delta)}{(\log N)^{\frac{1}{2}(K-r)}} \times |V_N^\delta| \times \prod_{j=r+1}^K e_j. \quad (3.16)$$

Now, we compare this with the second moment :

$$\mathbb{E}[(\mathcal{N}_\varepsilon^*)^2] = \sum_{v, v' \in V_N^\delta} \mathbb{P}(A_v \cap A_{v'}).$$

We divide the sum depending on the correlations between  $\psi_v$  and  $\psi_{v'}$ . More precisely, recall the definition of the branching scale in (2.1) :

$$\rho(v, v') \doteq \max\{\lambda \in [0, 1] : [v]_\lambda \cap [v']_\lambda \neq \emptyset\}, \quad v, v' \in V_N.$$

Write the second moment as

$$\sum_{\substack{v, v' \in V_N^\delta \\ \rho(v, v') < \alpha_r}} \mathbb{P}(A_v \cap A_{v'}) + \sum_{k=r+1}^{K-1} \sum_{\substack{v, v' \in V_N^\delta \\ \alpha_{k-1} \leq \rho(v, v') < \alpha_k}} \mathbb{P}(A_v \cap A_{v'}) + \sum_{\substack{v, v' \in V_N^\delta \\ \rho(v, v') \geq \alpha_{K-1}}} \mathbb{P}(A_v \cap A_{v'}). \quad (3.17)$$

In particular, the first term in (3.17) is equal to

$$\sum_{\substack{v, v' \in V_N^\delta \\ \rho(v, v') < \alpha_r}} \mathbb{P}(A_v) \mathbb{P}(A_{v'}) \leq \sum_{v, v' \in V_N^\delta} \mathbb{P}(A_v) \mathbb{P}(A_{v'}) = (\mathbb{E}[\mathcal{N}_\varepsilon^*])^2. \quad (3.18)$$

It remains to show that the second and third term in (3.17) are negligible compared to  $(\mathbb{E}[\mathcal{N}_\varepsilon^*])^2$ . We write the details for the second term since the last term is done similarly and is easier. By Lemma A.1 (following the Markov property of the GFF), note that if  $\alpha_{k-1} \leq \rho(v, v') < \alpha_k$  for some  $k \geq r + 1$ , then  $\nabla\psi_{v'}(\alpha_{j'})$ ,  $j' \geq k + 1$ , is independent of  $\nabla\psi_v(\alpha_j)$  for  $j \leq k - 2$  and  $j \geq k + 1$ . Therefore, for  $v, v' \in V_N^\delta$  such that  $\alpha_{k-1} \leq \rho(v, v') < \alpha_k$ , we have

$$\mathbb{P}(A_v \cap A_{v'}) \leq \prod_{j=r+1}^{k-2} p_{v,j} \prod_{j=k+1}^K p_{v,j} p_{v',j} \leq \left( \prod_{j=r+1}^K e_j^2 \right) \left( \frac{\prod_{j=1}^r e_j}{e_{k-1} e_k^2} \right) \left( \prod_{j=1}^{k-1} e_j \right)^{-1}$$

where we dropped the conditions on  $j \in \{k - 1, k\}$  for  $v$  as well as the conditions on  $j \leq k$  for  $v'$  in the first inequality. We simply rearranged the probabilities and eliminated the log terms to get the last inequality. The number of pairs  $v, v' \in V_N^\delta$  such that  $\alpha_{k-1} \leq \rho(v, v') < \alpha_k$  is at most  $|V_N^\delta| \times N^{2(1-\alpha_{k-1})}$ . Therefore, by (3.16),

$$\sum_{\substack{v, v' \in V_N^\delta \\ \alpha_{k-1} \leq \rho(v, v') < \alpha_k}} \mathbb{P}(A_v \cap A_{v'}) \leq \frac{(\mathbb{E}[\mathcal{N}_\varepsilon^*])^2}{(\log N)^{-K}} \times N^{-2\alpha_{k-1}(1-(1-\varepsilon)^2)} \left( \frac{\prod_{j=1}^r e_j}{e_{k-1} e_k^2} \right) \times \frac{N^{-2\alpha_{k-1}(1-\varepsilon)^2}}{\prod_{j=1}^{k-1} e_j} \quad (3.19)$$

The right-hand side of (3.19) is separated in three factors by  $\times$ . The third factor is bounded by 1 because

$$\int_0^t \frac{\sigma^2(s)}{\bar{\sigma}^2(s)} ds \leq t, \quad t \in (0, 1],$$

by definition of  $\bar{\sigma}$ . To bound the second factor, set  $r \geq 3$  independently of any other variable. Note that if  $r$  depended on  $K$ , the bound in (3.13) would not necessarily tend to 0. There are two cases to consider :  $\alpha_k \leq \lambda_1$  and  $\alpha_k > \lambda_1$ . When  $\alpha_k \leq \lambda_1$ , the ratio of exponentials is bounded by 1 because  $e_1 e_2 e_3 = e_{k-1} e_k^2$  and we have



$N^{-2\alpha_{k-1}(1-(1-\varepsilon)^2)} \leq N^{-\frac{1}{K}(1-(1-\varepsilon)^2)}$  since  $\alpha_{k-1} \geq \alpha_r \geq 1/(2K)$ . When  $\alpha_k > \lambda_1$ , the ratio of exponentials is bounded by  $N^{\lambda_1(1-(1-\varepsilon)^2)}$  by choosing  $K$  large enough for a fixed  $\varepsilon$  and we have  $N^{-2\alpha_{k-1}(1-(1-\varepsilon)^2)} \leq N^{-2\lambda_1(1-(1-\varepsilon)^2)}$  because  $\alpha_{k-1} \geq \lambda_1$ . Since  $\lambda_1 \geq 1/K$ , the right-hand side of (3.19) is always bounded by

$$(\mathbb{E}[\mathcal{N}_\varepsilon^*])^2 (\log N)^K \times N^{-\frac{1}{K}(1-(1-\varepsilon)^2)}.$$

With (3.18), this shows (3.14) and concludes the proof of the lemma.  $\square$

**3.3. Log-number of high points.** The proof of the upper bound for the log-number of high-points uses an argument based on the path at every scale  $\lambda^l$  similar to the one in Lemma 3.1. Recall the definition of the critical levels  $\gamma^l$  and the entropy  $\mathcal{E}_\gamma$  in Theorem 1.3.

**Lemma 3.4** (Upper bound on the log-number of high points). *Let  $\{\psi_v\}_{v \in V_N}$  be the  $(\sigma, \lambda)$ -GFF on  $V_N$  of Definition 1.1 and  $\gamma^*$  as defined in Theorem 1.2. Also, let  $\gamma^{l-1} < \gamma \leq \gamma^l$  for some  $l \in \{1, \dots, m\}$ . For all  $0 < \varepsilon < (\gamma - \gamma^{l-1})/m$ , there exists a constant  $c = c(\gamma, \varepsilon, \sigma, \lambda) > 0$  such that*

$$\mathbb{P}(|\mathcal{H}_N^\gamma| \geq N^{2\mathcal{E}_\gamma + \varepsilon}) \leq N^{-c} \tag{3.20}$$

for  $N$  large enough.

*Proof:* Recall the definition of the optimal path  $L_N^\gamma$  from (2.6) and the notation  $R_{\lambda^j}$  for the set of representatives at scale  $\lambda^j$ . Consider

$$\mathcal{H}_{N,j}^{\gamma,\varepsilon} \stackrel{\circ}{=} \left\{ v \in R_{\lambda^j} \mid \psi_v(\lambda^j) \geq L_N^{\gamma+j\varepsilon}(\lambda^j) \right\}, \quad 1 \leq j \leq m.$$

Since  $R_{\lambda^m} = V_N$ , note that

$$\mathcal{H}_N^\gamma = \mathcal{H}_{N,m}^{\gamma,0} = \mathcal{H}_{N,m}^{\gamma-m\varepsilon,\varepsilon}.$$

This is useful because the hypothesis  $\varepsilon < (\gamma - \gamma^{l-1})/m$  implies  $\gamma^{l-1} < \gamma - j\varepsilon \leq \gamma^l$ , which means (in particular) that for all  $j \in \{1, \dots, l-1\}$ ,

$$\text{the paths } L_N^* \text{ and } L_N^{\gamma-j\varepsilon} \text{ coincide on the interval } [0, \lambda^j]. \tag{3.21}$$

The idea is to split the probability that at least  $N^{2\mathcal{E}_\gamma + \varepsilon}$  points in  $V_N$  reach the optimal height by looking at the first scale  $\lambda^j$ ,  $1 \leq j \leq l-1$ , where the set  $\mathcal{H}_{N,j}^{\gamma-j\varepsilon,\varepsilon}$  is not empty. As for the maximum, this yields the appropriate constraints along the scales to get the correct upper bound. A union bound in (3.20) gives

$$\begin{aligned} \mathbb{P}(|\mathcal{H}_N^\gamma| \geq N^{2\mathcal{E}_\gamma + \varepsilon}) &= \mathbb{P}\left(|\mathcal{H}_{N,m}^{\gamma-m\varepsilon,\varepsilon}| \geq N^{2\mathcal{E}_\gamma + \varepsilon}\right) \\ &\leq \mathbb{P}\left(\begin{array}{l} |\mathcal{H}_{N,1}^{\gamma-1\varepsilon,\varepsilon}| = \dots = |\mathcal{H}_{N,l-1}^{\gamma-(l-1)\varepsilon,\varepsilon}| = 0 \\ \text{and } |\mathcal{H}_{N,m}^{\gamma-m\varepsilon,\varepsilon}| \geq N^{2\mathcal{E}_\gamma + \varepsilon} \end{array}\right) + \sum_{j=1}^{l-1} \mathbb{P}\left(|\mathcal{H}_{N,j}^{\gamma-j\varepsilon,\varepsilon}| \geq 1\right). \end{aligned} \tag{3.22}$$

Because of (3.21), the probabilities in the sum are bounded by  $N^{-c(\varepsilon)}$  in exactly the same manner as  $\mathbb{P}(|\mathcal{H}_{N,m}^{\gamma-m\varepsilon,\varepsilon}| \geq 1)$  in Lemma 3.1. The first probability in (3.22) is bounded by

$$\begin{aligned} &\mathbb{P}\left(\left|\left\{ v \in V_N \mid \begin{array}{l} \psi_v \geq L_N^\gamma(1) \text{ and } \psi_{v_{\lambda^j}}(\lambda^j) < L_N^\gamma(\lambda^j) \\ \text{for all } 1 \leq j \leq l-1 \end{array} \right\}\right| \geq N^{2\mathcal{E}_\gamma + \varepsilon}\right) \\ &\leq C e^{-c(\eta_\varepsilon)(\log N)^2} + N^{-\varepsilon} N^{-2\mathcal{E}_\gamma} N^2 \max_{v \in V_N^o} \mathbb{P}\left(\begin{array}{l} \psi_v \geq L_N^{\gamma-\eta_\varepsilon}(1) \text{ and} \\ \psi_v(\lambda^j) < L_N^{\gamma+j\eta_\varepsilon}(\lambda^j) \\ \text{for all } 1 \leq j \leq l-1 \end{array}\right) \end{aligned} \tag{3.23}$$

using Markov's inequality and using the event  $A_\varepsilon$  as in (3.4), where we impose

$$0 < \eta_\varepsilon < \min\{\gamma, \mathcal{J}_{\sigma^2}(1)\varepsilon/(4\gamma), \bar{\sigma}_{l-1}\varepsilon/(4lc_\gamma), \varepsilon/m\}$$

this time around. See (3.26) for the definition of  $c_\gamma$ . See just below and also (3.27) for the justification of the constraints on  $\eta_\varepsilon$ . When  $l = 1$ , a Gaussian estimate and the bound  $\max_{v \in V_N} \mathbb{V}(\psi_v) \leq \mathcal{J}_{\sigma^2}(1) \log N + C$  from (3.2) yield

$$\mathbb{P}(\psi_v \geq L_N^{\gamma-\eta_\varepsilon}(1)) \leq \frac{\sqrt{\mathbb{V}(\psi_v)}}{L_N^{\gamma-\eta_\varepsilon}(1)} \exp\left(-\frac{(L_N^{\gamma-\eta_\varepsilon}(1))^2}{2\mathbb{V}(\psi_v)}\right) \leq \frac{C(\gamma)N^{-2+2\mathcal{E}_\gamma}}{\sqrt{\log N}} N^{\frac{4\gamma}{\mathcal{J}_{\sigma^2}(1)}\eta_\varepsilon}$$

because  $L_N^{\gamma-\eta_\varepsilon}(1) = 2(\gamma - \eta_\varepsilon) \log N$  and  $\mathcal{E}_\gamma = 1 - \gamma^2/\mathcal{J}_{\sigma^2}(1)$  in this case. This proves that the second term in (3.23) decays like  $N^{-c(\gamma, \varepsilon)}$ , as needed.

It remains to show a similar estimate for a fixed  $l \in \{2, \dots, m\}$ . To simplify the notation, denote  $(X_v^1, \dots, X_v^{l-1}, X_v^m) \doteq (\psi_v(\lambda^1), \dots, \psi_v(\lambda^{l-1}), \psi_v)$ . By conditioning on the value of the vector  $\mathbf{X} \doteq (X_v^1, \dots, X_v^{l-1})$ , the probability in (3.23) is equal to

$$\int_{-\infty}^{L_N^{\gamma+1\eta_\varepsilon}(\lambda^1)} \dots \int_{-\infty}^{L_N^{\gamma+(l-1)\eta_\varepsilon}(\lambda^{l-1})} \mathbb{P}(X_v^m \geq L_N^{\gamma-\eta_\varepsilon}(1) \mid \mathbf{X} = \mathbf{x}) f_v(\mathbf{x}) d\mathbf{x}$$

where  $f_v$  is the density function of  $\mathbf{X}$ . By independence of the increments, the last integral is equal to

$$\int_{-\infty}^{L_N^{\gamma+1\eta_\varepsilon}(\lambda^1)} \dots \int_{-\infty}^{L_N^{\gamma+(l-1)\eta_\varepsilon}(\lambda^{l-1})} \mathbb{P}(X_v^m - X_v^{l-1} \geq L_N^{\gamma-\eta_\varepsilon}(1) - x_{l-1}) f_v(\mathbf{x}) d\mathbf{x}. \quad (3.24)$$

The bound  $\max_{v \in V_N} \mathbb{V}(X_v^m - X_v^{l-1}) \leq \mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1) \log N + C$  from (3.2) and a Gaussian estimate show that

$$\begin{aligned} & \mathbb{P}(X_v^m - X_v^{l-1} \geq L_N^{\gamma-\eta_\varepsilon}(1) - x_{l-1}) \\ &= \mathbb{P}(X_v^m - X_v^{l-1} \geq L_N^\gamma(1) - L_N^\gamma(\lambda^{l-1}) + L_N^{\gamma-\eta_\varepsilon}(\lambda^{l-1}) - x_{l-1}) \\ &\leq \frac{C(\gamma)}{\sqrt{\log N}} N^{-2\frac{(\gamma - \mathcal{J}_{\sigma^2}/\bar{\sigma}(\lambda^{l-1}))^2}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)}} \exp\left(-2\frac{(\gamma - \mathcal{J}_{\sigma^2}/\bar{\sigma}(\lambda^{l-1}))}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)}(L_N^{\gamma-\eta_\varepsilon}(\lambda^{l-1}) - x_{l-1})\right) \end{aligned}$$

where we introduced  $L_N^\gamma(\lambda^{l-1})$  and used (2.6). By definition of  $\mathcal{E}_\gamma$  and the definition of  $\gamma^{l-1}$  in (1.5), this is equal to

$$\begin{aligned} & \frac{C(\gamma)N^{-2+2\mathcal{E}_\gamma}}{\sqrt{\log N}} N^{2\lambda^{l-1}} \exp\left(-2\left[\frac{(\gamma - \gamma^{l-1})}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} + \frac{1}{\bar{\sigma}_{l-1}}\right](L_N^{\gamma-\eta_\varepsilon}(\lambda^{l-1}) - x_{l-1})\right) \\ &= \frac{C(\gamma)N^{-2+2\mathcal{E}_\gamma}}{\sqrt{\log N}} N^{2\lambda^{l-1}} N^{\frac{4lc_\gamma}{\bar{\sigma}_{l-1}}\eta_\varepsilon} \exp\left(-2\frac{c_\gamma}{\bar{\sigma}_{l-1}}(L_N^{\gamma+(l-1)\eta_\varepsilon}(\lambda^{l-1}) - x_{l-1})\right) \end{aligned} \quad (3.25)$$

where

$$c_\gamma \doteq \frac{(\gamma - \gamma^{l-1})\bar{\sigma}_{l-1}}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} + 1 > 1. \quad (3.26)$$

Putting the bound (3.25) in (3.24) and in (3.23), we get that the first term in (3.22) decays like

$$N^{-\left(\varepsilon - \frac{4lc_\gamma}{\bar{\sigma}_{l-1}}\eta_\varepsilon\right)} \quad (3.27)$$

provided that

$$\int_0^\infty \int_{-y_1}^\infty \dots \int_{-\sum_{j=1}^{l-2} y_j}^\infty \prod_{j=1}^{l-1} e^{-2\frac{c_\gamma}{\bar{\sigma}_{l-1}} y_j} \frac{e^{-\frac{((y_j - \eta_\varepsilon \log N^2) - \nabla L_N^\gamma(\lambda^j))^2}{2\mathbb{V}(Y_{v,j})}}}{\sqrt{2\pi\mathbb{V}(Y_{v,j})}} d\mathbf{y} \leq \tilde{C}N^{-2\lambda^{l-1}},$$

where  $Y_{v,j} \doteq \eta_\varepsilon \log N^2 + \nabla L_N^\gamma(\lambda^j) - \nabla X_v^j$ . Similarly to (3.7), the integral has the right decay as a consequence of Lemma 3.2, with  $a_l \doteq c_\gamma/\bar{\sigma}_{l-1} > 1/\bar{\sigma}_{l-1}$ , because  $L_N^*$  and  $L_N^\gamma$  coincide on the interval  $[0, \lambda^{l-1}]$ .  $\square$

**Lemma 3.5** (Lower bound on the log-number of high points). *Let  $\{\psi_v\}_{v \in V_N}$  be the  $(\sigma, \lambda)$ -GFF on  $V_N$  of Definition 1.1 and  $\gamma^*$  as in Theorem 1.2. Let  $\gamma > 0$  be such that  $\gamma^{l-1} \leq \gamma < \gamma^l$  for some  $l \in \{1, \dots, m\}$ . For all  $0 < \varepsilon < \min\{1/4, (\gamma^l - \gamma)/(4\gamma)\}$ , there exists a constant  $c = c(\gamma, \varepsilon, \sigma, \lambda) > 0$  such that*

$$\mathbb{P}(|\mathcal{H}_N^\gamma| < N^{2\varepsilon\gamma - \tilde{\varepsilon}}) \leq N^{-c}$$

for  $N$  large enough, where  $\tilde{\varepsilon} \doteq \frac{24(\gamma^*)^2}{\bar{\sigma}_m^2 \nabla \lambda^m} \varepsilon$ .

We use the same notations as in the proof of Lemma 3.3. As before, we can assume, without loss of generality, that  $\lambda_i K \in \mathbb{N}_0$  for all  $\{0, \dots, M\}$  so that the  $\alpha_k$ 's form a finer partition of  $[0, 1]$  than the  $\lambda_i$ 's. The parameter  $K \in \mathbb{N}$  will be chosen large enough depending on  $\gamma$  and  $\varepsilon$  during the proof. Again, we restrict ourselves to  $V_N^\delta$  to ensure that for all  $k \in \{1, \dots, K\}$  and for all  $v \in V_N^\delta$ :

$$|\mathbb{V}(\nabla \psi_v(\alpha_k)) - \sigma^2(\alpha_k) \nabla \alpha_k \log N| \leq C(\delta). \tag{3.28}$$

The parameter  $\delta \in (0, 1/2)$  remains fixed to an arbitrary value in the remainder of this section. Next, define the path:

$$L_{N,\varepsilon}^\gamma(s) \doteq (1 - \varepsilon)L_N^{\gamma(1+4\varepsilon)}(s), \quad s \in [0, 1].$$

Since  $\varepsilon < (\gamma^l - \gamma)/(4\gamma)$  by hypothesis, we have  $\gamma^{l-1} \leq \gamma < \gamma(1 + 4\varepsilon) < \gamma^l$ . This condition implies that the increments of the path  $L_{N,\varepsilon}^\gamma$  are always bounded by the increments of the sub-optimal path  $L_{N,\varepsilon}^*$  (see Figure 2.3), namely

$$L_{N,\varepsilon}^\gamma(s_2) - L_{N,\varepsilon}^\gamma(s_1) \leq L_{N,\varepsilon}^*(s_2) - L_{N,\varepsilon}^*(s_1), \quad 0 \leq s_1 \leq s_2 \leq 1. \tag{3.29}$$

Indeed, the paths  $L_N^{\gamma(1+4\varepsilon)}$  and  $L_N^*$  coincide on the interval  $[0, \lambda^{l-1}]$ . Moreover, when  $s \in (\lambda^{l-1}, 1]$ , we have by the definition of the critic levels  $\gamma^l$  in (1.5) and the optimal path  $L_N^{\gamma(1+4\varepsilon)}$  in (2.6):

$$\begin{aligned} \frac{d}{ds} \frac{(L_N^{\gamma(1+4\varepsilon)}(s) - L_N^*(s))}{\log N^2} &= \frac{d}{ds} \int_{\lambda^{l-1}}^s \left[ \sigma^2(u) \frac{(\gamma(1 + 4\varepsilon) - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} - \frac{\sigma^2(u)}{\bar{\sigma}(u)} \right] du \\ &\leq \frac{\sigma^2(s)}{\bar{\sigma}_l} - \frac{\sigma^2(s)}{\bar{\sigma}(s)} \quad \text{since } \gamma(1 + 4\varepsilon) < \gamma^l \\ &\leq 0 \quad \text{since } \bar{\sigma} \text{ is non-increasing.} \end{aligned}$$

This proves inequality (3.29). By hypothesis, we also have  $\varepsilon < 1/4$ , which yields

$$L_{N,\varepsilon}^\gamma(1) = (1 - \varepsilon)(1 + 4\varepsilon)\gamma \log N^2 > (1 + 2\varepsilon)\gamma \log N^2. \tag{3.30}$$

The proof again relies on the Paley-Zygmund inequality applied to a *modified number of exceedances* where we consider only points in  $V_N^\delta$  whose increments are almost optimal. We drop the first  $r$  increments to allow more independence which is needed for the second-moment method to work. We can choose  $r \geq 3$  independently of any other variable as in the proof of Lemma 3.3. The case  $l = 1$  is easier to deal with, so we omit the details. Assume  $l \in \{2, \dots, m\}$  and define

$$\mathcal{N}_\varepsilon^\gamma \doteq \sum_{v \in V_N^\delta} 1_{A_v} \quad \text{where } A_v \doteq \{\nabla \psi_v(\alpha_j) \geq \nabla L_{N,\varepsilon}^\gamma(\alpha_j) \quad \forall j \in \{r + 1, \dots, K\}\}.$$

Note that for a fixed  $\varepsilon > 0$ , there is the following inequality for  $c = c(\gamma, \varepsilon) > 0$  :

$$\mathbb{P}(|\mathcal{H}_N^\gamma| \geq N^{2\varepsilon\gamma - \tilde{\varepsilon}}) \geq \mathbb{P}(\mathcal{N}_\varepsilon^\gamma \geq N^{2\varepsilon\gamma - \tilde{\varepsilon}}) - O(N^{-c}). \tag{3.31}$$

Indeed, the probability  $\mathbb{P}(\mathcal{N}_\varepsilon^\gamma \geq N^{2\varepsilon\gamma - \tilde{\varepsilon}})$  is equal to

$$\begin{aligned} & \mathbb{P}\left(\mathcal{N}_\varepsilon^\gamma \geq N^{2\varepsilon\gamma - \tilde{\varepsilon}}, \min_{v \in V_N^\delta} \psi_v(\alpha_r) > -\varepsilon\gamma \log N^2\right) \\ & + \mathbb{P}\left(\mathcal{N}_\varepsilon^\gamma \geq N^{2\varepsilon\gamma - \tilde{\varepsilon}}, \min_{v \in V_N^\delta} \psi_v(\alpha_r) \leq -\varepsilon\gamma \log N^2\right) \end{aligned} \tag{3.32}$$

To simplify the argument, assume from now on that  $K$  is large enough to ensure  $\alpha_r \leq \lambda^{l-1}$ . The first probability in (3.32) is smaller than  $\mathbb{P}(|\mathcal{H}_N^\gamma| \geq N^{2\varepsilon\gamma - \tilde{\varepsilon}})$  because the points  $v \in V_N^\delta$  that are contributing to the sum  $\mathcal{N}_\varepsilon^\gamma$ , on the event  $\{\min_{v \in V_N^\delta} \psi_v(\alpha_r) > -\varepsilon\gamma \log N^2\}$ , are also in  $\mathcal{H}_N^\gamma$ . Indeed, when  $1_{A_v} = 1$ ,

$$\begin{aligned} \psi_v - \psi_v(\alpha_r) & \geq (1 - \varepsilon)L_N^{\gamma(1+4\varepsilon)}(1) - (1 - \varepsilon)L_N^*(\alpha_r) \\ & = (1 - \varepsilon)(1 + 4\varepsilon)\gamma \log N^2 - (1 - \varepsilon)\mathcal{J}_{\sigma^2/\bar{\sigma}}(\alpha_r) \log N^2 \\ & \geq (1 + \varepsilon)\gamma \log N^2 \end{aligned} \tag{3.33}$$

where we take  $K$  large enough that  $(1 - \varepsilon)\mathcal{J}_{\sigma^2/\bar{\sigma}}(\alpha_r) < \varepsilon\gamma$  and use (3.30) to obtain the last inequality in (3.33). The distribution of  $\psi_v(\alpha_r)$  is symmetric, so the second probability in (3.32) is smaller than

$$\mathbb{P}\left(\max_{v \in V_N^\delta} \psi_v(\alpha_r) \geq \varepsilon\gamma \log N^2\right) \leq N^2 \exp\left(-\frac{(\varepsilon\gamma)^2 \log N^2}{\max_i \sigma_i^2 \alpha_r}\right)$$

where we used a union bound, a Gaussian estimate and (3.2) to get the inequality. This is  $O(N^{-c})$  by choosing  $K$  large enough for a fixed  $\varepsilon$  and  $r$ . Therefore, we have (3.31) as claimed.

*Proof of Lemma 3.5:* In view of (3.31), it suffices to show that  $\mathbb{P}(\mathcal{N}_\varepsilon^\gamma \geq N^{2\varepsilon\gamma - \tilde{\varepsilon}}) = 1 - O(N^{-c})$ . The Paley-Zygmund inequality (Lemma A.8) implies

$$\mathbb{P}(\mathcal{N}_\varepsilon^\gamma \geq N^{2\varepsilon\gamma - \tilde{\varepsilon}}) \geq \left(1 - \frac{N^{2\varepsilon\gamma - \tilde{\varepsilon}}}{\mathbb{E}[\mathcal{N}_\varepsilon^\gamma]}\right)^2 \frac{(\mathbb{E}[\mathcal{N}_\varepsilon^\gamma])^2}{\mathbb{E}[(\mathcal{N}_\varepsilon^\gamma)^2]}. \tag{3.34}$$

First, we make sure that  $N^{2\varepsilon\gamma - \tilde{\varepsilon}}/\mathbb{E}[\mathcal{N}_\varepsilon^\gamma] \rightarrow 0$  as  $N \rightarrow \infty$ . By independence of the increments and the variance estimate (3.28), Gaussian estimates yield for some constant  $c = c(\gamma, \varepsilon, \delta) > 0$  :

$$\begin{aligned} \mathbb{E}[\mathcal{N}_\varepsilon^\gamma] & = \sum_{v \in V_N^\delta} \mathbb{P}(A_v) = \sum_{v \in V_N^\delta} \prod_{j=r+1}^K \mathbb{P}\left(\nabla \psi_v(\alpha_j) \geq \nabla L_{N,\varepsilon}^\gamma(\alpha_j)\right) \\ & \geq c \cdot (\log N)^{-\frac{1}{2}(K-r)} N^{2(1-(1-\varepsilon)^2)+2(1-\varepsilon)^2\varepsilon\gamma(1+4\varepsilon)+2(1-\varepsilon)^2 \int_0^{\alpha_r} \frac{\sigma^2(s)}{\bar{\sigma}^2(s)} ds} \\ & \geq N^{2(1-(1-\varepsilon)^2)+2(1-\varepsilon)^2\varepsilon\gamma(1+4\varepsilon)}. \end{aligned} \tag{3.35}$$

By the definition of  $\mathcal{E}_\gamma$  in Theorem 1.3, and because  $\gamma^{l-1} \leq \gamma < \gamma(1+4\varepsilon) < \gamma^l$ ,

$$\begin{aligned} |\mathcal{E}_{\gamma(1+4\varepsilon)} - \mathcal{E}_\gamma| &= \frac{(\gamma(1+4\varepsilon) - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))^2 - (\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))^2}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} \\ &= \frac{16\varepsilon^2\gamma^2 + 8\varepsilon\gamma(\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} \leq \frac{12(\gamma^*)^2}{\bar{\sigma}_m^2 \nabla \lambda^m} \varepsilon \doteq \tilde{\varepsilon}/2 \end{aligned} \quad (3.36)$$

where we used  $\varepsilon < 1/4$ ,  $\gamma < \gamma^*$  and  $\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1) \geq \mathcal{J}_{\sigma^2}(\lambda^{m-1}, 1) = \bar{\sigma}_m^2 \nabla \lambda^m$  to obtain the inequality. By inserting the bound (3.36) in (3.35), we get

$$\mathbb{E}[\mathcal{N}_\varepsilon^\gamma] \geq N^{2(1-(1-\varepsilon)^2)+2(1-\varepsilon)^2(\mathcal{E}_\gamma-\tilde{\varepsilon}/2)} = N^{2\mathcal{E}_\gamma-\tilde{\varepsilon}} N^{2(1-(1-\varepsilon)^2)(1-\mathcal{E}_\gamma+\tilde{\varepsilon}/2)}.$$

Since  $(1-\varepsilon)^2 < 1$  and  $\mathcal{E}_\gamma \leq 1$ , it proves the assertion that  $N^{2\mathcal{E}_\gamma-\tilde{\varepsilon}}/\mathbb{E}[\mathcal{N}_\varepsilon^\gamma] \rightarrow 0$  and also justify the use of the Paley-Zygmund inequality. In view of (3.34), it suffices to show, like in Lemma 3.3, that

$$\mathbb{E}[(\mathcal{N}_\varepsilon^\gamma)^2] \leq (1 + O(N^{-\frac{1}{2K}(1-(1-\varepsilon)^2)})) (\mathbb{E}[\mathcal{N}_\varepsilon^\gamma])^2$$

to prove the lemma. The proof is almost identical to the proof of Lemma 3.3. Indeed, by Gaussian estimates and the variance estimates in (3.28), the probabilities on the increments in  $A_v$  are for every  $j$  and  $v \in V_N^\delta$ :

$$\tilde{p}_{v,j} \doteq \mathbb{P}\left(\nabla\psi_v(\alpha_j) \geq \nabla L_{N,\varepsilon}^\gamma(\alpha_j)\right) \asymp \frac{\tilde{e}_j}{\sqrt{\log N}}$$

where the  $\tilde{e}_j$ 's are the corresponding exponential factors. The proof is exactly the same up to (3.19) with  $\tilde{e}_j$ 's instead of  $e_j$ 's. From there, the third factor in the decomposition is still bounded by 1 because of property (3.29), and the rest of the argument follows if we choose  $K$  large enough for a fixed  $\varepsilon$  and  $\gamma$ . This ends the proof of the lemma.  $\square$

### Appendix A. Technical lemmas

The Markov property of the GFF, which is a consequence of the strong Markov property of the simple random walk (in the covariance function in (1.1)), implies that the value of the field inside a neighborhood is independent of the field outside given the boundary, see e.g. Dynkin (1980). In particular, for the neighborhood  $[v]_\lambda$ , where  $\lambda \in [0, 1]$ , this implies

$$\phi_v(\lambda) \doteq \mathbb{E}\left[\phi_v \mid \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}\right] = \mathbb{E}\left[\phi_v \mid \mathcal{F}_{\partial[v]_\lambda}\right]. \quad (A.1)$$

Let  $v, v' \in V_N$ ,  $\lambda < \lambda'$  and  $\mu < \mu'$ . Another direct consequence is the fact that for  $\lambda, \mu > \rho(v, v')$  or  $\lambda > \rho(v, v') > \mu'$ ,

$$\phi_v(\lambda') - \phi_v(\lambda) \text{ is independent of } \phi_{v'}(\mu') - \phi_{v'}(\mu). \quad (A.2)$$

This is because the shell  $[v]_\lambda \cap [v]_{\lambda'}^c$  does not intersect the shell  $[v']_\mu \cap [v']_{\mu'}^c$  in both cases, see Figure 2.2. We stress that, in general, the field  $\psi$  does not have the Markov property. However, by working with increments of the field  $\psi$ , the property analogous to (A.2) can be proved.

**Lemma A.1.** *Let  $v, v' \in V_N$ ,  $\lambda < \lambda'$  and  $\mu < \mu'$ . If we have  $\lambda, \mu > \rho(v, v')$  or  $\lambda > \rho(v, v') > \mu'$ , then*

$$\psi_v(\lambda') - \psi_v(\lambda) \text{ is independent of } \psi_{v'}(\mu') - \psi_{v'}(\mu).$$

*Proof:* Let  $v \in V_N$  and  $\lambda < \lambda'$ . By Definition 1.1 of the field  $\psi$  and its conditional expectation, we have

$$\psi_v(\lambda) = \sum_{1 \leq i \leq M} \sigma_i \mathbb{E} \left[ \nabla \phi_v(\lambda_i) \mid \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c} \right] = \sum_{\substack{1 \leq i \leq M: \\ \lambda_{i-1} < \lambda}} \sigma_i (\phi_v(\lambda \wedge \lambda_i) - \phi_v(\lambda_{i-1})). \quad (\text{A.3})$$

For the last equality, note that, when  $\lambda_{i-1} < \lambda$ , the increments  $\phi_v(\lambda \wedge \lambda_i) - \phi_v(\lambda_{i-1})$  are linear combinations of variables inside the set  $\partial[v]_\lambda \cup [v]_\lambda^c$  and, when  $\lambda_i > \lambda$ , we have  $\mathbb{E}[\phi_v(\lambda_i) - \phi_v(\lambda \vee \lambda_{i-1}) \mid \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}] = 0$  by the tower property of conditional expectations. By applying the same argument to  $\psi_v(\lambda')$ , we get

$$\psi_v(\lambda') - \psi_v(\lambda) = \sum_{\substack{1 \leq i \leq M: \\ \lambda \leq \lambda_{i-1} < \lambda' \text{ or } \lambda < \lambda_i \leq \lambda' \\ \text{or } \lambda_{i-1} \leq \lambda < \lambda' \leq \lambda_i}} \sigma_i (\phi_v(\lambda' \wedge \lambda_i) - \phi_v(\lambda \vee \lambda_{i-1})). \quad (\text{A.4})$$

The conclusion of the lemma follows from (A.2).  $\square$

In the remainder of this section, we always assume, without loss of generality, that  $N = 2^n$  for some  $n \in \mathbb{N}$  and  $\lambda n, \lambda' n, \lambda_i n \in \mathbb{N}_0$  for all  $i \in \{0, \dots, M\}$ .

**Lemma A.2.** *Let  $\delta \in (0, 1/2]$  and  $\lambda_{i-1} \leq \lambda < \lambda' \leq \lambda_i$  for some  $i \in \{1, \dots, M\}$ , then*

$$-C_1(\delta, \sigma_i) \leq \mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) - (\lambda' - \lambda) \sigma_i^2 \log N \leq C_2(\sigma_i) \quad (\text{A.5})$$

for all  $v \in V_N^\delta$  and  $N$  large enough depending on  $\delta$ . The constant  $C_1$  only depends on  $\delta$  when  $\lambda = 0$ .

*Proof:* The Markov property (A.1) yields  $\mathbb{E}[\phi_v - \phi_v(\lambda) \mid \mathcal{F}_{\partial[v]_{\lambda'}}] = \phi_v(\lambda') - \phi_v(\lambda)$ . Using the conditional variance formula and  $\mathbb{V}(X \mid \mathcal{F}) \doteq \mathbb{E}[(X - \mathbb{E}[X \mid \mathcal{F}])^2 \mid \mathcal{F}]$ , we can compute the variance of (A.4) in the special case  $\lambda_{i-1} \leq \lambda < \lambda' \leq \lambda_i$ :

$$\begin{aligned} \mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) &= \sigma_i^2 \mathbb{V}(\mathbb{E}[\phi_v - \phi_v(\lambda) \mid \mathcal{F}_{\partial[v]_{\lambda'}}]) \\ &= \sigma_i^2 (\mathbb{V}(\phi_v - \phi_v(\lambda)) - \mathbb{E}[\mathbb{V}(\phi_v - \phi_v(\lambda) \mid \mathcal{F}_{\partial[v]_{\lambda'}})]) \\ &= \sigma_i^2 (\mathbb{V}(\phi_v - \phi_v(\lambda)) - \mathbb{V}(\phi_v - \phi_v(\lambda'))). \end{aligned} \quad (\text{A.6})$$

But, it is well known that  $\{\phi_u - \mathbb{E}[\phi_u \mid \mathcal{F}_{\partial B}]\}_{u \in B}$  is a GFF when  $B \subseteq \mathbb{Z}^2$  is a finite box, see e.g. Zeitouni (2014). Simply choose  $B = [v]_s$ ,  $s = \lambda, \lambda'$ , in (A.6), then by the variance definition in (1.1),

$$\mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) = \sigma_i^2 (G_{[v]_\lambda}(v, v) - G_{[v]_{\lambda'}}(v, v)). \quad (\text{A.7})$$

Using standard estimates for the discrete Green function, we can now evaluate the last expression. For every finite box  $B \subseteq \mathbb{Z}^2$ , Proposition 1.6.3 of Lawler (1991) shows that (keeping in mind our choice of normalization by  $\pi/2$  in (1.1)):

$$G_B(x, y) = \left[ \sum_{z \in \partial B} \mathcal{P}_x(W_{\tau_{\partial B}} = z) a(z - y) \right] - a(y - x), \quad x, y \in B, \quad (\text{A.8})$$

where

$$a(w) = \begin{cases} \log(\|w\|_2) + \text{const.} + O(\|w\|_2^{-2}) & \text{if } w \in \mathbb{Z}^2 \setminus \{\mathbf{0}\} \\ 0 & \text{if } w = \mathbf{0} \end{cases} \quad (\text{A.9})$$

and  $\mathcal{P}_x$  is the law of the simple random walk starting at  $x \in \mathbb{Z}^2$ . Using (A.8), we can rewrite the difference of Green functions in (A.7) as

$$\sum_{z \in \partial[v]_\lambda} \mathcal{P}_v \left( W_{\tau_{\partial[v]_\lambda}} = z \right) a(z-v) - \sum_{z \in \partial[v]_{\lambda'}} \mathcal{P}_v \left( W_{\tau_{\partial[v]_{\lambda'}}} = z \right) a(z-v). \quad (\text{A.10})$$

When  $\lambda' = 1$ , we have  $\|z-v\|_2 = 0$  for  $z \in \partial[v]_{\lambda'}$ . Otherwise, we assumed  $v \in V_N^\delta$ , so take  $N$  large enough (depending on  $\delta$ ) that  $[v]_{\lambda'}$  is not cut off by  $\partial V_N$ . We have  $\|z-v\|_2 \leq \sqrt{2}N^{1-\lambda}$  for  $z \in \partial[v]_\lambda$  in general and  $\|z-v\|_2 \geq \frac{1}{2}N^{1-\lambda'}$  for  $z \in \partial[v]_{\lambda'}$  when  $\lambda' \neq 1$ . We deduce the following bound on the variance in (A.7) :

$$\begin{aligned} \max_{v \in V_N^\delta} \mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) &\leq \sigma_i^2((1-\lambda) - (1-\lambda')) \log N + \sigma_i^2 C \\ &= (\lambda' - \lambda)\sigma_i^2 \log N + C_2(\sigma_i). \end{aligned}$$

Similarly, we have  $\|z-v\|_2 \geq \delta N$  for  $z \in \partial[v]_\lambda$  when  $\lambda = 0$ . Otherwise, take  $N$  large enough (depending on  $\delta$ ) that  $[v]_\lambda$  is not cut off by the boundary of  $V_N$ . We have  $\|z-v\|_2 \geq \frac{1}{2}N^{1-\lambda}$  for  $z \in \partial[v]_\lambda$  when  $\lambda \neq 0$  and  $\|z-v\|_2 \leq \frac{1}{\sqrt{2}}N^{1-\lambda'}$  for  $z \in \partial[v]_{\lambda'}$  in general. We deduce the following bound on the variance in (A.7) :

$$\begin{aligned} \min_{v \in V_N^\delta} \mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) &\geq \sigma_i^2((1-\lambda) - (1-\lambda')) \log N - \sigma_i^2 C(\delta) \\ &= (\lambda' - \lambda)\sigma_i^2 \log N - C_1(\delta, \sigma_i). \end{aligned}$$

This ends the proof of the lemma.  $\square$

Since the upper bound in Lemma A.2 is only valid for  $N$  large enough depending on  $\delta$ , we cannot immediately conclude that it holds for all  $v \in V_N$ . We show in the next lemma how to extend the bound.

**Lemma A.3.** *Let  $\lambda_{i-1} \leq \lambda < \lambda' \leq \lambda_i$  for a certain  $i \in \{1, \dots, M\}$ , then*

$$\max_{v \in V_N} \mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) \leq (\lambda' - \lambda)\sigma_i^2 \log N + C(\sigma_i) \quad (\text{A.11})$$

for  $N$  large enough.

*Proof:* When  $v \in \partial V_N$ , the bound is trivial because  $\psi_v = 0$ . Therefore, let  $v \in V_N^o$ . To obtain the upper bound on the difference of Green functions in (A.7), we only used the fact that  $[v]_{\lambda'}$  was not cut off by  $\partial V_N$  for  $N$  large enough depending on  $\delta$ . Hence, we only need to show that when  $[v]_{\lambda'}$  is cut off, there exists  $u \in V_N^o$  such that  $[u]_{\lambda'}$  is not cut off and for which

$$G_{[v]_\lambda}(v, v) - G_{[v]_{\lambda'}}(v, v) \leq G_{[u]_\lambda}(u, u) - G_{[u]_{\lambda'}}(u, u) + \tilde{C}(\sigma_i). \quad (\text{A.12})$$

Assume that  $[v]_{\lambda'}$  is cut off by  $\partial V_N$  and choose  $u$  to be the center of  $V_N$ . Clearly, the neighborhood  $[u]_{\lambda'}$  is not cut off by the boundary of  $V_N$ . When  $\lambda' = 1$ , inequality (A.12) is trivial because  $G_{[v]_{\lambda'}}(v, v) = G_{[u]_{\lambda'}}(u, u) = 0$  and  $G_{[v]_\lambda}(v, v) \leq G_{[u]_\lambda}(u, u)$  since  $[v]_\lambda$  is cut off and  $[u]_\lambda$  is not. Now, assume  $\lambda' < 1$ . Denote  $\theta(x) \doteq x + u - v$  the translation function that moves  $v$  to  $u$ , see Figure A.4.

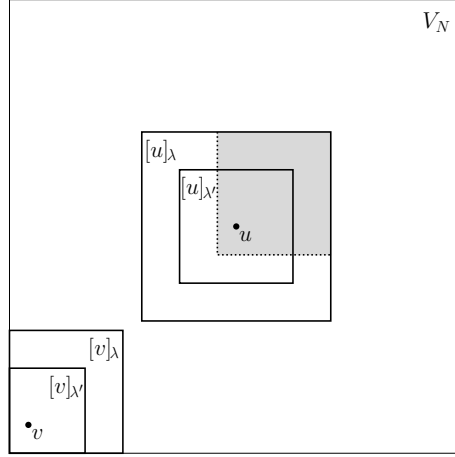


FIGURE A.4. The grey area  $\theta([v]_\lambda)$  is the translation of  $[v]_\lambda$ .

For the rest of the proof, redefine  $[v]_0$  as the square box of side length  $N$  centered at  $v$  that has been cut off by  $\partial V_N$ . Since  $\theta([v]_\lambda) \subseteq [u]_\lambda$ , we have

$$\begin{aligned} G_{[v]_\lambda}(v, v) - G_{[v]_{\lambda'}}(v, v) &= \frac{\pi}{2} \cdot \mathcal{E}_v \left[ \sum_{k=\tau_{\partial[v]_{\lambda'}}}^{\tau_{\partial[v]_\lambda}-1} \mathbf{1}_{\{W_k=v\}} \mathbf{1}_{\{\tau_{\partial[v]_{\lambda'}} < \tau_{\partial[v]_\lambda} \cap \partial V_N\}} \right] \\ &= \frac{\pi}{2} \cdot \mathcal{E}_u \left[ \sum_{k=\tau_{\partial[u]_{\lambda'}}}^{\tau_{\partial\theta([v]_\lambda)}-1} \mathbf{1}_{\{W_k=u\}} \mathbf{1}_{\{\tau_{\partial[u]_{\lambda'}} < \tau_{\partial\theta([v]_\lambda) \cap \partial V_N\}} \right] \\ &\leq \frac{\pi}{2} \cdot \mathcal{E}_u \left[ \sum_{k=\tau_{\partial[u]_{\lambda'}}}^{\tau_{\partial[u]_\lambda}-1} \mathbf{1}_{\{W_k=u\}} \right] = G_{[u]_\lambda}(u, u) - G_{[u]_{\lambda'}}(u, u) . \end{aligned}$$

This proves (A.12) when  $\lambda \neq 0$ . Since  $[v]_0 \doteq V_N$  throughout the article and we defined  $[v]_0$  differently in this proof, it remains to show that

$$\max_{v \in V_N} G_{V_N}(v, v) - G_{[v]_0}(v, v) \leq \tilde{C}(\sigma_i) \tag{A.13}$$

for (A.12) to be true when  $\lambda = 0$ . By the strong Markov property and (A.8) :

$$\begin{aligned} G_{V_N}(v, v) - G_{[v]_0}(v, v) &= \sum_{z \in \partial[v]_0 \cap V_N^o} \mathcal{P}_v \left( W_{\tau_{\partial[v]_0}} = z \right) G_{V_N}(z, v) \\ &= \sum_{z \in \partial[v]_0 \cap V_N^o} \mathcal{P}_v \left( W_{\tau_{\partial[v]_0}} = z \right) \sum_{w \in \partial V_N} \mathcal{P}_z \left( W_{\tau_{\partial V_N}} = w \right) (a(w - v) - a(v - z)) . \end{aligned}$$

But  $\|w - v\|_2 \leq \sqrt{2}N$  for all  $w \in \partial V_N$  and  $\|v - z\|_2 \geq \frac{1}{2}N$  for all  $z \in \partial[v]_0 \cap V_N^o$ . We get the desired conclusion using (A.9).  $\square$

In order to approximate the branching structure of the  $(\sigma, \lambda)$ -GFF in Lemma 3.1 and Lemma 3.4, we need to show that the variance of  $\psi_v(\lambda) - \psi_{v_\lambda}(\lambda)$  is bounded by a constant, where  $v_\lambda$  denotes any representative in  $R_\lambda$  that is closest to  $v$ . Our final goal here is to show Lemma A.6. We start by proving a more general version



of Lemma 12 found in Bolthausen et al. (2001). We define

$$\phi_v(A) \doteq \mathbb{E}[\phi_v \mid \mathcal{F}_{\partial(A \cap V_N)}]$$

and  $d(z, A) \doteq \min_{w \in A} \|z - w\|_2$  for any non-empty set  $A \subseteq \mathbb{Z}^2$ .

**Lemma A.4.** *Let  $B \subseteq \mathbb{Z}^2$  be a square box of width smaller or equal to  $N/2$  such that  $B \cap V_N \neq \emptyset$ . Moreover, let  $0 \leq \eta < 1$  and  $L \in \{1, 2, \dots, N/4\}$ , then there exists a constant  $C = C(\eta) > 0$  such that*

$$\max_{\substack{u, v \in B \cap V_N \\ d(u, \partial B) = L \\ \|u - v\|_2 \leq \eta L}} \mathbb{V}(\phi_u(B) - \phi_v(B)) \leq C. \tag{A.14}$$

*Proof:* Let  $u, v \in B \cap V_N$  be such that  $d(u, \partial B) = L$  and  $\|u - v\|_2 \leq \eta L$ . Denote  $\bar{B} \doteq B \cap V_N$ . Using the conditional variance formula as in (A.6), we have

$$\begin{aligned} \mathbb{V}(\phi_u(B) - \phi_v(B)) &= \mathbb{V}(\mathbb{E}[\phi_u - \phi_v \mid \mathcal{F}_{\partial \bar{B}}]) \\ &= \mathbb{V}(\phi_u - \phi_v) - \mathbb{E}[\mathbb{V}(\phi_u - \phi_v \mid \mathcal{F}_{\partial \bar{B}})] \\ &= (G_{V_N}(u, u) - G_{V_N}(u, v) + G_{V_N}(v, v) - G_{V_N}(v, u)) \\ &\quad - (G_{\bar{B}}(u, u) - G_{\bar{B}}(u, v) + G_{\bar{B}}(v, v) - G_{\bar{B}}(v, u)). \end{aligned} \tag{A.15}$$

For this proof, redefine  $[u]_0$  as the square box of side length  $N$  centered at  $u$  that has been cut off by  $\partial V_N$ . From (A.13), we know  $\max_{u \in V_N} G_{V_N}(u, u) - G_{[u]_0}(u, u) \leq C$ . Using the exact same method, we can also easily show that

$$\max_{\substack{v \in V_N \\ \|u - v\|_2 \leq \eta N/2}} G_{V_N}(v, v) - G_{[u]_0}(v, v) \leq C(\eta)$$

because we would have  $\|v - z\|_2 \geq (1 - \eta)N/2$  for all  $z \in \partial[u]_0 \cap V_N^o$  in the reasoning below (A.13), where  $\eta < 1$  by hypothesis. Finally,  $-G_{V_N}(u, v) \leq -G_{[u]_0}(u, v)$ , so proving (A.14) boils down to the proof of the following inequality :

$$(\clubsuit) \doteq \left\{ (G_{[u]_0}(u, u) - G_{[u]_0}(u, v)) - (G_{\bar{B}}(u, u) - G_{\bar{B}}(u, v)) \right. \\ \left. + (G_{[u]_0}(v, v) - G_{[u]_0}(v, u)) - (G_{\bar{B}}(v, v) - G_{\bar{B}}(v, u)) \right\} \leq \tilde{C}(\eta). \tag{A.16}$$

To show (A.16), we consider two cases :  $d(u, \partial V_N) \leq L$  and  $d(u, \partial V_N) > L$ .

Case 1 :  $d(u, \partial V_N) \leq L$

Since  $\bar{B} \subseteq [u]_0$  (recall that  $B$  is a square box of width smaller or equal to  $N/2$  and contains  $u$ ), then we always have

$$(\clubsuit) \leq (G_{[u]_0}(u, u) - G_{\bar{B}}(u, u)) + (G_{[u]_0}(v, v) - G_{\bar{B}}(v, v)). \tag{A.17}$$

Note that the box  $B$  is cut off by  $\partial V_N$  in Case 1. By translating  $u, v, B$  together in such a way that  $u$  doesn't get closer to  $\partial V_N$  with respect to both axes, each difference of Green functions in (A.17) can only increase (see the argument below Figure A.4). Therefore, it is sufficient to bound (A.17) when  $d(u, \partial V_N) = L$ . Assume  $d(u, \partial V_N) = L$  for the rest of Case 1. Since  $d(u, \partial B) = L$  by hypothesis, we have  $d(u, \partial \bar{B}) = L$  and we get  $d(v, \partial \bar{B}) \geq \lceil (1 - \eta)L \rceil \geq 1$  by the triangle inequality. Consequently,

$$(\clubsuit) \leq G_{[u]_0}(u, u) + G_{[u]_0}(v, v) - 2 \log L + C(\eta) \tag{A.18}$$

using (A.8) and (A.9).

By the symmetries of the square, we can assume, without loss of generality, that the minimum in  $d(u, \partial V_N) = L$  is achieved on the bottom edge of  $V_N$  (which lies on the  $x$ -axis). Define the half-space  $\mathcal{H} \doteq \{z = (z_1, z_2) \in \mathbb{Z}^2 \mid z_2 \geq 0\}$ . Since we have  $[u]_0 \subseteq \mathcal{H}$  and  $d(v, \partial \mathcal{H}) \leq (1 + \eta)L$ , by the triangle inequality, then

$$(\clubsuit) \leq 2 \max_{\substack{z \in \mathcal{H} \\ \lceil(1-\eta)L \rceil \leq d(z, \partial \mathcal{H}) \leq (1+\eta)L}} G_{\mathcal{H}}(z, z) - 2 \log L + C(\eta) . \tag{A.19}$$

From Proposition 8.1.1 of [Lawler and Limic \(2010\)](#),

$$G_{\mathcal{H}}(z, z) = a(z - \bar{z}) \stackrel{(A.9)}{=} \log(\|z - \bar{z}\|_2) + \text{const.} + O(\|z - \bar{z}\|_2^{-2}) \tag{A.20}$$

where  $z = (z_1, z_2)$  and  $\bar{z} \doteq (z_1, -z_2)$ . The conclusion for Case 1 follows from (A.19) because  $2 \leq 2\lceil(1 - \eta)L\rceil \leq \|z - \bar{z}\|_2 \leq 2(1 + \eta)L$  in (A.20).

Case 2 :  $d(u, \partial V_N) > L$

For Case 2, we follow the argument from [Bolthausen et al. \(2001\)](#). We give the details for convenience. For all  $k \in \mathbb{N}_0$ , define  $[u]_0^k \subseteq \mathbb{Z}^2$  the square box of side length  $2^k N$  centered at  $u$  (not cut off by anything). For instance,  $[u]_0 = [u]_0^0 \cap V_N$  in this proof. Note that  $[u]_0 \subseteq [u]_0^1 \subseteq [u]_0^2 \subseteq \dots$  and  $[u]_0 \cup \bigcup_{k=1}^{\infty} [u]_0^k = \mathbb{Z}^2$ , so

$$\begin{aligned} (\clubsuit) &\leq \left\{ (G_{[u]_0}(u, u) - G_{[u]_0}(u, v)) - (G_{\bar{B}}(u, u) - G_{\bar{B}}(u, v)) \right\} \\ &\quad + \left\{ (G_{[u]_0}(v, v) - G_{[u]_0}(v, u)) - (G_{\bar{B}}(v, v) - G_{\bar{B}}(v, u)) \right\} \\ &\quad + \sum_{k=1}^{\infty} \left\{ (G_{[u]_0^k}(u, u) - G_{[u]_0^k}(u, v)) - (G_{[u]_0^{k-1}}(u, u) - G_{[u]_0^{k-1}}(u, v)) \right\} \\ &\quad + \left\{ (G_{[u]_0^k}(v, v) - G_{[u]_0^k}(v, u)) - (G_{[u]_0^{k-1}}(v, v) - G_{[u]_0^{k-1}}(v, u)) \right\} \\ &= \frac{\pi}{2} \cdot \mathcal{E}_u \left[ \sum_{k=\tau_{\partial \bar{B}}}^{\infty} (1_{\{W_k=u\}} - 1_{\{W_k=v\}}) \right] + \frac{\pi}{2} \cdot \mathcal{E}_v \left[ \sum_{k=\tau_{\partial \bar{B}}}^{\infty} (1_{\{W_k=v\}} - 1_{\{W_k=u\}}) \right]. \end{aligned}$$

The inequality comes from the fact that each pair of braces in the infinite sum is equal to  $\mathbb{V}_{[u]_0^k}(\mathbb{E}[\phi_u - \phi_v \mid \mathcal{F}_{\partial [u]_0^{k-1}}]) \geq 0$  by steps analogous to (A.15). The equality follows because the infinite sum is telescopic.

By conditioning on the point  $z \in \partial \bar{B}$  where the simple random walk starting at  $u$  or  $v$  will be when hitting the boundary of  $\bar{B}$ , and using the strong Markov property, we deduce

$$\begin{aligned} (\clubsuit) &\leq \sum_{z \in \partial \bar{B}} (\mathcal{P}_u(W_{\tau_{\partial \bar{B}}}=z) - \mathcal{P}_v(W_{\tau_{\partial \bar{B}}}=z)) \cdot \frac{\pi}{2} \cdot \mathcal{E}_z \left[ \sum_{k=0}^{\infty} (1_{\{W_k=u\}} - 1_{\{W_k=v\}}) \right] \\ &= \sum_{z \in \partial \bar{B}} (\mathcal{P}_u(W_{\tau_{\partial \bar{B}}}=z) - \mathcal{P}_v(W_{\tau_{\partial \bar{B}}}=z)) \cdot (a(v - z) - a(u - z)) \tag{A.21} \end{aligned}$$

where “ $a$ ”, the *potential kernel* (see p.37 in [Lawler, 1991](#)), is defined by

$$a(w) \doteq \frac{\pi}{2} \cdot \mathcal{E}_0 \left[ \sum_{k=0}^{\infty} (1_{\{W_k=0\}} - 1_{\{W_k=w\}}) \right].$$

Theorem 1.6.2 in [Lawler \(1991\)](#) shows that this is the same function as in (A.9). Therefore, we can evaluate (A.21) :

$$a(v - z) - a(u - z) = \log \left( \frac{\|v - z\|_2}{\|u - z\|_2} \right) + O(\|v - z\|_2^{-2}) - O(\|u - z\|_2^{-2}) . \tag{A.22}$$

By the triangle inequality, we have

$$\log \left( 1 - \frac{\|u - v\|_2}{\|u - z\|_2} \right) \leq \log \left( \frac{\|v - z\|_2}{\|u - z\|_2} \right) \leq \log \left( 1 + \frac{\|u - v\|_2}{\|u - z\|_2} \right). \tag{A.23}$$

Now, notice that

- $\|u - v\|_2 \leq \eta L$  by hypothesis ;
- $\|u - z\|_2 \geq L$  for all  $z \in \partial \bar{B}$  by the assumption of Case 2 ;
- $\|v - z\|_2 \geq \|u - z\|_2 - \|u - v\|_2 \geq \lceil (1 - \eta)L \rceil$  for all  $z \in \partial \bar{B}$ , from the first two bullets and the triangle inequality.

Using the three bullets in (A.22) and (A.23), we have

$$\log(1 - \eta) - \frac{C_1}{\lceil (1 - \eta)L \rceil^2} \leq \text{(A.22)} \leq \log(1 + \eta) + \frac{C_2}{\lceil (1 - \eta)L \rceil^2} \tag{A.24}$$

for appropriate constants  $C_1, C_2 > 0$ . Since  $L \geq 1$  and  $\lceil (1 - \eta)L \rceil \geq 1$ , inequality (A.16) follows by regrouping (A.21), (A.22) and (A.24).  $\square$

**Lemma A.5.** *Let  $0 \leq \lambda' < 1$  and  $d \geq 1/\sqrt{2}$ . For all  $v \in V_N$ , define  $S_{v,d}$  to be the set of finite boxes  $B \subseteq \mathbb{Z}^2$  such that  $[v]_{\lambda'} \subseteq B \cap V_N$  and  $\max_{z \in \partial B} \|v - z\|_2 \leq dN^{1-\lambda'}$ , then there exists a constant  $C = C(d) > 0$  such that*

$$\max_{v \in V_N} \max_{B \in S_{v,d}} \mathbb{V}(\phi_v(\lambda') - \phi_v(B)) \leq C$$

for  $N$  large enough.

*Proof:* This follows directly from the calculations in Lemma A.2 and Lemma A.3 where  $B \cap V_N$  plays the same role as  $[v]_\lambda$ .  $\square$

The next lemma is used in equation (3.4) of Lemma 3.1 and equation (3.23) of Lemma 3.4 to show that the error coming from the approximation of the branching structure of  $\psi$  is small enough that the problem of finding the upper bound for the maximum and the log-number of  $\gamma$ -high points is the same (modulo the additional hurdle caused by the decay of variance near the edges of  $V_N$ ) as in the context of branching random walks.

**Lemma A.6.** *Let  $\lambda_{j-1} < \lambda \leq \lambda_j$  for a certain  $j \in \{1, \dots, M\}$ , then there exists a constant  $C = C(\sigma_1, \dots, \sigma_j) > 0$  such that*

$$\max_{v \in V_N} \mathbb{V}(\psi_v(\lambda) - \psi_{v_\lambda}(\lambda)) \leq C$$

for  $N$  large enough.

*Proof:* The lemma is trivial when  $\lambda = 1$  since  $v = v_1$ . Therefore, assume  $0 < \lambda < 1$ . Choose  $v_\lambda \in R_\lambda$  any representative that is closest to  $v$  (there may be more than one). For all  $\mu \in (0, \lambda]$ , the square box  $B_\mu \subseteq \mathbb{Z}^2$  of width  $2\lceil N^{1-\mu} \rceil$  centered at  $v_\lambda$  contains both  $[v]_\mu$  and  $[v_\lambda]_\mu$  because  $\|v - v_\lambda\|_\infty \leq \frac{1}{2}N^{1-\lambda}$ . Then, by Jensen's inequality :

$$\mathbb{V}(\phi_v(\mu) - \phi_{v_\lambda}(\mu)) \leq 3 \cdot \left\{ \begin{array}{l} \mathbb{V}(\phi_v(\mu) - \phi_v(B_\mu)) \\ + \mathbb{V}(\phi_v(B_\mu) - \phi_{v_\lambda}(B_\mu)) \\ + \mathbb{V}(\phi_{v_\lambda}(B_\mu) - \phi_{v_\lambda}(\mu)) \end{array} \right\} \leq \tilde{C}. \tag{A.25}$$

To see the last inequality, bound the first and third variance term inside the braces using Lemma A.5 with  $d = 3/\sqrt{2}$  and bound the second variance term inside the braces using Lemma A.4 with  $\eta = 1/\sqrt{2}$  (since  $\|v - v_\lambda\|_2 \leq N^{1-\lambda}/\sqrt{2} \leq N^{1-\mu}/\sqrt{2}$ ),  $u = v_\lambda$  and  $L = \lceil N^{1-\mu} \rceil$ . Now, from (A.3) and Jensen's inequality, we get

$$\begin{aligned} \mathbb{V}(\psi_v(\lambda) - \psi_{v_\lambda}(\lambda)) &= \mathbb{V} \left( \begin{array}{l} \sigma_j(\phi_v(\lambda) - \phi_{v_\lambda}(\lambda)) \\ + \sum_{i=1}^{j-1} (\sigma_i - \sigma_{i+1})(\phi_v(\lambda_i) - \phi_{v_\lambda}(\lambda_i)) \end{array} \right) \\ &\leq j \cdot \left\{ \begin{array}{l} \sigma_j^2 \mathbb{V}(\phi_v(\lambda) - \phi_{v_\lambda}(\lambda)) \\ + \sum_{i=1}^{j-1} (\sigma_i - \sigma_{i+1})^2 \mathbb{V}(\phi_v(\lambda_i) - \phi_{v_\lambda}(\lambda_i)) \end{array} \right\}. \end{aligned}$$

Simply use (A.25) to bound each variance term inside the braces by a constant. This ends the proof of the lemma.  $\square$

**Lemma A.7** (Gaussian estimates, see e.g. [Adler and Taylor, 2007](#)). *Suppose that  $Z \sim \mathcal{N}(0, \sigma^2)$  where  $\sigma > 0$ , then for all  $z > 0$ ,*

$$\left(1 - \frac{\sigma^2}{z^2}\right) \frac{\sigma}{\sqrt{2\pi}z} \exp\left(-\frac{z^2}{2\sigma^2}\right) \leq \mathbb{P}(Z \geq z) \leq \frac{\sigma}{\sqrt{2\pi}z} \exp\left(-\frac{z^2}{2\sigma^2}\right).$$

**Lemma A.8** ([Paley and Zygmund, 1932](#) inequality). *Let  $0 \leq X \in L^2(\mathbb{P})$  be such that  $\mathbb{P}(X > 0) > 0$ , then for all  $0 \leq \theta \leq 1$ ,*

$$\mathbb{P}(X \geq \theta \mathbb{E}[X]) \geq (1 - \theta)^2 \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}.$$

## Appendix B. Karush-Kuhn-Tucker theorem and applications

In this section, we state the Karush-Kuhn-Tucker theorem and the solutions to the two optimization problems posed in Section 2. The optimal path for the maximum,  $\lambda \mapsto L_N^*(\lambda)$ , comes from the solution to the problem stated in Lemma B.2 while the optimal path for  $\gamma$ -high points,  $\lambda \mapsto L_N^\gamma(\lambda)$ , comes from the solution to the problem stated in Lemma B.3. The Karush-Kuhn-Tucker theorem only gives, a priori, necessary conditions for local optimality. However, the conditions are also sufficient for global optimality here because the objective function  $f_\gamma$  below and the constraint functions  $g_k$  are continuously differentiable and concave ( $f_{\gamma^*}$  is linear), see [Hanson \(1981\)](#). The proof of the two lemmas can be found in Appendix A of [Ouimet \(2014\)](#) and are direct applications of the theorem.

**Theorem B.1** (Karush-Kuhn-Tucker, see e.g. [Delfour, 2012](#)). *Let  $f : \mathbb{R}^{n_1} \rightarrow \mathbb{R}$  be an objective function and let*

$$\mathcal{U}^\geq \doteq \{\mathbf{x} \in \mathbb{R}^{n_1} \mid g_k(\mathbf{x}) \geq 0 \quad \forall k \in \{1, \dots, n_2\}\}$$

*be a set of constraints specified by the constraint functions  $g_k : \mathbb{R}^{n_1} \rightarrow \mathbb{R}$ ,  $1 \leq k \leq n_2$ . Furthermore, assume that*

- (a)  *$f$  attains a local maximum at  $\mathbf{x}^* \in \mathcal{U}^\geq$  with respect to  $\mathcal{U}^\geq$ ;*
- (b)  *$f$  is Fréchet differentiable at  $\mathbf{x}^*$ ;*
- (c) *the  $g_k$ 's are Fréchet differentiable at  $\mathbf{x}^*$ .*

*When the constraints qualify (they do in Lemma B.2 and Lemma B.3 because the  $g_k$ 's are concave and  $\mathbf{0} \in \mathcal{U}^\geq$ , see Slater's condition in [Delfour, 2012](#)), then there exists  $(\mu_1, \dots, \mu_{n_2}) \in \mathbb{R}^{n_2}$  such that the following points hold for all  $k \in \{1, \dots, n_2\}$  ( $\nabla$  is the gradient here) :*

- (1)  $\nabla f(\mathbf{x}^*) + \sum_{k=1}^{n_2} \mu_k \nabla g_k(\mathbf{x}^*) = 0$ ;
- (2)  $g_k(\mathbf{x}^*) \geq 0$ ;
- (3)  $\mu_k \geq 0$ ;
- (4)  $\mu_k g_k(\mathbf{x}^*) = 0$ .

**Lemma B.2.** (Optimal path for the maximum) Let

$$f_{\gamma^*}(x_1, \dots, x_M) \doteq \sum_{i=1}^M x_i$$

be the objective function to maximize under the constraints

$$g_k(x_1, \dots, x_M) \doteq \sum_{i=1}^k \left( \nabla \lambda_i - \frac{x_i^2}{\sigma_i^2 \nabla \lambda_i} \right) \geq 0, \quad 1 \leq k \leq M,$$

then there exists a unique global maximum. The solution is given by

$$x_i^* = \nabla \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda_i), \quad 1 \leq i \leq M,$$

and the maximum is given by

$$f_{\gamma^*}(x_1^*, \dots, x_M^*) = \mathcal{J}_{\sigma^2/\bar{\sigma}}(1) \doteq \gamma^*.$$

**Lemma B.3.** (Optimal path for  $\gamma$ -high points) Let  $\gamma^{l-1} \leq \gamma < \gamma^l$  for a certain  $l \in \{1, \dots, m\}$ , where the critical levels  $\gamma^l$  are defined in (1.5). Furthermore, let

$$f_{\gamma}(x_1, \dots, x_{M-1}) \doteq \sum_{i=1}^{M-1} \left( \nabla \lambda_i - \frac{x_i^2}{\sigma_i^2 \nabla \lambda_i} \right) + \left( \nabla \lambda_M - \frac{(\gamma - \sum_{i'=1}^{M-1} x_{i'})^2}{\sigma_M^2 \nabla \lambda_M} \right)$$

be the objective function to maximize under the constraints

$$g_k(x_1, \dots, x_{M-1}) \doteq \sum_{i=1}^k \left( \nabla \lambda_i - \frac{x_i^2}{\sigma_i^2 \nabla \lambda_i} \right) \geq 0, \quad 1 \leq k \leq M-1,$$

then there exists a unique global maximum. The solution is given by

$$x_i^* = \begin{cases} \nabla \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda_i) & \text{when } \lambda_i \leq \lambda^{l-1} \\ \frac{\nabla \mathcal{J}_{\sigma^2}(\lambda_i)}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} (\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1})) & \text{when } \lambda_i > \lambda^{l-1} \end{cases}$$

for all  $i \in \{1, \dots, M-1\}$  and the maximum is given by

$$f_{\gamma}(x_1^*, \dots, x_{M-1}^*) = (1 - \lambda^{l-1}) - \frac{(\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))^2}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} \doteq \mathcal{E}_{\gamma}.$$

## Acknowledgements

The authors would like to thank an anonymous referee for a careful review of the preliminary version. We also gratefully acknowledge insightful discussions with David Belius, Anton Bovier, Nicola Kistler and Olivier Zindy.

## References

- R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Monographs in Mathematics. Springer, New York (2007). ISBN 978-0-387-48112-8. [MR2319516](#).
- L.-P. Arguin and O. Zindy. Poisson-Dirichlet statistics for the extremes of the two-dimensional discrete Gaussian free field. *Electron. J. Probab.* **20**, no. 59, 19 (2015). [MR3354619](#).
- M. Biskup and O. Louidor. Extreme Local Extrema of Two-Dimensional Discrete Gaussian Free Field. *Comm. Math. Phys.* **345** (1), 271–304 (2016). [MR3509015](#).

- E. Bolthausen, J.-D. Deuschel and G. Giacomin. Entropic repulsion and the maximum of the two-dimensional harmonic crystal. *Ann. Probab.* **29** (4), 1670–1692 (2001). [MR1880237](#).
- A. Bovier and L. Hartung. The extremal process of two-speed branching Brownian motion. *Electron. J. Probab.* **19**, no. 18, 28 (2014). [MR3164771](#).
- A. Bovier and L. Hartung. Variable speed branching Brownian motion 1. Extremal processes in the weak correlation regime. *ALEA Lat. Am. J. Probab. Math. Stat.* **12** (1), 261–291 (2015). [MR3351476](#).
- A. Bovier and I. Kurkova. Derrida’s generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.* **40** (4), 481–495 (2004). [MR2070335](#).
- M. Bramson, J. Ding and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.* **69** (1), 62–123 (2016). [MR3433630](#).
- D. Capocaccia, M. Cassandro and P. Picco. On the existence of thermodynamics for the generalized random energy model. *J. Statist. Phys.* **46** (3-4), 493–505 (1987). [MR883541](#).
- O. Daviaud. Extremes of the discrete two-dimensional Gaussian free field. *Ann. Probab.* **34** (3), 962–986 (2006). [MR2243875](#).
- M. C. Delfour. *Introduction to optimization and semidifferential calculus*, volume 12 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2012). ISBN 978-1-611972-14-6. [MR2954022](#).
- B. Derrida. A generalization of the random energy model which includes correlations between energies. *J. Physique Lett.* **46** (9), 401–407 (1985).
- B. Duplantier and S. Sheffield. Liouville quantum gravity and KPZ. *Invent. Math.* **185** (2), 333–393 (2011). [MR2819163](#).
- E. B. Dynkin. Markov processes and random fields. *Bull. Amer. Math. Soc. (N.S.)* **3** (3), 975–999 (1980). [MR585179](#).
- M. Fang. Tightness for maxima of generalized branching random walks. *J. Appl. Probab.* **49** (3), 652–670 (2012). [MR3012090](#).
- M. Fang and O. Zeitouni. Branching random walks in time inhomogeneous environments. *Electron. J. Probab.* **17**, no. 67, 18 (2012a). [MR2968674](#).
- M. Fang and O. Zeitouni. Slowdown for time inhomogeneous branching Brownian motion. *J. Stat. Phys.* **149** (1), 1–9 (2012b). [MR2981635](#).
- M. A. Hanson. On sufficiency of the Kuhn-Tucker conditions. *J. Math. Anal. Appl.* **80** (2), 545–550 (1981). [MR614849](#).
- X. Hu, J. Miller and Y. Peres. Thick points of the Gaussian free field. *Ann. Probab.* **38** (2), 896–926 (2010). [MR2642894](#).
- N. Kistler. Derrida’s random energy models. From spin glasses to the extremes of correlated random fields. In *Correlated Random Systems: Five Different Methods*, volume 2143 of *Lecture Notes in Mathematics*. Springer (2015). [MR3380419](#).
- N. Kistler and M. A. Schmidt. From Derrida’s random energy model to branching random walks: from 1 to 3. *Elec. Comm. Prob.* **20**, no. 47, 12 (2015). [MR3358969](#).
- G. F. Lawler. *Intersections of random walks*. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA (1991). ISBN 0-8176-3557-2. [MR1117680](#).
- G. F. Lawler and V. Limic. *Random walk: a modern introduction*, volume 123 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (2010). ISBN 978-0-521-51918-2. [MR2677157](#).

- P. Maillard and O. Zeitouni. Slowdown in branching Brownian motion with inhomogeneous variance. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (3), 1144–1160 (2016). [MR3531703](#).
- B. Mallein. Maximal displacement of a branching random walk in time-inhomogeneous environment. *Stoch. Proc. Appl.* **125** (10), 3958–4019 (2015). [MR3373310](#).
- F. Ouimet. *Étude du maximum et des hauts points de la marche aléatoire branchante inhomogène et du champ libre gaussien inhomogène*. Master's thesis, Université de Montréal (2014). URL <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/11510>.
- F. Ouimet. Maxima of branching random walks with piecewise constant variance. *ArXiv Mathematics e-prints* (2015). [arXiv: 1509.08172](#).
- R. Paley and A. Zygmund. A note on analytic functions in the unit circle. *Math. Proc. Cambridge Philos. Soc.* **28** (3), 266–272 (1932). DOI: [10.1017/S0305004100010112](https://doi.org/10.1017/S0305004100010112).
- S. Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields* **139** (3-4), 521–541 (2007). [MR2322706](#).
- O. Zeitouni. Gaussian fields. *Notes for Lectures* (2014). URL <http://www.cims.nyu.edu/~zeitouni/notesGauss.pdf>.