A LIMIT LAW FOR THE MAXIMUM OF SUBCRITICAL DG-MODEL ON A HIERARCHICAL LATTICE

MAREK BISKUP AND HAIYU HUANG

Department of Mathematics, UCLA, Los Angeles, California, USA

ABSTRACT. We study the extremal properties of the "integer-valued Gaussian" a.k.a. DG-model on the hierarchical lattice $\Lambda_n := \{1,\dots,b\}^n$ (with $b \geqslant 2$) of depth n. This is a random field $\varphi \in \mathbb{Z}^{\Lambda_n}$ with law proportional to $\mathrm{e}^{\frac12\beta(\varphi,\Delta_n\varphi)}\prod_{x\in\Lambda_n}\#(\mathrm{d}\varphi_x)$, where Δ_n is the hierarchical Laplacian, β is the inverse temperature and # is the counting measure on $\mathbb{Z}$. Denoting $\beta_{\mathrm{c}} := 2\pi^2/\log b$ and $m_n := \beta^{-1/2}[(2\log b)^{1/2}n - \frac32(2\log b)^{-1/2}\log n]$, for $0 < \beta < \beta_{\mathrm{c}}$ we prove that, along increasing sequences of n such that the fractional part of m_n converges to an $s \in [0,1)$, the centered maximum $\max_{x\in\Lambda_n}\varphi_x - \lfloor m_n\rfloor$ tends (as $n\to\infty$) in law to a discrete variant of a randomly shifted Gumbel law with the shift depending non-trivially on s. The convergence extends to the extremal process whose law tends to a decorated Poisson point process with a random intensity measure. The proofs rely on renormalization-group analysis which enables a tight coupling of the DG-model to a Gaussian Free Field. The interval $(0,\beta_{\mathrm{c}}]$ marks the full range of values of β for which the renormalization-group iterations tend to a "trivial" fixed point.

1. INTRODUCTION

1.1 Background.

The extremal properties of logarithmically correlated random fields have been a subject of considerable interest in recent years. A picture that has emerged from the analysis of specific examples is that, in these systems, the suitably centered maximum tends in law to a randomly-shifted Gumbel random variable while the associated extremal process tends to a decorated (Gumbel) Poisson point process with a random intensity measure. The randomness of the shift/intensity arises from a dependence structure that, in these systems, persists all the way up to the macroscopic scale.

The salient examples where such limit theorems have been proved include the Branching Brownian motion (Arguin, Bovier and Kistler [7–9], Aïdekon, Berestycki, Brunet and Shi [5]), critical Branching Random Walks (Aïdekon [4], Madaule [37]) and the Gaussian Free Field on $\mathbb{Z}^2$ (Bramson, Ding and Zeitouni [23], Biskup and Louidor [18–20]). Further evidence of universality has arrived in the studies of more general logarithmically correlated Gaussian processes (Madaule [36], Ding, Roy and Zeitouni [25], Schweiger and Zeitouni [42]) including the four-dimensional membrane model (Schweiger [41]). Non-Gaussian processes have been treated as well, e.g., the characteristic polynomial of a random matrix ensemble (Paquette and Zeitouni [40]), the local time of simple random walk on a regular tree (Biskup and Louidor [22], Abe and Biskup [2]) and a class of $P(\varphi)_2$-models on a torus (Barashkov, Gunaratnam and Hofstetter [10]).

The goal of the present paper is to address the limit law of the maximum for a process that distinguishes itself from the above by the field taking only integer values. Specifically, we are interested in the extremal properties of the model of a random interface

© 2023 M. Biskup, H. Huang. Reproduction for non-commercial use is permitted.

that is referred to as the "integer-valued Gaussian field" or the "DG-model" (with DG standing for "Discrete Gaussian") in the literature.

Generally, the DG-model is a $\mathbb{Z}$-valued process $\{\varphi_x\colon x\in\Lambda\}$ indexed by vertices of a finite graph Λ whose law takes the form

$$\frac{1}{\Sigma(\beta)}\mathrm{e}^{\frac12\beta(\varphi,\Delta\varphi)}\prod_{x\in\Lambda}\#(\mathrm{d}\varphi_x),\tag{1.1}$$

where Δ is the Laplacian associated with Λ, the parameter $\beta>0$ is the inverse temperature, # is the counting measure on $\mathbb{Z}$ and $\Sigma(\beta)$ is a normalization constant. The inner product $(\cdot,\cdot)$ is that in $\ell^2(\Lambda)$. A suitable boundary condition must be imposed for the Laplacian to ensure that the normalization constant is finite.

When the counting measure is replaced by the Lebesgue measure in (1.1), the resulting law is that of the $1/\sqrt{\beta}$-multiple of the Gaussian Free Field (GFF); we can thus think of the DG-model as the GFF conditioned on taking integer values. (This is the reason why the word Gaussian appears in the name of this model, which is otherwise not Gaussian at all.) The GFF is a well studied process so a natural question is then: Under what conditions is the GFF a good approximation to the DG-model?

As it turns out, the answer to this varies depending on the underlying graph and the parameter β. This is witnessed particularly by the most interesting case of Λ being a large box in $\mathbb{Z}^2$ where the DG-model exhibits a so-called roughening transition (first established by Fröhlich and Spencer [29]): For β small the DG-model is asymptotically GFF-like at large spatial scales (Wirth [43], Bauerschmidt, Park and Rodriguez [13, 14]) but not so at all for β large (Lubetzky, Martinelli and Sly [35]). When characterized by infinite vs finite limit variance of the field, the two regimes meet at a single value of β (Lammers [34], Aizenman, Harel, Peled and Shapiro [6]).

1.2 Overview.

While our prime interest rests with the behavior of the maximum and extremal values of the DG-model in finite subsets of $\mathbb{Z}^2$, extracting sharp results in the lattice model appears presently too hard. We will therefore resort to the hierarchical version of the DG-model that shares many important features with the model on $\mathbb{Z}^2$ but is more amenable to analysis. In a way, the hierarchical DG-model is the same approximation to the model on $\mathbb{Z}^2$ as the Branching Random Walk is to the lattice GFF. The understanding of the extremal behavior of Branching Random Walks has been instrumental for the corresponding results on the lattice GFF and we expect the same here as well.

Our control of the DG-model extends throughout the subcritical regime of inverse temperatures β. In the hierarchical model, this is defined as the regime in which the iterative approach we rely on converges to a "trivial" fixed point and the model thus behaves as the GFF at large spatial scales. In order to achieve this within the framework of the renormalization group theory, which is our main tool, we had to develop a novel way to control the iterations of the effective potential that avoids linearization.

Using the iterative analysis, we are then able to verify that the subcritical hierarchical DG-model belongs to the universality class of the continuum-valued GFF: the suitably centered maximum is asymptotically a randomly-shifted Gumbel and the extremal process is a (Gumbel) Poisson process with random intensity measure. The subcritical

DG-model is in fact so close to the GFF that the random shift/intensity is the *same* as for the GFF except for a deterministic correction that depends on the scale of the system and can be attributed to a "rounding error" caused by the integer-valued nature of the underlying field. Due to a scale-dependence of this correction, the stated limit laws only exist along particular subsequences.

Our analysis of the extremal behavior relies on a coupling of the DG-model to the GFF enabled by the iterative/renormalization group method. The idea to interpret renormalization iterations via coupling and use it to study extremal properties of the underlying field has appeared before (e.g., in [10] or in Bauerschmidt and Hofstetter [12] and Hofstetter [31]; see, however, Remark 3.7) but our approach here is different from these in that we derive the limit result directly from the corresponding conclusion for the GFF, rather than by mimicing a proof that did it for the GFF (which is what was done in [12] on which the relevant conclusion in [10,31] is based). In addition, we aim directly at the full extremal process.

2. Main results

Let us move to the specific details of the problem at hand, starting with definitions and statement of convergence of the maximum.

2.1 The model and convergence of the maximum.

Given naturals $n \geqslant 0$ and $b \geqslant 2$, the hierarchical DG-model is defined over the hierarchical lattice

$$\Lambda_n := \{1, \dots, b\}^n \tag{2.1}$$

of depth n. (For the connection with $\mathbb{Z}^d$ we take $b := L^d$ for some natural $L \geqslant 2$, in which case Λ_n can be identified with a box of side-length L^n.) The vertices of Λ_n are thus sequences taking values in $\{1, \dots, b\}$.

The hierarchical distance $d(x, y)$ between vertices $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ is the smallest $j \in \{1, \dots, n\}$ such that $x_i = y_i$ for all $i = 1, \dots, n - j$. This means that Λ_n splits, as a graph, into b copies of Λ_{n-1} that are distance one from one another, each of which then splits into b copies of Λ_{n-2} at unit distance from one another, etc. This geometric structure is the basis of iterative approaches to hierarchical models.

The hierarchical Laplacian Δ_n on Λ_n is naturally associated with the simple random walk on a b-ary tree of depth n killed when the walk exits the tree through the root. More precisely, Δ_n is the generator of the Markov chain induced by observing the walk only on the leaves of the tree which are naturally identified with Λ_n. The action of Δ_n on $f \colon \Lambda_n \to \mathbb{R}$ takes the explicit form

$$\Delta_n f(x) = \sum_{k=1}^{n} \Biggl[\biggl(\sum_{j=0}^{k-1} b^j \biggr)^{-1} - \biggl(\sum_{j=0}^{k} b^j \biggr)^{-1} \Biggr] \frac{1}{b^k} \sum_{y \in \mathcal{B}_k(x)} \bigl[f(y) - f(x) \bigr] - \biggl(\sum_{j=0}^{n} b^j \biggr)^{-1} f(x), \tag{2.2}$$

where $\mathcal{B}_k(x) := \{y \in \Lambda_n \colon d(x, y) \leqslant k\}$ is the set of vertices within distance k from x. As is readily checked, Δ_n is symmetric and, since the killing mechanism generates a non-trivial mass term, negative definite on $\ell^2(\Lambda_n)$.

The law of the hierarchical DG-model on Λ_n is a probability measure $P_{n,\beta}$ on $\mathbb{Z}^{\Lambda_n}$ that takes the form

$$P_{n,\beta}(\mathrm{d}\varphi) := \frac{1}{\Sigma_n(\beta)}\mathrm{e}^{\frac{1}{2}\beta(\varphi,\Delta_n\varphi)}\prod_{x\in\Lambda_n}\#(\mathrm{d}\varphi_x). \tag{2.3}$$

Here $(\cdot,\cdot)$ denotes the canonical inner product in $\ell^2(\Lambda_n)$ and # is the counting measure on $\mathbb{Z}$. The measure $P_{n,\beta}$ is well defined for all $\beta > 0$ thanks to the fact that, in light of Δ_n being symmetric and negative definite, the normalization constant $\Sigma_n(\beta)$ is finite. Two constants will play an important role throughout; namely,

$$\beta_{\mathrm{c}} := \frac{2\pi^2}{\log b} \quad \text{and} \quad \alpha := \sqrt{2\log b}. \tag{2.4}$$

Our first result is then:

Theorem 2.1 (Limit law of the maximum) *Given $\beta > 0$, let*

$$m_n := \frac{1}{\sqrt{\beta}}\left[\sqrt{2\log b}\,n - \frac{3}{2}\frac{1}{\sqrt{2\log b}}\log n\right]. \tag{2.5}$$

There exists a random variable $\mathcal{Z}$ with $\mathcal{Z} \in (0,\infty)$ a.s. and, for all $\beta \in (0,\beta_{\mathrm{c}})$, there exists a constant $\hat{c}_\beta(0) \in (0,\infty)$ with the following property: For all $\beta \in (0,\beta_{\mathrm{c}})$ and all sequences $\{n_k\}_{k\in\mathbb{N}}$ of naturals such that $n_k \to \infty$ and $s := \lim_{k\to\infty}(m_{n_k} - \lfloor m_{n_k}\rfloor)$ exists,

$$P_{n_k,\beta}\Big(\max_{x\in\Lambda_{n_k}}\varphi_x \leqslant \lfloor m_{n_k}\rfloor + u\Big) \underset{k\to\infty}{\longrightarrow} E\Big(\mathrm{e}^{-\hat{c}_\beta(s)\mathcal{Z}\mathrm{e}^{-\alpha\sqrt{\beta}u}}\Big), \quad u \in \mathbb{Z}, \tag{2.6}$$

where $\hat{c}_\beta(s) := \hat{c}_\beta(0)\mathrm{e}^{\alpha\sqrt{\beta}s}$.

The non-constancy of $s \mapsto \hat{c}_\beta(s)$ reflects on the discrete nature of the DG-model. Apart from that, the result is formally identical to that for the associated GFF which, in this case, is simply a (tree-indexed) Branching Random Walk with step distribution $\mathcal{N}(0,1/\beta)$. Indeed, the limit law of the GFF can be obtained from the above formally by replacing $\lfloor m_n\rfloor$ by m_n and $\hat{c}_\beta(s)$ by $(\alpha\sqrt{\beta})^{-1}$, see (7.57); the conclusions will then hold for all $u \in \mathbb{R}$. The random variable $\mathcal{Z}$, which translates to the random shift of the Gumbel law mentioned above, is the same for both processes and is independent of β and s (and the subsequence achieving s). These considerations imply:

Corollary 2.2 *Let $P'_{n,\beta}$ be the law of the GFF obtained by replacing the counting measure in* (2.3) *by the Lebesgue measure. Write φ' for the samples of the GFF under $P'_{n,\beta}$. For all $\beta \in (0,\beta_{\mathrm{c}})$ there exists $a > 0$ such that*

$$\sup_{u\in\mathbb{R}}\inf_{r\in[-a,a]}\left|P_{n,\beta}\big(\max_{x\in\Lambda_n}\varphi_x \leqslant u\big) - P'_{n,\beta}\Big(\max_{x\in\Lambda_n}\varphi'_x \leqslant u + r\Big)\right| \underset{n\to\infty}{\longrightarrow} 0. \tag{2.7}$$

In particular, for each $\beta \in (0,\beta_{\mathrm{c}})$, the maxima of the DG-model and GFF-model can be coupled to within a bounded distance, with probability tending to one as $n \to \infty$.

2.2 Extremal process.

As already alluded to, our proof of Theorem 2.1 is based on a coupling of the DG-model to the GFF that naturally arises from the renormalization-group analysis of the

DG-model. Unfortunately, the coupling does not necessarily preserve the maximizer(s) and so we will have to work with the full extremal process of the DG-model for which we thus get a limit law as well.

In order to make precise statements, we need more definitions. First, with each vertex $x = (x_1, \dots, x_n) \in \Lambda_n$ we can associate the real number

$$[x]_n := \sum_{i=1}^{n} (x_i - 1) b^{-i}. \tag{2.8}$$

The resulting map $x \to [x]_n$ images Λ_n injectively into $[0,1]$.

Second, in order to talk about point processes and their weak convergence, let $\mathcal{M}(\mathscr{X})$ denote the space of Radon measures on a topological space $\mathscr{X}$. We endow $\mathcal{M}(\mathscr{X})$ with the topology of vague convergence which is generated by integrals of the measure against continuous, compactly-supported functions. This in turn permits us to talk about random elements of $\mathcal{M}(\mathscr{X})$ and weak convergence thereof. We will also write $\mathcal{M}_{\mathbb{N}}(\mathscr{X})$ for the set of integer-valued measures in $\mathcal{M}(\mathscr{X})$ which, assuming $\mathscr{X}$ is Hausdorff, are necessarily concentrated on a locally-finite set of points.

Third, naturally associated with each sample φ of the DG-model is a random element of $\mathcal{M}_{\mathbb{N}}([0,1] \times \mathbb{R})$ called the extremal process

$$\eta_n := \sum_{x \in \Lambda_n} \delta_{[x]_n} \otimes \delta_{\varphi_x - \lfloor m_n \rfloor}, \tag{2.9}$$

where m_n is as in (2.5). Due to our reliance on vague topology, this process effectively records the position (in the parametrization $x \mapsto [x]_n$) and values of the DG-field at points where it is "close" to m_n.

Fourth and finally, let $\mathrm{PPP}(\nu)$, for a given σ-finite measure ν, denote the Poisson point process with intensity ν which, if ν is itself random, is sampled conditionally on ν. (The process is defined on a space that carries the measure ν.) We then have:

Theorem 2.3 (Limit extremal process) *There exists an a.s.-finite random Borel measure Z on $[0,1]$ with $Z(A) > 0$ a.s. for all non-empty (relatively) open $A \subseteq [0,1]$ and, for all $\beta \in (0, \beta_c)$ a probability measure ν_β on $\mathcal{M}_{\mathbb{N}}(\mathbb{Z})$ such that the following holds: For all $\beta \in (0, \beta_c)$ and all increasing sequences $\{n_k\}_{k \in \mathbb{N}}$ of naturals for which $s := \lim_{k\to\infty}(m_{n_k} - \lfloor m_{n_k} \rfloor)$ exists,*

$$\eta_{n_k} \underset{k\to\infty}{\overset{\text{law}}{\longrightarrow}} \sum_{i,j \geqslant 1} \delta_{x_i} \otimes \delta_{h_i + t_j^{(i)}}, \tag{2.10}$$

where $\{(x_i, h_i)\}_{i \geqslant 1}$ enumerates the points in a sample from

$$\mathrm{PPP}\Big(\tilde{c}\, \mathrm{e}^{\alpha\sqrt{\beta}\, s}\, Z \otimes \sum_{n \in \mathbb{Z}} \mathrm{e}^{-\alpha\sqrt{\beta}\, n} \delta_n\Big), \tag{2.11}$$

where $\tilde{c} := \alpha^{-1}(1 - \mathrm{e}^{-\alpha\sqrt{\beta}})$, and $\{t_j^{(i)}\}_{j \geqslant 1}$ enumerates the points in the i-th member of the sequence $\{t^{(i)}\}_{i \geqslant 1}$ of i.i.d. samples from ν_β that are independent of $\{(x_i, h_i)\}_{i \geqslant 1}$.

Theorem 2.1 will actually be derived from this result by noting that the event that $\max_{x \in \Lambda_n} \varphi_x \leqslant \lfloor m_n \rfloor + u$ with $u \in \mathbb{Z}$ translates to the extremal process η_n not charging the set $[0,1] \times (u, \infty)$. A simple calculation then shows that the objects in (2.6) are given by

$$\mathcal{Z} := Z\big([0,1]\big) \tag{2.12}$$

and

$$\hat{c}_\beta(s) := \tilde{c}\,\mathrm{e}^{\alpha\sqrt{\beta}s}\sum_{n\in\mathbb{Z}}\mathrm{e}^{-\alpha\sqrt{\beta}n}\,\nu_\beta\Bigl(\Bigl\{\zeta\in\mathcal{M}_{\mathbb{N}}(\mathbb{Z})\colon\zeta\bigl((-n,\infty)\bigr)\geqslant 1\Bigr\}\Bigr), \tag{2.13}$$

where the sum on the right converges. (A detailed derivation will be given in the proof of Theorem 2.3 and Theorem 2.1.)

As our proofs show, the Z measure coincides with the random intensity governing the scaling limit of the extremal process of the normalized GFF (i.e., that with $\beta := 1$); see Theorem 7.1. The cluster law ν_β of the DG-model is derived from the corresponding cluster law ν of the normalized GFF, albeit through a non-constructive limit procedure that makes the connection inexplicit.

2.3 Discussion.

We proceed with a discussion of the above statements. As noted numerous times earlier, our results demonstrate a very close relationship between the extremal properties of the DG-model and those of the GFF throughout the parameter regime $\beta\in(0,\beta_c)$. This should be regarded as a statement of universality. Notwithstanding, the reader may naturally wonder what happens when β does not lie in this interval.

In our proofs, β_c marks the largest value of β at which the coarse-graining/renormalization-group iterations converge to a "trivial" fixed point. This is usually expressed via the behavior of the two-point correlation function or the so-called fractional charge. We instead articulate it in much stronger terms by constructing a coupling of the DG-model and the GFF that, at typical points of Λ_n, keeps the two fields within a tight distance of each other. (If a suitably discretized version of the GFF is used, the two fields even agree on a positive fraction of all vertices.)

At $\beta = \beta_c$ the convergence to a "trivial" fixed point still takes place but only at a polynomial rate which precludes control of the coupling at the desired level. For $\beta > \beta_c$ it is expected that the iterations converge to a "non-trivial" fixed point but the proof of this remains elusive. If this convergence indeed does take place, then the "correct" approximation of the DG-model is not the Branching Random Walk with step distribution $\mathcal{N}(0,1/\beta)$ but rather a Branching Markov chain whose steps have a Gaussian density of $\mathcal{N}(0,1/\beta)$ multiplied by a positive 1-periodic function that depends on the current state of the chain.

Unlike Branching Random Walks, the extremal behavior of Branching Markov Chains has been controlled only in a few cases (e.g., the local time of simple random walk on regular tree [22]) and no overall theory exists one can rely on in generic situations. We can therefore offer only a sophisticated guess concerning the maximum of the DG-model for $\beta > \beta_c$: The limit law of the maximum remains formally the same discrete-Gumbel like (i.e., (2.6) applies), albeit with the various objects — namely, the centering m_n, the random variable $\mathcal{Z}$ and the tail exponent $\alpha\sqrt{\beta}$ — now having a different β-dependence than in the subcritical regime.

Our guess is based on the assumption that the Markov chain underlying the approximating Branching Markov Chain process behaves closely to a random walk albeit now with the process at time n having variance $\sigma(\beta)^2 n$, where $\sigma(\beta)^2$ is generally different from the plain factor $1/\beta$. (The variance can be computed from the step distribution by tools from periodic homogenization theory but it involves the unknown step distribution and also an expression for the invariant measure which both depend non-trivially on β.) Borrowing on arguments from the theory of Branching Random Walks, this should mean that the probability that the maximum exceeds value r is asymptotic to

$$b^n \frac{1}{n} \frac{c}{\sqrt{n}} \mathrm{e}^{-\frac{r^2}{2\sigma(\beta)^2 n}}, \tag{2.14}$$

where b^n accounts for the possible positions of the maxima, the factor $1/n$ arises from a "ballot theorem" that accounts for the entropic-repulsion effect caused by conditioning the remaining vertices to have value below r and the remainder of the term (with c a positive factor depending on u below) is the Gaussian density with variance $\sigma(\beta)^2 n$.

For the specific choice

$$r := \sigma(\beta)\left[\sqrt{2\log b}\, n - \frac{3}{2}\frac{1}{\sqrt{2\log b}} \log n\right] + u \tag{2.15}$$

the quantity in (2.14) is proportional to $\mathrm{e}^{-\tilde{\alpha}(\beta)u}$ with

$$\tilde{\alpha}(\beta) := \frac{\sqrt{2\log b}}{\sigma(\beta)}. \tag{2.16}$$

We thus expect that the sole effect on the centering and the tail exponent to be the replacement of β by $\sigma(\beta)^{-2}$. The situation with the random variable $\mathcal{Z}$ is less clear as it arises directly, via the so-called derivative martingale, from the approximating Branching Markov Chain and so it likely depends non-trivially on β. (The lack of this dependence for $\beta < \beta_c$ is due to the β-dependent factors scaling out explicitly.)

The reader unfamiliar with hierarchical models may wonder why the extremal behavior should remain so close in the supercritical and subcritical regimes. After all, the subcritical regime marks the delocalized phase of the model and the supercritical regime thus presumably corresponds to the localized one. Unfortunately, while this should be the case in the lattice setting, it is not in the hierarchical version. Indeed, in the hierarchical DG-model the correlations are expected to decay polynomially and the variance of the field diverges (as $n \to \infty$) at *all* $\beta > 0$. This can be attributed to the hierarchical model being effectively long-range.

We close the discussion by noting that the coupling argument presented here is capable of handling other extremal questions one may consider for the hierarchical DG-model. One of these is that of "intermediate" level sets, a.k.a., the thick points, which are those where the field exceeds $\lambda\beta^{-1/2}\sqrt{2\log b}\, n$ for $\lambda \in (0,1)$. Precise control of these has been achieved in a number of contexts; e.g., for the lattice GFF in Biskup and Louidor [21] and random-walk local time in Jego [33] and Abe, Biskup and Lee [1,3].

2.4 Outline.

The remainder of this paper is organized as follows: In Section 3 we explain the two main technical inputs underlying our results; namely, the iterative (a.k.a. renormalization group) approach to the DG-model and the ensuing coupling of the DG-model to the GFF. These steps are where the restriction to $\beta \in (0, \beta_c)$ originates from. Details of the technical inputs are worked out in Sections 4 (iterations) and 5 (coupling). In the remaining two sections we then address the proof of the above theorems starting from a proof of tightness (Section 6) and then convergence of the extremal process (Section 7).

3. TWO MAIN NEW IDEAS

We proceed by discussing the novel ideas of the proofs. While we rely on renormalization group analysis of the hierarchical DG-model that drives much of the earlier work on this problem, two novel aspects are worthy of attention. The first one is our control of the renormalization group flow where instead of linearization we address directly the full non-linear iteration. This is what allows us to work all the way up to the critical value β_c of parameter β. Second, rather than as a tool for computing correlation functions, we use the renormalization-group flow to build a coupling of the DG-model to the GFF which then allows us to link sample-path properties of one process to the other.

3.1 Iterative description of hierarchical DG-model.

We begin by explaining the structure of the hierarchical DG-model that drives the iterative description thereof. This will allow us to give several good reasons for the particular choice of the hierarchical Laplacian in (2.2).

The study of hierarchical models was initiated in Dyson [26]. The hierarchical setting is generally more friendly to coarse-graining arguments than their lattice counterpart which is why hierarchical models have served as a useful preliminary tool in real-space renormalization group theory (Brydges [24]). While the hierarchical version is for many models just a convenient approximation, other models — such as the two-dimensional GFF or long-range percolation (see, e.g., Biskup and Krieger [17] or Hutchcroft [32]) — seem to exhibit a hierarchical structure naturally.

In the past, the hierarchical DG-model has been studied mainly in its "soft" variant called the Sine-Gordon model (Marchetti and Perez [39], Guidi and Marchetti [30], Benfatto and Renn [15]) with calculations and results often presented for its dual version of the lattice Coulomb gas. We will work solely with the DG-model although the forthcoming derivations apply to a whole class of hierarchical models including the Sine-Gordon model. (We will give appropriate definitions in Section 3.4.)

The iterative description of the DG-model is based on an integral identity whose statement requires additional notation. First, let us add suffix n to the notation of the canonical inner product in $\ell^2(\Lambda_n)$ and write it as $(\cdot,\cdot)_n$. Recall also our notation

$$\mathcal{B}_k(x) := \{y \in \Lambda_n \colon d(x,y) \leqslant k\} \tag{3.1}$$

for the ball of (hierarchical) radius k centered at x. Note that $\mathcal{B}_k(x)$ has b^k elements. Introduce the map $m\colon \Lambda_n \to \Lambda_{n-1}$ defined by

$$m(x) := (x_1,\dots,x_{n-1}) \quad \text{when} \quad x = (x_1,\dots,x_n). \tag{3.2}$$

We then have:

Lemma 3.1 *Given a natural $n \geqslant 1$ and reals $\lambda_0, \dots, \lambda_n > 0$, consider the linear operator L_n on $\ell^2(\Lambda_n)$ that acts on $f\colon \Lambda_n \to \mathbb{R}$ as*

$$L_n f(x) := \sum_{k=1}^{n} \frac{\lambda_{k-1}^{-1} - \lambda_k^{-1}}{b^k} \sum_{y \in \mathcal{B}_k(x)} \bigl[f(y) - f(x) \bigr] - \lambda_n^{-1} f(x). \tag{3.3}$$

For each $n \geqslant 1$ there exists $\mathfrak{z}_n \in (0,\infty)$ such that for all $\varphi\colon \Lambda_n \to \mathbb{R}$,

$$\mathrm{e}^{\frac12 (\varphi, L_n \varphi)_n} = \mathfrak{z}_n \int_{\mathbb{R}^{\Lambda_{n-1}}} \mathrm{e}^{-\frac{1}{2\lambda_0} \sum_{x\in\Lambda_n} (\varphi_x - \psi_{m(x)})^2} \mathrm{e}^{\frac12 (\psi, L'_{n-1}\psi)_{n-1}} \prod_{z\in\Lambda_{n-1}} \mathrm{d}\psi_z, \tag{3.4}$$

where L'_{n-1} is defined using reals $\lambda'_0, \dots, \lambda'_{n-1}$ that are uniquely determined from

$$\lambda'_0 = \frac{\lambda_1 - \lambda_0}{b} \tag{3.5}$$

and

$$\lambda'_k - \lambda'_{k-1} = \frac{\lambda_{k+1} - \lambda_k}{b}, \quad k = 1, \dots, n-1. \tag{3.6}$$

Proof. The formula (3.4) can be understood informally as a convolution identity for multivariate Gaussians except that interpreting these Gaussians on the same space makes some of them singular and so talking about densities is difficult. We will instead use the fact that it suffices prove equality under integration of (3.4) against $\mathrm{e}^{(\varphi,f)_n}$ with some $\mathfrak{z}_n$ independent of $f \in \ell^2(\Lambda_n)$, for all such functions f.

Using the substitution $\widetilde{\varphi} := \varphi + L_n^{-1} f$, which is well defined as L_n^{-1} will be shown to be negative definite (and thus invertible), on the left-hand side we get

$$\int_{\mathbb{R}^{\Lambda_n}} \mathrm{e}^{\frac12 (\varphi, L_n\varphi)_n + (\varphi, f)_n} \prod_{x\in\Lambda_n} \mathrm{d}\varphi_x = \mathfrak{z}' \, \mathrm{e}^{-\frac12 (f, L_n^{-1} f)_n}, \tag{3.7}$$

where $\mathfrak{z}'$ is a positive and finite quantity independent of f. Letting $f'\colon \Lambda_{n-1} \to \mathbb{R}$ be the function defined by $f'(y) := \sum_{x\in\Lambda_n} f(x) 1_{\{m(x)=y\}}$, on the right-hand side the substitutions $\widetilde{\varphi} := \varphi - \psi_{m(\cdot)} - \lambda_0 f$ followed by $\widetilde{\psi} := \psi + (L'_{n-1})^{-1} f'$ in turn give

$$\begin{aligned} \int_{\mathbb{R}^{\Lambda_n}\times\mathbb{R}^{\Lambda_{n-1}}} \mathrm{e}^{-\frac{1}{2\lambda_0} \sum_{x\in\Lambda_n} (\varphi_x - \psi_{m(x)})^2} \mathrm{e}^{\frac12 (\psi, L'_{n-1}\psi)_{n-1} + (\varphi, f)_n} \prod_{z\in\Lambda_n} \mathrm{d}\varphi_z \prod_{z\in\Lambda_{n-1}} \mathrm{d}\psi_z \\ = \mathfrak{z}'' \, \mathrm{e}^{\frac12 \lambda_0 (f,f)_n - \frac12 (f', (L'_{n-1})^{-1} f')_{n-1}}, \end{aligned} \tag{3.8}$$

where $\mathfrak{z}''$ is again independent of f and where we used that $(\psi_{m(\cdot)}, f)_n = (\psi, f')_{n-1}$. To prove (3.4), we thus have to show that

$$(f, L_n^{-1} f)_n = -\lambda_0 (f,f)_n + \bigl(f', (L'_{n-1})^{-1} f'\bigr)_{n-1} \tag{3.9}$$

holds for all $f \in \ell^2(\Lambda_n)$.

Define the operator Q_k by $Q_k f(x) := \frac{1}{b^k} \sum_{y\in\mathcal{B}_k(x)} f(y)$ for $k = 1, \dots, n$ and let Q_0 be the identity and $Q_{n+1} := 0$. Note that Q_k is, for each $k = 0, \dots, n+1$, an orthogonal projection (in $\ell^2(\Lambda_n)$) such that $Q_k Q_\ell = Q_\ell Q_k = Q_k$ whenever $\ell \leqslant k$. In particular, for all $k, \ell = 0, \dots, n$ we have

$$(Q_\ell - Q_{\ell+1})(Q_k - Q_{k+1}) = \delta_{k,\ell} (Q_k - Q_{k+1}) \tag{3.10}$$

and so $\{Q_k - Q_{k+1} : k = 0, \dots, n\}$ are orthogonal projectors on orthogonal subspaces that, in light of $\sum_{k=0}^n (Q_k - Q_{k+1})$ being the identity, span all of $\ell^2(\Lambda_n)$.

With the help of the convention $(\lambda_{-1})^{-1} := 0$, the operator L_n now takes the form

$$\begin{aligned} L_n &= \sum_{k=0}^n (\lambda_{k-1}^{-1} - \lambda_k^{-1})(Q_k - Q_0) - \lambda_n^{-1} Q_0 \\ &= \sum_{k=0}^n (\lambda_{k-1}^{-1} - \lambda_k^{-1}) Q_k = -\sum_{k=0}^n \lambda_k^{-1}(Q_k - Q_{k+1}). \end{aligned} \tag{3.11}$$

This along with (3.10) imply

$$L_n^{-1} = -\sum_{k=0}^n \lambda_k (Q_k - Q_{k+1}). \tag{3.12}$$

In particular, L_n is invertible and negative definite.

The previous formula gives

$$L_n^{-1} = -\lambda_0 Q_0 - \sum_{k=1}^n (\lambda_k - \lambda_{k-1}) Q_k. \tag{3.13}$$

Observing that, for each $k \geqslant 1$ we have $(f, Q_k f)_n = b^{-1}(f', Q_{k-1} f')_{n-1}$, using (3.5–3.6) we now readily verify (3.9) and thus conclude the proof. $\square$

3.2 What makes the hierarchical Laplacian special?.

The operator defined in (3.3) is an instance of a hierarchical operator; i.e., one that commutes with all the Q_j defined in the previous proof. As noted above, L_n is invertible and negative definite for all $\lambda_0, \dots \lambda_n > 0$. If in fact $0 < \lambda_0 \leqslant \dots \leqslant \lambda_n$, then we can write

$$\lambda_k^{-1} := \lambda_0^{-1} P(\tau > k), \quad k = 0, \dots, n, \tag{3.14}$$

for a random variable τ taking values in $\{1, 2 \dots, n+1\}$. As $\lambda_{k-1}^{-1} - \lambda_k^{-1} = \lambda_0^{-1} P(\tau = k)$, we can then view L_n as the generator of a Markov chain on Λ_n induced (by observing only visits to the leaves) by a Markov chain on the full b-ary tree $\mathbb{T}_n := \bigcup_{k=0}^n \Lambda_k$ of depth n whose transition probabilities have all the symmetries of $\mathbb{T}_n$ and τ represents the maximal height reached by that chain between two successive visits to the leaves; namely, the set Λ_n. The event $\tau > n$ marks the event that the chain exits $\mathbb{T}_n$ through the root and "dies." For L_n, this imitates the effect of Dirichlet boundary conditions.

Any Markov chain respecting the symmetries of $\mathbb{T}_n$ is determined by a collection of numbers $\{(p_k, q_k)\}_{k=0}^n$, with $q_0 = 0$ and $p_k > 0$ and $p_k + b q_k = 1$ for each k, that have the following meaning: p_k is the probability that the chain at a site at height k above the leaves takes a step towards the root while $b q_k$ is the (total) probability that it takes a step away from the root. In this parametrization we have

$$P(\tau > k) = \left[\sum_{\ell=0}^k \prod_{j=1}^{\ell} \frac{b q_j}{p_j} \right]^{-1}, \quad k = 0, \dots, n, \tag{3.15}$$

and the generator associated with the process induced by observing the walk only on the leaves is then L_n with λ_k as in (3.14). (The prefactor λ_0 just changes the speed with which the process moves in continuous time.)

The above representations now permit us to explain the special role that the hierarchical Laplacian Δ_n plays among the operators of the form (3.3). Indeed, using the simple random walk on $\mathbb{T}_n$ in place of the above Markov chain boils down to $p_k = q_k$ for each $k = 1, \dots, n$ and so, using (3.14–3.15),

$$\lambda_k = \lambda_0 \sum_{j=0}^{k} b^j, \quad k = 0, \dots, n. \tag{3.16}$$

It follows that, in this case, L_n coincides with a $1/\lambda_0$-multiple of the hierarchical Laplacian Δ_n defined in (2.2).

Another way to see the relevance of the hierarchical Laplacian stems from the following question: What specific choice (if any) of the coefficients $\lambda_0, \lambda_1, \dots$ is preserved by the transformation (3.5–3.6)? A calculation reveals that this happens when

$$\lambda_k - \lambda_{k-1} = \lambda_0 b^k, \quad k = 1, \dots, n, \tag{3.17}$$

which is again solved uniquely by (3.16). It follows that the hierarchical Laplacian reproduces itself under the convolution identity (3.4).

Finally, yet another fact singling out the particular form of Δ_n is that plugging (3.16) with $\lambda_0 := 1$ in (3.13) yields

$$(\delta_x, -\Delta_n^{-1}\delta_y) = \lambda_0 \delta_{xy} + \sum_{k=d(x,y)\vee 1}^{n} (\lambda_k - \lambda_{k-1}) \frac{1}{b^k} = n + 1 - d(x,y), \tag{3.18}$$

where we used that $(\delta_x, Q_k \delta_y)$ vanishes when $d(x,y) > k$ and equals b^{-k} otherwise. In light of the hierarchical distance translating into the logarithm of the Euclidean distance in the identification of Λ_n for $b := L^d$ with a square box of side-length L^n in $\mathbb{Z}^d$, this shows that Δ_n^{-1} behaves much like the inverse of the Laplacian in the box with Dirichlet boundary conditions.

3.3 Representation as a tree-indexed Markov chain.

In light of the preceding discussion, we will henceforth focus on the case when the sequence $\lambda_0, \dots, \lambda_n$ takes the form (3.16) with $\lambda_0 := \beta^{-1}$, i.e., $L_n = \beta\Delta_n$. The integral formula (3.4) then allows for iterative computation of the partition function

$$\Sigma_n(\beta) := \sum_{\varphi \in \mathbb{Z}^{\Lambda_n}} \mathrm{e}^{\frac{\beta}{2}(\varphi, \Delta_n \varphi)} \tag{3.19}$$

normalizing the measure (2.3). As a starting point we observe

$$\begin{aligned} \Sigma_n(\beta) &= \mathfrak{z}_n \int_{\mathbb{R}^{\Lambda_{n-1}}} \Bigl(\sum_{\varphi \in \mathbb{Z}^{\Lambda_n}} \mathrm{e}^{-\frac{\beta}{2} \sum_{x \in \Lambda_n} (\varphi_x - \varphi'_{m(x)})^2} \Bigr) \mathrm{e}^{\frac{\beta}{2}(\varphi', \Delta_{n-1}\varphi')_{n-1}} \prod_{z \in \Lambda_{n-1}} \mathrm{d}\varphi'_z \\ &= \mathfrak{z}_n \int_{\mathbb{R}^{\Lambda_{n-1}}} \mathrm{e}^{-\sum_{z \in \Lambda_{n-1}} b v_0(\varphi'_z)} \mathrm{e}^{\frac{\beta}{2}(\varphi', \Delta_{n-1}\varphi')_{n-1}} \prod_{z \in \Lambda_{n-1}} \mathrm{d}\varphi'_z, \end{aligned} \tag{3.20}$$

where we have set

$$\mathrm{e}^{-v_0(z)} := \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\frac{\beta}{2}(z-n)^2} \tag{3.21}$$

and used that $L'_{n-1} = \beta\Delta_{n-1}$ when $L_n = \beta\Delta_n$. Defining $v_1, \dots, v_n$ recursively via

$$\mathrm{e}^{-v_{k+1}(z)} := E\big(\mathrm{e}^{-bv_k(z+\zeta)}\big) \quad \text{for} \quad \zeta \overset{\text{law}}{=} \mathcal{N}(0,1/\beta). \tag{3.22}$$

Further applications of (3.4) then yield

$$\begin{aligned}\Sigma_n(\beta) &= \mathfrak{z}_n\mathfrak{z}_{n-1}(2\pi/\beta)^{|\Lambda_{n-1}|/2} \int_{\mathbb{R}^{\Lambda_{n-2}}} \mathrm{e}^{-\sum_{z\in\Lambda_{n-2}} bv_1(\varphi''_z)} \mathrm{e}^{\frac{\beta}{2}(\varphi'',\Delta_{n-2}\varphi'')_{n-2}} \prod_{z\in\Lambda_{n-2}} \mathrm{d}\varphi''_z \\ &= \dots = \mathfrak{z}_n\Big(\prod_{k=0}^{n-1} \mathfrak{z}_k (2\pi/\beta)^{|\Lambda_k|/2}\Big)\mathrm{e}^{-v_n(0)},\end{aligned} \tag{3.23}$$

where the factors $(2\pi/\beta)^{|\Lambda_k|/2}$ arise from the normalization of the Gaussian measure in (3.22) which is not included in (3.4).

The calculation (3.23) is not of much use in its own right — after all, the factors $\mathfrak{z}_k$ are not very explicit and the value of the partition function is not particularly illuminating. What people in mathematical physics usually do is to adapt the procedure to calculate expectations of suitable observables. A minor problem is that the observables keep chaining through the iterations and so one has to follow their "flow" separately. In all of this, not much attention is paid or meaning assigned to the auxiliary fields one integrates over; indeed, they are just tools in a massive inductive argument.

Our aim here is different in that we will keep track of the auxiliary fields and, in fact, use them to give a convenient representation of the DG-field. This is the content of:

Lemma 3.2 *Writing $\mathfrak{g}_{1/\beta}$ for the law of $\mathcal{N}(0,1/\beta)$, for each $k \geqslant 0$ and $\varphi \in \mathbb{R}$, let $\mathfrak{q}_k(\cdot|\varphi)$ be the Borel probability measure on $\mathbb{R}$ given in infinitesimal form by*

$$\mathfrak{q}_k(\mathrm{d}\zeta|\varphi) := \begin{cases} \mathrm{e}^{v_k(\varphi)-bv_{k-1}(\varphi+\zeta)}\, \mathfrak{g}_{1/\beta}(\mathrm{d}\zeta), & \text{if } k \geqslant 1, \\ \mathrm{e}^{v_0(\varphi)-\frac{\beta}{2}\zeta^2}\, \#(\varphi+\mathrm{d}\zeta), & \text{if } k = 0. \end{cases} \tag{3.24}$$

Given $n \geqslant 1$, let $\{\zeta_k(x)\colon x \in \Lambda_k,\ k = 1, \dots, n\}$ be a family of random variables with joint law given (again in infinitesimal form) by

$$\bigotimes_{k=1}^{n} \bigotimes_{x\in\Lambda_k} \mathfrak{q}_{n-k}\big(\mathrm{d}\zeta_k(x)\,\big|\,\varphi_{k-1}(m(x))\big), \tag{3.25}$$

where $\varphi_0 := 0$ and, for $k = 1, \dots, n$,

$$\varphi_k(x) := \sum_{j=1}^{k} \zeta_j\big(m^{k-j}(x)\big), \quad x \in \Lambda_k. \tag{3.26}$$

Then $\{\varphi_n(x)\colon x \in \Lambda_n\}$ has the law $P_{n,\beta}$ of the DG-model as defined in (2.3).

Proof. After some careful bookkeeping we find that the prefactors in (3.24) mostly cancel out and the measure in (3.25) can be cast as

$$\mathrm{e}^{bv_{n-1}(0)-\frac{\beta}{2}\sum_{x\in\Lambda_n}\zeta_n(x)^2}\bigotimes_{x\in\Lambda_n}\#\bigl(\varphi_{n-1}(m(x))+\mathrm{d}\zeta_n(x)\bigr)\otimes\bigotimes_{k=1}^{n-1}\bigotimes_{z\in\Lambda_k}\mathfrak{g}_{1/\beta}\bigl(\mathrm{d}\zeta_k(z)\bigr),\tag{3.27}$$

where φ_{n-1} is derived from $\{\zeta_k(x)\colon x\in\Lambda_k, k=1,\dots,n-1\}$ via (3.26). Under the second part of the product law on the right, the law of φ_{n-1} is that of a GFF on Λ_{n-1}. Integrating over the variables $\{\zeta_k(x)\colon x\in\Lambda_k, k=1,\dots,n-1\}$ with φ_{n-1} fixed thus reduces this to

$$Z_{n-1}^{-1}\mathrm{e}^{-\frac{\beta}{2}\sum_{x\in\Lambda_n}\zeta_n(x)^2}\mathrm{e}^{\frac{\beta}{2}(\varphi_{n-1},\Delta_{n-1}\varphi_{n-1})_{n-1}}\tag{3.28}$$

times the measure

$$\bigotimes_{x\in\Lambda_n}\#\bigl(\varphi_{n-1}(m(x))+\mathrm{d}\zeta_n(x)\bigr)\otimes\bigotimes_{z\in\Lambda_{n-1}}\mathrm{d}\varphi_{n-1}(z),\tag{3.29}$$

where Z_{n-1} is a suitable normalization constant. Substituting $\varphi_n(x):=\varphi_{n-1}(m(x))+\zeta_n(x)$ and performing the integrals over φ_{n-1} with the help of (3.4) then reduces this to

$$Z_{n-1}^{-1}\mathfrak{z}_n^{-1}\mathrm{e}^{\frac{\beta}{2}(\varphi_n,\Delta_n\varphi_n)_n}\bigotimes_{x\in\Lambda_n}\#\bigl(\mathrm{d}\varphi_n(x)\bigr).\tag{3.30}$$

Modulo the form of the normalization constant, this is the law $P_{n,\beta}$ from (2.3). □

A key property of the law (3.25) is that the collection of random variables

$$\bigl\{\varphi_k(x)\colon x\in\Lambda_k, k=1,\dots,n\bigr\}\tag{3.31}$$

is tree-indexed Markovian since conditioning on $\{\varphi_j(x)\colon x\in\Lambda_j, j=1,\dots,k-1\}$ determines the law of φ_k in such a way that $\varphi_k(x)$ depends on $\varphi_{k-1}(m(x))$ only. The transition probabilities vary with k and so the chain is "time"-inhomogeneous.

The observation we just made also permits us to think of the collection (3.31) as a branching Markov chain. Indeed, the value $\varphi_k(x)$ at "time" k splits into b values $\varphi_{k+1}(x^{(1)}),\dots,\varphi_{k+1}(x^{(b)})$ at "time" $k+1$ (where $x^{(1)},\dots,x^{(b)}$ denotes the "children" of x in Λ_{k+1}) that are independent samples from a law that depends only on $\varphi_k(x)$. (In the absence of this dependence, the process would be a Branching Random Walk.)

3.4 Control of effective potentials.

In order to make a good use of the representation (3.26), we need to control the "evolution" of the effective potentials $v_0, v_1,\dots$ from (3.22). Although we are primarily interested in v_0 arising for the DG-model (see (3.21)), other choices of the "initial value" may be considered as well. One of these is the Sine-Gordon model where

$$v_0(z):=-\kappa\cos(2\pi z),\quad z\in\mathbb{R}.\tag{3.32}$$

This case is interesting because, by varying κ over the positive reals, it allows us to smoothly interpolate between the GFF ($\kappa=0$) and the DG-model ($\kappa=+\infty$).

The available analyses of the iterations (3.22) seem invariably to be based on linearization (see, e.g., Bauerschmidt and Bodineau [11]). With the eyes on the "trivial" fixed point, this amounts to replacement of the map $v_k\mapsto v_{k+1}$ by its "infinitesimal" version represented by the linear operator

$$\mathcal{L}f(z):=bE\bigl(f(z+\zeta)\bigr)\tag{3.33}$$

that can now be studied using methods of linear algebra. Relying on the 1-periodicity dictated by the setting of the problem, this is achieved by passing to the Fourier representation $f(z) = \sum_{n\in\mathbb{Z}} \hat{f}(n)\mathrm{e}^{2\pi\mathrm{i}nz}$ as

$$\mathcal{L}f(z) = \sum_{n\in\mathbb{Z}} \hat{f}(n)\, b\theta^{n^2}\mathrm{e}^{2\pi\mathrm{i}nz}, \tag{3.34}$$

where

$$\theta := \mathrm{e}^{-2\pi^2/\beta}. \tag{3.35}$$

Denoting the dual L^1-norm by $\|f\|_1^* := \sum_{n\in\mathbb{Z}} |\hat{f}(n)|$, hereby we get

$$\big\|\mathcal{L}(f-\hat{f}(0))\big\|_1^* \leqslant b\theta\big\|f-\hat{f}(0)\big\|_1^* \tag{3.36}$$

showing that the linear map is contractive on mean-zero functions whenever $b\theta < 1$. In terms of β, this regime translates into

$$0 < \beta < \beta_{\mathrm{c}} := \frac{2\pi^2}{\log b}. \tag{3.37}$$

As is readily checked from (3.34), the map is not contractive when $\beta > \beta_{\mathrm{c}}$.

In order for the linearized flow (3.33) to serve as a good approximation of the non-linear iterations (3.22), one needs to control the higher-order terms. A standard approach (used, e.g., in [11]) is to absorb these into a small change of the first order term. Unfortunately, this forces us to start the iterations the closer to the desired limit point the larger is the norm of $\mathcal{L}$, thus leading to a loss of uniformity.

One way to overcome the lack of uniformity is by tracking the "evolution" of some of the higher order terms as well. This has its merits and has been made to work for, e.g., the lattice Sine-Gordon model at uniformly small κ (Falco [27,28]). But, even in this approach, one needs a "small parameter" to ensure that the remaining high-order terms can be dealt with perturbatively. Unless we are willing (as is sometimes done) to turn b into a variable, no such "small parameter" seems to be there for the DG-model.

As a consequence, none of the existing approaches [11, 15, 30, 39] seem to extend throughout the subcritical regime of the DG-model. As our desire is to work all the way up to β_{c}, we will proceed non-perturbatively from the outset. Namely, we will work directly with the non-linear problem (3.22) relying on Fourier representation and combinatorial arguments for the Fourier coefficients. Our analysis applies to a rather general class of "initial values" v_0; definitely, those conforming to:

Definition 3.3 *Let $\mathcal{V}$ denote the set of even functions $v\colon \mathbb{R}\to\mathbb{R}$ for which there exists a positive sequence $\{a(n)\}_{n\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$ such that*

$$\mathrm{e}^{-v(z)} = \sum_{n\in\mathbb{Z}} a(n)\mathrm{e}^{2\pi\mathrm{i}nz}, \quad z\in\mathbb{R}, \tag{3.38}$$

and

$$\sup_{n\geqslant 0} \frac{a(n+1)}{a(n)} < \infty. \tag{3.39}$$

As we will see, the set $\mathcal{V}$ includes the function v_0 from (3.21) for the DG-model as well as that in (3.32) corresponding to the Sine-Gordon model. We note that the sequence $\{a(n)\}_{n\in\mathbb{Z}}$ has a meaning in its own right as it represents the *a priori* charge intensity in the Coulomb gas model dual to the GFF with on-site potential v.

For "initial values" in $\mathcal{V}$ we then get:

Theorem 3.4 (Subcritical regime) *Let $v_0 \in \mathcal{V}$ and suppose that $\theta := \mathrm{e}^{-2\pi^2/\beta}$ obeys $b\theta < 1$. Denoting $c_0 := \sup_{n \geqslant 0} \frac{a_0(n+1)}{a_0(n)}$ for $\{a_0(n)\}_{n\in\mathbb{Z}}$ related to v_0 via* (3.38),

$$\sup_{z,z' \in [0,1]} \bigl| v_{k+1}(z) - bv_k(z') \bigr| \leqslant 8(b\theta)^k bc_0 \tag{3.40}$$

then holds for all $k \in \mathbb{N}$ satisfying $(b\theta)^k bc_0 \leqslant 1/8$.

This theorem, which we will prove in Section 4, will serve as a key input in all of the remaining derivations in this paper. The statement pertains to $(z, z') \mapsto v_{k+1}(z) - bv_k(z')$ as this is what enters the definition of the law (3.24). The function v_k does not actually converge by itself which can be attributed to the zero mode (i.e., $n = 0$ term) in (3.34) being expansive — meaning, having eigenvalue $b > 1$.

In order to demonstrate the strength of our method, we will show that it is capable of handling the critical case as well:

Theorem 3.5 (Critical regime) *Let $v_0 \in \mathcal{V}$ and suppose $\theta := \mathrm{e}^{-2\pi^2/\beta}$ obeys $b\theta = 1$. Denoting $c_0 := \sup_{n \geqslant 0} \frac{a_0(n+1)}{a_0(n)}$ for $\{a_0(n)\}_{n\in\mathbb{Z}}$ related to v_0 via* (3.38)*, there exists $\gamma > 0$ such that*

$$\sup_{z,z' \in [0,1]} \bigl| v_{k+1}(z) - bv_k(z') \bigr| \leqslant 8[1 + \gamma k]^{-1/2} bc_0 \tag{3.41}$$

holds for all $k \in \mathbb{N}$ satisfying $[1 + \gamma k]^{-1/2} bc_0 \leqslant 1/8$.

The constant γ depends only on v_0; see (4.39) for an explicit definition. The fact that the bound on the decay in (3.41) is only polynomial reflects an actual phenomenon. Indeed, it is not hard to prove a corresponding lower bound as well. Similarly one can check that, for $\beta > \beta_{\mathrm{c}}$, the difference $v_{k+1}(z) - bv_k(z')$ does not converge to zero as $k \to \infty$; that is, unless v_0 is $1/m$-periodic for some natural $m \geqslant 2$. Notwithstanding, we expect the difference to converge to a nice function when $\beta > \beta_{\mathrm{c}}$ but are unable to prove this, along with a convergence rate, at the desired level of generality.

3.5 Coupling.

With the iterations under control throughout the subcritical regime, we now address the second key ingredient of our proofs. For easier formulations later, we introduce the following notation. Given $n \geqslant 1$ and a sample $\{\zeta_k(x) \colon x \in \Lambda_k,\, k = 1, \dots, n\}$ from the measure (3.24), for each $x \in \Lambda_n$ and $k = 0, \dots, n-1$ denote

$$\xi_k^{\mathrm{DG}}(x) := \zeta_{n-k}\bigl(m^k(x)\bigr). \tag{3.42}$$

By Lemma 3.2, the field

$$\varphi_x^{\mathrm{DG}} := \sum_{k=0}^{n-1} \xi_k^{\mathrm{DG}}(x) \tag{3.43}$$

on Λ_n is then distributed as the DG-model on Λ_n.

Similarly, given $n \geqslant 1$ and a sample $\{\zeta_k'(x) \colon x \in \Lambda_k,\, k = 1, \dots, n\}$ of i.i.d. $\mathcal{N}(0, 1/\beta)$, for each $x \in \Lambda_n$ and $k = 0, \dots, n-1$ denote

$$\xi_k^{\mathrm{GFF}}(x) := \zeta_{n-k}'\bigl(m^k(x)\bigr). \tag{3.44}$$

The field $\varphi^{\rm GFF}$ on Λ_n defined by

$$\varphi_x^{\rm GFF} := \sum_{k=0}^{n-1} \zeta_k^{\rm GFF}(x) \tag{3.45}$$

is then distributed as the GFF on Λ_n at inverse temperature β. Our statement of the coupling of these two fields is then as follows:

Theorem 3.6 (Coupling) *For all $\beta \in (0, \beta_{\rm c})$ there exist a constant $C > 0$, a positive sequence $\{R_k\}_{k\geqslant 1}$ with $\limsup_{k\to\infty} k^{-1}\log R_k < 0$ and, for all natural $n \geqslant 1$, there exists a coupling of*

$$\bigl\{\zeta_k^{\rm DG}(x)\colon k = 0, \dots, n-1,\, x \in \Lambda_n\bigr\}, \tag{3.46}$$

$$\bigl\{\zeta_k^{\rm GFF}(x)\colon k = 0, \dots, n-1,\, x \in \Lambda_n\bigr\}, \tag{3.47}$$

and a family of zero-one valued random variables

$$\bigl\{B_k(x)\colon x \in \Lambda_{n-k},\, k = 1, \dots, n-1\bigr\} \tag{3.48}$$

such that the following holds:

(1) *the random variables in* (3.48) *are independent with*

$$P\bigl(B_k(x) = 1\bigr) = {\rm e}^{-R_k}, \quad x \in \Lambda_{n-k},\, k = 1, \dots, n-1, \tag{3.49}$$

(2) *the families* (3.47) *and* (3.48) *are independent of each other,*

(3) *for all $k = 1, \dots, n-1$ and all $x \in \Lambda_n$,*

$$\zeta_k^{\rm DG}(x) = \zeta_k^{\rm GFF}(x) \quad \text{on } \bigl\{B_k(m^k(x)) = 1\bigr\}, \tag{3.50}$$

(4) *for all $k = 0, \dots, n-1$ and all $x \in \Lambda_n$,*

$$P\Bigl(\bigl|\zeta_k^{\rm DG}(x) - \zeta_k^{\rm GFF}(x)\bigr| > C\Bigr) = 0. \tag{3.51}$$

That a joint coupling of the increments exists is perhaps not surprising due to their Markovian structure and the relative closeness of (3.25) to the product normal law. Considerably less obvious is the fact that the increments can be kept within a uniformly bounded distance of each other; cf (3.51). This is quite useful in estimates of the difference of the sums (3.43) and (3.45) which can thus be controlled by the sheer number of places where equality on the left of (3.50) fails. In light of (3.50), this number is dominated by the number of Bernoulli's in (3.48) that equal zero.

The sequence $\{R_k\}_{k\geqslant 1}$ governing the errors in (3.49) is closely related to the error in (3.40); see (5.58). An exponential decay is key but the rate of that decay is not, which is another reason why our approach works all the way up to $\beta_{\rm c}$. Theorem 3.6 will serve as the main input in our control of the extremal behavior; besides this the proofs require only some standard facts from analysis of the GFF extrema and technical estimates. The proof of Theorem 3.6 comes at the end of Section 5.

Remark 3.7 The idea that the renormalization group iterations define a useful coupling was presented by the first author to R. Bauerschmidt and P.-F. Rodriguez in late 2018 along with a proposal to use this coupling to study the extremal properties of the DG-model. The ensuing discussion identified the hierarchical DG-model as a reasonable first test case to try. Unfortunately, the project then stalled for several years and the key idea was instead taken as a basis of the works [12] and [10].

4. Renormalization group flow

The formal presentation of our proofs opens up by careful control of the iterations (3.22) resulting in proofs of Theorem 3.4 and 3.5. We start with the former theorem as it deals with the regime that the rest of this paper is focused on.

4.1 Subcritical regime.

The argument will be based on Fourier techniques that are enabled by 1-periodicity of the relevant functions. A starting point is the following representation:

Lemma 4.1 *Suppose $a_0 = \{a_0(n)\}_{n\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$ is such that v_0 obeys*

$$e^{-v_0(z)} = \sum_{n\in\mathbb{Z}} a_0(n)e^{2\pi i nz}, \quad z \in \mathbb{R}. \tag{4.1}$$

Then the iterates $\{v_k\}_{k\in\mathbb{N}}$ defined from v_0 via (3.22) *admit the representation*

$$e^{-v_k(z)} = \sum_{n\in\mathbb{Z}} a_k(n)e^{2\pi i nz}, \quad z \in \mathbb{R}, \tag{4.2}$$

with $a_k = \{a_k(n)\}_{n\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$, for each $k \in \mathbb{N}$, where the a_k's are defined recursively by

$$a_{k+1}(n) := \sum_{\substack{\ell_1,\dots,\ell_b\in\mathbb{Z}\\ \ell_1+\dots+\ell_b=n}} \left[\prod_{i=1}^{b} a_k(\ell_i)\right]\theta^{n^2}, \quad n \in \mathbb{Z}, \tag{4.3}$$

for θ as in (3.35).

Proof. The fact that the a_k's obey (4.3) is checked directly via (3.22) provided the sums in (4.3) converge absolutely. To see the latter, note that (4.3) implies $a_{k+1}(n) \leqslant \|a_k\|_1^b\theta^{n^2}$ which in light of $\theta < 1$ shows that $a_k \in \ell^1(\mathbb{Z})$ implies $a_{k+1} \in \ell^1(\mathbb{Z})$ for all $k \geqslant 1$. The summability assumption imposed on a_0 thus propagates through the iterations. □

We note that for the initial "value" of the DG-model (3.21) gives

$$a_0(n) = \theta^{n^2}\sqrt{\frac{2\pi}{\beta}}, \quad n \in \mathbb{Z}, \tag{4.4}$$

while for the Sine-Gordon model (3.32) we obtain

$$a_0(n) = \sum_{\ell=0}^{\infty} \frac{(\kappa/2)^{2\ell+|n|}}{(\ell+|n|)!\ell!}, \quad n \in \mathbb{Z}. \tag{4.5}$$

Both of these lie in $\mathcal{V}$ as desired.

We will assume $v_0 \in \mathcal{V}$ throughout the rest of this section. Observe that, as is checked from (3.22), all the v_k's are real-valued and even and so

$$0 < a_k(-n) = a_k(n) \leqslant a_k(0), \quad n, k \geqslant 0, \tag{4.6}$$

where the strict positivity of a_k is implied by the strict positivity of a_0 and (4.3). The key estimate driving our control of the sequence $v_0, v_1, \dots$ now comes in:

Lemma 4.2 *For all naturals $k \geqslant 0$ and $n \geqslant 0$,*

$$\frac{a_{k+1}(n+1)}{a_{k+1}(n)} \leqslant b\,\theta^{(n+1)^2-n^2} \sup_{\ell\geqslant 0} \frac{a_k(\ell+1)}{a_k(\ell)}. \tag{4.7}$$

In particular, $v_0 \in \mathcal{V}$ implies $v_k \in \mathcal{V}$ for all $k \geqslant 1$.

Proof. The proof is based on a combinatorial argument that is reminiscent of the Peierls argument from statistical mechanics. Denote

$$\Xi(n) := \bigl\{(\ell_1,\dots,\ell_b) \in \mathbb{Z}^b \colon \ell_1+\dots+\ell_b = n\bigr\} \tag{4.8}$$

and, for $n \geqslant 0$, construct a map $\phi\colon \Xi(n+1) \to \Xi(n)$ by setting

$$\phi\bigl((\ell_1,\dots,\ell_b)\bigr) := (\ell_1,\dots,\ell_{i-1},\ell_i-1,\ell_{i+1},\dots,\ell_b), \tag{4.9}$$

where

$$i = i(\ell_1,\dots,\ell_b) := \min\{j=1,\dots,b\colon \ell_j > 0\}. \tag{4.10}$$

(This is well defined because $n+1>0$ and so, by the pigeon-hole principle, the set on the right is non-empty.) Note that, since the map $(i(\cdot),\phi(\cdot))\colon \Xi(n+1) \to \{1,\dots,b\}\times\Xi(n)$ is injective, its projection to ϕ is at most b-fold degenerate and so

$$\bigl|\,\phi^{-1}(\{(\ell_1,\dots,\ell_b)\})\bigr| \leqslant b, \quad (\ell_1,\dots,\ell_b) \in \Xi(n). \tag{4.11}$$

Denoting the supremum on the right of (4.7) by c_k, observe also that

$$\prod_{i=1}^b a_k(\ell_i) \leqslant c_k \prod_{i=1}^b a_k(\ell_i') \tag{4.12}$$

holds for each $(\ell_1',\dots,\ell_b') \in \phi(\Xi(n+1))$ and each $(\ell_1,\dots,\ell_b) \in \phi^{-1}(\{(\ell_1',\dots,\ell_b')\})$.

Abbreviating $\bar\ell := (\ell_1,\dots,\ell_b)$ and $\bar\ell' := (\ell_1',\dots,\ell_b')$, the inductive definition (4.3) of a_{k+1} from a_k now shows

$$\begin{aligned} a_{k+1}(n+1) &= \sum_{\bar\ell'\in\phi(\Xi(n+1))}\ \sum_{\bar\ell\in\phi^{-1}(\{\bar\ell'\})} \Bigl[\prod_{i=1}^b a_k(\ell_i)\Bigr]\theta^{(n+1)^2} \\ &\leqslant c_k \sum_{\bar\ell'\in\phi(\Xi(n+1))}\ \sum_{\bar\ell\in\phi^{-1}(\{\bar\ell'\})} \Bigl[\prod_{i=1}^b a_k(\ell_i')\Bigr]\theta^{(n+1)^2} \\ &\leqslant c_k b\theta^{(n+1)^2-n^2} \sum_{\bar\ell'\in\phi(\Xi(n+1))} \Bigl[\prod_{i=1}^b a_k(\ell_i')\Bigr]\theta^{n^2}, \end{aligned} \tag{4.13}$$

where the first inequality follows from (4.12) and the second inequality from (4.11). By (4.3) again, the sum on the right is at most $a_{k+1}(n)$. □

Iterations of (4.7) yield:

Lemma 4.3 *Suppose $b\theta \leqslant 1$ and let $c_0 := \sup_{n\geqslant 0}\frac{a_0(n+1)}{a_0(n)}$. Then*

$$\sup_{\ell\geqslant 0}\frac{a_k(\ell+1)}{a_k(\ell)} \leqslant (b\theta)^k c_0, \quad k \geqslant 0, \tag{4.14}$$

and, consequently,

$$a_k(n) \leqslant \big[(b\theta)^{k-1}bc_0\big]^n\theta^{n^2}a_k(0), \quad n,k \geqslant 1. \tag{4.15}$$

Proof. Continuing to write c_k for the supremum in (4.7), the fact that $(n+1)^2 - n^2 \geqslant 1$ for all $n \geqslant 0$ gives $c_{k+1} \leqslant (b\theta)c_k$. This shows $c_k \leqslant (b\theta)^k c_0$, proving (4.14). Plugging this in (4.7) yields

$$a_k(n+1) \leqslant \theta^{(n+1)^2-n^2}(b\theta)^{k-1}bc_0\, a_k(n), \quad k \geqslant 1,\, n \geqslant 0, \tag{4.16}$$

and so, by induction, we get (4.15). □

Returning to the original setting of the problem, hereby we conclude:

Proof of Theorem 3.4. The definition of the coefficients shows

$$\mathrm{e}^{-v_{k+1}(z)} = \sum_{n\in\mathbb{Z}} a_{k+1}(n)\mathrm{e}^{2\pi\mathrm{i}nz} \tag{4.17}$$

while (4.3) gives

$$\mathrm{e}^{-bv_k(z')} = \sum_{n\in\mathbb{Z}} a_{k+1}(n)\theta^{-n^2}\mathrm{e}^{2\pi\mathrm{i}nz'}. \tag{4.18}$$

The main point is to show that these sums are dominated by the $n=0$ term for k large. To this end we write (4.17) as $a_{k+1}(0)[1+e_1]$ and (4.18) as $a_{k+1}(0)[1+e_2]$. Assuming k to be so large that $\eta := (b\theta)^k bc_0 < 1$, (4.15) shows

$$|e_1| \leqslant 2\sum_{n\geqslant 1}\eta^n\theta^{n^2} \leqslant 2\theta\frac{\eta}{1-\eta} \tag{4.19}$$

while

$$|e_2| \leqslant 2\sum_{n\geqslant 1}\eta^n \leqslant 2\frac{\eta}{1-\eta}. \tag{4.20}$$

Now

$$v_{k+1}(z) - bv_k(z') = \log\frac{1+e_2}{1+e_1} = \log\Big[1+\frac{e_2-e_1}{1+e_1}\Big] \tag{4.21}$$

and

$$\Big|\frac{e_2-e_1}{1+e_1}\Big| \leqslant \frac{|e_1|+|e_2|}{1-|e_1|}. \tag{4.22}$$

Under the (stronger) assumption that $\eta \leqslant 1/8$ we then have $|e_2| \leqslant 2/7$ and $|e_1| \leqslant (2/7)\theta$ which by the fact that $\theta \leqslant b^{-1} \leqslant 1/2$ gives $|e_1| \leqslant 1/7$. Then the right-hand side of (4.22) is less than 1/2 and so we may use that $|\log(1+x)| \leqslant 2|x|$ holds for all $|x| \leqslant 1/2$. A calculation (still under $\eta \leqslant 1/8$) shows

$$\frac{|e_1|+|e_2|}{1-|e_1|} \leqslant 4\eta \tag{4.23}$$

and so the right-hand side of (4.21) is at most 8η. □

Remark 4.4 The superexponential decay of $n \mapsto a_k(n)$ indicates that a similar control extends to all derivatives of the function $(z,z') \mapsto v_{k+1}(z) - bv_k(z')$.

4.2 Critical regime.

Our next item of concern is the critical regime $b\theta = 1$. Here the sequence of coefficients still concentrates to a "point mass" at $n = 0$ or, more precisely, $(z, z') \mapsto v_{k+1}(z) - bv_k(z')$ tends to zero as $k \to \infty$, but the convergence is no longer exponentially fast. Notwithstanding, with suitable refinements, this can still be proved using the same ideas that drove the proofs in the subcritical regime. The key is to control the ratio $a_k(1)/a_k(0)$ separately from the other ratios $a_k(\ell+1)/a_k(\ell)$ with $\ell \geqslant 1$:

Lemma 4.5 *For each $k \geqslant 0$,*

$$\frac{a_{k+1}(1)}{a_{k+1}(0)} \leqslant \theta \frac{\frac{a_k(1)}{a_k(0)}}{1 + \binom{b}{2}\left(\frac{a_k(1)}{a_k(0)}\right)^2} + (b-1)\theta \sup_{\ell \geqslant 0} \frac{a_k(\ell+1)}{a_k(\ell)}. \tag{4.24}$$

Proof. We proceed similarly as in (4.13) albeit with additional care paid to certain particular terms. Recall the definition of the map ϕ from (4.9). Denoting $\bar{0} := (0, \dots, 0)$, each $\bar{\ell} \in \phi^{-1}(\{\bar{0}\})$ has exactly one index equal to one and the others equal to zero. This shows that the contribution of $\bar{\ell} := \bar{0}$ to $a_{k+1}(1)$ on the first line of (4.13) equals

$$\sum_{\bar{\ell}' \in \phi^{-1}(\{\bar{0}\})} \left[\prod_{i=1}^{b} a_k(\ell_i)\right] \theta = b\theta a_k(1) a_k(0)^{b-1}. \tag{4.25}$$

Next consider $\bar{\ell} \in \Xi(0) \smallsetminus \{\bar{0}\}$ and observe that each $\bar{\ell}' \in \phi^{-1}(\{\bar{\ell}\})$ has at least two non-zero entries with total sum of the entries equal to one. This forces at least one of the entries to be negative; i.e., $A := \{i = 1, \dots, b \colon \ell_i' < 0\} \neq \emptyset$. But the definition of ϕ implies $\ell_i = \ell_i'$ for each $i \in A$ and, since $\bar{\ell}$ and $\bar{\ell}'$ differ only at one index, say i, where $\ell_i' = \ell_i + 1$, we have

$$\left|\phi^{-1}(\{\bar{\ell}\})\right| \leqslant b - 1, \quad \bar{\ell} \in \Xi(0) \smallsetminus \{\bar{0}\}. \tag{4.26}$$

The calculation (4.12–4.13) then shows

$$\begin{aligned} &\sum_{\bar{\ell}' \in \phi(\Xi(1)) \smallsetminus \{\bar{0}\}} \; \sum_{\bar{\ell} \in \phi^{-1}(\{\bar{\ell}'\})} \left[\prod_{i=1}^{b} a_k(\ell_i)\right] \theta \\ &\qquad \leqslant c_k \theta \sum_{\bar{\ell}' \in \phi(\Xi(1)) \smallsetminus \{\bar{0}\}} \; \sum_{\bar{\ell} \in \phi^{-1}(\{\bar{\ell}'\})} \left[\prod_{i=1}^{b} a_k(\ell_i')\right] \\ &\qquad \leqslant c_k (b-1) \theta \sum_{\bar{\ell}' \in \Xi(0) \smallsetminus \{\bar{0}\}} \left[\prod_{i=1}^{b} a_k(\ell_i')\right] \leqslant c_k (b-1)\theta \big[a_{k+1}(0) - a_k(0)^b\big], \end{aligned} \tag{4.27}$$

where c_k again denotes the supremum in (4.7) and where (4.3) was used in the last step.

Combining (4.25) and (4.27) along with the fact that $c_k a_k(0) \geqslant a_k(1)$ yields

$$\begin{aligned} a_{k+1}(1) &\leqslant b\theta a_k(1) a_k(0)^{b-1} + c_k (b-1)\theta \big[a_{k+1}(0) - a_k(0)^b\big] \\ &= b\theta a_k(1) a_k(0)^{b-1} - (b-1)\theta c_k a_k(0) a_k(0)^{b-1} + c_k (b-1)\theta a_{k+1}(0) \\ &\leqslant \theta a_k(1) a_k(0)^{b-1} + c_k (b-1)\theta a_{k+1}(0). \end{aligned} \tag{4.28}$$

Restricting the sum in (4.3) for $n := 0$ to b-tuples of indices that are either all zeros or contain one one and one negative one and are zero otherwise shows

$$a_{k+1}(0) \geqslant a_k(0)^b + \binom{b}{2} a_k(0)^{b-2} a_k(1)^2. \tag{4.29}$$

(We could write $b(b-1)$ instead of $\binom{b}{2}$ but the latter is more compact so we keep that.) Using this jointly with (4.28) then gives (4.24). □

The recursive inequalities (4.24) and (4.7) do not have the same structure but we can put them together with the help of a suitable parameter. This yields:

Lemma 4.6 *Suppose $b\theta = 1$ and write c_k for the supremum in* (4.24)*. For each $k \geqslant 0$, define $\alpha_k \in (0,1)$ to be the unique number such that*

$$\alpha_k = \frac{1}{1 + \binom{b}{2} c_k^2 \alpha_k^2}. \tag{4.30}$$

Then

$$c_{k+1} \leqslant \frac{b^{-1}\, c_k}{1 + \binom{b}{2} c_k^2 \alpha_k^2} + (1 - b^{-1}) c_k. \tag{4.31}$$

The sequence $\{c_k\}_{k\geqslant 0}$ is non-increasing while $\{\alpha_k\}_{k\geqslant 0}$ is non-decreasing.

Proof. The proof is based on the fact that the definition of α_k gives

$$(1 - b^{-1} + b^{-1}\alpha_k) c_k = \frac{b^{-1} c_k}{1 + \binom{b}{2} c_k^2 \alpha_k^2} + (1 - b^{-1}) c_k. \tag{4.32}$$

It thus suffices to dominate the ratio $a_{k+1}(n+1)/a_{k+1}(n)$ by either side of this equality for all $n \geqslant 0$.

We start with a bound on $a_{k+1}(1)/a_{k+1}(0)$. In the situation when $a_k(1)/a_k(0) \leqslant \alpha_k c_k$, dropping the second term in the denominator in (4.24) while bounding the numerator by $\alpha_k c_k$ shows

$$\frac{a_{k+1}(1)}{a_{k+1}(0)} \leqslant b^{-1}\alpha_k c_k + (1 - b^{-1}) c_k = (1 - b^{-1} + b^{-1}\alpha_k) c_k. \tag{4.33}$$

If in turn $a_k(1)/a_k(0) \geqslant \alpha_k c_k$, then we bound the denominator of (4.24) by $1 + \binom{b}{2} c_k^2 \alpha_k^2$ from below and dominate the ratio in the numerator by c_k to get

$$\frac{a_{k+1}(1)}{a_{k+1}(0)} \leqslant \frac{b^{-1} c_k}{1 + \binom{b}{2} c_k^2 \alpha_k^2} + (1 - b^{-1}) c_k. \tag{4.34}$$

Hence $a_{k+1}(1)/a_{k+1}(0)$ is bounded by the right-hand side of (4.31) in both cases.

Moving to the terms $a_{k+1}(n+1)/a_{k+1}(n)$ with $n \geqslant 1$, here (4.7) gives

$$\frac{a_{k+1}(n+1)}{a_{k+1}(n)} \leqslant b^{n^2 - (n+1)^2 + 1} c_k. \tag{4.35}$$

For $n \geqslant 1$ the prefactor is at most $b^{-2} \leqslant 1/4$ while $(1 - b^{-1} + b^{-1}\alpha_k) \geqslant 1 - b^{-1} \geqslant \frac{1}{2}$. Hence, this is less than the right-hand side of (4.31) in this case as well.

In order to prove the monotonicity statements, note that (4.7) along with $\theta < 1$ and $(n+1)^2 - n^2 \geqslant 1$ for $n \geqslant 0$ gives $c_{k+1} \leqslant b\theta c_k$. For $b\theta \leqslant 1$ this shows that $k \mapsto c_k$ is non-increasing. Using this on the right of (4.30) then rules out that $\alpha_{k+1} < \alpha_k$, proving that $k \mapsto \alpha_k$ is non-decreasing. □

We are now ready to give:

Proof of Theorem 3.5. Assume $b\theta = 1$. We start by processing the bound (4.31) a bit further. Putting the two terms on the right together we get

$$c_{k+1} \leqslant c_k \frac{b^{-1} + (1-b^{-1}) + (1-b^{-1})\binom{b}{2}c_k^2\alpha_k^2}{1+\binom{b}{2}c_k^2\alpha_k^2} \leqslant c_k \frac{1+r\binom{b}{2}c_k^2\alpha_k^2}{1+\binom{b}{2}c_k^2\alpha_k^2}, \tag{4.36}$$

where we abbreviated $r := 1 - b^{-1}$. Taking reciprocals and squaring gives

$$\frac{1}{c_{k+1}^2} \geqslant \frac{1}{c_k^2}\left[1 + \frac{[1+\binom{b}{2}c_k^2\alpha_k^2]^2 - [1+r\binom{b}{2}c_k^2\alpha_k^2]^2}{[1+r\binom{b}{2}c_k^2\alpha_k^2]^2}\right] \geqslant \frac{1}{c_k^2} + \frac{2\binom{b}{2}(1-r)\alpha_k^2}{[1+r\binom{b}{2}c_k^2\alpha_k^2]^2}. \tag{4.37}$$

Invoking $1 - r = b^{-1}$ along with $\alpha_k^2 \geqslant \alpha_0^2$ in the numerator while using that $r \leqslant 1$ along with $c_k^2 \leqslant c_0^2$ and $\alpha_k^2 \leqslant 1$ in the denominator then yields

$$\frac{1}{c_{k+1}^2} \geqslant \frac{1}{c_k^2} + \frac{\gamma}{c_0^2}, \quad k \geqslant 0, \tag{4.38}$$

where

$$\gamma := c_0^2 \frac{2\binom{b}{2}b^{-1}\alpha_0^2}{[1+\binom{b}{2}c_0^2]^2} \tag{4.39}$$

is a constant expressed only using the properties of a_0.

Iterating (4.38) shows $c_k^{-2} \geqslant c_0^{-2} + c_0^{-2}\gamma k = c_0^{-2}[1+\gamma k]$ and so

$$c_k \leqslant c_0[1+\gamma k]^{-1/2}, \quad k \geqslant 0. \tag{4.40}$$

Using this in (4.7), for the Fourier coefficients we get

$$a_{k+1}(n+1) \leqslant b^{n^2-(n+1)^2} b c_0 \big[1+\gamma k\big]^{-1/2} a_{k+1}(n), \quad n,k \geqslant 0 \tag{4.41}$$

which implies

$$a_{k+1}(n) \leqslant b^{-n^2}\big(bc_0[1+\gamma k]^{-1/2}\big)^n a_{k+1}(0), \quad n,k \geqslant 0. \tag{4.42}$$

As an inspection of (4.15) reveals, the right-hand side of this bound has a very similar structure as for the subcritical case; indeed, θ^{n^2} is replaced by b^{-n^2}, which is its value at criticality, while the error term $(b\theta)^k bc_0$ is replaced by $[1+\gamma k]^{-1/2}bc_0$. Proceeding as in the proof of Theorem 3.4, this implies the claim. □

Remark 4.7 Also in this case the control extends to all derivatives.

5. COUPLING

Here we will construct the coupling between the DG-model and the GFF stated in Theorem 3.6. Due to the discrete nature of the DG-field and, simultaneously, the continuum nature of the intermediate fields during iterations, the arguments split into two parts depending on whether we are dealing with continuum-valued or discrete-valued steps.

5.1 Continuum-valued steps.

Let $\|X\|_\infty$ denote the L^∞-norm of real-valued random variable X. We start with some general considerations pertaining to coupling of two random variables that admit a density with respect to a normal law.

Proposition 5.1 *Let μ be the law of $\mathcal{N}(0,\sigma^2)$ with $\sigma \in (0,\infty)$ and let X and Y have laws*

$$P\bigl(X \in A\bigr) = \int_A f_1(t)\mu(\mathrm{d}t) \quad \textit{and} \quad P\bigl(Y \in A\bigr) = \int_A f_2(t)\mu(\mathrm{d}t) \tag{5.1}$$

for some measurable $f_1, f_2 \colon \mathbb{R} \to (0,\infty)$. Denote the associated CDFs by $F(t) := P(X \leqslant t)$ and $G(t) := P(Y \leqslant t)$ and let $h\colon \mathbb{R} \to \mathbb{R}$ be defined by

$$h(t) := G^{-1}\bigl(F(t)\bigr) \tag{5.2}$$

Then

$$h(X) \overset{\text{law}}{=} Y \tag{5.3}$$

and, if f_1, $1/f_1$, f_2 and $1/f_2$ are bounded, then

$$\bigl\|h(X) - X\bigr\|_\infty < 3\sigma\sqrt{2\log\Bigl(\max\bigl\{\|f_1\|_\infty\|1/f_2\|_\infty,\, \|f_2\|_\infty\|1/f_1\|_\infty\bigr\}\Bigr)}. \tag{5.4}$$

The pair $(X, h(X))$ thus defines a coupling of X and Y with $X - Y$ bounded in L^∞.

The assumption that $f_2 > 0$ implies that G is strictly increasing and so G^{-1} is a proper inverse. A calculation shows

$$P\bigl(h(X) \leqslant t\bigr) = P\bigl(F(X) \leqslant G(t)\bigr) = G(t) \tag{5.5}$$

and so $h(X)$ has the law of Y. It remains to prove the bound (5.4) which we will do by proving that $h(t) - t$ is bounded. The argument comes in two lemmas dealing with various separate cases.

Lemma 5.2 *Suppose that $1/f_1$ and f_2 are bounded. Then*

$$h(t) \leqslant t + \frac{2}{t}\,\sigma^2\log\bigl(\,\|f_2\|_\infty\|1/f_1\|_\infty\bigr), \quad t > 0, \tag{5.6}$$

and

$$h(t) \geqslant t - \frac{2}{|t|}\,\sigma^2\log\bigl(\,\|f_2\|_\infty\|1/f_1\|_\infty\bigr), \quad t < 0. \tag{5.7}$$

Proof. First observe that $\tilde{X} := X/\sigma$ and $\tilde{Y} := Y/\sigma$ has density $\sigma f_1(\sigma\cdot)$ and $\sigma f_2(\sigma\cdot)$ with respect to the law of $\mathcal{N}(0,1)$. This means that, once we prove the above inequalities in the case $\sigma^2 = 1$, the case of general σ^2 follows the fact that $\tilde{h}(t) := G_{\tilde{Y}}^{-1}(F_{\tilde{X}}(t)) = \frac{1}{\sigma}h(\sigma t)$. So henceforth we assume that $\sigma^2 = 1$.

Writing g for the probability density of $\mathcal{N}(0,1)$, the definition (5.2) translates into

$$\int_{-\infty}^{t} f_1(x)g(x)\mathrm{d}x = \int_{-\infty}^{h(t)} f_2(x)g(x)\mathrm{d}x. \tag{5.8}$$

This shows that h is differentiable and obeys the ODE

$$f_1(t)g(t) = h'(t)f_2\bigl(h(t)\bigr)g\bigl(h(t)\bigr), \tag{5.9}$$

which takes the explicit form

$$h'(t) = \frac{f_1(t)}{f_2(h(t))} \mathrm{e}^{\frac{1}{2}(h(t)-t)(h(t)+t)}. \tag{5.10}$$

Suppose now that $t > 0$ and $\epsilon > 0$ are such that

$$h(t) \geqslant t + \epsilon \ \wedge \ \bigl(\|f_2\|_\infty \|1/f_1\|_\infty\bigr)^{-1} \mathrm{e}^{\epsilon t/2} > 1. \tag{5.11}$$

The first inequality implies

$$\frac{f_1(t)}{f_2(h(t))} \mathrm{e}^{(h(t)-t)(h(t)+t)/2} \geqslant \bigl(\|f_2\|_\infty \|1/f_1\|_\infty\bigr)^{-1} \mathrm{e}^{\epsilon t/2} \mathrm{e}^{\epsilon h(t)/2} \tag{5.12}$$

and the second shows that the right-hand side exceeds one. This in turn shows that $u \mapsto h(u) - u$ is non-decreasing for $u - t$ small positive. As this reinforces the conditions in (5.11), we readily infer that

$$(5.11) \ \Rightarrow \ h(u) \geqslant u + \epsilon \ \text{ when } \ u \geqslant t. \tag{5.13}$$

Abbreviating $c := (\|f_2\|_\infty \|1/f_1\|_\infty)^{-1}$, this shows

$$h'(u) \mathrm{e}^{-\epsilon h(u)/2} \geqslant c \mathrm{e}^{\epsilon u/2}, \quad u \geqslant t, \tag{5.14}$$

which integrates to

$$\mathrm{e}^{-\epsilon h(t)/2} - \mathrm{e}^{-\epsilon h(u)/2} \geqslant c\bigl(\mathrm{e}^{\epsilon u/2} - \mathrm{e}^{\epsilon t/2}\bigr), \quad u \geqslant t. \tag{5.15}$$

But this is absurd because the left-hand side is bounded while the right-hand side diverges as $u \to \infty$. No $t > 0$ and $\epsilon > 0$ satisfying (5.11) may thus exist and so $h(t) < t + \epsilon$ once $c\mathrm{e}^{\epsilon t/2} > 1$. Solving for ϵ yields (5.6).

Suppose now that $t < 0$ and $\epsilon > 0$ are such that

$$h(t) \leqslant t - \epsilon \ \wedge \ \bigl(\|f_2\|_\infty \|1/f_1\|_\infty\bigr)^{-1} \mathrm{e}^{-\epsilon t/2} > 1. \tag{5.16}$$

Then

$$\frac{f_1(t)}{f_2(h(t))} \mathrm{e}^{(h(t)-t)(h(t)+t)/2} \geqslant \bigl(\|f_2\|_\infty \|1/f_1\|_\infty\bigr)^{-1} \mathrm{e}^{-\epsilon t/2} \mathrm{e}^{-\epsilon h(t)/2}. \tag{5.17}$$

This again forces $h(u) \leqslant u - \epsilon$ as well as $h'(u) > 1$ for all $u \leqslant t$ and thus proves

$$h'(u) \mathrm{e}^{\epsilon h(u)/2} \geqslant c \mathrm{e}^{-\epsilon u/2}, \quad u \leqslant t, \tag{5.18}$$

where c is as above. Integrating we get

$$\mathrm{e}^{\epsilon h(t)/2} - \mathrm{e}^{\epsilon h(u)/2} \geqslant c\bigl(\mathrm{e}^{-\epsilon u/2} - \mathrm{e}^{-\epsilon t/2}\bigr), \quad u \leqslant t, \tag{5.19}$$

which is again absurd when $u \to -\infty$. Hence we conclude (5.7) as well. □

Next we need to control the difference $h(t) - t$ in the regimes not covered by Lemma 5.2. Explicitly, we need an upper bound when $t < 0$ and a lower bound when $t > 0$. This is the content of:

Lemma 5.3 *Suppose that $1/f_2$ and f_1 are bounded. Then*

$$0 \leqslant h(t) \ \Rightarrow \ h(t) \geqslant t - \frac{2}{t} \sigma^2 \log\bigl(\|f_1\|_\infty \|1/f_2\|_\infty\bigr), \quad t > 0, \tag{5.20}$$

and

$$h(t) \leqslant 0 \;\Rightarrow\; h(t) \leqslant t + \frac{2}{|t|}\,\sigma^2 \log\bigl(\,\|f_1\|_\infty \|1/f_2\|_\infty\bigr), \quad t < 0. \tag{5.21}$$

Proof. We again assume without loss of generality that $\sigma^2 = 1$. Suppose that, for some $\epsilon > 0$ and $t > \epsilon$,

$$0 \leqslant h(t) \leqslant t - \epsilon \;\wedge\; \|f_1\|_\infty \|1/f_2\|_\infty \,\mathrm{e}^{-\epsilon t/2} < 1. \tag{5.22}$$

Then

$$\frac{f_1(t)}{f_2(h(t))}\mathrm{e}^{(h(t)-t)(h(t)+t)/2} \leqslant \|f_1\|_\infty \|1/f_2\|_\infty \,\mathrm{e}^{-\epsilon t/2}\mathrm{e}^{-\epsilon h(t)/2}. \tag{5.23}$$

This shows $h'(t) < 1$ and thus that $u \mapsto h(u) - u$ is non-increasing for $u - t$ small positive. As this reinforces the conditions in (5.22) and h is non-decreasing by definition, we deduce that

$$0 \leqslant h(u) \leqslant u - \epsilon, \quad u \geqslant t. \tag{5.24}$$

But then

$$h'(u)\mathrm{e}^{\epsilon h(u)/2} \leqslant \|f_1\|_\infty \|1/f_2\|_\infty \,\mathrm{e}^{-\epsilon u/2}, \quad u \geqslant t. \tag{5.25}$$

which integrates into

$$\mathrm{e}^{\epsilon h(u)/2} - \mathrm{e}^{\epsilon h(t)/2} \leqslant \|f_1\|_\infty \|1/f_2\|_\infty \left(\mathrm{e}^{-\epsilon t/2} - \mathrm{e}^{-\epsilon u/2}\right), \quad u \geqslant t. \tag{5.26}$$

This contradicts the fact that $h(u) \to \infty$ as $u \to \infty$. No $t > 0$ and $\epsilon > 0$ satisfying (5.22) can thus exist, proving (5.20).

Next suppose that, for some $\epsilon > 0$ and $t < -\epsilon$,

$$t + \epsilon \leqslant h(t) \leqslant 0 \;\wedge\; \|f_1\|_\infty \|1/f_2\|_\infty \,\mathrm{e}^{\epsilon t/2} < 1. \tag{5.27}$$

Then

$$\frac{f_1(t)}{f_2(h(t))}\mathrm{e}^{(h(t)-t)(h(t)+t)/2} \leqslant \|f_1\|_\infty \|1/f_2\|_\infty \,\mathrm{e}^{\epsilon t/2}\mathrm{e}^{\epsilon h(t)/2}. \tag{5.28}$$

This shows $h'(t) < 1$ and thus that $u \mapsto h(u) - u$ is non-increasing for $u - t$ small negative. As this again reinforces the conditions in (5.27) and h is non-decreasing by definition, we deduce that

$$u + \epsilon \leqslant h(u) \leqslant 0, \quad u \leqslant t. \tag{5.29}$$

But then

$$h'(u)\mathrm{e}^{-\epsilon h(u)/2} \leqslant \|f_1\|_\infty \|1/f_2\|_\infty \,\mathrm{e}^{\epsilon u/2}, \quad u \leqslant t, \tag{5.30}$$

which integrates into

$$\mathrm{e}^{-\epsilon h(u)/2} - \mathrm{e}^{-\epsilon h(t)/2} \leqslant \|f_1\|_\infty \|1/f_2\|_\infty \left(\mathrm{e}^{\epsilon t/2} - \mathrm{e}^{\epsilon u/2}\right), \quad u \leqslant t. \tag{5.31}$$

This contradicts the fact that $h(u) \to -\infty$ as $u \to -\infty$. No $t < 0$ and $\epsilon > 0$ satisfying (5.27) can thus exist, proving (5.21). □

With Lemmas 5.2 and 5.3 in hand, we now give:

Proof of Proposition 5.1. First observe that $\|f_1\|_\infty \|1/f_2\|_\infty \geqslant 1$ for otherwise $f_1(t) < f_2(t)$ for all $t \in \mathbb{R}$, contradicting $\int_{\mathbb{R}} f_1(t)\mu(dt) = \int_{\mathbb{R}} f_2(t)\mu(dt) = 1$. Similarly, $\|f_2\|_\infty \|1/f_1\|_\infty \geqslant 1$ holds as well. This allows us to define

$$A := 2\sigma^2 \log\Bigl(\,\max\bigl\{\|f_1\|_\infty \|1/f_2\|_\infty, \|f_2\|_\infty \|1/f_1\|_\infty\bigr\}\Bigr). \tag{5.32}$$

Note that, since h is continuous, (5.20–5.21) forces $h(t) > 0$ on $[\sqrt{A},\infty)$ and $h(t) < 0$ on $(-\infty,-\sqrt{A}]$. From (5.6–5.7) and (5.20–5.21) we then get

$$\sup_{t\colon |t|\geqslant\sqrt{A}} \bigl|h(t)-t\bigr| \leqslant \sqrt{A}. \tag{5.33}$$

For $t\in[-\sqrt{A},\sqrt{A}]$ the monotonicity of h shows

$$-2\sqrt{A}\leqslant h(-\sqrt{A})\leqslant h(t)\leqslant h(\sqrt{A})\leqslant 2\sqrt{A} \tag{5.34}$$

and so

$$\sup_{t\colon |t|\leqslant\sqrt{A}} \bigl|h(t)-t\bigr| \leqslant 3\sqrt{A}. \tag{5.35}$$

Along with (5.33), this gives the desired claim. □

5.2 Discrete-valued steps and proof of Theorem 3.6.

We will now move to the proof of Theorem 3.6. Here we note that Proposition 5.1 gives a bound on the coupling distance but does not provide any bound on the probability of equality. This is fixed in:

Lemma 5.4 *Let $\mu=\mathcal{N}(0,\sigma^2)$ and $f\colon\mathbb{R}\to(0,\infty)$ be such that $\int f\,\mathrm{d}\mu=1$. Let X have the law μ and let Y have the law $f\mu$. Assuming f and $1/f$ to be bounded and $\mu(f\neq 1)>0$, for all strictly positive $\eta\geqslant\log\|1/f\|_\infty$ there exists a coupling of X, Y and a zero-one valued random variable B such that*

$$\|X-Y\|_\infty\leqslant 3\sigma\sqrt{2\log\biggl(\max\biggl\{\frac{\mathrm{e}^{2\eta}\|f\|_\infty-1}{\mathrm{e}^{2\eta}-1},\,\mathrm{e}^{\eta}+1\biggr\}\biggr)}, \tag{5.36}$$

$$X=Y \text{ on } \{B=1\}, \tag{5.37}$$

$$B \text{ and } X \text{ are independent} \tag{5.38}$$

and

$$P(B=1)=\mathrm{e}^{-2\eta}. \tag{5.39}$$

(When $\mu(f\neq1)=0$, we have $X\overset{\text{law}}{=}Y$ and we can use $Y:=X$ and $B:=1$.)

Proof. Let $\eta\geqslant\log\|1/f\|_\infty$ be strictly positive. Sample X from μ and let B be an independent Bernoulli with $P(B=1):=\mathrm{e}^{-2\eta}$. Then set $Y:=X$ when $B=1$ and $Y:=h(X)$ when $B=0$, for h as in (5.2) with F being the CDF of μ and G the CDF of $\tilde f\mu$ where

$$\tilde f:=\frac{f-\mathrm{e}^{-2\eta}}{1-\mathrm{e}^{-2\eta}}. \tag{5.40}$$

As is readily checked, this defines a coupling of (X,Y,B) with the given marginals and such that (5.37–5.39) hold.

In order to prove (5.36), note that the bound in Proposition 5.1 gives

$$\|X-Y\|_\infty\leqslant 3\sigma\sqrt{2\log\bigl(\max\{\,\|\tilde f\|_\infty,\,\|1/\tilde f\|_\infty\}\bigr)}. \tag{5.41}$$

A computation shows

$$\|\tilde f\|_\infty\leqslant\frac{\|f\|_\infty-\mathrm{e}^{-2\eta}}{1-\mathrm{e}^{-2\eta}} \tag{5.42}$$

while

$$\|1/\tilde{f}\|_\infty \leqslant \frac{1-\mathrm{e}^{-2\eta}}{\mathrm{e}^{-\eta}-\mathrm{e}^{-2\eta}} = \mathrm{e}^{\eta}+1. \tag{5.43}$$

Inserting these in (5.41), we get (5.36). □

Lemma 5.4 is sufficient to handle the coupling of the increments $\zeta_k^{\mathrm{DG}}(x)$ and $\zeta_k^{\mathrm{GFF}}(x)$ for all $k = 1, \dots, n-1$. For the discrete-valued step (the case $k = 0$ of the DG-field) we will instead need a different argument for uniform closeness. This comes in:

Lemma 5.5 *For each $\beta > 0$ there exists $C_0 > 0$ such that the following holds for all $z \in \mathbb{R}$: If $X \overset{\text{law}}{=} \mathcal{N}(0, 1/\beta)$ and Y takes values in $z + \mathbb{Z}$ with probabilities*

$$P\big(Y = z+n\big) = \mathrm{e}^{v_0(z) - \frac{\beta}{2}(z+n)^2}, \quad n \in \mathbb{Z}, \tag{5.44}$$

for v_0 as in (3.21)*, then there exists a coupling of X and Y such that*

$$\|X - Y\|_\infty \leqslant C_0. \tag{5.45}$$

Proof. Since the law of Y does not change if z is changed by an integer (this also uses the periodicity of v_0), we may and will henceforth assume $z \in [0, 1)$. The proof uses similar arguments as those underlying Proposition 5.1 albeit adapted to the discrete setting.

We start by noting that, in light of the strict positivity of the Gaussian density, there exists a unique non-decreasing $\{a_k\}_{k\in\mathbb{Z}}$ of reals such that

$$\int_{z+a_k}^{\infty} \frac{1}{\sqrt{2\pi/\beta}}\, \mathrm{e}^{-\frac{\beta}{2}x^2} \mathrm{d}x = P(Y \geqslant z+k), \quad k \in \mathbb{Z}. \tag{5.46}$$

The positivity also implies $a_k \to \pm\infty$ as $k \to \pm\infty$. Now let $h\colon \mathbb{R} \to \mathbb{R}$ be such that

$$h(x) := z+k \;\text{ if }\; x \in [z+a_k, z+a_{k+1}) \tag{5.47}$$

for $k \in \mathbb{Z}$. The definition (5.46) then gives $h(X) \overset{\text{law}}{=} Y$.

In order to show (5.45), we thus have to prove that $\sup_{k\in\mathbb{Z}} |a_k - k| < \infty$. To this end we employ the explicit form of the law of Y along with a substitution and some algebra to turn (5.46) into

$$\int_0^\infty \mathrm{e}^{-\frac{\beta}{2}(x+a_k+z)^2} \mathrm{d}x = \mathrm{e}^{v_0(z)} \sqrt{2\pi/\beta}\, \mathrm{e}^{-\frac{\beta}{2}(z+k)^2} \left(1 + \sum_{j>k} \mathrm{e}^{-\frac{\beta}{2}(z+j)^2 + \frac{\beta}{2}(z+k)^2}\right). \tag{5.48}$$

Abbreviating $c_\beta := \mathrm{e}^{v_0(z)}\sqrt{2\pi/\beta}$ and $A_\beta(k) := \sum_{j>k} \mathrm{e}^{-\frac{\beta}{2}(z+j)^2+\frac{\beta}{2}(z+k)^2}$, this becomes

$$\int_0^\infty \mathrm{e}^{-\frac{\beta}{2}x^2 - \beta(z+a_k)x} \mathrm{d}x = c_\beta\, \mathrm{e}^{-\frac{\beta}{2}(z+k)^2 + \frac{\beta}{2}(z+a_k)^2}\, \big(1 + A_\beta(k)\big) \tag{5.49}$$

whereby we obtain

$$\int_0^\infty \mathrm{e}^{-\frac{x^2}{2\beta(z+a_k)} - x} \mathrm{d}x = \beta c_\beta (z+a_k) \mathrm{e}^{-\frac{\beta}{2}(z+k)^2 + \frac{\beta}{2}(z+a_k)^2}\, \big(1 + A_\beta(k)\big) \tag{5.50}$$

by employing another substitution under the integral.

We will now use (5.50) to control $a_k - k$ for k large positive. First note that $a_k \to \infty$ as $k \to \infty$ tells us there exists $k_1 \in \mathbb{Z}$ such that $a_k \geqslant 1$ for all $k \geqslant k_1$. Suppose we had $a_k \leqslant k-1$ for some $k \geqslant k_1$. Then (5.50) gives

$$\int_0^\infty \mathrm{e}^{-\frac{x^2}{2\beta(1+z)-x}\mathrm{d}x} \leqslant \beta c_\beta(k+z-1)\mathrm{e}^{-\frac{\beta}{2}(2(k+z)+1)}\bigl(1+A_\beta(k)\bigr). \tag{5.51}$$

But

$$A_\beta(k) = \sum_{j>k} \mathrm{e}^{-\frac{\beta}{2}(j-k)(k+j+2z)} = \sum_{l=1}^\infty \mathrm{e}^{-\frac{\beta}{2}l(2k+l+2z)} \leqslant \sum_{l=1}^\infty \mathrm{e}^{-\frac{\beta}{2}(l^2+2kl)} < \infty \tag{5.52}$$

tells us that the right side of (5.51) vanishes in the limit as $k \to \infty$, which is absurd since the left-hand side remains uniformly positive. It follows that there exists $K_1 \geqslant k_1$ so that $k \geqslant K_1$ implies $a_k > k-1$.

Next assume $a_k \geqslant k+1$. Then the equality in (5.50) gives

$$1 = \int_0^\infty \mathrm{e}^{-x}\mathrm{d}x \geqslant \beta c_\beta(k+z+1)\mathrm{e}^{\frac{\beta}{2}(2(k+z)+1)} \tag{5.53}$$

which is again absurd for k large because the right hand side diverges as $k \to \infty$. It follows that there exists K_1' such that $k \geqslant K_1'$ implies $a_k < k+1$. In particular, $k-1 < a_k < k+1$ holds for $k \geqslant K := \max\{K_1, K_1'\}$.

Similar estimates work for k negative. Indeed, (5.50) becomes

$$\int_{-\infty}^0 \mathrm{e}^{-\frac{x^2}{2\beta(-z-a_k)}+x}\mathrm{d}x = \beta c_\beta(-z-a_k)\mathrm{e}^{-\frac{\beta}{2}(z+k)^2+\frac{\beta}{2}(z+a_k)^2}\Bigl(1+A'_\beta(k)\Bigr), \tag{5.54}$$

where

$$A'_\beta(k) := \sum_{j<k} \mathrm{e}^{-\frac{\beta}{2}(z+j)^2+\frac{\beta}{2}(z+k)^2} = \sum_{l=1}^\infty \mathrm{e}^{-\frac{\beta}{2}l(-2k+l-2z)} \leqslant \sum_{l=1}^\infty \mathrm{e}^{-\frac{\beta}{2}(l^2-2kl-2l)} < \infty. \tag{5.55}$$

As before, $a_k \to -\infty$ as $k \to -\infty$ tells us there exists $k_2 > 0$ such that $a_k \leqslant -1$ holds for all $k \leqslant -k_2$. If we had $-a_k \leqslant -k-1$ for some $k \leqslant -k_2$, then (5.54) would give

$$\int_{-\infty}^0 \mathrm{e}^{-\frac{x^2}{2\beta(1-z)}-x}\mathrm{d}x \leqslant \beta c_\beta(-k-z-1)\mathrm{e}^{-\frac{\beta}{2}(2(-z-k)+1)}\Bigl(1+A'_\beta(k)\Bigr). \tag{5.56}$$

which is again absurd for k large negative because the right hand side vanishes as $k \to -\infty$. It follows that there exists $K_2 \geqslant k_2$ so that $k \leqslant -K_2$ implies $-a_k > -k-1$. To rule out that $-a_k \geqslant -k+1$, note that (5.54) implies

$$1 = \int_{-\infty}^0 \mathrm{e}^{x}\mathrm{d}x \geqslant \beta c_\beta(-k-z+1)\mathrm{e}^{\frac{\beta}{2}(2(-z-k)+1)} \tag{5.57}$$

where the right-hand side diverges to $+\infty$ as $k \to -\infty$. Hence there exists $K_2' > 0$ such that $k \leqslant -K_2'$ forces $-a_k < -k+1$. In particular, $k-1 < a_k < k+1$ once $-k \geqslant K' := \max\{K_2, K_2'\}$.

To summarize the above inequalities, note that we have proved that $|a_k - k| < 1$ when $|k| \geqslant \max\{K, K'\}$ while the monotonicity of $k \mapsto a_k$ then tells us that $|a_k| \leqslant 1+\max\{K, K'\}$ when $|k| \leqslant \max\{K, K'\}$. It follows that $|a_k - k| \leqslant C_0 := 1+\max\{K, K'\}$ for all $k \in \mathbb{Z}$, as desired. $\square$

We are now ready to give:

Proof of Theorem 3.6. Fix $\beta \in (0, \beta_c)$ and $n \geqslant 1$. For $k = 1, \dots, n-1$, define R_k by

$$R_k := 2 \sup_{z,z' \in [0,1]} \big| v_k(z) - b v_{k-1}(z') \big|. \tag{5.58}$$

Then sample i.i.d. random variables (3.47) with common law $\mathcal{N}(0, 1/\beta)$ and independent Bernoulli random variables (3.48) with

$$P\big(B_k(x) = 1\big) = \mathrm{e}^{-R_k}, \quad x \in \Lambda_{n-k},\, k = 1, \dots, n-1. \tag{5.59}$$

Noting that the density governing the definition of $\xi_k^{\mathrm{DG}}(x)$ from $\xi_k^{\mathrm{GFF}}(x)$ takes the form

$$f_{k,x}(z) := \mathrm{e}^{v_k(\varphi_k(x)) - b v_{k-1}(z + \varphi_k(x))} \tag{5.60}$$

where $\varphi_k(x) := \sum_{j=k+1}^{n-1} \xi_j^{\mathrm{DG}}(x)$ and, in particular, $\varphi_{n-1}(x) := 0$. Lemma 5.4 with $\eta := R_k/2$ (which exceeds $\log \|1/f_{k,x}\|_\infty$) allows us to recursively construct $\xi_k^{\mathrm{DG}}(x)$ for $x \in \Lambda_n$ and $k = n-1, \dots, 1$ as a deterministic function of the GFF increments (3.47) and the Bernoulli's (3.48). For $k = 0$ we in turn use Lemma 5.5.

The bound (5.36) now shows

$$\big\| \xi_k^{\mathrm{DG}}(x) - \xi_k^{\mathrm{GFF}}(x) \big\|_\infty \leqslant 3\beta^{-1/2} \sqrt{2 \log(\mathrm{e}^{R_k} + 1)} \tag{5.61}$$

for all $k = 1, \dots, n-1$ and $x \in \Lambda_n$, while (5.45) bounds the $k = 0$ case by a constant C_0. Theorem 3.4 asserts that $R_k \to 0$ as $k \to \infty$ and so the right-hand side of (5.61) is bounded uniformly on k. Moreover, the exponential decay of R_k in subcritical regime proven in Theorem 3.4 gives $\limsup_{k\to\infty} k^{-1} R_k < 0$. Denoting

$$C_1 := 3\beta^{-1/2} \sqrt{2 \log(\mathrm{e}^{\sup_{k \geqslant 1} R_k} + 1)} \tag{5.62}$$

the claim thus holds with $C := \max\{C_0, C_1\}$. $\square$

6. Tightness of DG extrema

With the needed tools assembled, we now finally turn our attention to the extremal behavior of the DG-model. The goal of this section is to establish tightness of the maximum (centered by m_n) as well as the level sets (above $m_n - \lambda$ with $\lambda > 0$ fixed). These will be useful in the proofs in Section 7.

6.1 Statements and preliminaries.

Recall that $P_{n,\beta}$ denotes the law of the DG-model at inverse temperature β and depth n. We start with the statement of the relevant results to be proved in this section:

Theorem 6.1 *For all $\beta \in (0, \beta_c)$ and with m_n as in (2.5),*

$$\lim_{u\to\infty} \sup_{n \geqslant 1} P_{n,\beta}\Big(\big| \max_{x \in \Lambda_n} \varphi_x^{\mathrm{DG}} - m_n \big| \geqslant u \Big) = 0. \tag{6.1}$$

Theorem 6.2 *Given $\lambda \in \mathbb{R}$ and a sample φ^{DG} from $P_{n,\beta}$, let*

$$G_n(\lambda) := \big| \{ x \in \Lambda_n \colon \varphi_x^{\mathrm{DG}} \geqslant m_n - \lambda \} \big|. \tag{6.2}$$

For all $\beta \in (0, \beta_c)$ and all $\lambda > 0$,

$$\lim_{u\to\infty} \sup_{n\geqslant 1} P_{n,\beta}\big(G_n(\lambda) \geqslant u\big) = 0. \tag{6.3}$$

The proofs will rely on some standard results about the GFF which we will state next. Recall that the GFF at inverse temperature β on the hierarchical lattice is nothing but a Branching Random Walk with steps distributed according to $\mathcal{N}(0, 1/\beta)$. The tightness (and, in fact, limit laws of the maxima) of general Branching Random Walks has been studied extensively (e.g., by Aïdekon [4] and Madaule [37]). For Gaussian step distributions, the calculations become very explicit. Writing $P'_{n,\beta}$ for the law of the GFF of Λ_n, we summarize what we need in:

Lemma 6.3 *There exists $a > 0$ such that for all $t > 0$ and all $\beta > 0$,*

$$\sup_{n\geqslant 1} P'_{n,\beta}\Big(\big|\max_{x\in\Lambda_n} \varphi^{\text{GFF}}_x - m_n\big| > t\Big) \leqslant \frac{1}{a}\,\mathrm{e}^{-a\sqrt{\beta}\,t}. \tag{6.4}$$

Proof. This is proved, modulo scaling by β and quotations of "ballot theorems," in the review article [16, Lemmas 7.3 and 7.15]. Alternatively, consult Mallein [38, Theorem 4.1] which gives upper tail for general Branching Random Walks as well as the key technical input $\inf_{n\geqslant 1} P'_{n,\beta}(\max_{x\in\Lambda_n} \varphi^{\text{GFF}}_x \geqslant m_n) > 0$ for the lower tail. □

Lemma 6.4 *There exist $c', C' > 0$ and, for all $\beta > 0$, all $t > 0$ and all $u < t$, there exists $n_0 \geqslant 1$ such that for all $n \geqslant n_0$ and all $z \in \Lambda_n$,*

$$P'_{n,\beta}\left(\varphi^{\text{GFF}}_z \geqslant m_n + u, \max_{x\in\Lambda_n} \varphi^{\text{GFF}}_x \leqslant m_n + t\right) \leqslant C' b^{-n}\big(1 + (t \vee (t-u))^2\beta\big)\mathrm{e}^{-\sqrt{\beta}c'u}. \tag{6.5}$$

Proof. This follows by conditioning on φ^{GFF}_z and applying a "ballot estimate." See, e.g., the proof of [20, Lemma 6.9] for the (harder) case of lattice GFF. □

We note that n_0 in Lemma 6.4 can and will be chosen so that $m_{n_0} + u$ is a large positive number. We also record a standard estimate on the tail of sum of Bernoulli random variables based on the Chernoff bound:

Lemma 6.5 *Let $\{Z_k\}_{k\geqslant 1}$ be independent zero-one valued random variables. Abbreviate $p_k := P(Z_k = 0)$. Then for all natural $k \leqslant l$ and $r > p_{k,l} := \sum_{i=k}^{l} p_i$,*

$$P\bigg(\sum_{i=k}^{l} 1_{Z_i=0} > r\bigg) \leqslant \exp\Big\{r - p_{k,l} + r\log p_{k,l} - r\log r\Big\}. \tag{6.6}$$

In particular, if $\sum_{k=1}^{\infty} p_k < \infty$ then there exists $r_0 \geqslant 1$ such that for all $r \geqslant r_0$

$$P\bigg(\sum_{i=k}^{l} 1_{Z_i=0} > r\bigg) \leqslant \exp\Big\{-\frac{1}{2} r\log r\Big\} \tag{6.7}$$

holds for all $l \geqslant k \geqslant 1$.

Proof. Given any $\lambda > 0$, the Chernoff bound shows

$$\begin{aligned} P\Big(\sum_{i=k}^{l} 1_{Z_i=0} > r\Big) &\leqslant \mathrm{e}^{-r\lambda} E\Big[\exp\Big(\lambda \sum_{i=k}^{l} 1_{Z_i=0}\Big)\Big] \\ &= \exp\Big\{-r\lambda + \sum_{i=k}^{l} \log\big(1+(\mathrm{e}^{\lambda}-1)p_i\big)\Big\} \leqslant \exp\Big\{-r\lambda + \sum_{i=k}^{l}(\mathrm{e}^{\lambda}-1)p_i\Big\}. \end{aligned} \tag{6.8}$$

Now set $\lambda := \log(r/p_{k,l})$ and observe that $r \geqslant 1 > p_{k,\ell}$ gives $\lambda > 0$ to get

$$P\Big(\sum_{i=k}^{l} 1_{Z_i=0} > r\Big) \leqslant \exp\Big\{-r\log r + r\log p_{k,l} + r - p_{k,l}\Big\}. \tag{6.9}$$

The bound (6.7) follows immediately from

$$\exp\Big\{-\frac{1}{2} r\log r + r\log p_{k,l} + r - p_{k,l}\Big\} \leqslant \exp\Big\{-\frac{1}{2} r\log r + r\log\Big(\sum_{k=1}^{\infty} p_k\Big) + r\Big\} \tag{6.10}$$

and the fact that the right-hand side tends to zero as $r \to \infty$. □

6.2 Key proposition.

We now move to the statement and proof of a technical proposition that provides the key step in the proof of both Theorem 6.1 and 6.2.

Proposition 6.6 *For each $\beta \in (0, \beta_c)$ there exists $C'' > 0$ and, for all $t > 0$ and all $u < t$, there exists $n_0 \geqslant 1$ such that for all $n \geqslant n_0$ and all $z \in \Lambda_n$,*

$$P\left(\varphi_z^{\mathrm{DG}} \geqslant m_n + u, \max_{x\in\Lambda_n} \varphi_x^{\mathrm{GFF}} \leqslant m_n + t\right) \leqslant C'' b^{-n}\big(1 + (t \vee (t-u+1))^2\beta\big)\mathrm{e}^{-\sqrt{\beta}\,c' u}, \tag{6.11}$$

where c' is as in Lemma 6.4 and P is the coupling law from Theorem 3.6.

The proof is based on swapping φ_z^{DG} for φ_z^{GFF} which effectively reduces the claim to Lemma 6.4. To control the error incurred by the swap, we have to bound the probability that the difference $\varphi_z^{\mathrm{DG}} - \varphi_z^{\mathrm{GFF}}$ is large on the background of the event that φ_z^{DG} is itself large (as this provides the much needed b^{-n} term).

The argument relies on the coupling from Theorem 3.6 but its use is complicated by the restriction on the GFF-maximum (which is needed to beat a polynomial term in (6.12)). Indeed, this maximum is a global event that involves, in principle, all the coupling variables. That being said, this restriction is not needed if the difference $\varphi_x^{\mathrm{DG}} - \varphi_x^{\mathrm{GFF}}$ is sufficiently large. This will be handled using:

Lemma 6.7 *Let $\beta \in (0, \beta_c)$ and let C and $\{R_k\}_{k\geqslant 1}$ be as in Theorem 3.6. There exists $C''' > 0$ such that for all $n \geqslant 1$, all $x \in \Lambda_n$ and all $k \geqslant 1$,*

$$P\big(\varphi_x^{\mathrm{DG}} \geqslant m_n + u, \varphi_x^{\mathrm{DG}} - \varphi_x^{\mathrm{GFF}} > Ck\big) \leqslant C''' n b^{-n} \mathrm{e}^{-\sqrt{\beta}\,c' u} \sum_{j=k}^{\infty} R_j, \tag{6.12}$$

where c' is as in Lemma 6.4.

Proof. Fix $x \in \Lambda_n$ and note that, since the event involves only the coupling variables on the path from x to the root, we may write $\xi_j^{\rm DG}$, $\xi_j^{\rm GFF}$, and B_i instead of $\xi_j^{\rm DG}(x)$, $\xi_j^{\rm GFF}(x)$, and $B_i(x)$, respectively. Define

$$\tau = \tau(x) := n \wedge \sup\{j = 1,\dots,n-1 : B_j = 0\} \tag{6.13}$$

to be the height that (as far as the Bernoulli's (3.48) are concerned) the coupling fails for the first time along the path from the root to x. By Theorem 3.6 and our choice of C, the event $\varphi_x^{\rm DG} - \varphi_x^{\rm GFF} > Ck$ implies $\tau \geqslant k$. The probability in (6.12) is thus bounded by

$$P(\varphi_x^{\rm DG} \geqslant m_n + u, \tau \geqslant k) = \sum_{i=k}^{n-1} P(\varphi_x^{\rm DG} \geqslant m_n + u, \tau = i), \tag{6.14}$$

where the default case $\tau = n$ is ruled out by the fact that, on the event $\{\varphi_x^{\rm DG} - \varphi_x^{\rm GFF} > C\}$, not all B_j can be one.

Next we note that, since only the variables on the "path" to a single $x \in \Lambda_n$ are involved, the DG-process and the GFF-process are mutually absolutely continuous. More precisely, let $\mathcal{F}_k := \sigma(\xi_i^{\rm DG}, \xi_i^{\rm GFF}, B_i \colon i \geqslant k)$ and abbreviate $\xi_{\geqslant k} := \sum_{i=k}^{n-1} \xi_i$, with the superscript "DG" or "GFF" added depending on the context. Recall the notation $\mathfrak{q}_k(\cdot|\varphi)$ from (3.24) and let μ denote the law of $\mathcal{N}(0, 1/\beta)$. Then for each Borel $A \subseteq \mathbb{R}^k$, (3.25) along with (3.42) and (3.44) give

$$\begin{aligned} P\biggl((\xi_0^{\rm DG},\dots,\xi_{k-1}^{\rm DG}) \in A \,\bigg|\, \mathcal{F}_k \biggr) &= \int_A \bigotimes_{i=0}^{k-1} \mathfrak{q}_i(\mathrm{d}\xi_i | \xi_{\geqslant i+1}) \\ &\leqslant \int_A \Bigl(\prod_{j=0}^{k-1} \mathrm{e}^{R_j/2} \Bigr) \bigotimes_{i=0}^{k-1} \mu(\mathrm{d}\xi_i) \leqslant cP\biggl((\xi_0^{\rm GFF},\dots,\xi_{k-1}^{\rm GFF}) \in A \,\bigg|\, \mathcal{F}_k \biggr), \end{aligned} \tag{6.15}$$

where $R_0 := 2\mathrm{e}^{\|v_0\|}$ and $c := \exp\{\frac12 \sum_{j=0}^{\infty} R_j\} < \infty$. On the event $\{|\xi_{\geqslant k}^{\rm DG} - \xi_{\geqslant k}^{\rm GFF}| \leqslant C\}$, which lies in $\mathcal{F}_k$, we thus get

$$\begin{aligned} P\bigl(\varphi_x^{\rm DG} \geqslant m_n + u \,\big|\, \mathcal{F}_k\bigr) &= P\bigl(\varphi_x^{\rm DG} - \xi_{\geqslant k}^{\rm DG} \geqslant m_n + u - \xi_{\geqslant k}^{\rm DG} \,\big|\, \mathcal{F}_k\bigr) \\ &= P\biggl(\sum_{j=0}^{k-1} \xi_j^{\rm DG} \geqslant m_n + u - \xi_{\geqslant k}^{\rm DG} \,\bigg|\, \mathcal{F}_k \biggr) \\ &\leqslant cP\biggl(\sum_{j=0}^{k-1} \xi_j^{\rm GFF} \geqslant m_n + u - \xi_{\geqslant k}^{\rm DG} \,\bigg|\, \mathcal{F}_k \biggr) \\ &\leqslant cP\biggl(\sum_{j=0}^{k-1} \xi_j^{\rm GFF} \geqslant m_n + u - \xi_{\geqslant k}^{\rm GFF} - C \,\bigg|\, \mathcal{F}_k \biggr) \\ &= cP\bigl(\varphi_x^{\rm GFF} \geqslant m_n + u - C \,\big|\, \mathcal{F}_k\bigr), \end{aligned} \tag{6.16}$$

where we use that $\xi_{\geqslant k}^{\rm DG}$ and $\xi_{\geqslant k}^{\rm GFF}$ are $\mathcal{F}_k$-measurable to treat them as constants under the conditional probability.

Since $\zeta^{\text{DG}}_{\geqslant i+1} = \zeta^{\text{GFF}}_{\geqslant i+1}$ on $\{\tau = i\}$, Theorem 3.6 shows $\{\tau = i\} \subseteq \{|\zeta^{\text{DG}}_{\geqslant i} - \zeta^{\text{GFF}}_{\geqslant i}| \leqslant C\}$. This enables the inequality in (6.16) which then yields

$$\begin{aligned} P\big(\varphi^{\text{DG}}_x \geqslant m_n + u, \tau = i\big) &= E\Big[P\big(\varphi^{\text{DG}}_x \geqslant m_n + u \,\big|\, \mathcal{F}_i\big) 1_{\{\tau=i\}} 1_{\{|\zeta^{\text{DG}}_{\geqslant i} - \zeta^{\text{GFF}}_{\geqslant i}| \leqslant C\}}\Big] \\ &\leqslant cE\Big[P\big(\varphi^{\text{GFF}}_x \geqslant m_n + u - C \,\big|\, \mathcal{F}_i\big) 1_{\{\tau=i\}}\Big] \\ &= cP\big(\varphi^{\text{GFF}}_x \geqslant m_n + u - C\big) P(\tau = i), \end{aligned} \tag{6.17}$$

where in the final step we used that the Bernoulli's (3.48) are independent of the GFF increments, as stated in Theorem 3.6(2).

From (6.14) and and explicit calculation we now deduce

$$P\big(\varphi^{\text{DG}}_x \geqslant m_n + u, \varphi^{\text{DG}}_x - \varphi^{\text{GFF}}_x > Ck\big) \leqslant C''' n b^{-n} \mathrm{e}^{-\sqrt{\beta}\, c' u} P(\tau \geqslant k), \tag{6.18}$$

where C''' absorbs the constant c along with various constants that arise in the calculation. The claim then follows from $P(\tau \geqslant k) = 1 - \prod_{j=k}^{n-1} \mathrm{e}^{-R_k} \leqslant \sum_{j=k}^{n-1} R_j \leqslant \sum_{j=k}^{\infty} R_j$. □

We are now ready to give:

Proof of Proposition 6.6. Given integers $\ell, \gamma \geqslant 1$ to be chosen later, we write

$$\begin{aligned} \big\{\varphi^{\text{DG}}_z \geqslant m_n + u\big\} \subseteq \big\{\varphi^{\text{GFF}}_z \geqslant m_n + u - \ell\big\} \\ \cup \big\{\varphi^{\text{DG}}_z \geqslant m_n + u,\, \ell < \varphi^{\text{DG}}_z - \varphi^{\text{GFF}}_z \leqslant C\gamma \log n\big\} \\ \cup \big\{\varphi^{\text{DG}}_z \geqslant m_n + u,\, \varphi^{\text{DG}}_z - \varphi^{\text{GFF}}_z > C\gamma \log n\big\}. \end{aligned} \tag{6.19}$$

For first event on the right, Lemma 6.4 gives

$$\begin{aligned} P\Big(\max_{x\in\Lambda_n} \varphi^{\text{GFF}}_x \leqslant m_n + t, \varphi^{\text{GFF}}_z \geqslant m_n + u - \ell\Big) \\ \leqslant C' b^{-n} \big(1 + (t \vee (t - u + \ell))^2 \beta\big) \mathrm{e}^{-\sqrt{\beta} c' (u-\ell)} \\ \leqslant C' \ell^2 \mathrm{e}^{\sqrt{\beta} c' \ell} b^{-n} \big(1 + (t \vee (t - u + 1))^2 \beta\big) \mathrm{e}^{-\sqrt{\beta}\, c' u}. \end{aligned} \tag{6.20}$$

For the third event, we can drop the restriction on the maximum and bound the result via Lemma 6.7 to get

$$P\big(\varphi^{\text{DG}}_z \geqslant m_n + u, \varphi^{\text{DG}}_z - \varphi^{\text{GFF}}_z > C\gamma \log n\big) \leqslant C''' n b^{-n} \mathrm{e}^{-\sqrt{\beta} c' u} \sum_{j=\lfloor \gamma \log n \rfloor}^{\infty} R_j. \tag{6.21}$$

Since $\sum_{j=k}^{\infty} R_j$ is exponentially decaying in k, we can choose γ so large that

$$\sum_{j=\lfloor \gamma \log n \rfloor}^{\infty} R_j \leqslant \frac{1}{n^2} \tag{6.22}$$

once n is sufficiently large. Under such circumstances, this term contributes at most a constant times $n^{-1} b^{-n} \mathrm{e}^{-\sqrt{\beta} c' u}$.

It thus remains to estimate the contribution of the second event on the right of (6.19) to the probability in question. Here we first partition the event as

$$\begin{aligned}
&P\Bigl(\max_{x\in\Lambda_n}\varphi_x^{\rm GFF}\leqslant m_n+t,\varphi_z^{\rm DG}\geqslant m_n+u,\ell<\varphi_z^{\rm DG}-\varphi_z^{\rm GFF}\leqslant C\gamma\log n\Bigr)\\
&\qquad\leqslant\sum_{j=0}^{\lfloor C\gamma\log n\rfloor}P\Bigl(\max_{x\in\Lambda_n}\varphi_x^{\rm GFF}\leqslant m_n+t,\varphi_z^{\rm DG}\geqslant m_n+u,\varphi_z^{\rm DG}-\varphi_z^{\rm GFF}-\ell-j\in(0,1]\Bigr)
\end{aligned}\tag{6.23}$$

Invoking the Bernoulli's in the coupling, the probability on the right is then bounded as

$$\begin{aligned}
&P\Bigl(\max_{x\in\Lambda_n}\varphi_x^{\rm GFF}\leqslant m_n+t,\varphi_z^{\rm GFF}\geqslant m_n+u-(\ell+j+1),\varphi_z^{\rm DG}-\varphi_z^{\rm GFF}>\ell+j\Bigr)\\
&\qquad\leqslant P\Bigl(\max_{x\in\Lambda_n}\varphi_x^{\rm GFF}\leqslant m_n+t,\varphi_z^{\rm GFF}\geqslant m_n+u-(\ell+j+1),\sum_{k=0}^{n-1}C\cdot 1_{B_k(z)=0}>\ell+j\Bigr)
\end{aligned}\tag{6.24}$$

where we set $B_0(z):=0$ and invoked the bound (3.51).

The independence stated in Theorem 3.6(2) then reduces the last probability to the product

$$P\Bigl(\max_{x\in\Lambda_n}\varphi_x^{\rm GFF}\leqslant m_n+t,\varphi_z^{\rm GFF}\geqslant m_n+u-(\ell+j+1)\Bigr)P\Bigl(\sum_{k=0}^{n-1}1_{B_k(z)=0}>\frac{\ell+j}{C}\Bigr).\tag{6.25}$$

For ℓ large, (6.7) gives

$$P\Bigl(\sum_{k=0}^{n-1}1_{B_k(z)=0}>\frac{\ell+j}{C}\Bigr)\leqslant\exp\Bigl(-\frac12C^{-1}(\ell+j)\log(C^{-1}(\ell+j))\Bigr).\tag{6.26}$$

Lemma 6.4 in turn bounds the probability on the left of (6.25) by

$$C'b^{-n}\bigl(1+(t\vee(t-u+\ell+j+1))^2\beta\bigr)e^{\sqrt{\beta}\,c'(\ell+j+1)}e^{-\sqrt{\beta}\,c'u},\tag{6.27}$$

provided that n is large enough so that $m_n+u-(l+C\gamma\log n+1)$ is a large positive number, which guarantees that Lemma 6.4 is applicable to the probability on the left of (6.25) for all j between 0 and $\lfloor C\gamma\log n\rfloor$.

Combining these observations (and extending, as a bound, the sum over j to infinity), the probability on the left of (6.23) is thus bounded by $C'b^{-n}e^{-\sqrt{\beta}\,c'u}$ times

$$\begin{aligned}
&\sum_{j=0}^{\infty}(1+(t\vee(t-u+\ell+j+1))^2\beta)\,e^{\sqrt{\beta}\,c'(\ell+j+1)-\frac12C^{-1}(\ell+j)\log(C^{-1}(\ell+j))}\\
&\qquad\leqslant(1+(t\vee(t-u+1))^2\beta)\sum_{j=\ell+1}^{\infty}j^2e^{\sqrt{\beta}c'j-\frac12C^{-1}(j-1)\log(C^{-1}(j-1))}.
\end{aligned}\tag{6.28}$$

The sum on the right converges and can thus be absorbed into the definition of C''. Together with (6.20–6.22), this implies the assertion. □

6.3 Proofs of tightness.

With the technical inputs settled, we are ready to prove the main results of this section. We start with the tightness of the centered DG-maximum.

Proof of Theorem 6.1. We first address the tightness of the upper tail. (This is the difficult part but we have already done most of the work.) Indeed, for any $t > 0$ we have

$$\begin{aligned} &P\Big(\max_{x\in\Lambda_n}\varphi_x^{\mathrm{DG}} \geqslant m_n + u\Big) \\ &\quad\leqslant P\Big(\max_{x\in\Lambda_n}\varphi_x^{\mathrm{GFF}} \leqslant m_n + t, \max_{x\in\Lambda_n}\varphi_x^{\mathrm{DG}} \geqslant m_n + u\Big) + P\Big(\max_{x\in\Lambda_n}\varphi_x^{\mathrm{GFF}} > m_n + t\Big). \end{aligned} \tag{6.29}$$

By the union bound and Lemma 6.3,

$$\begin{aligned} &P\Big(\max_{x\in\Lambda_n}\varphi_x^{\mathrm{DG}} \geqslant m_n + u\Big) \\ &\qquad\leqslant \sum_{z\in\Lambda_n} P\Big(\max_{x\in\Lambda_n}\varphi_x^{\mathrm{GFF}} \leqslant m_n + t, \varphi_z^{\mathrm{DG}} \geqslant m_n + u\Big) + \frac{1}{a}\mathrm{e}^{-a\sqrt{\beta}\,t}. \end{aligned} \tag{6.30}$$

Proposition 6.6 then shows that, for sufficiently large n,

$$P\left(\max_{x\in\Lambda_n}\varphi_x^{\mathrm{DG}} \geqslant m_n + u\right) \leqslant C''\big(1 + (t \vee (t - u + 1))^2\beta\big)\mathrm{e}^{-\sqrt{\beta}c'u} + \frac{1}{a}\mathrm{e}^{-a\sqrt{\beta}\,t}. \tag{6.31}$$

Setting $t := u/2$ then shows that the supremum over $n \geqslant 1$ of the left-hand side tends to zero as $u \to \infty$.

It remains to prove the tightness of the lower tail. Let $k \in \{1, \dots, n-1\}$ and recall that $\zeta_{\geqslant k}^{\mathrm{DG}}(x)$ and $\zeta_{\geqslant k}^{\mathrm{GFF}}(x)$ denote the sums of $\zeta_i^{\mathrm{DG}}(x)$ and $\zeta_i^{\mathrm{GFF}}(x)$ for $i = k, \dots, n-1$. (These are still parametrized by $x \in \Lambda_n$.) Then for any $x \in \Lambda_n$,

$$P\Big(\zeta_{\geqslant k}^{\mathrm{DG}}(x) = \zeta_{\geqslant k}^{\mathrm{GFF}}(x)\Big) \geqslant P\Big(B_j(x) = 1, j = k, \dots, n-1\Big) \geqslant \mathrm{e}^{-\sum_{j\geqslant k} R_j}. \tag{6.32}$$

Since $\zeta_{\geqslant k}^{\mathrm{GFF}}$ has the law of GFF on Λ_{n-k}, Lemma 6.3 shows that, for any $t > 0$,

$$P\Big(\max_{x\in\Lambda_n}\zeta_{\geqslant k}^{\mathrm{GFF}}(x) \geqslant m_{n-k} - t\Big) \geqslant 1 - \frac{1}{a}\mathrm{e}^{-a\sqrt{\beta}\,t}. \tag{6.33}$$

Writing X_k for the maximizer of $x \mapsto \zeta_{\geqslant k}^{\mathrm{GFF}}(x)$ that is minimal in a natural ordering of Λ_n, we then have

$$\begin{aligned} &P\Big(\max_{x\in\Lambda_n}\zeta_{\geqslant k}^{\mathrm{DG}}(x) < m_{n-k} - t\Big) \\ &\qquad\leqslant P\Big(\max_{x\in\Lambda_n}\zeta_{\geqslant k}^{\mathrm{GFF}}(x) < m_{n-k} - t\Big) + P\Big(\zeta_{\geqslant k}^{\mathrm{DG}}(X_k) \neq \zeta_{\geqslant k}^{\mathrm{GFF}}(X_k)\Big). \end{aligned} \tag{6.34}$$

Using (6.32) along with the fact that, since X_k is determined by the GFF increments, it is independent of the Bernoulli's and so

$$P\Big(\zeta_{\geqslant k}^{\mathrm{DG}}(X_k) \neq \zeta_{\geqslant k}^{\mathrm{GFF}}(X_k)\Big) \leqslant 1 - \mathrm{e}^{-\sum_{j=k}^{\infty} R_j} \leqslant \sum_{j=k}^{\infty} R_j. \tag{6.35}$$

Then

$$P\Big(\max_{x\in\Lambda_n}\zeta^{\mathrm{DG}}_{\geqslant k}(x) < m_{n-k}-t\Big) \leqslant \frac{1}{a}\mathrm{e}^{-a\sqrt{\beta}\,t}+\sum_{j=k}^{\infty}R_j \tag{6.36}$$

follows with the help of (6.33).

Now for any $x\in\Lambda_n$ and any $\lambda>0$, the standard estimate for the Gaussian distribution gives

$$P\bigg(\sum_{j=0}^{k-1}\zeta^{\mathrm{GFF}}_j(x) < -\lambda\bigg) \leqslant \mathrm{e}^{-\frac{\beta\lambda^2}{2k}}. \tag{6.37}$$

Let Y_k be the maximizer of $x\mapsto\zeta^{\mathrm{DG}}_{\geqslant k}(x)$ that is again minimal in the natural ordering of Λ_n. Since Y_k is measurable with respect to the increments "above" level $k-1$, the bound (6.37) applies even with x replaced by Y_k. Combining (6.36) and (6.37) using the union bound, along with $\{\sum_{j=0}^{k-1}\zeta^{\mathrm{DG}}_j(x) < -\lambda - Ck\}\subseteq\{\sum_{j=0}^{k-1}\zeta^{\mathrm{GFF}}_j(x) < -\lambda\}$ almost surely, we see

$$P\left(\max_{x\in\Lambda_n}\varphi^{\mathrm{DG}}_x < m_{n-k}-t-\lambda-Ck\right) \leqslant \frac{1}{a}\mathrm{e}^{-a\sqrt{\beta}\,t}+\sum_{j=k}^{\infty}R_j+\mathrm{e}^{-\frac{\beta\lambda^2}{2k}}. \tag{6.38}$$

Note that, for $n>k\geqslant 1$,

$$m_{n-k}=m_n-c_1\,k+c_2\log\left(\frac{n}{n-k}\right)\geqslant m_n-c_1\,k \tag{6.39}$$

for some constants $c_1,c_2>0$ that can be gleaned from (2.5). Setting $\lambda:=c_1k$ and $t:=c_1k$ above, we obtain

$$P\left(\max_{x\in\Lambda_n}\varphi^{\mathrm{DG}}_x < m_n-(3c_1+C)k\right) \leqslant \frac{1}{a}\mathrm{e}^{-a\sqrt{\beta}\,c_1k}+\sum_{j=k}^{\infty}R_j+\mathrm{e}^{-\frac{\beta c_1^2}{2}k}. \tag{6.40}$$

By Theorem 3.4, $R_k\to 0$ exponentially fast when $\beta\in(0,\beta_c)$. Taking $k\to\infty$ (after $n\to\infty$) thus yields the lower-tail tightness and thus proves the desired claim. $\square$

Next we will prove tightness of the size of level sets:

Proof of Theorem 6.2. With the help of Proposition 6.6 we get for n large enough,

$$\begin{aligned} E\big[G_n(\lambda)1_{\{\max_{x\in\Lambda_n}\varphi^{\mathrm{GFF}}_x\leqslant m_n+t\}}\big] &= E\left[\sum_{z\in\Lambda_n}1_{\{\varphi^{\mathrm{DG}}_z\geqslant m_n-\lambda\}}1_{\{\max_{x\in\Lambda_n}\varphi^{\mathrm{GFF}}_x\leqslant m_n+t\}}\right] \\ &= \sum_{z\in\Lambda_n}P\left(\varphi^{\mathrm{DG}}_z\geqslant m_n-\lambda,\ \max_{x\in\Lambda_n}\varphi^{\mathrm{GFF}}_x\leqslant m_n+t\right) \\ &\leqslant C''\big(1+(t\vee(t+\lambda+1))^2\beta\big)\mathrm{e}^{\sqrt{\beta}\,c'\lambda}. \end{aligned} \tag{6.41}$$

Then by Markov inequality and Lemma 6.3 for n sufficiently large we get

$$\begin{aligned} P\big(G_n(\lambda) \geqslant u\big) &\leqslant P\Big(G_n(\lambda) \geqslant u, \max_{x\in\Lambda_n} \varphi_x^{\rm GFF} \leqslant m_n + t\Big) + P\Big(\max_{x\in\Lambda_n} \varphi_x^{\rm GFF} > m_n + t\Big) \\ &\leqslant \frac{1}{u} E\big[G_n(\lambda) 1_{\{\max_{x\in\Lambda_n} \varphi_x^{\rm GFF} \leqslant m_n+t\}}\big] + \frac{1}{a} \mathrm{e}^{-a\sqrt{\beta}\,t} \\ &\leqslant \frac{1}{u} C''\big(1 + (t \vee (t+\lambda+1))^2\beta\big) \mathrm{e}^{\sqrt{\beta}\,c'\lambda} + \frac{1}{a}\mathrm{e}^{-a\sqrt{\beta}\,t}. \end{aligned} \tag{6.42}$$

Taking $u \to \infty$ followed by $t \to \infty$ now yields the desired conclusion. □

7. Limit law

Here we will finally prove our main results concerning the extremal behavior of the hierarchical DG-model. As noted earlier, the proofs will be deduced from the corresponding statements about the extremal behavior of the GFF, a.k.a. Branching Random Walk.

7.1 GFF extremal process.

Recall our notation $P'_{n,\beta}$ for the law of the GFF on Λ_n at inverse temperature β and depth n. We will write $E'_{n,\beta}$ for the associated expectation. Recall also the notation m_n for the centering sequence from (2.5) which, we note, depends on β, and the notation $[x]_n$ from (2.8) for the image of x under the natural embedding of Λ_n into the unit interval $[0, 1]$. We then have:

Theorem 7.1 *There exists an a.s.-finite random Borel measure Z on $[0, 1]$ with $Z(A) > 0$ a.s. for all non-empty (relatively) open $A \subseteq [0, 1]$ and a probability measure ν on $\mathcal{M}_{\mathbb{N}}(\mathbb{R})$ such that for all $\beta > 0$ and all continuous $f \colon [0, 1] \times \mathbb{R} \to [0, \infty)$ with compact support,*

$$\begin{aligned} &E'_{n,\beta}\bigg(\exp\Big\{ -\sum_{x\in\Lambda_n} f\big([x]_n, \varphi_x^{\rm GFF} - m_n\big)\Big\}\bigg) \\ &\quad \underset{n\to\infty}{\longrightarrow} E\bigg(\exp\Big\{-\int Z(\mathrm{d}x)\otimes \mathrm{e}^{-\alpha h}\mathrm{d}h \otimes \nu(\mathrm{d}\chi)\big[1-\mathrm{e}^{-\int f(x,\beta^{-1/2}(h+\cdot))\mathrm{d}\chi}\big]\Big\}\bigg), \end{aligned} \tag{7.1}$$

where $\alpha := \sqrt{2\log b}$ and where the expectation on the right is over the law of Z. Almost-every sample χ from ν obeys $\operatorname{supp}(\chi) \subseteq (-\infty, 0]$ with $\sup\operatorname{supp}(\chi) = 0$.

For functions not depending on the "spatial" coordinate, this is a special case of Madaule [37, Theorem 1.1]. The inclusion of the "spatial" positions is fairly straightforward but doing that here would detract from the main line of the argument. We give a detailed proof at the very end of this section.

Noting that that each sample of the process χ can be written as $\sum_{i\geqslant 1} \delta_{t_i}$ for a real-valued sequence $\{t_i\}_{i\geqslant 1}$, the integral inside the exponential boils down to

$$\int f\big(x, \beta^{-1/2}(h+\cdot)\big)\mathrm{d}\chi = \sum_{i\geqslant 1} f\big(x, \beta^{-1/2}(h + t_i)\big). \tag{7.2}$$

The fact that ν is a law on the space of Radon measures entails that the sum is finite for ν-a.e. sample χ. The conditions on the support of a.e.-sample from ν in turn ensure that the representation of the limit process is unique.

In the process notation, Theorem 7.1 can be restated as

$$\sum_{x\in\Lambda_n}\delta_{[x]_n}\otimes\delta_{\varphi_x^{\text{GFF}}-m_n}\ \underset{n\to\infty}{\overset{\text{law}}{\longrightarrow}}\ \sum_{i,j\geqslant 1}\delta_{x_i}\otimes\delta_{\beta^{-1/2}(h_i+t_j^{(i)})},\tag{7.3}$$

where $\{(x_i,h_i)\}_{i\geqslant1}$ enumerates the points in a sample from the Poisson point process

$$\text{PPP}\big(Z(\mathrm{d}x)\otimes \mathrm{e}^{-\alpha h}\mathrm{d}h\big)\tag{7.4}$$

and $\{t_j^{(i)}\}_{j\geqslant1}$ enumerates the points in the i-th member of the sequence $\{t^{(i)}\}_{i\geqslant1}$ of i.i.d. samples from ν that are independent of $\{(x_i,h_i)\}_{i\geqslant1}$. Observe that neither Z nor ν depend on β which reflects on the GFF at inverse temperature β being just a $\beta^{-1/2}$-multiple of the GFF at unit inverse temperature.

The weak convergence of probability laws (7.3) is relative to the vague topology on Radon measures on $[0,1]\times\mathbb{R}$ which is why it suffices to state (7.1) for continuous functions with compact support. Unfortunately, this restriction is too strict for our purposes. We therefore state and prove:

Lemma 7.2 *Let $f\colon[0,1]\times\mathbb{R}\to[0,\infty)$ be continuous and such that $f(x,h)\leqslant A\mathrm{e}^{-ah^2}$, for some $A,a>0$ and all $x\in[0,1]$ and $h\in\mathbb{R}$. Then* (7.1) *holds.*

Proof. Let $\rho\colon[0,\infty)\to[0,1]$ be non-increasing and continuous with $\rho=1$ on $[0,1]$ and $\rho=0$ outside $[0,2]$. Let $f_r(x,h):=f(x,h)\rho(|h|/r)$ and observe that $f_r\leqslant f$ with f_r increasing monotonically to f as r increases to infinity.

Next let us write

$$G_n'(\lambda):=\big|\{x\in\Lambda_n\colon\varphi_x^{\text{GFF}}\geqslant m_n-\lambda\}\big|\tag{7.5}$$

for the size of an extremal GFF level set and note that the same argument as in the proof of Theorem 6.2 using Lemma 6.4 gives

$$E_{n,\beta}'\big[G_n'(\lambda)1_{\{\max_{x\in\Lambda_n}\varphi_x^{\text{GFF}}\leqslant m_n+t\}}\big]\leqslant C'(1+(t\vee(t+\lambda))^2\beta)\,\mathrm{e}^{\sqrt{\beta}\,c'\lambda}\tag{7.6}$$

for large enough n and $\lambda>0$. Noting $|f(x,h)-f_r(x,h)|\leqslant A\mathrm{e}^{-ah^2}1_{\{|h|>r\}}$, for all $\epsilon>0$ and $r>t>0$ a simple use of Markov inequality yields

$$\begin{aligned}
P_{n,\beta}'\bigg(\Big|\sum_{x\in\Lambda_n}f\big([x]_n,\varphi_x^{\text{GFF}}-m_n\big)-\sum_{x\in\Lambda_n}f_r\big([x]_n,\varphi_x^{\text{GFF}}-m_n\big)\Big|>\epsilon\bigg)&\\
\leqslant P_{n,\beta}'\Big(\max_{x\in\Lambda_n}\varphi_x^{\text{GFF}}>m_n+t\Big)&\\
+A\epsilon^{-1}E_{n,\beta}'\bigg(\sum_{x\in\Lambda_n}\mathrm{e}^{-a(\varphi_x^{\text{GFF}}-m_n)^2}1_{\{\varphi_x^{\text{GFF}}<m_n-r\}}1_{\{\max_{x\in\Lambda_n}\varphi_x^{\text{GFF}}\leqslant m_n+t\}}\bigg)&,
\end{aligned}\tag{7.7}$$

where the values with $\varphi_x^{\text{GFF}}>m_n+r$ do not enter due to our restriction on the maximum and the assumption that $r>t$. Assuming r to be a natural, the last expectation is bounded using $1_{\{\varphi_x^{\rm GFF}<m_n-r\}}=\sum_{\ell=r+1}^\infty 1_{\{-\ell\leqslant\varphi_x^{\rm GFF}-m_n<-\ell+1\}}$ by

$$\begin{aligned}\sum_{\ell=r+1}^\infty \mathrm{e}^{-a\ell^2}\,E'_{n,\beta}\bigl[G'_n(\ell)1_{\{\max_{x\in\Lambda_n}\varphi_x^{\rm GFF}\leqslant m_n+t\}}\bigr]\\ \leqslant\sum_{\ell=r+1}^\infty C'(1+(t\vee(t+\ell))^2\beta)\,\mathrm{e}^{\sqrt{\beta}c'\ell-a\ell^2},\end{aligned}\tag{7.8}$$

where (7.6) was used to get the inequality.

The sum on the right of (7.8) vanishes as $r\to\infty$ and, taking $t\to\infty$ afterwards with the help of Lemma 6.3, so does the probability on the left of (7.7). Hence we get

$$\lim_{r\to\infty}\limsup_{n\to\infty}\left|E'_{n,\beta}\bigl(\mathrm{e}^{-\sum_{x\in\Lambda_n}f([x]_n,\varphi_x^{\rm GFF}-m_n)}\bigr)-E'_{n,\beta}\bigl(\mathrm{e}^{-\sum_{x\in\Lambda_n}f_r([x]_n,\varphi_x^{\rm GFF}-m_n)}\bigr)\right|=0.\tag{7.9}$$

Since f_r does have compact support, (7.1) applies to it. It then suffices to show that the right-hand side of (7.1) for f_r tends to that for f as $r\to\infty$. Thanks to the monotonicity of $r\mapsto f_r$, this follows by the Monotone Convergence Theorem for the inner integral and then the Bounded Convergence Theorem for the expectation. □

7.2 Conversion to GFF extremal process.

Next we will develop an argument that reduces observables involving the extremal process of the DG-model to those of a GFF. This is what we will need to deduce Theorem 2.3 directly from Theorem 7.1.

Pick $\beta\in(0,\beta_{\rm c})$ and a continuous function $f\colon[0,1]\times\mathbb{R}\to[0,\infty)$ with compact support; we will keep these fixed throughout the rest of this subsection. We are interested in the $n\to\infty$ weak limit of the random variable

$$Y_n:=\sum_{x\in\Lambda_n}f\bigl([x]_n,\varphi_x^{\rm DG}-\lfloor m_n\rfloor\bigr).\tag{7.10}$$

We will work in the coupling from Theorem 3.6. Pick k with $1\leqslant k<n-1$ and recall the notation $\zeta_{\geqslant k}^{\rm GFF}(x)$ for the sum $\sum_{j=k}^{n-1}\zeta_j^{\rm GFF}(x)$ which we for later convenience simplify as follows: Given $z\in\Lambda_{n-k}$, pick any $x\in\Lambda_n$ with $m^k(x)=z$ and abbreviate

$$\varphi_z^{\rm GFF}:=\zeta_{\geqslant k}^{\rm GFF}(x).\tag{7.11}$$

Similarly, using $\zeta_{\geqslant k}^{\rm DG}(x):=\sum_{j=k}^{n-1}\zeta_j^{\rm DG}(x)$, let

$$\varphi_z^{\rm DG}:=\zeta_{\geqslant k}^{\rm DG}(x).\tag{7.12}$$

The definition of the coupling ensures that these objects do not depend on the choice of x and so we may think of them as indexed by Λ_{n-k}. Note that $\{\varphi_z^{\rm DG}\colon z\in\Lambda_{n-k}\}$ is continuously distributed and $\{\varphi_z^{\rm GFF}\colon z\in\Lambda_{n-k}\}$ is a GFF on Λ_{n-k}.

Using these objects we now define $g_k\colon[0,1]\times\mathbb{R}\to[0,\infty)$ via

$$\mathrm{e}^{-g_k(v,h)}:=E\Bigl(\mathrm{e}^{-\sum_{x\in\Lambda_k(z)}f(v,\varphi_x^{\rm DG}-\varphi_z^{\rm DG}+h)}\,\Big|\,\varphi_z^{\rm GFF}=\varphi_z^{\rm DG}=h\Bigr),\tag{7.13}$$

where $\Lambda_k(z):=\{x\in\Lambda_n\colon m^k(x)=z\}$ and where the conditioning on the (apparently singular) event $\varphi_z^{\rm GFF}=\varphi_z^{\rm DG}=h$ is to be interpreted as setting an initial value in the tree-indexed Markov chain defining the coupling. The desired conversion will be facilitated by:

Proposition 7.3 *We have*

$$\lim_{k\to\infty}\limsup_{n\to\infty}\left|E_{n,\beta}(\mathrm{e}^{-Y_n})-E'_{n-k,\beta}\biggl(\exp\Bigl\{-\sum_{z\in\Lambda_{n-k}}g_k\bigl([z]_{n-k},\varphi_z^{\rm GFF}-\lfloor m_n\rfloor\bigr)\Bigr\}\biggr)\right|=0. \tag{7.14}$$

The proof of Proposition 7.3 will require a few lemmas. Let k be a natural with $1\leqslant k<n-1$ and let $\tau(x)$ be the random "time" defined in (6.13). We then modify Y_n into

$$Y'_{n,k}:=\sum_{x\in\Lambda_n}f\bigl([m^k(x)]_{n-k},\varphi_x^{\rm DG}-\lfloor m_n\rfloor\bigr)1_{\{\tau(x)<k\}}. \tag{7.15}$$

As we show next, these modifications are not very significant provided n and k are taken sufficiently large:

Lemma 7.4 *For each $\epsilon>0$,*

$$\lim_{k\to\infty}\limsup_{n\to\infty}P\bigl(|Y_n-Y'_{n,k}|>\epsilon\bigr)=0, \tag{7.16}$$

where P denote the coupling law from Theorem 2.6.

Proof. Let $Y''_{n,k}$ be given by the same formula as $Y'_{n,k}$ except with the first argument of f replaced by $[x]_n$. Let $\lambda>0$ be such that f is supported in $[0,1]\times[-\lambda,\lambda]$ and let $\omega(r):=\sup\{|f(x,h)-f(y,h)|\colon|x-y|<r,\,h\in[-\lambda,\lambda]\}$ be the modulus of continuity of f in the first variable. In light of $|[m^k(x)]_{n-k}-[x]_n|<b^{-n+k}$ we then have

$$|Y'_{n,k}-Y''_{n,k}|\leqslant\omega(b^{-n+k})G_n(\lambda) \tag{7.17}$$

which tends to zero in probability as $n\to\infty$ thanks to Theorem 6.2.

It thus suffices to prove the claim with $Y'_{n,k}$ replaced by $Y''_{n,k}$. Using the coupling measure P and writing A for a constant that bounds $|f|$ uniformly, here we get

$$P\bigl(|Y_n-Y''_{n,k}|>\epsilon\bigr)\leqslant P\biggl(\sum_{x\in\Lambda_n}1_{\{\varphi_x^{\rm DG}\geqslant\lfloor m_n\rfloor-\lambda\}}1_{\{\tau(x)\geqslant k\}}>\epsilon/A\biggr). \tag{7.18}$$

The probability on the right is bounded via Markov inequality as

$$\begin{aligned}P\Bigl(&\max_{x\in\Lambda_n}\varphi_x^{\rm GFF}>m_n+t\Bigr)\\&+A\epsilon^{-1}\sum_{x\in\Lambda_n}P\Bigl(\varphi_x^{\rm DG}\geqslant\lfloor m_n\rfloor-\lambda,\,\tau(x)\geqslant k,\,\max_{x\in\Lambda_n}\varphi_x^{\rm GFF}\leqslant m_n+t\Bigr).\end{aligned} \tag{7.19}$$

Similarly as in (6.19), the probability under the sum can be bounded by splitting the event $\{\varphi_x^{\rm DG}\geqslant\lfloor m_n\rfloor-\lambda\}$ into three events. For each of these three events, we need to show the same type of bound as in right side of (6.11) with additional exponential decay in k induced by $\{\tau(x)\geqslant k\}$.

The first event only involves GFF so the exponential decay in k from $\{\tau(x)\geqslant k\}$ can be extracted immediately from the independence of the GFF and the Bernoulli's ensured by Theorem 3.6(2). For the third event, the probability vanishes already in the limit as $n\to\infty$ due to the n^{-1} decay facilitated by the choice of γ in (6.22) so exponential decay in k is not required. For the second event, we follow the estimates up to (6.25) where instead of the second probability on the right we get

$$P\Big(\sum_{k=0}^{n-1} 1_{B_k(z)=0} > \frac{\ell+j}{C},\, \tau(x) \geqslant k\Big). \tag{7.20}$$

To bound this we write $\eta := \frac{\ell+j}{C}$ and note that, for η large enough, Lemma 6.5 gives

$$\begin{aligned} P\Big(&\sum_{j=0}^{n-1} 1_{B_j(x)=0} > \eta,\, \tau(x) \geqslant k\Big) \\ &\leqslant P\Big(\sum_{j=0}^{k-1} 1_{B_j(x)=0} > \frac{\eta}{2}\Big)\, P\big(\tau(x) \geqslant k\big) + P\Big(\sum_{j=k}^{n-1} 1_{B_j(x)=0} > \frac{\eta}{2}\Big) \\ &\leqslant \exp\Big\{-\frac{1}{2}\frac{\eta}{2}\log\Big(\frac{\eta}{2}\Big)\Big\}\Big(1-\prod_{j\geqslant k} \mathrm{e}^{-R_j}\Big) \\ &\qquad + \exp\Big\{\frac{\eta}{2}+\frac{\eta}{2}\log\Big(\sum_{j=k}^{n-1} 1-\mathrm{e}^{-R_j}\Big) - \frac{\eta}{2}\log\Big(\frac{\eta}{2}\Big)\Big\}, \end{aligned} \tag{7.21}$$

where the case $\{\tau(x) = n\}$ reduces to $P(\eta < 1)$, which vanishes for large η, and so is excluded. Both terms above exhibit simultaneously super-exponential decay in η as well as exponential decay in k and so the probability on the right of (7.19) vanishes in the limit as $n \to \infty$ and then as $k \to \infty$. Taking $t \to \infty$, we get Lemma 7.4. □

We also need the following "reverse-Jensen" type of inequality:

Lemma 7.5 *Given a sequence $\{p_k\}_{k=1}^n$ of positive probabilities, let $\{Y_x\colon x \in \bigcup_{k=1}^n \Lambda_k\}$ be independent zero-one valued random variables such that $\mathbb{P}(Y_x = 1) = p_k$ for all $k = 1, \dots, n$ and all $x \in \Lambda_k$. Writing $x \in \Lambda_n$ as $x = (x_1, \dots, x_n)$, set $Z_x := \prod_{i=1}^n Y_{x_i}$. Then for all $h\colon \Lambda_n \to [0,\infty)$ and all $\lambda \in (0,1)$,*

$$\mathbb{E}\bigg(\exp\Big\{-\sum_{x\in\Lambda_n} h(x)Z_x\Big\}\bigg) \leqslant \frac{1}{\lambda^2}\frac{1-q}{q} + \exp\Big\{-(1-\lambda)q\sum_{x\in\Lambda_n} h(x)\Big\}, \tag{7.22}$$

where $q := \prod_{k=1}^n p_k$.

Proof. Assume h is not identically zero for otherwise there is nothing to prove. We will use that the Chebyshev inequality for non-negative random variable X gives

$$\mathbb{P}\big(X \leqslant (1-\lambda)\mathbb{E}X\big) \leqslant \mathbb{P}\big(|X-\mathbb{E}X| \geqslant \lambda\mathbb{E}X\big) \leqslant \frac{1}{\lambda^2}\frac{\mathrm{Var}(X)}{[\mathbb{E}X]^2} \tag{7.23}$$

whenever $\mathbb{E}X > 0$. For this we compute the first moment to be

$$\mathbb{E}\Big(\sum_{x\in\Lambda_n} h(x)Z_x\Big) = q\sum_{x\in\Lambda_n} h(x) \tag{7.24}$$

As to the variance, for $x,y\in\Lambda_n$ with the nearest common ancestor in $\Lambda_{n-\ell}$ we use

$$\text{Cov}(Z_xZ_y)=\Bigl(\prod_{k=1}^{\ell-1}p_k\Bigr)^2\prod_{i=\ell}^{n}p_i-\Bigl(\prod_{k=1}^{n}p_k\Bigr)^2=q^2\biggl(\Bigl(\prod_{k=\ell}^{n}p_k\Bigr)^{-1}-1\biggr)\leqslant q^2[q^{-1}-1] \tag{7.25}$$

to get

$$\text{Var}\biggl(\sum_{x\in\Lambda_n}h(x)Z_x\biggr)=\sum_{x,y\in\Lambda_n}h(x)h(y)\text{Cov}(Z_xZ_y)\leqslant q^2\,\frac{1-q}{q}\Bigl(\sum_{x\in\Lambda_n}h(x)\Bigr)^2. \tag{7.26}$$

The above Chebyshev bound then gives

$$\mathbb{P}\biggl(\sum_{x\in\Lambda_n}h(x)Z_x\leqslant(1-\lambda)q\sum_{x\in\Lambda_n}h(x)\biggr)\leqslant\frac{1}{\lambda^2}\frac{1-q}{q} \tag{7.27}$$

when now readily implies the claim. □

We are now fully equipped to give:

Proof of Proposition 7.3. We will work in the coupling measure P throughout. With the help of Lemma 7.4, it suffices to prove (7.14) where $E_{n,\beta}(\text{e}^{-Y_n})$ is replaced by $E(\text{e}^{-Y'_{n,k}})$. Denote

$$\mathcal{F}_k:=\sigma\Bigl(\zeta_j^{\text{DG}}(x),\zeta_j^{\text{GFF}}(x),B_j(x)\colon j\geqslant k,\,x\in\Lambda_n\Bigr). \tag{7.28}$$

The Markovian definition of the coupling along with the rewrite

$$\varphi_x^{\text{DG}}-\lfloor m_n\rfloor=\varphi_x^{\text{DG}}-\varphi_z^{\text{GFF}}+\varphi_z^{\text{GFF}}-\lfloor m_n\rfloor \tag{7.29}$$

then casts $E(\text{e}^{-Y'_{n,k}}|\mathcal{F}_k)$ as

$$\prod_{z\in\Lambda_{n-k}}E\biggl(\exp\Bigl\{-\sum_{x\in\Lambda_k(z)}1_{\{\tau(x)<k\}}f\bigl([z]_{n-k},\varphi_x^{\text{DG}}-\varphi_z^{\text{GFF}}+\varphi_z^{\text{GFF}}-\lfloor m_n\rfloor\bigr)\Bigr\}\biggm|\mathcal{F}_k\biggr). \tag{7.30}$$

The event $\{\tau(x)<k\}$ is $\mathcal{F}_k$ measurable and so it either occurs for all $x\in\Lambda_k(z)$ or none of them. In the former case we have $\varphi_z^{\text{DG}}=\varphi_z^{\text{GFF}}$. Writing $\tilde g_{n,k}$ for the function defined by

$$\text{e}^{-\tilde g_{n,k}(v,h)}:=E\Bigl(\text{e}^{-\sum_{x\in\Lambda_k(z)}f(v,\varphi_x^{\text{DG}}-\varphi_z^{\text{DG}}+h)}\Bigm|\varphi_z^{\text{GFF}}=\varphi_z^{\text{DG}}=h+\lfloor m_n\rfloor\Bigr), \tag{7.31}$$

the expectation in (7.30) can thus be contracted to

$$E(\text{e}^{-Y'_{n,k}}|\mathcal{F}_k)=\exp\Bigl\{-\sum_{z\in\Lambda_{n-k}}\tilde g_{n,k}\bigl([z]_{n-k},\varphi_z^{\text{GFF}}-\lfloor m_n\rfloor\bigr)1_{F(z)}\Bigr\}, \tag{7.32}$$

where $F(z):=\bigcap_{j=k}^{n-1}\{B_j(x)=1\}$ for some (and any) $x\in\Lambda_n$ such that $z=m^k(x)$.

Next observe a key fact that the law of $\varphi_x^{\text{DG}}-\varphi_z^{\text{DG}}$ "initiated" from $\varphi_x^{\text{DG}}=m+h$ is the same for all $m\in\mathbb{Z}$ due to the 1-periodicity of the effective potential $\{v_k\}_{k\geqslant0}$. This implies

$$\tilde g_{n,k}=g_k \tag{7.33}$$

and so

$$
\begin{aligned}
E(\mathrm{e}^{-Y'_{n,k}}) &= E\big(E(\mathrm{e}^{-Y'_{n,k}}|\mathcal{F}_k)\big) \\
&= E\bigg(\exp\Big\{-\sum_{z\in\Lambda_{n-k}} \tilde{g}_{n,k}\big([z]_{n-k},\varphi_z^{\mathrm{GFF}}-\lfloor m_n\rfloor\big)1_{F(z)}\Big\}\bigg) \\
&\geqslant E\bigg(\exp\Big\{-\sum_{z\in\Lambda_{n-k}} g_k\big([z]_{n-k},\varphi_z^{\mathrm{GFF}}-\lfloor m_n\rfloor\big)\Big\}\bigg).
\end{aligned}
\tag{7.34}
$$

where we used $g_k \geqslant 0$ to drop the indicators of $F(z)$.

For the complementary upper bound we note that the random variable $1_{F(z)}$ can be thought of as product of tree-indexed Bernoulli's on the path from z to the tree root. For each $\lambda \in (0,1)$, Lemma 7.5 applied to $E(\mathrm{e}^{-Y'_{n,k}}|\mathcal{F}_k)$ along with Jensen's inequality gives

$$
E(\mathrm{e}^{-Y'_{n,k}}) \leqslant \frac{1}{\lambda^2}\frac{1-q_k}{q_k} + \left[E\bigg(\exp\Big\{-\sum_{z\in\Lambda_{n-k}} g_k\big([z]_{n-k},\varphi_z^{\mathrm{GFF}}-\lfloor m_n\rfloor\big)\Big\}\bigg)\right]^{(1-\lambda)q_k}. \tag{7.35}
$$

Setting $\lambda := (1-q_k)^{1/3}$, the first term on the right tends to zero as $k \to \infty$ while the exponent $(1-\lambda)q_k$ tends to one. Taking limit as $n \to \infty$ and then $k \to \infty$ in (7.35), combined with the lower bound in (7.34), we get (7.14). □

7.3 Proofs of the main results.

We are now finally in a position to give formal proofs of our main results. We start with the convergence of the extremal process.

Proof of Theorem 2.3. Let $\lambda > 0$ be such that $\operatorname{supp}(f) \subseteq [0,1]\times[-\lambda,\lambda]$. The function g_k defined in (7.13) is bounded and, thanks to the continuity of f and v_k, continuous. While g_k does not have compact support, we claim that it has Gaussian tails in the second variable. To see this note that $f(v,\varphi_x^{\mathrm{DG}}-\varphi_z^{\mathrm{DG}}+h) > 0$ implies $\varphi_x^{\mathrm{DG}}-\varphi_z^{\mathrm{DG}}+h \in [-\lambda,\lambda]$. Then $\varphi_x^{\mathrm{GFF}}-\varphi_z^{\mathrm{GFF}}+h \in [-\lambda-Ck,\lambda+Ck]$ a.s. for C as in Theorem 3.6(4). Since $\varphi_x^{\mathrm{GFF}}-\varphi_z^{\mathrm{GFF}}+h$ is independent of φ_z^{GFF} with law $\mathcal{N}(h,k/\beta)$, for any $x \in \Lambda_k(z)$ the standard Gaussian estimate yields

$$
\begin{aligned}
&P\Big(f(v,\varphi_x^{\mathrm{DG}}-\varphi_z^{\mathrm{DG}}+h) > 0 \,\Big|\, \varphi_z^{\mathrm{GFF}} = \varphi_z^{\mathrm{DG}} = h\Big) \\
&\qquad\leqslant P\Big(\big|\mathcal{N}(0,k/\beta)\big| \geqslant |h|-\lambda-Ck\Big) \leqslant 2\mathrm{e}^{-\frac{\beta(|h|-\lambda-Ck)^2}{2k}}
\end{aligned}
\tag{7.36}
$$

whenever $|h| \geqslant \lambda + Ck$. Since $\Lambda_k(z)$ has cardinality b^k, the union bound then shows

$$
1-\mathrm{e}^{-g_k(v,h)} \leqslant 2b^k\mathrm{e}^{-\frac{\beta(|h|-\lambda-Ck)^2}{2k}}. \tag{7.37}
$$

It follows that $g_k(v,h)$ has Gaussian tails.

With g_k conforming to the condition of Lemma 7.2 (which feeds into Theorem 7.1), for any sequence $n_j \to \infty$ such that

$$
s := \lim_{j\to\infty}\big(m_{n_j}-\lfloor m_{n_j}\rfloor\big) \quad \text{exists,} \tag{7.38}
$$

we claim to get

$$
\begin{aligned}
&E'_{n_j-k,\beta}\biggl(\exp\Bigl\{-\sum_{z\in\Lambda_{n_j-k}} g_k\bigl([z]_{n_j-k},\varphi_z^{\text{GFF}}-\lfloor m_{n_j}\rfloor\bigr)\Bigr\}\biggr)\\
&\quad\underset{j\to\infty}{\longrightarrow}\;E\biggl(\exp\Bigl\{-\int Z(\mathrm{d}x)\otimes \mathrm{e}^{-\alpha h}\mathrm{d}h\otimes\nu(\mathrm{d}\chi)\bigl(1-\mathrm{e}^{-\int g_k(x,s+\beta^{-1/2}(h-\alpha k+\cdot))\,\mathrm{d}\chi}\bigr)\Bigr\}\biggr).
\end{aligned}
\tag{7.39}
$$

This would follow from Theorem 7.1 (or, more precisely, Lemma 7.2) if $\varphi_z^{\text{GFF}}-\lfloor m_{n_j}\rfloor$ were replaced by $\varphi_z^{\text{GFF}}-m_{n_j-k}+s-\beta^{-1/2}\alpha k$, so a key point to address here is the implicit limit $m_{n_j-k}-s+\beta^{-1/2}\alpha k-\lfloor m_{n_j}\rfloor\to 0$ inside the second argument of g_k. This is handled by setting $g_{k,\delta}^-(v,h)$, resp, $g_{k,\delta}^+(v,h)$ to be the infimum, resp., supremum of $g_k(v,h')$ over $h'\in[h-\delta,h+\delta]$ and noting that then

$$
\begin{aligned}
g_{k,\delta}^-\bigl([z]_{n_j-k},\varphi_z^{\text{GFF}}&-m_{n_j-k}+s-\beta^{-1/2}\alpha k\bigr)\\
&\leqslant g_k\bigl([z]_{n_j-k},\varphi_z^{\text{GFF}}-\lfloor m_{n_j}\rfloor\bigr)\\
&\qquad\leqslant g_{k,\delta}^+\bigl([z]_{n_j-k},\varphi_z^{\text{GFF}}-m_{n_j-k}+s-\beta^{-1/2}\alpha k\bigr)
\end{aligned}
\tag{7.40}
$$

as soon as j is so large that $|m_{n_j-k}-s+\beta^{-1/2}\alpha k-\lfloor m_{n_j}\rfloor|\leqslant\delta$. As $g_{k,\delta}^\pm$ are both continuous with a Gaussian upper bound, Lemma 7.2 implies sub-subsequential convergence of the type (7.39) with g_k replaced by the corresponding $g_{k,\delta}^\pm$ in (7.40) and thus bound the *limes superior*, resp., *inferior* of the expectations on the left of (7.39) as $j\to\infty$ by the right-hand side with g_k replaced by $g_{k,\delta}^+$, resp., $g_{k,\delta}^-$. Using that $g_{k,\delta}^\pm\to g_k$ monotonically as $\delta\downarrow 0$, (7.39) is concluded with the help of the Dominated Convergence Theorem.

With (7.39) in place, a simple change of variables allows us to get rid of the shift by s at the cost of multiplying the integral by $\mathrm{e}^{\alpha\sqrt{\beta}s}$. Next note that each sample χ from ν can be written as

$$
\chi=\sum_{i\geqslant 1}\delta_{t_i}
\tag{7.41}
$$

for a sequence of real-valued points $t_1\geqslant t_2\geqslant\dots$. The definition of g_k then shows

$$
\begin{aligned}
\mathrm{e}^{-\int g_k(v,\beta^{-1/2}(h+\cdot))\mathrm{d}\chi}&=\prod_{i\geqslant 1}\mathrm{e}^{-\int g_k(v,\beta^{-1/2}(h+t_i))}\\
&=\prod_{i\geqslant 1}E\Bigl(\mathrm{e}^{-\sum_{x\in\Lambda_k(z)}f(v,\varphi_x^{\text{DG}})}\,\Big|\,\varphi_z^{\text{GFF}}=\varphi_z^{\text{DG}}=\beta^{-1/2}(h+t_i)\Bigr),
\end{aligned}
\tag{7.42}
$$

where the conditional event was used to bring in the second argument of f to the stated form. Observe that the second argument of f receives an integer value, as it should.

We now want to cast the right-hand side of (7.42) as exponential of an integral but for that we will need to borrow part of the integral with respect to h from (7.39). Writing the conditional expectation on the right of (7.42) as $F_k(\beta^{-1/2}(h+t_i))$, elementary manipulations with integrals show

$$\begin{aligned}\int \mathrm{d}h\, \mathrm{e}^{-\alpha h}\Bigl(1-\prod_{i\geqslant 1} F_k\bigl(\beta^{-1/2}(h+t_i)\bigr)\Bigr)\\ &= \sqrt{\beta}\int \mathrm{d}h\, \mathrm{e}^{-\alpha\sqrt{\beta}\,h}\Bigl(1-\prod_{i\geqslant 1} F_k\bigl(h+\beta^{-1/2}t_i\bigr)\Bigr)\\ &= \sqrt{\beta}\sum_{n\in\mathbb{Z}} \mathrm{e}^{-\alpha\sqrt{\beta}\,n}\int_0^1 \mathrm{d}u\, \mathrm{e}^{-\alpha\sqrt{\beta}\,u}\Bigl(1-\prod_{i\geqslant 1} F_k\bigl(n+u+\beta^{-1/2}t_i\bigr)\Bigr)\\ &= \alpha^{-1}(1-\mathrm{e}^{-\alpha\sqrt{\beta}})\sum_{n\in\mathbb{Z}} \mathrm{e}^{-\alpha\sqrt{\beta}\,n}\, E\Bigl(1-\prod_{i\geqslant 1} F_k\bigl(n+U+\beta^{-1/2}t_i\bigr)\Bigr),\end{aligned} \tag{7.43}$$

where the expectation is with respect to the random variable U with the law

$$1_{[0,1)}(u)\frac{\alpha\sqrt{\beta}\,\mathrm{e}^{-\alpha\sqrt{\beta}\,u}}{1-\mathrm{e}^{-\alpha\sqrt{\beta}}}\mathrm{d}u. \tag{7.44}$$

In order to rewrite this expectation further, given independent samples of U and χ, let $\{\varphi^{(i)}\}_{i\geqslant 1}$ be independent samples of the DG-model in Λ_k with $\varphi^{(i)}$ drawn, for each $i\geqslant 1$, from the law of $\{\varphi^{\mathrm{DG}}_x\colon x\in\Lambda_k\}$ induced by

$$P\bigl(\,\cdot\,\big|\,\varphi^{\mathrm{GFF}}_z=\varphi^{\mathrm{DG}}_z=U+\beta^{-1/2}(t_i-\alpha k)\bigr). \tag{7.45}$$

where $\{t_i\}_{i\geqslant 1}$ is associated with χ as in (7.41). Then let $\nu_{\beta,k}$ be the law of

$$\zeta:=\sum_{i\geqslant 1}\sum_{x\in\Lambda_k(z)}\delta_{\varphi^{(i)}_x} \tag{7.46}$$

on $\mathcal{M}_{\mathbb{N}}(\mathbb{Z})$ which, we note, depends non-trivially on β both through the conditioning and the β-dependence of the DG-field.

Using the conditional independence of $\{\varphi^{(i)}\}_{i\geqslant 1}$ we now readily check

$$\int \nu(\mathrm{d}\chi)E\Bigl(1-\prod_{i\geqslant 1}F_k\bigl(n+U+\beta^{-1/2}t_i\bigr)\Bigr)=\int \nu_{\beta,k}(\mathrm{d}\zeta)\,\bigl(1-\mathrm{e}^{-\int f(x,n+\cdot)\mathrm{d}\zeta}\bigr) \tag{7.47}$$

thus producing the desired integral form. Combining the above observations, the right-hand side of (7.39) becomes

$$E\left(\exp\Bigl\{-\tilde{c}\,\mathrm{e}^{\alpha\sqrt{\beta}s}\int Z(\mathrm{d}x)\otimes\nu_{\beta,k}(\mathrm{d}\zeta)\sum_{n\in\mathbb{Z}}\mathrm{e}^{-\alpha\sqrt{\beta}\,n}\bigl(1-\mathrm{e}^{-\int f(x,n+\cdot)\mathrm{d}\zeta}\bigr)\Bigr\}\right), \tag{7.48}$$

where

$$\tilde{c}:=\alpha^{-1}(1-\mathrm{e}^{-\alpha\sqrt{\beta}}) \tag{7.49}$$

has no dependency on s.

Our next task is to extract a weak (subsequential) limit as $k\to\infty$. For this we claim

$$\bigl\{\nu_{\beta,k}\colon k\geqslant 1\bigr\}\text{ is tight.} \tag{7.50}$$

In light of our reliance on the vague topology, this will follow once we show that, given any $C \subseteq \mathbb{R}$ compact, the events $A_\epsilon := \{\zeta \in \mathcal{M}_{\mathbb{N}}(\mathbb{R})\colon \zeta(C) > 1/\epsilon\}$ obey

$$\limsup_{k\to\infty} \nu_{k,\beta}(A_\epsilon) \underset{\epsilon\downarrow 0}{\longrightarrow} 0. \tag{7.51}$$

To this end we take a continuous non-negative function f with compact support such that $f = 1$ on C and note that the integral inside the expectation in (7.48) with f replaced by ϵf is then at least $Z([0,1])(1-\mathrm{e}^{-1})\nu_{k,\beta}(A_\epsilon)$. For Y_n associated with f, Proposition 7.3 along with (7.39) then give

$$\liminf_{n\to\infty} E_{n,\beta}\bigl(\mathrm{e}^{-\epsilon Y_n}\bigr) \leqslant \liminf_{k\to\infty} E\bigl(\mathrm{e}^{-\tilde c Z([0,1])(1-\mathrm{e}^{-1})\nu_{k,\beta}(A_\epsilon)}\bigr). \tag{7.52}$$

The fact that $\{Y_n\}_{n\geqslant 1}$ is tight (by Theorem 6.2) implies that, as $\epsilon \downarrow 0$, the left-hand side tends to one. Since $Z([0,1]) > 0$ a.s., we must therefore have (7.51) as desired.

The tightness then permits us to extract a subsequential weak limit $\nu_{\beta,k_j} \overset{\text{law}}{\to} \nu_\beta$ (all relative to the vague topology) and get that, along the subsequence $n_j \to \infty$ such that (7.38) holds, $E\bigl(\mathrm{e}^{-Y_{n_j}}\bigr)$ converges to

$$E\left(\exp\Bigl\{-\tilde c\, \mathrm{e}^{\alpha\sqrt{\beta}\, s} \int Z(\mathrm{d}x)\otimes\nu_\beta(\mathrm{d}\zeta) \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\alpha\sqrt{\beta}\, n}\bigl(1-\mathrm{e}^{-\int f(x,n+\cdot)\mathrm{d}\zeta}\bigr)\Bigr\}\right). \tag{7.53}$$

Rewriting this in process language, this now gives the desired claim. □

It remains to give:

Proof of Theorem 2.1. This is a relatively straightforward consequence of Theorem 2.3 and tightness of the maximum. Indeed, for any integers $u < t$

$$\begin{aligned} 0 \leqslant E_{n,\beta}\Bigl(\mathrm{e}^{-\lambda\sum_{x\in\Lambda_n} 1_{(u,t]}(\varphi_x^{\mathrm{DG}}-\lfloor m_n\rfloor)}\Bigr) &- P_{n,\beta}\Bigl(\max_{x\in\Lambda_n}\varphi_x^{\mathrm{DG}} \leqslant \lfloor m_n\rfloor + u\Bigr) \\ &\leqslant \mathrm{e}^{-\lambda} + P_{n,\beta}\Bigl(\max_{x\in\Lambda_n}\varphi_x^{\mathrm{DG}} > \lfloor m_n\rfloor + t\Bigr) \end{aligned} \tag{7.54}$$

which uses the observation that the sum is at least one when the maximum lies in the interval $\lfloor m_n\rfloor + (u,t]$. Since $\lambda 1_{(u,t]}$ coincides with a bounded continuous function on the integers, along the sequence $\{n_j\}_{j\geqslant 1}$ such that (7.38) holds, Theorem 2.3 shows that the expectation in (7.54) tends to

$$E\biggl(\exp\Bigl\{-\tilde c\, \mathrm{e}^{\alpha\sqrt{\beta}\, s} Z\bigl([0,1]\bigr) \int \nu_\beta(\mathrm{d}\zeta) \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\alpha\sqrt{\beta}\, n}\bigl(1-\mathrm{e}^{-\lambda\int 1_{(u,t]}(n+\cdot)\mathrm{d}\zeta}\bigr)\Bigr\}\biggr). \tag{7.55}$$

Taking $t\to\infty$ followed by $\lambda\to\infty$, this becomes

$$E\biggl(\exp\Bigl\{-\tilde c\, \mathrm{e}^{\alpha\sqrt{\beta}\, s} Z\bigl([0,1]\bigr) \int \nu_\beta(\mathrm{d}\zeta) \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\alpha\sqrt{\beta}\, n} 1_{\{\zeta((-n+u,\infty))\geqslant 1\}}\bigr)\Bigr\}\biggr). \tag{7.56}$$

Thanks to Theorem 6.1, all the errors in (7.54) are wiped out in these limits, so this expectations is also the limit of the CDF of the centered maximum.

The tightness from Theorem 6.1 shows that the expectation (7.56) tends to one in the limit as $u\to\infty$. Since $Z([0,1]) < \infty$ a.s., the sum over n in the exponent is thus finite for all u sufficiently large and thus for all $u \in \mathbb{Z}$. Invoking the notations (2.12–2.13), we now get the claim. □

For completeness' sake, we also give:

Proof of Corollary 2.2. We will adhere to the notation in the statement in which φ' denotes the GFF and φ the DG-model. As will be discussed in the proof of Theorem 7.1, a routine calculation along with (2.12) shows

$$P'_{n,\beta}\Bigl(\max_{x\in\Lambda_n}\varphi'_x \leqslant m_n+u\Bigr) \underset{n\to\infty}{\longrightarrow} E\bigl(\mathrm{e}^{-(\alpha\sqrt{\beta})^{-1}\mathcal{Z}\mathrm{e}^{-\alpha\sqrt{\beta}u}}\bigr), \quad u\in\mathbb{R}, \tag{7.57}$$

where $\mathcal{Z} := Z([0,1])$. Given $r\in\mathbb{R}$ and $u\in\mathbb{R}$, (2.6) then gives

$$\begin{aligned}\liminf_{n\to\infty}\biggl[P_{n,\beta}\Bigl(\max_{x\in\Lambda_n}\varphi_x \leqslant \lfloor m_n\rfloor+u\Bigr) - P'_{n,\beta}\Bigl(\max_{x\in\Lambda_n}\varphi'_x \leqslant m_n+u+r\Bigr)\biggr] \\ \geqslant E\Bigl(\mathrm{e}^{-\hat{c}_\beta(s)\mathcal{Z}\mathrm{e}^{-\alpha\sqrt{\beta}\lfloor u\rfloor}}\Bigr) - E\Bigl(\mathrm{e}^{-(\alpha\sqrt{\beta})^{-1}\mathcal{Z}\mathrm{e}^{-\alpha\sqrt{\beta}(u+r)}}\Bigr),\end{aligned} \tag{7.58}$$

where s is a value obtained by taking a sequence $n_k\to\infty$ achieving the *limes inferior* such that $s := \lim_{k\to\infty}(m_{n_k}-\lfloor m_{n_k}\rfloor)$ exists. Thanks to the explicit form of $\hat{c}_\beta(s)$, the right-hand side is non-negative regardless of s once r obeys $(\alpha\sqrt{\beta})^{-1}\mathrm{e}^{-\alpha\sqrt{\beta}(r+1)} \geqslant \hat{c}_\beta(0)\mathrm{e}^{\alpha\sqrt{\beta}}$. Solving for r, we get $r\leqslant -2-(\alpha\sqrt{\beta})^{-1}\log[\hat{c}_\beta(0)\alpha\sqrt{\beta}]$.

A completely analogous argument shows that

$$\begin{aligned}\limsup_{n\to\infty}\biggl[P_{n,\beta}\Bigl(\max_{x\in\Lambda_n}\varphi_x \leqslant \lfloor m_n\rfloor+u\Bigr) - P'_{n,\beta}\Bigl(\max_{x\in\Lambda_n}\varphi'_x \leqslant m_n+u+r\Bigr)\biggr] \\ \leqslant E\Bigl(\mathrm{e}^{-\hat{c}_\beta(0)\mathcal{Z}\mathrm{e}^{-\alpha\sqrt{\beta}\lceil u\rceil}}\Bigr) - E\Bigl(\mathrm{e}^{-(\alpha\sqrt{\beta})^{-1}\mathcal{Z}\mathrm{e}^{-\alpha\sqrt{\beta}(u+r)}}\Bigr) \leqslant 0\end{aligned} \tag{7.59}$$

once $\hat{c}_\beta(0) \geqslant (\alpha\sqrt{\beta})^{-1}\mathrm{e}^{-\alpha\sqrt{\beta}r}$. Solving for r, we get $r\geqslant -(\alpha\sqrt{\beta})^{-1}\log[\hat{c}_\beta(0)\alpha\sqrt{\beta}]$. Replacing $\lfloor m_n\rfloor+u$ by u and using that $|m_n-\lfloor m_n\rfloor|\leqslant 1$ along with the monotonicity of the involved CDFs, the claim follows with $a := 1+(\alpha\sqrt{\beta})^{-1}|\log[\hat{c}_\beta(0)\alpha\sqrt{\beta}]|$. ☐

7.4 Proof of Theorem 7.1.

As our final task, we need to provide the details how to add the spatial component to convergence of the extremal process of the Branching Random Walk with normal step distribution. The argument is based on manipulations that are likely standard in the theory of Mandelbrot's multiplicative cascades; unfortunately, we are not aware of a treatment that could be cited without proof. To reduce clutter, we will assume that $\beta=1$ throughout. (The scaling by β can be added to the statement at the very end.) Departing from our earlier notation, we will also write $\varphi^{(n)}$ for the GFF on Λ_n.

Let us start by recalling Madaule's result [37] that we cast in the form of convergence of Laplace transforms: There exists an a.s.-positive and finite random variable $\mathfrak{Z}$ and a law ν on $\mathcal{M}_{\mathbb{N}}(\mathbb{R})$ with

$$\nu\Bigl(\chi\in\mathcal{M}_{\mathbb{N}}(\mathbb{R})\colon \chi\bigl((-\infty,0]\bigr)=\infty,\, \chi\bigl(\{0\}\bigr)\geqslant 1,\, \chi\bigl((0,\infty)\bigr)=0\Bigr)=1 \tag{7.60}$$

such that for all continuous $f\colon \mathbb{R}\to[0,\infty)$ with compact support,

$$E'_{n,1}\biggl(\exp\Bigl\{-\sum_{x\in\Lambda_n} f\bigl(\varphi^{(n)}_x-m_n\bigr)\Bigr\}\biggr) \underset{n\to\infty}{\longrightarrow} E\biggl(\exp\Bigl\{-\mathfrak{Z}\int \mathrm{e}^{-\alpha h}\mathrm{d}h\otimes\nu(\mathrm{d}\chi)(1-\mathrm{e}^{-\int f(h+\cdot)\mathrm{d}\chi})\Bigr\}\biggr), \tag{7.61}$$

where $E'_{n,1}$ is expectation with respect to the law $P'_{n,1}$ of GFF on Λ_n with $\beta:=1$ and $\alpha:=\sqrt{2\log b}$. The expectation on the right is with respect to the law of $\mathfrak{Z}$. (We deliberately use a different letter than $\mathcal{Z}$ as a connection with a random measure yet needs to be shown.) We start with the observation that $\mathfrak{Z}$ obeys a so called cascade relation:

Lemma 7.6 (Cascade relation) *Given a natural $n\geqslant 1$ be natural, let $\varphi^{(n)}$ be a GFF on Λ_n and let $\{\mathfrak{Z}_{n,x}\colon x\in\Lambda_n\}$ be i.i.d. copies of $\mathfrak{Z}$ that we assume are independent of $\varphi^{(n)}$. Then*

$$\mathfrak{Z} \overset{\text{law}}{=} \sum_{x\in\Lambda_n} \mathrm{e}^{\alpha\varphi^{(n)}_x-\alpha^2 n}\,\mathfrak{Z}_{n,x}. \tag{7.62}$$

Proof. Using suitable approximations along with the fact that the integral on the right of (7.61) is insensitive to changes of f at a single point the limit (7.61) applies to the function $f(h):=\lambda 1_{(u,\infty)}(h)$, for any $u\in\mathbb{R}$. Since the difference of the left-hand side of (7.61) with f as above and its limit as $\lambda\to\infty$ is bounded by $\mathrm{e}^{-\lambda}$ independent of n, we may take $\lambda\to\infty$ and invoke the properties (7.60) to get

$$P'_{n,1}\Bigl(\max_{x\in\Lambda_n}\varphi^{(n)}_x\leqslant m_n+u\Bigr) \underset{n\to\infty}{\longrightarrow} E\bigl(\mathrm{e}^{-\alpha^{-1}\mathrm{e}^{-\alpha u}\mathfrak{Z}}\bigr) \tag{7.63}$$

thus recovering the main result of Aïdekon [4] on which [37] draws heavily.

Let M_n be a random variable with $M_n \overset{\text{law}}{=} \max_{x\in\Lambda_n}\varphi^{(n)}_x$ and let us write $\zeta_1,\dots,\zeta_b$ for i.i.d. copies of $\mathcal{N}(0,1)$ and $M^{(1)}_{n-1},\dots,M^{(b)}_{n-1}$ for i.i.d. copies of M_{n-1} that are assumed to be independent of $\zeta_1,\dots,\zeta_b$. Using that the GFF is a Branching Random Walk, decomposing according to the first "step" shows

$$M_n \overset{\text{law}}{=} \max_{i=1,\dots,b}\bigl(\zeta_i+M^{(i)}_{n-1}\bigr). \tag{7.64}$$

This equates the probability on the left of (7.63) with b-th power of the probability that $\zeta+M_{n-1}\leqslant m_n+u$, for $\zeta=\mathcal{N}(0,1)$ independent of M_{n-1}. Passing this through the $n\to\infty$ limit while using that $m_n=m_{n-1}+\alpha+o(1)$ gives

$$E\bigl(\mathrm{e}^{-\alpha^{-1}\mathrm{e}^{-\alpha u}\mathfrak{Z}}\bigr)=\Bigl[E\bigl(\mathrm{e}^{-\alpha^{-1}\mathrm{e}^{-\alpha(u-\zeta+\alpha)}\mathfrak{Z}}\bigr)\Bigr]^b, \tag{7.65}$$

where $\zeta=\mathcal{N}(0,1)$ is independent of $\mathfrak{Z}$ on the right hand side. Interpreting the right-hand side expectation with respect to a product measure, the fact that the Laplace transform determines the law yields

$$\mathfrak{Z} \overset{\text{law}}{=} \sum_{i=1}^b \mathrm{e}^{\alpha\zeta_i-\alpha^2}\mathfrak{Z}_i, \tag{7.66}$$

where $\zeta_1,\dots,\zeta_b$ are i.i.d. $\mathcal{N}(0,1)$ and $\mathfrak{Z}_1,\dots,\mathfrak{Z}_b$ are i.i.d. copies of $\mathfrak{Z}$ that are independent of $\zeta_1,\dots,\zeta_b$. This is the desired cascade relation (7.62) for $n=1$. The general case is now readily proved by induction. □

The argument underlying the proof of the previous lemma can be bolstered to construct a full multiplicative cascade associated with random variable $\mathfrak{Z}$:

Lemma 7.7 (Multiplicative cascade) *Let Λ_0 be a singleton designating an empty sequence. There exists random variables $\{\mathfrak{Z}_{n,x},\zeta_{n,x}\colon n\geqslant 0,\ x\in\Lambda_n\}$ such that*

(1) $\{\zeta_{n,x}\colon n\geqslant 0,\ x\in\Lambda_n\}$ *are i.i.d.* $\mathcal{N}(0,1)$,
(2) *for each* $n\geqslant 0$, $\{\mathfrak{Z}_{n,x}\colon x\in\Lambda_n\}$ *are i.i.d. copies of* $\mathfrak{Z}$ *that are independent of random variables* $\{\zeta_{k,x}\colon k=0,\dots,n,\ x\in\Lambda_k\}$,
(3) *for all* $n\geqslant 0$ *and all* $x\in\Lambda_n$,

$$\mathfrak{Z}_{n,x}=\sum_{\substack{z\in\Lambda_{n+1}\\ m(z)=x}}\mathrm{e}^{\alpha\zeta_{n+1,z}-\alpha^2}\mathfrak{Z}_{n+1,z} \tag{7.67}$$

holds pointwise almost surely.

Proof. The one-step cascade relation (7.66) can be realized as an almost sure identity by simply declaring $\mathfrak{Z}$ to be the right-hand side of (7.66). Given $n\geqslant 0$ and i.i.d. standard normals $\{\zeta_{k,x}\colon k\geqslant 0,\dots,n,\ x\in\Lambda_k\}$, we now construct the joint law of random variables $\{\mathfrak{Z}_{k,x},\zeta_{n,x}\colon k=0,\dots,n,\ x\in\Lambda_k\}$ by taking independent i.i.d. copies $\{\mathfrak{Z}_{n,x}\colon x\in\Lambda_n\}$ of $\mathfrak{Z}$ and using (7.67) iteratively to define $\{\mathfrak{Z}_{k,x}\colon x\in\Lambda_k\}$ for all $k<n$. Noting that the cascade relation from Lemma 7.6 makes these laws consistent, the Kolmogorov Extension Theorem extends this to a law on the full infinite collection. □

Using the family $\{\mathfrak{Z}_{n,x},\zeta_{n,x}\colon n\geqslant 0,\ x\in\Lambda_n\}$, we can couple all of the GFFs on the same probability space by setting

$$\varphi_x^{(n)}:=\sum_{k=1}^{n}\zeta_{k,m^{n-k}(x)},\quad x\in\Lambda_n. \tag{7.68}$$

We now represent the random variables using a random Borel measure:

Lemma 7.8 *There exists a unique random Borel measure Z on $\mathbb{R}$ concentrated on $[0,1)$ such that for each Borel function $g\colon[0,1]\to[0,\infty)$ and each $n\geqslant 0$,*

$$\int Z(\mathrm{d}x)g\big([x]_n\big)=\sum_{x\in\Lambda_n}g\big([x]_n\big)\mathrm{e}^{\alpha\varphi_x^{(n)}-\alpha^2 n}\mathfrak{Z}_{n,x} \tag{7.69}$$

holds almost surely.

Proof. The identity (7.67) upgrades the equality in law (7.62) to the pointwise identity

$$\mathrm{e}^{\alpha\varphi_x^{(n)}-\alpha^2 n}\mathfrak{Z}_{n,x}=\sum_{\substack{z\in\Lambda_{n+k}\\ m^k(z)=x}}\mathrm{e}^{\alpha\varphi_z^{(n+k)}-\alpha^2(n+k)}\mathfrak{Z}_{n+k,z},\quad x\in\Lambda_n, \tag{7.70}$$

for each $n\geqslant 0$. Let $\mathcal{A}$ be the set of finite disjoint unions of sets from the collection

$$\mathcal{A}_0:=\big\{[kb^{-n},(k+1)b^{-n})\colon k=0,\dots,b^n-1,\ n\geqslant 0\big\}. \tag{7.71}$$

Define $Z\colon \mathcal{A}_0 \to [0,\infty)$ by putting

$$Z\big([kb^{-n},(k+1)b^{-n})\big) := \mathrm{e}^{\alpha\varphi^{(n)}_x - \alpha^2 n}\mathcal{Z}_{n,x} \tag{7.72}$$

for the unique $x \in \Lambda_n$ such that $[x]_n \in [kb^{-n},(k+1)b^{-n})$. Since (7.70) makes Z finitely additive on $\mathcal{A}_0$, a standard argument extends Z uniquely to an additive set function on $\mathcal{A}$. But $\mathcal{A}$ is an algebra such that any decreasing sequence $\{A_n\}_{n\geqslant 1}$ of elements from $\mathcal{A}$ with $A_n \downarrow \varnothing$ is eventually empty. This shows that Z is also trivially countably subadditive on $\mathcal{A}$ and so, by the Carathéodory Extension Theorem, extends uniquely to a Borel measure on $[0,1)$. Setting the measure to zero in the complement of $[0,1)$, the claim follows directly from (7.72). □

With these in hand, we are ready to give:

Proof of Theorem 7.1. Let $\lambda > 0$ and let $f\colon [0,1]\times\mathbb{R} \to [0,\infty)$ be continuous with support in $[0,1]\times[-\lambda,\lambda]$. For $k \geqslant 0$ define $f_k\colon [0,1]\times\mathbb{R}\to[0,\infty)$ by

$$f_k(x,h) := f\bigg(\sum_{j=1}^k x_j b^{-j}, h\bigg). \tag{7.73}$$

where $x_j := \lfloor b^j x\rfloor \bmod b$ for each $j = 1,\dots,k$. Note that since f is uniformly continuous and $m_n - (m_{n-k} + \alpha k) \to 0$ as $n \to \infty$, the quantity

$$\epsilon_k := \sup_{n\geqslant 2k}\ \sup_{x\in[0,1]}\ \sup_{h\in\mathbb{R}} \Big| f(x,h) - f_k(x, h + m_n - m_{n-k} - \alpha k)\Big| \tag{7.74}$$

obeys $\epsilon_k \to 0$ as $k\to\infty$. Using that also $\operatorname{supp}(f_k) \subseteq [0,1]\times[-\lambda,\lambda]$, we have

$$\begin{aligned}\bigg|\sum_{x\in\Lambda_n} f\big([x]_n, \varphi^{(n)}_x - m_n\big) - \sum_{x\in\Lambda_n} f_k\big([x]_n,\varphi^{(n)}_x - m_{n-k} - \alpha k\big)\bigg| \\ \leqslant \epsilon_k \sum_{x\in\Lambda_n} 1_{[-\lambda-1,\infty)}\big(\varphi^{(n)}_x - m_n\big)\end{aligned} \tag{7.75}$$

once $n \geqslant 2k$ and k is so large that $|m_n - m_{n-k} - \alpha k| \leqslant 1$. Since the convergence (7.61) entails that the level sets of the GFF are tight, we can swap the first sum on the left for the second sum provided we take the limit $k\to\infty$ after taking $n\to\infty$.

Invoking the coupling (7.68) and using that $f_k([x]_n,\cdot) = f([x]_k,\cdot)$ when $n \geqslant k$, for n in excess of k we get

$$\begin{aligned}&\sum_{x\in\Lambda_n} f_k\big([x]_n,\varphi^{(n)}_x - m_{n-k} - \alpha k\big) \\ &\qquad = \sum_{x\in\Lambda_k}\ \sum_{\substack{z\in\Lambda_n\\ m^{n-k}(z)=x}} f\big([x]_k, \varphi^{(n)}_z - \varphi^{(k)}_x - m_{n-k} + \varphi^{(k)}_x - \alpha k\big).\end{aligned} \tag{7.76}$$

Noting that the sets of values $\{\varphi^{(n)}_z - \varphi^{(k)}_x\colon z\in\Lambda_n,\ m^{n-k}(z) = x\}$ are independent for distinct $x\in\Lambda_k$, and are all independent of $\varphi^{(k)}$ with law of $\{\varphi^{(n-k)}_z\colon z\in\Lambda_{n-k}\}$, Madule's limit result (7.61) along with the Bounded Convergence Theorem show

$$\begin{aligned}&\lim_{n\to\infty} E'_{n,1}\Bigl(\exp\Bigl\{-\sum_{x\in\Lambda_n} f_k\bigl([x]_n,\varphi^{(n)}_x-m_{n-k}-\alpha k\bigr)\Bigr\}\Bigr)\\&=E'_{k,1}\otimes E\Bigl(\exp\Bigl\{-\sum_{x\in\Lambda_k}\mathfrak{Z}_{k,x}\int \mathrm{e}^{-\alpha h}\mathrm{d}h\otimes\nu(\mathrm{d}\chi)(1-\mathrm{e}^{-\int f([x]_k,h+\varphi^{(k)}_x-\alpha k+\cdot)\mathrm{d}\chi})\Bigr\}\Bigr),\end{aligned}\tag{7.77}$$

where the product expectation indicates that the law of $\varphi^{(k)}$ is independent of the independent copies $\{\mathfrak{Z}_{k,x}\colon x\in\Lambda_k\}$ of $\mathfrak{Z}$. Performing a routine change of variables, the latter expectation equals

$$\begin{aligned}&E'_{k,1}\otimes E\Bigl(\exp\Bigl\{-\int \mathrm{e}^{-\alpha h}\mathrm{d}h\otimes\nu(\mathrm{d}\chi)\sum_{x\in\Lambda_k}\mathrm{e}^{\alpha\varphi^{(k)}_x-\alpha^2 k}\mathfrak{Z}_{k,x}(1-\mathrm{e}^{-\int f([x]_k,h+\cdot)\mathrm{d}\chi})\Bigr\}\Bigr)\\&\qquad=E\Bigl(\exp\Bigl\{-\int Z(\mathrm{d}x)\otimes\mathrm{e}^{-\alpha h}\mathrm{d}h\otimes\nu(\mathrm{d}\chi)(1-\mathrm{e}^{-\int f_k(x,h+\cdot)\mathrm{d}\chi})\Bigr\}\Bigr),\end{aligned}\tag{7.78}$$

where Lemma 7.8 was used to rewrite the expression inside exponential as an integral with respect to Z-measure. Since $f_k\to f$ pointwise and the indicator $1_{[-\lambda,\infty)}(h)$ can be added to the integral over h thanks to the restriction on $\operatorname{supp}(f)$, the Bounded Convergence Theorem shows that the expectation tends to that for f as $k\to\infty$. In combination with (7.75), this proves the desired claim. □

Acknowledgments

This project has been supported in part by the NSF award DMS-1954343 and the BSF award 2018330.

References

[1] Y. Abe and M. Biskup (2022). Exceptional points of two-dimensional random walks at multiples of the cover time. *Probab. Theory Rel. Fields* **183**, 1–55.

[2] Y. Abe and M. Biskup (2023). A limit law for the extremal process of the local time of simple random walk on *b*-ary tree. In preparation.

[3] Y. Abe, M. Biskup and S. Lee (2023). Exceptional points of discrete-time random walks in planar domains. *Electron. J. Probab.* **28**, 1–45.

[4] E. Aïdékon (2013). Convergence in law of the minimum of a branching random walk. *Ann. Probab.* **41**, no. 3A, 1362–1426.

[5] E. Aïdékon, J. Berestycki, E. Brunet and Z. Shi (2013). The branching Brownian motion seen from its tip. *Probab. Theory Rel. Fields* **157**, no. 1–2, 405–451.

[6] M. Aizenman, M. Harel, R. Peled and J. Shapiro (2021). Depinning in integer-restricted Gaussian Fields and BKT phases of two-component spin models. arXiv:2110.09498

[7] L.-P. Arguin, A. Bovier, and N. Kistler (2011). Genealogy of extremal particles of branching Brownian motion. *Commun. Pure Appl. Math.* **64**, 1647–1676.

[8] L.-P. Arguin, A. Bovier, and N. Kistler (2012). Poissonian statistics in the extremal process of branching brownian motion. *Ann. Appl. Probab.* **22**, no. 4, 1693–1711.

[9] L.-P. Arguin, A. Bovier, and N. Kistler (2011). The extremal process of branching Brownian motion. *Probab. Theory Rel. Fields* **157**, no. 3–4, 535–574.

[10] N. Barashkov, T.S. Gunaratnam and M. Hofstetter (2022). Multiscale coupling and the maximum of $P(\varphi)_2$ models on the torus. arXiv:2207.13673

[11] R. Bauerschmidt and T. Bodineau (2020). Spectral gap critical exponent for Glauber dynamics of hierarchical spin models. *Commun. Math. Phys.* **373**, 1167–1206.

[12] R. Bauerschmidt and M. Hofstetter (2022). Maximum and coupling of the sine-Gordon field. *Ann. Probab.* **50**, no. 2, 455–508.
[13] R. Bauerschmidt, J. Park and P.-F. Rodriguez (2022). The Discrete Gaussian model, I. Renormalisation group flow at high temperature. arXiv:2202.02286
[14] R. Bauerschmidt, J. Park and P.-F. Rodriguez (2022). The Discrete Gaussian model, II. Infinite-volume scaling limit at high temperature. arXiv:2202.02287
[15] G. Benfatto and J. Renn (1992). Nontrivial fixed points and screening in the hierarchical two-dimensional Coulomb gas. *J. Statist. Phys.* **92**, 957–980.
[16] M. Biskup (2020). Extrema of the two-dimensional Discrete Gaussian Free Field. In: M. Barlow and G. Slade (eds.): Random Graphs, Phase Transitions, and the Gaussian Free Field. SSPROB 2017. Springer Proceedings in Mathematics & Statistics, 304 (2020) 163–407. Springer, Cham.
[17] M. Biskup A. Krieger (2021). Arithmetic oscillations of the chemical distance in long-range percolation on $\mathbb{Z}^d$. *Ann. Applied Probab.* (to appear) arXiv:2112.12365
[18] M. Biskup and O. Louidor (2016). Extreme local extrema of two-dimensional discrete Gaussian free field. *Commun. Math. Phys.* **345** 271–304.
[19] M. Biskup and O. Louidor (2020). Conformal symmetries in the extremal process of two-dimensional discrete Gaussian free field. *Commun. Math. Phys.* **375**, no. 1, 175–235.
[20] M. Biskup and O. Louidor (2018). Full extremal process, cluster law and freezing for two-dimensional discrete Gaussian free field. *Adv. Math.* **330** 589–687.
[21] M. Biskup and O. Louidor (2019). On intermediate level sets of two-dimensional discrete Gaussian free field. *Ann. Inst. Henri Poincaré* **55**, no. 4, 1948–1987.
[22] M. Biskup and O. Louidor (2021). A limit law for the most favorite point of simple random walk on a regular tree. *Ann. Probab.* (to appear) arXiv:2111.09513
[23] M. Bramson, J. Ding and O. Zeitouni (2016). Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Commun. Pure Appl. Math* **69**, no. 1, 62–123.
[24] D.C. Brydges (2009). *Lectures on the renormalisation group.* In: Statistical mechanics, IAS/Park City Math. Ser., vol. 16, pages 7–93. Amer. Math. Soc., Providence, RI.
[25] J. Ding, R. Roy, and O. Zeitouni (2017). Convergence of the centered maximum of log-correlated Gaussian fields. *Ann. Probab.* **45**, no 6A, 3886–3928.
[26] F. Dyson (1969). Existence of a phase transition in a one dimensional Ising ferromagnet, *Commun. Math. Phys.* **12**, 91–107.
[27] P. Falco (2012). Kosterlitz-Thouless transition line for the two dimensional Coulomb gas. *Commun. Math. Phys.* **312**, no. 2, 559–609.
[28] P. Falco (2013). Critical exponents of the two dimensional Coulomb gas at the Berezinskii-Kosterlitz-Thouless transition. arXiv:1311.2237.
[29] J. Fröhlich and T. Spencer (1981). The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas. *Commun. Math. Phys.* **81**, no. 4, 527–602.
[30] L.F. Guidi and D.H.U. Marchetti (2001). Renormalization group flow of the two-dimensional hierarchical Coulomb gas. *Commun. Math. Phys.* **219** 671–702.
[31] M. Hofstetter (2021). Extremal process of the sine-Gordon field. arXiv:2111.04842
[32] T. Hutchcroft (2022). Critical cluster volumes in hierarchical percolation. arXiv:2211.05686
[33] A. Jego (2023). Characterisation of planar Brownian multiplicative chaos. *Commun. Math. Phys.* **399** 971–1019.
[34] P. Lammers (2022). Height function delocalisation on cubic planar graphs. *Probab. Theory Rel. Fields* **82** 531–550.
[35] E. Lubetzky, F. Martinelli and A. Sly (2016). Harmonic pinnacles in the discrete Gaussian model. *Commun. Math. Phys.* **344** 673–717.
[36] T. Madaule (2015). Maximum of a log-correlated Gaussian field. *Ann. Inst. H. Poincaré Probab. Statist.* **51**, no 4, 1369–1431.
[37] T. Madaule (2017). Convergence in law for the branching random walk seen from its tip. *J. Theor. Probab.* **30** 27–63.
[38] B. Mallein (2016). Asymptotic of the maximal displacement in a branching random walk. *Graduate J. Math.* **1**, 92–104.

[39] D.H.U. Marchetti and J.F. Perez (1989). The Kosterlitz-Thouless phase transition in two-dimensional hierarchical Coulomb gases. *J. Statist. Phys.* **55**, no. 1-2, 141–156.
[40] E. Paquette and O. Zeitouni (2022). The extremal landscape for the CβE ensemble. arXiv:2209.06743
[41] F. Schweiger (2020). The maximum of the four-dimensional membrane model. *Ann. Probab.* **48**, no. 2, 714–741.
[42] F. Schweiger and O. Zeitouni (2022). The maximum of log-correlated Gaussian fields in random environments. arXiv:2205.07210
[43] M. Wirth (2019). Maximum of the integer-valued Gaussian free field. arXiv:1907.08868