Université de Montréal

Extremes of log-correlated random fields and the Riemann zeta function, and some asymptotic results for various estimators in statistics

par

Frédéric Ouimet

Département de mathématiques et de statistique
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures et postdoctorales
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en mathématiques

Orientation mathématiques appliquées

Mai 2019

© Frédéric Ouimet, 2019

Université de Montréal

Faculté des études supérieures et postdoctorales

Cette thèse intitulée

Extremes of log-correlated random fields and the Riemann zeta function, and some asymptotic results for various estimators in statistics

présentée par

Frédéric Ouimet

a été évaluée par un jury composé des personnes suivantes :

Louis-Pierre Arguin
(directeur de recherche)

Louigi Addario-Berry
(examinateur externe)

Alexander Fribergh
(co-directeur)

François Perron
(membre du jury)

Sabin Lessard
(président-rapporteur)

Marine Carrasco
(rep. du doyen de la FESP)

Thèse acceptée le :
29 mai 2019

Summary

In this thesis, we study the extreme values of certain log-correlated random fields that are Gaussian (the scale-inhomogeneous Gaussian free field and the time-inhomogeneous branching random walk) or approximatively Gaussian (the log-modulus of the Riemann zeta function on the critical line and a randomized toy model of it), as well as asymptotic properties of various estimators in statistics. Apart from the introduction and conclusion, the thesis is divided in three parts, each containing three articles.

The first part contains three articles on log-correlated Gaussian fields. The first article shows the first order convergence of the maximum and the number of high points for the scale-inhomogeneous Gaussian free field on its full domain. The second article uses the results from the first article to show that the limiting law of the Gibbs measure is a Ruelle probability cascade with a certain number of effective scales (a tree of Poisson-Dirichlet processes). The third article shows the tightness of the recentered maximum for the time-inhomogeneous branching random walk.

The second part contains three articles on the Riemann zeta function. The first article shows that, at low temperature, the limiting law of the Gibbs measure for a randomized toy model of the log-modulus of zeta on the critical line is a Poisson-Dirichlet process. The second article deals with the open problem of the tightness of the recentered maximum for this toy model on an interval of length $O(1)$. We simplify the problem by showing that the continuous maximum is at the order of constant away from a discrete maximum over $O(\log T \sqrt{\log \log T})$ points. The third article shows the first order of convergence of the maximum and the free energy for the log-modulus of the Riemann zeta function on short intervals of length $O(\log^{\theta} T)$, $\theta > -1$, on the critical line.

The third part contains three articles treating various topics in asymptotic statistics. The first article shows the complete monotonicity of multinomial probabilities and opens the door to the study of the asymptotic properties of Bernstein estimators on the simplex. The second article shows a uniform law of large numbers for sums containing terms that "blow up". The third article finds the limiting law of a modified score statistic when we test a given member of the exponential power distribution family against the family of asymmetric power distributions.

The thesis contains nine articles of which seven are already published in peer-reviewed journals. All the information is gathered on my personal website :

https://sites.google.com/site/fouimet26/research.

Keywords : probability, statistics, extreme value theory, log-correlated fields, Gaussian fields, Gaussian free field, branching random walk, inhomogeneous environment, Riemann zeta function, Gibbs measure, Ghirlanda-Guerra identities, ultrametricity, large deviations, asymptotic statistics, complete monotonicity, multinomial probabilities, Bernstein estimators, uniform law of large numbers, Laplace distribution, goodness-of-fit tests.

Sommaire

Dans cette thèse, nous étudions les valeurs extrêmes de certains champs aléatoires log-corrélés qui sont gaussiens (le champ libre gaussien inhomogène et la marche aléatoire branchante inhomogène) ou approximativement gaussiens (le log-module de la fonction zêta de Riemann sur la ligne critique et un modèle-jouet randomisé de celui-ci), ainsi que les propriétés asymptotiques de divers estimateurs en statistique. Outre l'introduction et la conclusion, la thèse est divisée en trois parties, chacune contenant trois articles.

La première partie contient trois articles sur les champs gaussiens log-corrélés. Le premier article montre le premier ordre de convergence du maximum et du nombre de hauts points pour le champ libre gaussien inhomogène sur tout son domaine. Le deuxième article utilise les résultats du premier article pour montrer que la loi limite de la mesure de Gibbs est une cascade de Ruelle avec un certain nombre d'échelles effectives (un arbre de processus de Poisson-Dirichlet). Le troisième article montre la tension du maximum recentré pour la marche aléatoire branchante inhomogène.

La deuxième partie contient trois articles sur la fonction zêta de Riemann. Le premier article montre que, à basse température, la loi limite de la mesure de Gibbs d'un modèle-jouet randomisé du log-module de zêta sur la ligne critique est un processus de Poisson-Dirichlet. Le deuxième article concerne le problème ouvert de la tension du maximum recentré pour ce modèle-jouet sur un intervalle de longueur $O(1)$. Nous simplifions le problème en montrant que le maximum continue se situe à une constante près d'un maximum discret sur $O(\log T \sqrt{\log\log T})$ points. Le troisième article montre le premier ordre de convergence du maximum et de l'énergie libre pour le log-module de la fonction zêta de Riemann sur des intervalles courts de longueur $O(\log^\theta T)$, $\theta > -1$, de la ligne critique.

La troisième partie contient trois articles traitant de sujets divers en statistique asymptotique. Le premier article montre la monotonicité complète des probabilités multinomiales et ouvre la porte sur l'étude des propriétés asymptotiques des estimateurs de Bernstein sur le simplexe. Le deuxième article prouve une loi uniforme des grands nombres pour les sommes contenant des termes qui « explosent ». Le troisième article trouve la loi limite d'une statistique de score modifiée lorsqu'on teste un membre donné de la famille des lois exponentielles de puissances contre la famille des lois de puissances asymétriques.

La thèse contient neuf articles dont sept sont déjà publiés dans des journaux évalués par les pairs. Toute l'information se trouve sur mon site web personnel :

`https://sites.google.com/site/fouimet26/research`.

Mots clés : probabilité, statistique, théorie des valeurs extrêmes, champs log-corrélés, champs gaussiens, champ libre gaussien, marche aléatoire branchante, environnements inhomogènes, fonction zêta de Riemann, mesure de Gibbs, identités de Ghirlanda-Guerra, ultramétricité, grandes déviations, statistique asymptotique, monotonicité complète, probabilités multinomiales, estimateurs de Bernstein, loi uniforme des grands nombres, loi de Laplace, tests d'ajustements.

Contents

List of acronyms and abbreviations

APD	Asymmetric Power Distribution
a.s.	almost surely
BBM	Branching Brownian Motion
BRW	Branching Random Walk
c.d.f.	cumulative distribution function
cf.	confer (*compare with* or *consult*)
Const.	Constant (*denotes a generic constant*)
CREM	Continuous Random Energy Model
CUE	Circular Unitary Ensemble
e.g.	exempli gratia (*for example*)
EPD	Exponential Power Distribution
et al.	et alia or et alii (*and others*)
etc.	et cetera (*and other similar things*)
GFF	Gaussian Free Field
GG	Ghirlanda-Guerra
GIP	Gaussian Integration by Parts
GMC	Gaussian Multiplicative Chaos
GREM	Generalized Random Energy Model
GTE	Gaussian Tail Estimate
IBRW	Time-Inhomogeneous Branching Random Walk
i.e.	id est (*that is*)
IGFF	Scale-Inhomogeneous Gaussian Free Field
i.i.d.	independent and identically distributed

LCP-CUE	Log-Characteristic Polynomials of the CUE field
LM-RZF	Log-Modulus of the RZF
LQG	Liouville Quantum Gravity
MM	Membrane Model
PNT	Prime Number Theorem
PPP	Poisson Point Process
REM	Random Energy Model
resp.	respectively
RH	Riemann Hypothesis
RLM-RZF	Randomized Log-Modulus of the RZF
RMF	Random Magnetic Field
ROSt	Random Overlap Structure
r.v.	random variable
RZF	Riemann Zeta Function
SRW	Simple Random Walk
s.t.	such that
VSBBM	Variable Speed Branching Brownian Motion

Some notation

Symbol	Meaning
$\circeq$	A definition or an equality that holds by definition.
$f(N) = o(g(N))$	$\lim_{N\to\infty} \frac{f(N)}{g(N)} = 0$.
$f(N) = O(g(N))$	$\lvert f(N)\rvert \le Cg(N)$ for N large enough, where $C > 0$ is a generic positive constant.
$f(N) \ll g(N)$	$f(N) = O(g(N))$.
$f(N) \gg g(N)$	$g(N) = O(f(N))$, where g is non-negative.
$f(N) \asymp g(N)$	$f(N) \ll g(N)$ and $f(N) \gg g(N)$.
$f(N) \sim g(N)$	$\lim_{N\to\infty} \frac{f(N)}{g(N)} = 1$.
$\mathcal{J}_f$	$\mathcal{J}_f(s) \circeq \int_0^s f(r)dr$ for any non-negative measurable function f.
$a \wedge b$	$\min\{a, b\}$.
$a \vee b$	$\max\{a, b\}$.

List of Figures

Acknowledgements

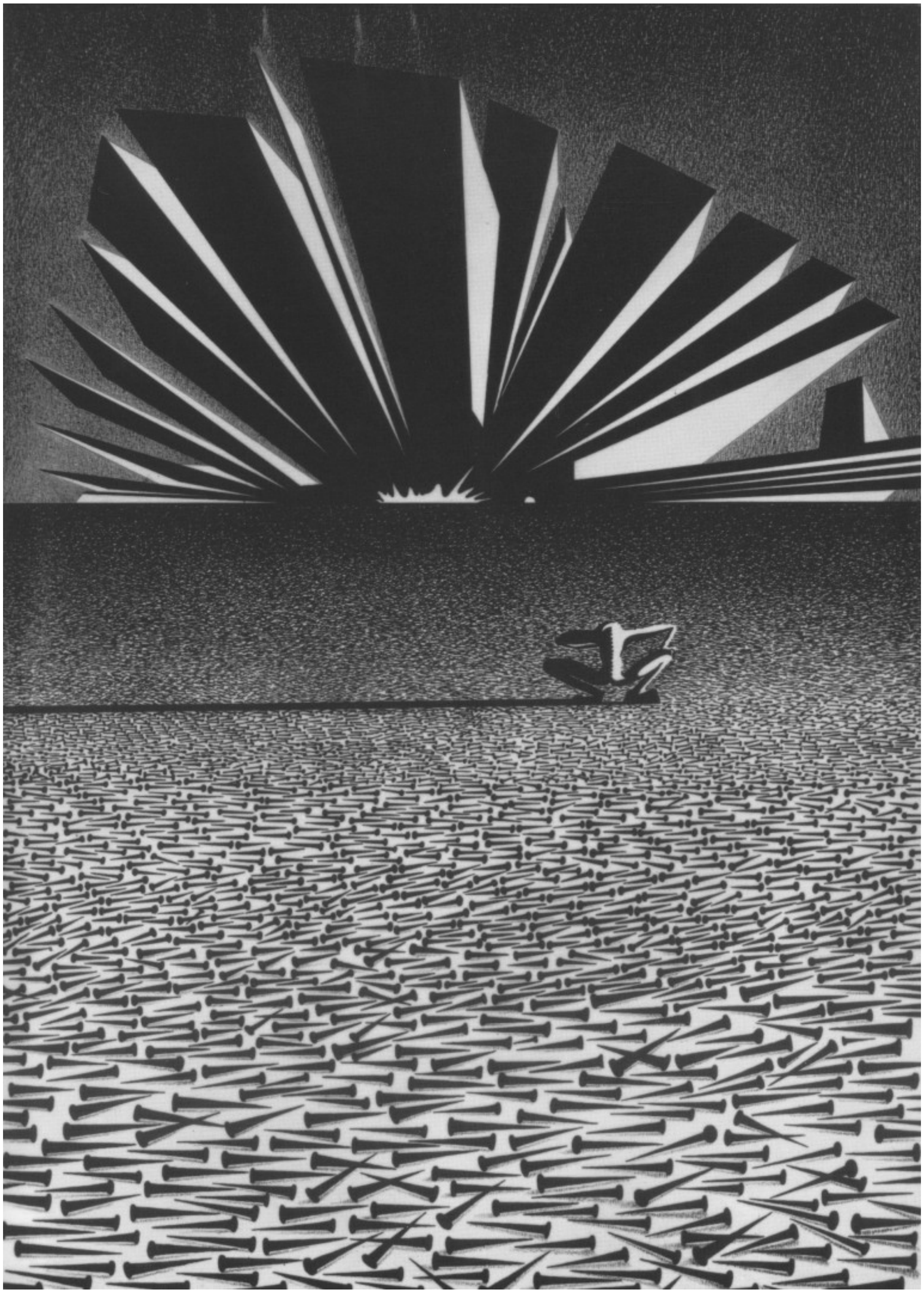
First, I want to thank Louis-Pierre Arguin for his support (moral and financial) throughout the years. I feel very blessed to have him as my advisor. You allowed me to pursue some of my research interests independently while keeping me on track on our collaborations. I thank you especially for letting me contribute to Article 6. It is definitely a game changer for me. I can only hope to become half as talented, both as a writer and as a mathematician.

I thank my co-advisor, Alexander Fribergh, (and also Louis-Pierre) for helping me dealing with the endless stream of bureaucratic paperwork and recommendation letters.

Finally, I thank everyone else who contributed, directly or indirectly, to conversations or email exchanges that led to bits and pieces of the material presented in this thesis : David Belius, Alessandra Cipriani, Yan Fyodorov, Nicola Kistler, Pierre Lafaye de Micheaux, Bastien Mallein, Roberto Persechino, Maksym Radziwiłł and Olivier Zindy.

This work is supported by a NSERC Doctoral program Alexander Graham Bell scholarship.

To my brother.



From chaos to order

(Courtesy of Professor A. T. Fomenko of the Moscow State University)

Introduction

0.1. Introduction

Over the past three decades (over the last 15 years in particular), there has been a push by physicists and mathematicians to extend certain universality results from classical extreme value theory to correlated random fields. On the physics side, the motivation stems mainly, on the one hand, from statistical mechanics where physicists/mathematicians are interested in understanding the behavior of spin glasses (see e.g. Mézard et al. (1987); Bovier (2006); Talagrand (2011a,b); Panchenko (2013b); Bovier (2017) and references therein), and on the other hand, from theoretical physics where they want to understand the behavior of subatomic particles in the framework of quantum field theory. In the latter case, the Gaussian free field (GFF) has been an important toy model to describe the properties of Liouville quantum gravity (LQG) measures and more generally Gaussian multiplicative chaos (GMC) measures (see e.g. Robert and Vargas (2010); Duplantier and Sheffield (2011); Rhodes and Vargas (2011); Barral et al. (2013); Garban (2013); Chen and Jakobson (2014); Duplantier et al. (2014a,b); Rhodes and Vargas (2014); Shamov (2016); Berestycki (2016); Garban et al. (2016); David et al. (2016); Berestycki (2017); Rhodes and Vargas (2017); Junnila and Saksman (2017), and references therein), the study of which goes back to Kahane's seminal work : Kahane (1985, 1986).

The appeal for mathematicians is often one of beauty, where they aim to generalize and extend as much as possible the properties and phenomena they observe for very simple models to models with more complex correlation structures. The subclass of log-correlated random fields (and log-correlated Gaussian fields especially) has emerged as particularly appropriate in the pursuit of this goal. As we will see in this thesis, the properties that are characteristic of this class of models appear far beyond the obvious examples. For instance, in Part 2, we will see that extreme values of the log-modulus of the Riemann zeta function on the critical line (and random toy models of it) behave approximately, quite surprisingly, as the extreme values of log-correlated Gaussian fields.

Before going further, here is how the thesis is organized.

0.2. Organization

The thesis is divided into five parts : the Introduction, the three main parts (Part 1, Part 2 and Part 3), and the Conclusion. Each of the main parts contains three articles (seven are published). Part 1 and Part 2 prove certain asymptotic results for the extreme values of models in the class of log-correlated random fields, whereas Part 3 collects miscellaneous results of interest in the theory of asymptotic statistics. More specifically,

- Part 1 (called Log-correlated Gaussian fields) deals with the extreme values of the scale-inhomogeneous Gaussian free field (Article 1 and Article 2) and the time-inhomogeneous branching random walk (Article 3);
- Part 2 (called The Riemann zeta function) deals with the extreme values of a random toy model of the log-modulus of the Riemann zeta function on the critical line (Article 4 and Article 5) and the Riemann zeta function itself (Article 6).
- Part 3 (called Asymptotic statistics) starts by proving the complete monotonicity of multinomial probabilities (Article 7) and shows how it can be used to study the asymptotic properties of Bernstein estimators on the simplex. In Article 8, a new uniform law of large numbers for summands that blow up is proved, which is then used to prove the convergence in law of a modified score statistic in Article 9, when testing a given exponential power distribution against asymmetric alternatives.

For the remainder of the introduction,

- We motivate Part 1 and Part 2 in Section 0.3 by :
  - recalling the classical version of extreme value theory (Section 0.3.1),
  - listing various questions of interest for log-correlated random fields in the modern extreme value theory (Section 0.3.2),
  - reviewing the literature to answer the questions of Section 0.3.2 for a selection of 12 log-correlated models (Section 0.3.3);
- We summarize the new results and ideas of the thesis in Section 0.4.

In the Conclusion, the reader can find a list of conjectures (Section 10.1), a list of open problems (Section 10.2), and a small errata for the published articles (Section 10.3). In the Appendix, there are two useful lemmas (Section 11.1), simulation codes (Section 11.2), and permissions from the coauthors and the journal editors (Section 11.3).

0.3. Motivation for Part 1 and Part 2

0.3.1. Classical extreme value theory

Let $\{X_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. random variables. We are interested in the cumulative distribution function (c.d.f.) of the maximum

$$M_N \mathrel{\hat=} \max_{1\leq i\leq N} X_i, \tag{0.3.1}$$

as $N \to \infty$. Since

$$\lim_{N\to\infty} \mathbb{P}(M_N \leq x) = \lim_{N\to\infty} (\mathbb{P}(X_1 \leq x))^N \in \{0,1\} \tag{0.3.2}$$

for all $x \in \mathbb{R}$, a more suitable question (analogous to the central limit theorem) is to ask if there exist sequences $\{a_N\}_{N\in\mathbb{N}}$ and $\{b_N\}_{N\in\mathbb{N}}$ (with $b_N > 0$) such that

$$\frac{M_N - a_N}{b_N} \quad \text{has a non-trivial limiting c.d.f., denoted by } F. \tag{0.3.3}$$

The existence of such sequences is not guaranteed, but if they exist, then only three types of limiting distributions are possible.

Theorem 0.3.1 (Fisher-Tippett-Gnedenko theorem, Proposition 0.3 in Resnick (2008))**.** *If (0.3.3) is satisfied, then F belongs to one and only one of the following classes of distributions (for some $\alpha > 0$) :*

— *Gumbel distribution :* $F(x) = \Lambda(x) \mathrel{\hat=} \exp(-e^{-x})$, *where* $x \in \mathbb{R}$.

— *Fréchet distribution :* $F(x) = \Phi_\alpha(x) \mathrel{\hat=} \begin{cases} 0, & \textit{if } x < 0, \\ \exp(-x^{-\alpha}), & \textit{if } x \geq 0. \end{cases}$

— *Weibull distribution :* $F(x) = \Psi_\alpha(x) \mathrel{\hat=} \begin{cases} \exp(-(-x)^{\alpha}), & \textit{if } x < 0, \\ 1, & \textit{if } x \geq 0. \end{cases}$

If $F \in \{\Lambda, \Phi_\alpha, \Psi_\alpha\}$ for a certain $\alpha > 0$ and (0.3.3) is satisfied, we say that $\mathbb{P}(X_1 \leq \cdot\,)$ belongs to the domain of attraction of F. The parameter α will be determined by the tail behavior of $\mathbb{P}(X_1 \leq \cdot\,)$. We refer the interested reader to Chapter 1 of Resnick (2008) for characterizations of the sequences $\{a_N\}_{N\in\mathbb{N}}$ and $\{b_N\}_{N\in\mathbb{N}}$ when $\mathbb{P}(X_1 \leq \cdot\,)$ belongs to each domain of attraction.

0.3.2. Questions of interest

When we introduce correlations between the variables X_i, the problem of the convergence in law becomes much less obvious to solve. Still, in some cases where the correlation structure is not too complicated, we can often answer simpler questions. In Part 1 and Part 2 of this thesis, we are particularly interested in *log-correlated random fields*, meaning that the correlations between the variables of the model decrease logarithmically with respect to a given notion of distance between the indices of the variables. Below, we list some of the questions that are of interest for log-correlated random fields.

Remark 0.3.1. *Since this thesis is concerned with log-correlated Gaussian fields (Part 1) and log-correlated random fields that are approximately Gaussian (Part 2), the X_i's that we consider below are at least close to be (centered) Gaussian random variables and their variance is proportional to $\log N$ (the log-number of points in the model), so that $b_N \sim 1$ and the search described in Section 0.3.1 is only about $\{a_N\}_{N\in\mathbb{N}}$.*

Here are the questions of interest :

(Q1): Does there exists a sequence $\{v_N\}_{N\in\mathbb{N}}$ such that

$$\frac{M_N}{v_N} \xrightarrow{\mathbb{P}} \gamma^\star, \quad \text{as } N \to \infty, \tag{0.3.4}$$

for some constant $\gamma^\star$? This is called the *first order of the maximum*.

(Q2): If $(Q1)$ is answered in the affirmative, then for every $\gamma \in (0, \gamma^\star)$, does

$$\frac{\log |\{i \in \{1, \dots, N\} : X_i \geq \gamma v_N\}|}{\log N} \xrightarrow{\mathbb{P}} \mathcal{E}(\gamma), \quad \text{as } N \to \infty, \tag{0.3.5}$$

for some constant $\mathcal{E}(\gamma)$? This is called the *first order of the log-number of γ-high points*.

(Q3): If $(Q1)$ is answered in the affirmative, is there a sequence $\{w_N\}_{N\in\mathbb{N}}$ such that

$$\frac{M_N - \gamma^\star v_N}{w_N} \xrightarrow{\mathbb{P}} \lambda^\star, \quad \text{as } N \to \infty, \tag{0.3.6}$$

for some constant $\lambda^\star$? This is called the *second order of the maximum*.

(Q4): If $(Q3)$ is answered in the affirmative,

$$\text{is the sequence } \Big\{M_N - (\gamma^\star v_N + \lambda^\star w_N)\Big\}_{N\in\mathbb{N}} \quad \textit{tight ?} \tag{0.3.7}$$

(Q5): If $(Q4)$ is answered in the affirmative, then we can ask the question about the convergence in law. Does

$$\mathbb{P}\Big(M_N - (\gamma^\star v_N + \lambda^\star w_N) \leq \,\cdot\,\Big) \xrightarrow{\text{law}} F(\,\cdot\,), \quad \text{as } N \to \infty, \tag{0.3.8}$$

for some non-trivial c.d.f. F ?

Around the same level of analysis, another related question of importance can be formulated in terms of the *extremal process* :

$$\Xi_N(A) \mathrel{\hat{=}} \sum_{i=1}^{N} \delta_{X_i - (\gamma^\star v_N + \lambda^\star w_N)}(A), \quad A \in \mathcal{B}(\mathbb{R}), \tag{0.3.9}$$

where δ denotes the Dirac measure and $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$. The question is : Does the sequence of random measures $\{\Xi_N\}_{N\in\mathbb{N}}$ converges weakly to a non-trivial *point process* (random counting measure) on $\mathbb{R}$?

(Q6): A *dual question* to $(Q1)$ and $(Q2)$ is the following. Can we determine the limit (in probability) of the *free energy* (also called the *log-partition function*)

$$f_N(\beta) \mathrel{\hat{=}} \frac{1}{\log N} \log \sum_{1\leq i\leq N} e^{\beta X_i} \quad ? \tag{0.3.10}$$

If we bound every term in the summation by the maximal term for the upper bound and we only keep the maximal term for the lower bound, we clearly have

$$\beta\, \frac{\max_{1\leq i\leq N} X_i}{\log N} \leq f_N(\beta) \leq 1 + \beta\, \frac{\max_{1\leq i\leq N} X_i}{\log N}, \tag{0.3.11}$$

so the reader can see an explicit link between $(Q1)$ and $(Q6)$. For certain models, the *limiting free energy* can be expressed as the Fenchel-Legendre transform of $-\mathcal{E}(\gamma)$, so it is also linked to $(Q2)$. This is because, for N large and $v_N = \log N$,

$$f_N(\beta) \approx \frac{1}{\log N} \log\left(\int_{-\infty}^{\infty} e^{\beta\gamma v_N} d\pi(\gamma)\right) \approx \max_{\gamma\in[0,\gamma^\star]} \{\beta\gamma + \mathcal{E}(\gamma)\}, \tag{0.3.12}$$

where $\pi(\gamma) \mathrel{\hat{=}} \#\{i \in \{1, \dots, N\} : X_i \leq \gamma v_N\}$. The last approximation follows from (0.3.5) and the fact that only the highest values in the integral in (0.3.12) have non-negligible weight under $d\gamma$. This is referred to as *Laplace's method*.

One of the goals in studying the limit of $f_N(\beta)$, and perturbed versions of it (where variance perturbations are added to the model), is finding the limiting law of the *overlaps* (also called *correlation coefficients*)

$$\rho(i,j) \mathrel{\hat{=}} \mathrm{Corr}(X_i, X_j) \mathrel{\hat{=}} \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}(X_i)}\sqrt{\mathrm{Var}(X_j)}} \tag{0.3.13}$$

under the product of random *Gibbs measures*

$$\mathcal{G}_{\beta,N}(\{i\}) \mathrel{\hat{=}} \frac{e^{\beta X_i}}{\sum_{1 \leq i \leq N} e^{\beta X_i}}, \tag{0.3.14}$$

which sample the large values of the field $\{X_i\}_{i=1}^N$ when $\beta > 0$. For instance, one quantity of interest is the so-called *limiting two-overlap distribution*

$$q \mapsto \lim_{N\to\infty} \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\Big[\mathbf{1}_{\{\rho(i,j)\leq q\}}\Big], \tag{0.3.15}$$

which is a measure of relative distance between the extremes of the model. Let $h : [-1,1]^{s(s-1)/2} \to \mathbb{R}$ be a continuous function of the overlaps of s points, then the more general question is to describe

$$\lim_{N\to\infty} \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h\Big((\rho(i_\ell, i_{\ell'}))_{1\leq \ell,\ell'\leq s}\Big)\Big] \tag{0.3.16}$$

as a function of $\beta > 0$. For log-correlated Gaussian fields (and other close models), the (mean) weak limit of $\{\mathcal{G}_{\beta,N}^{\times\infty}\}_{N\in\mathbb{N}}$ has a tendency to satisfy certain universal identities known as the Ghirlanda-Guerra identities, introduced in Ghirlanda and Guerra (1998). One very important result, due to Panchenko (2013a), proves that a random measure on the unit ball of a separable Hilbert space that satisfies the extended Ghirlanda-Guerra identities must be *ultrametric* (i.e. have a hierarchical structure of the overlaps (scalar products) under the mean measure). One consequence of this theorem is that, if the extended Ghirlanda-Guerra identities are satisfied in the limit for a given model, then it can be shown that the general joint distribution of the overlaps in (0.3.16) is completely determined by the limiting two-overlap distribution (see e.g. Theorem 2.13 in Panchenko (2013b)). The fact that the Ghirlanda-Guerra identities together with the limiting two-overlap distribution characterize the law of the overlaps under the limiting mean Gibbs measure was known to be true for specific models well before Panchenko's proof. For instance, it was pointed out for the REM in (Talagrand, 2003, Chapter 1).

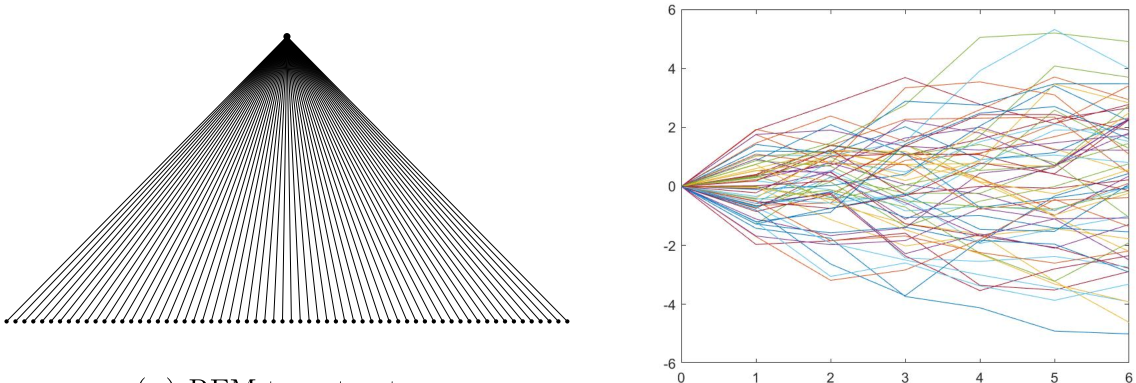
Apart from the motivations of statistical mechanics (see e.g. Bovier (2006)), the interest of the questions $(Q1)-(Q6)$ comes from the conjectured universality of the answers for the class of models that are log-correlated or close to it. Under certain regularity assumptions, it is for example expected that the sequence of recentered maxima for most branching models converges to a Gumbel distribution or the mean of randomly shifted Gumbel distributions. Similarly, it is expected that the Gibbs measures converge weakly towards a mixture of random measures called *Ruelle probability cascades* (see Ruelle (1987)) or, perhaps more generally, towards sampling measures of stochastically stable overlap structures (see Arguin and Aizenman (2009) and Arguin and Chatterjee (2013)).

0.3.3. Examples of log-correlated random fields : Old and new results

In this section, we present a list of 12 log-correlated random fields that are Gaussian or approximately Gaussian, some of which we will revisit in the articles of this thesis. We answer the questions of interest posed in the previous section by pointing to the relevant literature. All 12 fields are believed to belong to the REM class (resp. the GREM class, when the variances are macroscopically dependent on time or scale), meaning that the answers to the questions of interest $(Q1)$, $(Q2)$ and $(Q6)$ should all be the same as for the REM (see Section 0.3.3.1) (resp. the GREM, see Section 0.3.3.2) with the possible exception that some constants and critical levels could be model-specific (depending on the number of particles or the variances).

To give a sense of (some of) the techniques we will use in Part 1 and Part 2, we will answer the six questions directly (with proofs !) for the REM and state the answers for the GREM. For the other models, we will point directly to the literature.

0.3.3.1. *The random energy model (REM)*



(A) REM tree structure

(B) REM simulation ($N = 2^6$)

FIGURE 0.3.1. The random energy model

In the physics literature, the REM was first presented by Derrida (1980, 1981) as a toy model to study the properties of disordered systems (such as magnetic alloys at different temperatures). For an introduction, we refer the reader to Chapter 9 in Bovier (2006).

Definition 0.3.2 (REM)**.** *The REM consists of $N = 2^n$ i.i.d. r.v.s $X_i \sim \mathcal{N}(0, \sigma^2 \log N)$.*

Remark 0.3.2. *To see the analogy with the binary branching random walk, it can also be seen as the leaves of a tree of 2^n independent random walks on the time interval $[0, n]$, where each branch of length 1 is a standard Gaussian r.v., as shown in Figure 0.3.1b.*

The following proposition answers to $(Q1)$ and $(Q3)$.

Proposition 0.3.3 (First and second order of the maximum for the REM)**.** *Let $M_N \mathrel{\hat{=}} \max_{1 \leq i \leq N} X_i$ where the random field $\{X_i\}_{i=1}^N$ follows Definition 0.3.2, then*

$$\lim_{N \to \infty} \frac{M_N}{\log N} = g\sigma, \quad \textit{in probability,} \tag{0.3.17}$$

for the first order, where $g \mathrel{\hat{=}} \sqrt{2}$, and

$$\lim_{N \to \infty} \frac{M_N - g\sigma \log N}{\log \log N} = -\frac{1}{2} \cdot \frac{\sigma}{g}, \quad \textit{in probability,} \tag{0.3.18}$$

for the second order.

PROOF. Fix any $\varepsilon > 0$, and note that (0.3.17) follows from (0.3.18). To obtain the upper bound, use a union bound and a Gaussian tail estimate (Lemma 11.1.1) :

$$\begin{aligned}\mathbb{P}\left(M_N \geq g\sigma \log N - (1-\varepsilon)\frac{\sigma}{2g}\log\log N\right) &\leq N \cdot \frac{\sigma\sqrt{\log N}\cdot e^{-\frac{1}{2}\frac{(g\sigma\log N-(1-\varepsilon)\frac{\sigma}{2g}\log\log N)^2}{\sigma^2\log N}}}{(g\sigma\log N-(1-\varepsilon)\frac{\sigma}{2g}\log\log N)} \\ &\ll N\cdot\frac{1}{\sqrt{\log N}}\cdot N^{-1}(\log N)^{(1-\varepsilon)/2} \\ &\ll (\log N)^{-\varepsilon/2}, \end{aligned} \tag{0.3.19}$$

which tends to 0 as $N \to \infty$. For the lower bound, define

$$\mathcal{N}_N \mathrel{\hat{=}} \sum_{i=1}^N \mathbf{1}_{A_i}, \quad \text{where } A_i \mathrel{\hat{=}} \left\{X_i \geq g\sigma\log N - (1+\varepsilon)\frac{\sigma}{2g}\log\log N\right\}.$$

To conclude, we need to show that $\mathbb{P}(\mathcal{N}_N > 0) \to 1$ as $N \to \infty$. By applying the Paley-Zygmund inequality from Lemma 11.1.2 (with $\theta = 0$), we have

$$\mathbb{P}(\mathcal{N}_N > 0) \geq \frac{\mathbb{E}[\mathcal{N}_N]^2}{\mathbb{E}[\mathcal{N}_N^2]}. \tag{0.3.20}$$

It suffices to show that $\mathbb{E}[\mathcal{N}_N^2] = (1+o(1))\,\mathbb{E}[\mathcal{N}_N]^2$. Since the variables X_i and X_j are identically distributed and independent whenever $i \neq j$, we have the decomposition

$$\mathbb{E}\left[\mathcal{N}_N^2\right] = \sum_{\substack{i,j=1\\ i\neq j}}^N \mathbb{P}(A_i)^2 + \sum_{i=1}^N \mathbb{P}(A_i) = (1-N^{-1})\,\mathbb{E}[\mathcal{N}_N]^2 + \mathbb{E}[\mathcal{N}_N]\,, \tag{0.3.21}$$

because there are $N^2 - N$ terms in the first sum. By another Gaussian tail estimate (Lemma 11.1.1), note that

$$\mathbb{E}[\mathcal{N}_N] \asymp N\cdot\frac{1}{\sqrt{\log N}}\cdot N^{-1}(\log N)^{(1+\varepsilon)/2} \asymp (\log N)^{\varepsilon/2}. \tag{0.3.22}$$

Together with (0.3.21), this proves $\mathbb{E}[\mathcal{N}_N^2] = (1+o(1))\,\mathbb{E}[\mathcal{N}_N]^2$. $\square$

The following proposition answers to $(Q2)$.

Proposition 0.3.4 (Log-number of γ-high points for the REM). *Let $\gamma \in (0,\sigma)$ and define the set of points above the γ-level : $\mathcal{H}_N(\gamma) \mathrel{\hat{=}} \{1 \leq i \leq N : X_i \geq \gamma g \log N\}$. Then,*

$$\lim_{N\to\infty}\frac{\log|\mathcal{H}_N(\gamma)|}{\log N} = 1 - (\gamma/\sigma)^2, \quad \textit{in probability.} \tag{0.3.23}$$

Proof. By Markov's inequality and a Gaussian tail estimate (Lemma 11.1.1),

$$\begin{aligned}
\mathbb{P}\Big(|\mathcal{H}_N(\gamma)| \geq N^{1-(\gamma/\sigma)^2}\Big) &\leq N^{-1+(\gamma/\sigma)^2}\, \mathbb{E}\Big[|\mathcal{H}_N(\gamma)|\Big] \\
&\leq N^{-1+(\gamma/\sigma)^2} \cdot N \cdot \frac{\sigma\sqrt{\log N} \cdot e^{-\frac{1}{2}\frac{(\gamma g\, \log N)^2}{\sigma^2 \log N}}}{\gamma g\, \log N} \\
&\ll (\log N)^{-1/2}, \qquad (0.3.24)
\end{aligned}$$

which goes to 0 as $N \to \infty$.

For the lower bound, we want to apply (again) the Paley-Zygmund inequality :

$$\mathbb{P}\Big(|\mathcal{H}_N(\gamma)| \geq N^{1-(\gamma/\sigma)^2-\epsilon}\Big) \geq \left(1 - \frac{N^{1-(\gamma/\sigma)^2-\epsilon}}{\mathbb{E}\Big[|\mathcal{H}_N(\gamma)|\Big]}\right)^2 \frac{\mathbb{E}\Big[|\mathcal{H}_N(\gamma)|\Big]^2}{\mathbb{E}\Big[|\mathcal{H}_N(\gamma)|^2\Big]}. \qquad (0.3.25)$$

Using Gaussian tail estimates (Lemma 11.1.1), it is easy to see that

$$\mathbb{E}[|\mathcal{H}_N(\gamma)|] \asymp (\log N)^{-1/2} N^{1-(\gamma/\sigma)^2}, \qquad (0.3.26)$$

so it suffices to prove $\mathbb{E}\Big[|\mathcal{H}_N(\gamma)|^2\Big] = (1+o(1))\, \mathbb{E}\Big[|\mathcal{H}_N(\gamma)|\Big]^2$ to conclude the proof. Since the variables X_i and X_j are identically distributed and independent whenever $i \neq j$, we have the decomposition

$$\begin{aligned}
\mathbb{E}\Big[|\mathcal{H}_N(\gamma)|^2\Big] &= \sum_{\substack{i,j=1 \\ i\neq j}}^{N} \Big(\mathbb{P}\Big(X_i \geq \gamma g\, \log N\Big)\Big)^2 + \sum_{i=1}^{N} \mathbb{P}\Big(X_i \geq \gamma g\, \log N\Big) \\
&= (1 - N^{-1})\, \mathbb{E}\Big[|\mathcal{H}_N(\gamma)|\Big]^2 + \mathbb{E}\Big[|\mathcal{H}_N(\gamma)|\Big] \\
&= (1+o(1))\, \mathbb{E}\Big[|\mathcal{H}_N(\gamma)|\Big]^2, \qquad (0.3.27)
\end{aligned}$$

where the last equality follows from (0.3.26) and $\gamma \in (0, \sigma)$. This ends the proof. □

The following proposition answers the first part of ($Q6$) about the free energy of the REM. It is consequence of Proposition 0.3.4 and the first part of Proposition 0.3.3.

Proposition 0.3.5 (Limiting free energy of the REM)**.** *For $\beta > 0$,*

$$\lim_{N\to\infty} f_N(\beta) = \lim_{N\to\infty} \frac{1}{\log N} \log \sum_{i=1}^{N} e^{\beta X_i} = f(\beta;\sigma^2), \qquad (0.3.28)$$

where the limit holds in probability and in L^1, and where $g \circeq \sqrt{2}$ and

$$f(\beta;\sigma^2) \circeq \begin{cases} 1+\left(\frac{\beta}{\beta_c}\right)^2, & \text{if } \beta \leq \beta_c \circeq \frac{g}{\sigma}, \\ 2\frac{\beta}{\beta_c}, & \text{if } \beta > \beta_c. \end{cases} \tag{0.3.29}$$

Proof. We prove the upper bound first. For $M \in \mathbb{N}$ and a fixed $\varepsilon > 0$, consider $\widetilde{\varepsilon} \circeq \varepsilon/(2+\beta g)$, $\gamma_j \circeq \frac{j}{M}\sigma + \widetilde{\varepsilon}$ for $0 \leq j \leq M$, and the good event

$$\begin{aligned} U \circeq \bigcap_{j=1}^{M} &\left\{\left|\left\{1 \leq i \leq N : X_i > \gamma_{j-1} g \log N\right\}\right| \leq N^{1-(\gamma_{j-1}/\sigma)^2+\widetilde{\varepsilon}}\right\} \\ &\bigcap \left\{\max_{1\leq i\leq N} X_i \leq (\sigma+\widetilde{\varepsilon}) g \log N\right\}. \end{aligned} \tag{0.3.30}$$

By (0.3.17) and (0.3.23), we know that $\mathbb{P}(U^c) \to 0$ as $N \to \infty$. On the event U, we have

$$\begin{aligned} \sum_{i=1}^{N} e^{\beta X_i} &= \sum_{j=1}^{M}\sum_{i=1}^{N} e^{\beta X_i} \mathbf{1}_{\{\gamma_{j-1} g \log N < X_i \leq \gamma_j g \log N\}} + \sum_{i=1}^{N} e^{\beta X_i} \mathbf{1}_{\{X_i \leq \gamma_0 g \log N\}} \\ &\leq \sum_{j=1}^{M} N^{\beta\gamma_j g + 1 - (\gamma_{j-1}/\sigma)^2 + \widetilde{\varepsilon}} + N^{\beta\gamma_0 g + 1}, \end{aligned} \tag{0.3.31}$$

which implies that, for M large enough with respect to ε and β,

$$f_N(\beta) \circeq \frac{\log \sum_{i=1}^{N} e^{\beta X_i}}{\log N} \leq \max_{\gamma \in [0,\sigma+\widetilde{\varepsilon}]} \left\{\beta\gamma g + 1 - (\gamma/\sigma)^2\right\} + 2\widetilde{\varepsilon}. \tag{0.3.32}$$

In the case $\beta \leq \beta_c \circeq \frac{g}{\sigma}$, the maximum is attained at $\gamma = \beta\sigma/\beta_c$. Hence, the right-hand side of (0.3.32) is at most $1 + (\beta/\beta_c)^2 + \varepsilon$. In the other case $\beta > \beta_c$, we can choose $\varepsilon > 0$ small enough that $\beta > (\sigma + \widetilde{\varepsilon})\beta_c/\sigma$. The maximum is attained at $\gamma = \sigma + \widetilde{\varepsilon}$, in which case the right-hand side of (0.3.32) is smaller than $2(\beta/\beta_c) + \varepsilon$. This proves the upper bound in probability.

For the lower bound, consider the levels $\gamma_j \circeq \frac{j}{M}\sigma + \varepsilon$ for $0 \leq j \leq M$, and the good event

$$\begin{aligned} L \circeq \bigcap_{j=1}^{M} &\left\{\left|\left\{1 \leq i \leq N : X_i > \gamma_{j-1} g \log N\right\}\right| \geq N^{1-(\gamma_{j-1}/\sigma)^2-\varepsilon/2}\right\} \\ &\bigcap \left\{\max_{1\leq i\leq N} X_i \leq (\sigma+\varepsilon) g \log N\right\}. \end{aligned} \tag{0.3.33}$$

By (0.3.17) and (0.3.23), we know that $\mathbb{P}(L^c) \to 0$ as $N \to \infty$. On the event L, we have

$$\begin{aligned}\sum_{i=1}^{N} e^{\beta X_i} &= \sum_{j=1}^{M}\sum_{i=1}^{N} e^{\beta X_i}\mathbf{1}_{\{\gamma_{j-1} g \log N < X_i \leq \gamma_j g \log N\}} + \sum_{i=1}^{N} e^{\beta X_i}\mathbf{1}_{\{X_i \leq \gamma_0 g \log N\}} \\ &\geq \max_{1\leq j\leq M} N^{\beta\gamma_{j-1} g + 1 - (\gamma_{j-1}/\sigma)^2 - \varepsilon/2}, \end{aligned} \tag{0.3.34}$$

which implies that, for M large enough with respect to ε and β,

$$f_N(\beta) \mathrel{\hat{=}} \frac{\log \sum_{i=1}^{N} e^{\beta X_i}}{\log N} \geq \max_{\gamma\in[\varepsilon,\sigma]} \Big\{\beta\gamma g + 1 - (\gamma/\sigma)^2\Big\} - \varepsilon. \tag{0.3.35}$$

In the case $0 < \beta \leq \beta_c \mathrel{\hat{=}} \frac{g}{\sigma}$, take $\varepsilon > 0$ small enough that $\beta > \varepsilon\beta_c/\sigma$. The maximum is attained at $\gamma = \beta\sigma/\beta_c$, in which case the right-hand side of (0.3.35) is equal to $1 + (\beta/\beta_c)^2 - \varepsilon$. In the other case $\beta > \beta_c$, the maximum is attained at $\gamma = \sigma$, in which case the right-hand side of (0.3.35) is equal to $2(\beta/\beta_c) - \varepsilon$. This proves the lower bound in probability.

It remains to show the convergence of the free energy in L^1. It is a consequence of the uniform integrability of the sequence $\{f_N(\beta)\}_{N\in\mathbb{N}}$, namely

$$\lim_{\alpha\to\infty} \sup_{N\in\mathbb{N}} \mathbb{E}\Big[|f_N(\beta)|\mathbf{1}_{\{|f_N(\beta)|>\alpha\}}\Big] = 0. \tag{0.3.36}$$

To prove (0.3.36), let $\xi_N \mathrel{\hat{=}} M_N/\log N$. Then, from (0.3.11),

$$\beta\xi_N \leq f_N(\beta) \leq \beta\xi_N + 1. \tag{0.3.37}$$

Assume that $\alpha > 1$. By splitting the event $\{|f_N(\beta)| > \alpha\}$ in two parts : $\{f_N(\beta) > \alpha\}$ and $\{-f_N(\beta) > \alpha\}$, and then using (0.3.37), we deduce

$$\begin{aligned}&\mathbb{E}\Big[|f_N(\beta)|\mathbf{1}_{\{|f_N(\beta)|>\alpha\}}\Big] \\ &\quad\leq \mathbb{E}\Big[(\beta\xi_N + 1)\mathbf{1}_{\{\beta\xi_N+1>\alpha\}}\Big] + \mathbb{E}\Big[(-\beta\xi_N)\mathbf{1}_{\{-\beta\xi_N>\alpha\}}\Big] \\ &\quad= \sum_{\ell=1}^{\infty}\mathbb{E}\Big[(\beta\xi_N + 1)\mathbf{1}_{\{(\ell+1)\alpha\geq\beta\xi_N+1>\ell\alpha\}}\Big] + \sum_{\ell=1}^{\infty}\mathbb{E}\Big[(-\beta\xi_N)\mathbf{1}_{\{(\ell+1)\alpha\geq-\beta\xi_N>\ell\alpha\}}\Big] \\ &\quad\leq \sum_{\ell=1}^{\infty}(\ell+1)\,\alpha\,\mathbb{P}\Big(|\xi_N| > \frac{1}{\beta}(\ell\alpha - 1)\Big). \end{aligned} \tag{0.3.38}$$

Using a union bound and a Gaussian tail estimate (Lemma 11.1.1), we have

$$\begin{aligned}\mathbb{P}\Big(|\xi_N| > \frac{1}{\beta}(\ell\alpha - 1)\Big) &\le N \cdot \max_{1\le i\le N} 2\,\mathbb{P}\Big(X_i > \frac{1}{\beta}(\ell\alpha - 1)\log N\Big) \\ &\le N \cdot 2\, N^{-\frac{(\ell\alpha-1)^2}{2\beta^2\sigma^2}}.\end{aligned} \tag{0.3.39}$$

For α large enough with respect to β and σ, applying (0.3.39) in (0.3.38) implies (0.3.36). This ends the proof. $\square$

The following theorem answers to $(Q4)$ and $(Q5)$.

Theorem 0.3.6 (Convergence in law of the recentered maximum for the REM)**.** *Define*

$$a_N \circeq g\sigma \log N - \frac{\sigma}{2g}\log\log N - \frac{\sigma \log(g\sqrt{2\pi})}{g}. \tag{0.3.40}$$

Then, for $\beta_c \circeq \frac{g}{\sigma}$ and any $x \in \mathbb{R}$,

$$\lim_{N\to\infty} \mathbb{P}\Big(M_N - a_N \le x\Big) = \exp(-e^{-\beta_c x}). \tag{0.3.41}$$

In other words, the limiting law of the recentered maximum is a Gumbel distribution.

Proof. Fix any $x \in \mathbb{R}$. Since the X_i's are i.i.d. Gaussian r.v.s with variance $\sigma^2 \log N$,

$$\mathbb{P}\Big(M_N - a_N \le x\Big) = (1 - \Psi(z^\star))^N, \tag{0.3.42}$$

where Ψ denotes the survival function of the standard Gaussian distribution and

$$z^\star \circeq \frac{a_N + x}{\sigma\sqrt{\log N}} = g\sqrt{\log N} - \frac{\log\log N}{2g\sqrt{\log N}} - \frac{\log(g\sqrt{2\pi})}{g\sqrt{\log N}} + \frac{x}{\sigma\sqrt{\log N}}. \tag{0.3.43}$$

The Gaussian tail estimates from Lemma 11.1.1 tell us that, for $z > 0$,

$$\frac{1}{z}\left(1 - \frac{1}{z^2}\right) \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \le \Psi(z) \le \frac{1}{z} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}. \tag{0.3.44}$$

Thus, from (0.3.42), (0.3.43) and (0.3.44), the reader can easily verify that

$$\left[1 - \frac{(1 + O(\frac{(\log\log N)^2}{\log N}))e^{-\frac{g}{\sigma}x}}{N}\right]^N \le \mathbb{P}(M_N - a_N \le x) \le \left[1 - \frac{(1 + O(\frac{(\log\log N)^2}{\log N}))e^{-\frac{g}{\sigma}x}}{N}\right]^N$$

from which (0.3.41) directly follows. $\square$

To answer ($Q6$), one way to proceed would be to show that the extremal process converges weakly to a specific Poisson point process (PPP), and then deduce the limiting law of the Gibbs measure. This is done in Section 8.3 of Bolthausen and Sznitman (2002). Specifically, they show in Proposition 8.6, using Laplace functionals, that the extremal process $\sum_{i=1}^N \delta_{X_i - a_N}$ converges weakly to a PPP with intensity $t \mapsto \beta_c e^{-\beta_c t}$, and using the nice transformation properties of PPPs under continuous mappings (Proposition 8.5), they deduce that, for $\beta > \beta_c$ (at low temperature), the point process $\sum_{i=1}^N \delta_{\exp(\beta(X_i - a_N))}$ converges weakly to a PPP with intensity $t \mapsto (\beta_c/\beta) t^{-(\beta_c/\beta)-1}$. Now, note that the limiting Gibbs measure is defined by the normalized points of the last limiting point process; it can be shown that the normalized weights, when arranged in a decreasing order, form a Poisson-Dirichlet process of parameter β_c/β (see e.g. Proposition 2 in Tao (2013)).

In general, it is very very hard to find the limit of the extremal process for most models (like the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-IGFF and the RLM-RZF, for which we have new results on the limiting Gibbs measure in this thesis), so we will instead present a much longer (but more robust) alternative route to illustrate some of the methods used in Ouimet (2017) (Article 2) and Ouimet (2018b) (Article 4).

Here is our plan :

(a) Find the limiting free energy of a perturbed version of the REM;

(b) Use (a) to find the limiting two-overlap distribution for the REM;

(c) Prove the extended Ghirlanda-Guerra identities for the REM;

(d) Deduce from (b) and (c) the joint distribution of the overlaps under the limiting mean Gibbs measure in terms of Poisson-Dirichlet variables.

For $\alpha \in [0,1]$, $u \in (-1,1)$ and $1 \leq i \leq N$, define the *perturbed REM* by

$$\widetilde{X}_{i,\alpha,u} \mathrel{\hat{=}} X_i + u X_i^{(1)} = (1+u) X_i^{(1)} + X_i^{(2)}, \tag{0.3.45}$$

where $X_i^{(1)} \sim \mathcal{N}(0, \alpha\sigma^2 \log N)$, $X_i^{(2)} \sim \mathcal{N}(0, (1-\alpha)\sigma^2 \log N)$, $X_i \mathrel{\hat{=}} X_i^{(1)} + X_i^{(2)}$, and where all the $X_i^{(1)}$'s and $X_i^{(2)}$'s are independent. By convention, $X \sim \mathcal{N}(0,0)$ means $X = 0$.

Proposition 0.3.7 (Limiting free energy of the perturbed REM)**.** *Define the free energy of the perturbed REM by*

$$\widetilde{f}_{N,\alpha,u}(\beta) \mathrel{\mathring{=}} \frac{1}{\log N} \log \sum_{1 \leq i \leq N} e^{\beta \widetilde{X}_{i,\alpha,u}}, \quad \beta > 0. \tag{0.3.46}$$

Then, for any $\alpha \in [0,1]$ and $\beta > 0$,

$$\begin{aligned}
\lim_{N\to\infty} \widetilde{f}_{N,\alpha,u}(\beta) \mathrel{\mathring{=}} \widetilde{f}_{\alpha,u}(\beta) &= \max_{\gamma \in [0,\gamma^\star]} \Big\{\beta\gamma + \mathcal{E}_{\alpha,u}(\gamma)\Big\} \\
&= \begin{cases} f(\beta; V_{\alpha,u}), & \text{if } u < 0, \\ \alpha f(\beta; \sigma^2(1+u)^2) + (1-\alpha) f(\beta;\sigma^2), & \text{if } u \geq 0, \end{cases} \\
&= \begin{cases} 1 + \frac{\beta^2 V_{\alpha,u}}{g^2}, & \text{if } u < 0 \text{ and } \beta \leq \frac{g}{\sqrt{V_{\alpha,u}}}, \\ \beta g \sqrt{V_{\alpha,u}}, & \text{if } u < 0 \text{ and } \beta > \frac{g}{\sqrt{V_{\alpha,u}}}, \\ 1 + \frac{\beta^2 V_{\alpha,u}}{g^2}, & \text{if } u \geq 0 \text{ and } \beta \leq \frac{g}{\sigma(1+u)}, \\ \alpha\Big(\beta g \sigma(1+u)\Big) + (1-\alpha)\Big(1 + \frac{\beta^2\sigma^2}{g^2}\Big), & \text{if } u \geq 0 \text{ and } \frac{g}{\sigma(1+u)} < \beta \leq \frac{g}{\sigma}, \\ \beta g \sigma\Big[\alpha(1+u) + (1-\alpha)\Big], & \text{if } u \geq 0 \text{ and } \beta > \frac{g}{\sigma}, \end{cases}
\end{aligned} \tag{0.3.47}$$

where $V_{\alpha,u} \mathrel{\mathring{=}} \alpha(\sigma(1+u))^2 + (1-\alpha)\sigma^2$, $f(\beta;\sigma^2)$ is defined in (0.3.29), $\gamma^\star$ is defined in (0.3.49), and where the limit in (0.3.47) holds in probability and in L^1.

Proof. This proof is easy but somewhat long and tedious, so we will only sketch it. For the details, see e.g. Ouimet (2017) (Article 2) for a more general statement and proof in the context of the scale-inhomogeneous Gaussian free field or Section 4 in Arguin and Tai (2018) for a proof in the context of the randomized Riemann zeta function. The reader can also check Section 9 of Bolthausen and Sznitman (2002) for the proof of the limiting free energy in the context of the GREM with two levels. The idea is to apply second-moment methods as in the proof of Proposition 0.3.3 and Proposition 0.3.4 to find the first order of the maximum and the first order of the log-number of γ-high points for the perturbed REM. We would find

$$\lim_{N\to\infty} \frac{\max_{1\leq i \leq N} \widetilde{X}_{i,\alpha,u}}{\log N} = g\gamma^\star, \quad \text{in probability}, \tag{0.3.48}$$

where

$$\gamma^\star \mathrel{\hat{=}} \begin{cases} \sqrt{V_{\alpha,u}}, & \text{if } u \leq 0, \\ \alpha(\sigma(1+u)) + (1-\alpha)\sigma, & \text{if } u > 0, \end{cases} \tag{0.3.49}$$

and, for $\gamma \in (0, \gamma^\star)$, we would find

$$\lim_{N\to\infty} \frac{\log \left|\left\{1 \leq i \leq N : \widetilde{X}_{i,\alpha,u} \geq \gamma g \log N\right\}\right|}{\log N} = \mathcal{E}_{\alpha,u}(\gamma), \quad \text{in probability,} \tag{0.3.50}$$

where

$$\mathcal{E}_{\alpha,u}(\gamma) \mathrel{\hat{=}} \begin{cases} 1 - \frac{\gamma^2}{V_{\alpha,u}}, & \text{if } u < 0, \text{ or if } u \geq 0 \text{ and } \gamma < \gamma_c, \\ (1-\alpha) - \frac{(\gamma - \alpha\sigma(1+u))^2}{(1-\alpha)\sigma^2}, & \text{if } u \geq 0 \text{ and } \gamma \geq \gamma_c, \end{cases} \tag{0.3.51}$$

and where $\gamma_c \mathrel{\hat{=}} V_{\alpha,u}/(\sigma(1+u))$. Using (0.3.48) and (0.3.50), we could then apply Laplace's method like we did in the proof of Proposition 0.3.5 and solve the related optimization problem to conclude. □

The next proposition finds the limiting two-overlap distribution of the REM. The strategy of the proof is to link the derivative (with respect to u) of the limiting free energy of the perturbed REM to the limiting two-overlap distribution through Gaussian integration by parts. We refer to this strategy as the *Bovier-Kurkova technique* as it was successfully applied to find the limiting two-overlap distribution of the general GREM in Bovier and Kurkova (2004a,b).

The following proposition answers the second part of ($Q6$).

Proposition 0.3.8 (Limiting two-overlap distribution of the REM)**.** *Recall the definition of the overlaps $\rho(i,j)$ from (0.3.13). Then, for $\beta_c \mathrel{\hat{=}} \frac{g}{\sigma}$, and any $\beta > 0$,*

$$\begin{aligned} \beta \leq \beta_c : \qquad & \lim_{N\to\infty} \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\left[\mathbf{1}_{\{\rho(i,j)\leq q\}}\right] = \begin{cases} 0, & \text{if } q < 0, \\ 1, & \text{if } q \geq 0, \end{cases} \\ \beta > \beta_c : \qquad & \lim_{N\to\infty} \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\left[\mathbf{1}_{\{\rho(i,j)\leq q\}}\right] = \begin{cases} 0, & \text{if } q < 0, \\ \frac{\beta_c}{\beta}, & \text{if } 0 \leq q < 1, \\ 1, & \text{if } q \geq 1, \end{cases} \end{aligned} \tag{0.3.52}$$

Here is an intuitive explanation of the meaning of this proposition. When $\beta > \beta_c$, the Gibbs measure gives a lot of weight to the particles i that are near the maximum's height in the tree structure. The result simply says that if you sample two particles under the Gibbs measure, then, in the limit and on average, either the particles branched off "at the last moment" in the tree structure or they branched off in the beginning. They cannot branch at intermediate scales. Since the particles are independent by definition in the case of the REM, this means that β_c/β is the probability, in the limit $N \to \infty$, that the particles sampled under $\mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}$ differ.

When $\beta < \beta_c$, the weights in the Gibbs measure are more spread out so that most contributions to the free energy actually come from particles reaching heights that are well below the level of the maximum in the tree structure. Hence, when two particles are selected from this larger pool of contributors that are not clustering, it can be shown that, in the limit and on average, the particles necessarily branched off in the beginning of the tree. In other words, in the limit $N \to \infty$, the particles sampled under $\mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}$ are different with probability 1.

Proof of Proposition 0.3.8. From (0.3.47), we can verify that, for any $0 < \alpha < 1$ and $\beta > 0$, there exists $\delta = \delta(\alpha, \beta) > 0$ small enough that $\widetilde{f}_{\alpha,u}(\beta)$ is differentiable with respect to $u \in (-\delta, \delta)$ and

$$\frac{\beta_c^2}{2\beta^2} \frac{\partial}{\partial u} \widetilde{f}_{\alpha,u}(\beta)\bigg|_{u=0} = \begin{cases} \alpha, & \text{if } \beta \leq \beta_c, \\ \frac{\beta_c}{\beta}\alpha, & \text{if } \beta > \beta_c. \end{cases} \tag{0.3.53}$$

By Holder's inequality, note that $u \mapsto \widetilde{f}_{N,\alpha,u}(\beta)$ is convex (and thus $u \mapsto \mathbb{E}\big[\widetilde{f}_{N,\alpha,u}(\beta)\big]$ is also convex). Since pointwise limits preserve convexity and we have the mean convergence $\lim_{N\to\infty} \mathbb{E}\big[\widetilde{f}_{N,\alpha,u}(\beta)\big] = \widetilde{f}_{\alpha,u}(\beta)$ from Proposition 0.3.7, then $u \mapsto \widetilde{f}_{\alpha,u}(\beta)$ is convex. The fact that $u \mapsto \widetilde{f}_{\alpha,u}(\beta)$ is differentiable on $(-\delta, \delta)$ together with Theorem 25.7 in Rockafellar (1970) then implies

$$\frac{\partial}{\partial u} \widetilde{f}_{\alpha,u}(\beta)\bigg|_{u=0} = \lim_{N\to\infty} \frac{\partial}{\partial u} \mathbb{E}\big[\widetilde{f}_{N,\alpha,u}(\beta)\big]\Big|_{u=0}. \tag{0.3.54}$$

Also, using *Gaussian integration by parts* (see e.g. Lemma 1.1 in Panchenko (2013b)), we have the relation

$$\begin{aligned}
\frac{\beta_c^2}{2\beta^2}\frac{\partial}{\partial u}\,\mathbb{E}\Big[\widetilde{f}_{N,\alpha,u}(\beta)\Big]\Big|_{u=0} &= \frac{1}{\beta\sigma^2\log N}\mathbb{E}\mathcal{G}_{\beta,N}\Big[X_i^{(1)}\Big] \\
&= \frac{1}{\beta\sigma^2\log N}\mathbb{E}\left[\frac{\sum_{1\le i\le N} X_i^{(1)}\exp(\beta(X_i^{(1)}+X_i^{(2)}))}{\sum_{1\le j\le N}\exp(\beta(X_j^{(1)}+X_j^{(2)}))}\right] \\
&\overset{\text{GIP}}{=} \frac{1}{\sigma^2\log N}\Big\{\mathbb{E}\mathcal{G}_{\beta,N}\Big[\mathbb{E}[X_i^{(1)}X_i]\Big] - \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\Big[\mathbb{E}[X_i^{(1)}X_j]\Big]\Big\} \\
&= \alpha - \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\left[\frac{\mathbb{E}[X_i^{(1)}X_j]}{\sigma^2\log N}\right], \qquad (0.3.55)
\end{aligned}$$

where

$$\frac{\mathbb{E}[X_i^{(1)}X_j]}{\sigma^2\log N} = \rho(i,j)\mathbf{1}_{\{\rho(i,j)\le\alpha\}} + \alpha\mathbf{1}_{\{\rho(i,j)>\alpha\}} = \int_0^\alpha \mathbf{1}_{\{\rho(i,j)>y\}}dy. \qquad (0.3.56)$$

We deduce from (0.3.55) and (0.3.56) that

$$\frac{\beta_c^2}{2\beta^2}\frac{\partial}{\partial u}\,\mathbb{E}\Big[\widetilde{f}_{N,\alpha,u}(\beta)\Big]\Big|_{u=0} = \int_0^\alpha \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\Big[\mathbf{1}_{\{\rho(i,j)\le y\}}\Big]dy. \qquad (0.3.57)$$

Since $[0,1]\subseteq\mathbb{R}$ is compact, the space $\mathcal{M}_1([0,1])$ of probability measures on $[0,1]$ is compact under the weak topology. Thus, any subsequence of the cumulative distribution function on the left-hand side of (0.3.52) has a subsequence converging to a cumulative distribution function. Pick any such sub-subsequence and denote its limit by $q\mapsto x_\beta(q)$. From (0.3.53), (0.3.54) and (0.3.57), we have, for any $\alpha\in(0,1)$,

$$\int_0^\alpha x_\beta(y)dy = \begin{cases} \alpha, & \text{if } \beta\le\beta_c, \\ \frac{\beta_c}{\beta}\alpha, & \text{if } \beta>\beta_c. \end{cases} \qquad (0.3.58)$$

By Lebesgue's differentiation theorem and the fact that the c.d.f. x_β is right-continuous and concentrated on $[0,1]$ (since $\rho(i,j)\in[0,1]$ for all i,j), the conclusion follows. $\square$

By the representation theorem of Dovbysh and Sudakov (1982) for symmetric positive definite weakly exchangeable infinite random arrays (for a proof, see Panchenko (2010b)), we can show (since $[0,1]^{\mathbb{N}\times\mathbb{N}}$ being compact implies that the space of probability measures $\mathcal{M}_1([0,1]^{\mathbb{N}\times\mathbb{N}})$ is weakly compact) that there exists a subsequence $\{N_m\}_{m\in\mathbb{N}}$ converging to $+\infty$ such that for any $s\in\mathbb{N}$ and any continuous function $h:[0,1]^{s(s-1)/2}\to\mathbb{R}$, we have

$$\lim_{m\to\infty} \mathbb{E}\mathcal{G}_{\beta,N_m}^{\times\infty}\Big[h((\rho(i_\ell,i_{\ell'}))_{1\le \ell,\ell'\le s})\Big] = E\mu_\beta^{\times\infty}\Big[h((R_{\ell,\ell'})_{1\le \ell,\ell'\le s})\Big], \tag{0.3.59}$$

where R is a random element of some probability space with measure P (and expectation E), generated by the random matrix of scalar products

$$(R_{\ell,\ell'})_{\ell,\ell'\in\mathbb{N}} = \Big((\rho_\ell,\rho_{\ell'})_{\mathcal{H}}\Big)_{\ell,\ell'\in\mathbb{N}}, \tag{0.3.60}$$

where $(\rho_\ell)_{\ell\in\mathbb{N}}$ is an i.i.d. sample from some random measure μ_β concentrated a.s. on the unit sphere of a separable Hilbert space $\mathcal{H}$. In particular, from Proposition 0.3.8, we have, for all $A\in\mathcal{B}(\mathbb{R})$,

$$E\mu_\beta^{\times 2}\Big[\mathbf{1}_{\{R_{1,2}\in A\}}\Big] = \begin{cases} \mathbf{1}_A(0), & \text{if } \beta\le\beta_c,\\ \frac{\beta_c}{\beta}\mathbf{1}_A(0) + \Big(1-\frac{\beta_c}{\beta}\Big)\mathbf{1}_A(1), & \text{if } \beta>\beta_c. \end{cases} \tag{0.3.61}$$

The meaning behind (0.3.61) is the same as we explained below (0.3.52), where μ_β is simply a subsequential limit of $\{\mathcal{G}_{\beta,N}\}_{N\in\mathbb{N}}$ in the specific sense of (0.3.59).

Proposition 0.3.9 (Extended Ghirlanda-Guerra identities in the limit)**.** *Let $\beta>0$, and let μ_β be a subsequential limit of $\{\mathcal{G}_{\beta,N}\}_{N\in\mathbb{N}}$ in the sense of (0.3.59). For any $s\in\mathbb{N}$, any $k\in\{1,\ldots,s\}$, and any functions $g:\{0,1\}\to\mathbb{R}$ and $h:\{0,1\}^{s(s-1)/2}\to\mathbb{R}$, we have*

$$\begin{aligned} E\mu_\beta^{\times(s+1)}\Big[g(R_{k,s+1})h((R_{i,i'})_{1\le i,i'\le s})\Big] &= \frac{1}{s}E\mu_\beta^{\times 2}\Big[g(R_{1,2})\Big]E\mu_\beta^{\times s}\Big[h((R_{i,i'})_{1\le i,i'\le s})\Big]\\ &\quad+\frac{1}{s}\sum_{\ell\ne k}^{s} E\mu_\beta^{\times s}\Big[g(R_{k,\ell})h((R_{i,i'})_{1\le i,i'\le s})\Big]. \end{aligned} \tag{0.3.62}$$

The meaning behind this proposition is a bit more difficult to explain. In the special case $h\equiv 1$ and $g(x)=x$, the formula (0.3.62) becomes

$$E\mu_\beta^{\times(s+1)}\Big[R_{k,s+1}\Big] = \frac{1}{s}E\mu_\beta^{\times 2}\Big[R_{1,2}\Big] + \frac{1}{s}\sum_{\ell\ne k}^{s} E\mu_\beta^{\times s}\Big[R_{k,\ell}\Big]. \tag{0.3.63}$$

It says that, on average, given the particles $1,2,\ldots,s$ sampled under μ_β, the overlap between the next particle to be sampled (called $s+1$) and the k-th particle is a generic overlap between two particles (1 possibility) or it is the overlap between another particle that we already sampled and the k-th particle ($s-1$ possibilities), where each possibility has probability $1/s$. Now, the formula (0.3.62) describes more generally the joint behavior of the matrix of overlaps for the first s particles that are sampled and the overlap between the next particle to be sampled and the k-th one. As already mentioned below (0.3.16), it can be shown, in general, that (0.3.62) implies the ultrametricity of the overlaps (a tree-like hierarchical structure) and completely characterizes the joint law of the overlaps under $E\mu_\beta^{\times\infty}$ up to the knowledge of all $E\mu_\beta^{\times 2}[g(R_{1,2})]$'s, which themselves are determined by the limiting two-overlap distribution (0.3.61) (see Chapter 2 in Panchenko (2013b)).

Proof of Proposition 0.3.9. Let $\phi_N : \{1,2,\dots,N\}^s \to \mathbb{R}$ be any set of functions that satisfies $\sup_N \|\phi_N\|_\infty < \infty$. By Gaussian integration by parts (see e.g. Exercise 1.1 in Panchenko (2013b)), note that, for $\boldsymbol{i} \doteq (i_1, i_2, \dots, i_s) \in \{1,2,\dots,N\}^s$,

$$\begin{aligned}
\frac{\mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[X_{i_k}^{(1)}\phi_N(\boldsymbol{i})\Big]}{\beta\sigma^2\log N} &\overset{\text{GIP}}{=} \sum_{\ell=1}^s \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\left[\frac{\mathbb{E}[X_{i_k}^{(1)}X_{i_\ell}]}{\sigma^2\log N}\,\phi_N(\boldsymbol{i})\right] - s\,\mathbb{E}\mathcal{G}_{\beta,N}^{\times(s+1)}\left[\frac{\mathbb{E}[X_{i_k}^{(1)}X_{i_{s+1}}]}{\sigma^2\log N}\,\phi_N(\boldsymbol{i})\right] \\
&\overset{(0.3.56)}{=} \sum_{\ell=1}^s \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[\int_0^\alpha \mathbf{1}_{\{\rho(i_k,i_\ell)>y\}}dy\,\phi_N(\boldsymbol{i})\Big] \\
&\qquad - s\,\mathbb{E}\mathcal{G}_{\beta,N}^{\times(s+1)}\Big[\int_0^\alpha \mathbf{1}_{\{\rho(i_k,i_{s+1})>y\}}dy\,\phi_N(\boldsymbol{i})\Big].
\end{aligned}$$

On the other end, from (0.3.55), we have

$$\frac{\mathbb{E}\mathcal{G}_{\beta,N}\Big[X_{i_k}^{(1)}\Big]}{\beta\sigma^2\log N} = \mathbb{E}\mathcal{G}_{\beta,N}\Big[\int_0^\alpha \mathbf{1}_{\{\rho(i_k,i_k)>y\}}dy\Big] - \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\Big[\int_0^\alpha \mathbf{1}_{\{\rho(i,j)>y\}}dy\Big], \tag{0.3.64}$$

and a concentration argument (see e.g. Theorem 3.8 in Panchenko (2013b) or Proposition 5.6 in Ouimet (2018b) (Article 4)) shows that, as $N \to \infty$,

$$\left|\frac{\mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[X_{i_k}^{(1)}\phi_N(\boldsymbol{i})\Big]}{\beta\sigma^2\log N} - \frac{\mathbb{E}\mathcal{G}_{\beta,N}\Big[X_{i_k}^{(1)}\Big]}{\beta\sigma^2\log N}\mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[\phi_N(\boldsymbol{i})\Big]\right| = o(1). \tag{0.3.65}$$

Putting the last three equations together, we find that, as $N \to \infty$,

$$\begin{aligned}
&\Bigg|\mathbb{E}\mathcal{G}_{\beta,N}^{\times(s+1)}\Big[\int_0^\alpha \mathbf{1}_{\{\rho(i_k,i_{s+1})>y\}}dy\,\phi_N(\boldsymbol{i})\Big] \\
&\qquad - \left\{\begin{array}{l} \frac{1}{s}\mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\Big[\int_0^\alpha \mathbf{1}_{\{\rho(i,j)>y\}}dy\Big]\,\mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[\phi_N(\boldsymbol{i})\Big] \\ +\frac{1}{s}\sum_{\ell\neq k}^s \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[\int_0^\alpha \mathbf{1}_{\{\rho(i_k,i_\ell)>y\}}dy\,\phi_N(\boldsymbol{i})\Big]\end{array}\right\}\Bigg| = o(1).
\end{aligned} \tag{0.3.66}$$

If we take the limit (0.3.59) in the last equation with $\phi_N(\boldsymbol{i}) \doteq h^\star((\rho(i_\ell, i_{\ell'}))_{1\leq\ell,\ell'\leq s})$ where

$h^\star\colon [0,1]^{s(s-1)/2} \to \mathbb{R}$ is a continuous extension of h, we find that, for all $\alpha \in (0,1)$,

$$\begin{aligned} &E\mu_\beta^{\times(s+1)}\Big[\int_0^\alpha \mathbf{1}_{\{R_{k,s+1}>y\}} dy\, h((R_{i,i'})_{1\le i,i'\le s})\Big] \\ &\quad = \frac{1}{s} E\mu_\beta^{\times 2}\Big[\int_0^\alpha \mathbf{1}_{\{R_{1,2}>y\}} dy\Big] E\mu_\beta^{\times s}\Big[h((R_{i,i'})_{1\le i,i'\le s})\Big] \\ &\qquad + \frac{1}{s}\sum_{\ell\neq k}^{s} E\mu_\beta^{\times s}\Big[\int_0^\alpha \mathbf{1}_{\{R_{k,\ell}>y\}} dy\, h((R_{i,i'})_{1\le i,i'\le s})\Big]. \end{aligned} \tag{0.3.67}$$

From (0.3.61), we know that $\mathbf{1}_{\{R_{i,i'}>y\}}$ is $E\mu_\beta^{\times 2}$-a.s. constant in y on $[-1,0)$ and $[0,1)$, respectively. Therefore, for any $x \in \{-1,0\}$,

$$\begin{aligned} &E\mu_\beta^{\times(s+1)}\left[\mathbf{1}_{\{R_{k,s+1}>x\}}\, h((R_{i,i'})_{1\le i,i'\le s})\right] \\ &\quad = \frac{1}{s} E\mu_\beta^{\times 2}\left[\mathbf{1}_{\{R_{1,2}>x\}}\right] E\mu_\beta^{\times s}\Big[h((R_{i,i'})_{1\le i,i'\le s})\Big] \\ &\qquad + \frac{1}{s}\sum_{\ell\neq k}^{s} E\mu_\beta^{\times s}\left[\mathbf{1}_{\{R_{k,\ell}>x\}}\, h((R_{i,i'})_{1\le i,i'\le s})\right]. \end{aligned} \tag{0.3.68}$$

But, any function $g : \{0,1\} \to \mathbb{R}$ can be written as a linear combination of the indicator functions $\mathbf{1}_{\{\,\cdot\,>-1\}}$ and $\mathbf{1}_{\{\,\cdot\,>0\}}$, so we get the conclusion by the linearity of (0.3.68). □

Finally, we can answer the last part of ($Q6$).

Theorem 0.3.10. *Let $\beta > 0$. Let $\xi \doteq \{\xi_k\}_{k\in\mathbb{N}}$ be a Poisson-Dirichlet variable of parameter β_c/β, namely a random variable on the space of decreasing weights which has the same law as the decreasing rearrangement*

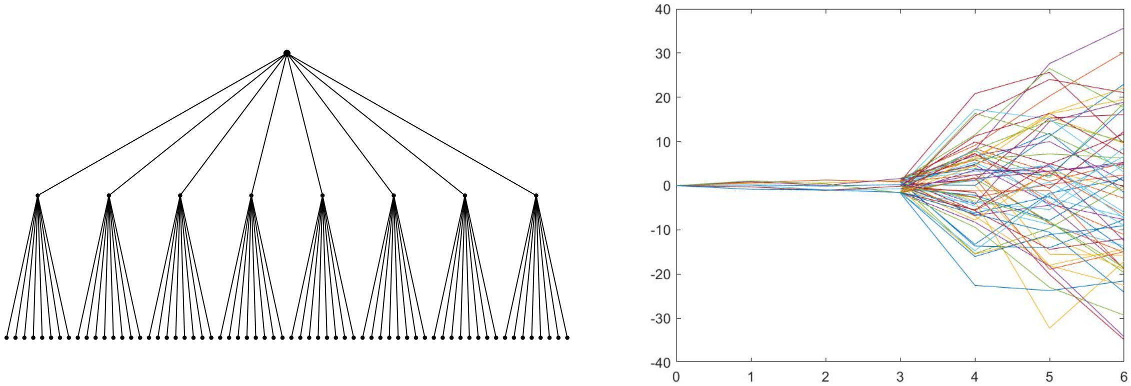
$$\left(\frac{\eta_k}{\sum_{j=1}^\infty \eta_j}, k \in \mathbb{N}\right)_{\downarrow}, \tag{0.3.69}$$

where $\{\eta_k\}_{k\in\mathbb{N}}$ denotes the atoms of a PPP on $(0,\infty)$ with intensity $t \mapsto (\beta_c/\beta)t^{(\beta_c/\beta)-1}$. Then, for any $s \in \mathbb{N}$ and any continuous function $h : [0,1]^{s(s-1)/2} \to \mathbb{R}$ of the overlaps of s points,

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h\Big((\rho(i_\ell, i_{\ell'}))_{1\le \ell,\ell'\le s}\Big)\Big] \\ &\quad = \begin{cases} h(\mathrm{Id}_s), & \text{if } \beta \le \beta_c, \\ E\left[\sum_{k_1,\dots,k_s\in\mathbb{N}} \xi_{k_1}\dots\xi_{k_s}\, h\Big(\Big(\mathbf{1}_{\{k_\ell = k_{\ell'}\}}\Big)_{1\le \ell,\ell'\le s}\Big)\right], & \text{if } \beta > \beta_c, \end{cases} \end{aligned} \tag{0.3.70}$$

where Id_s denotes the $s\times s$ identity matrix.

0.3.3.2. *The generalized random energy model (GREM)*



(A) GREM tree structure with 2 levels

(B) GREM simulation ($N = 2^6$) with 2 levels ($\lambda_1 = 1/2$, $\sigma_1 < \sigma_2$)

FIGURE 0.3.2. The generalized random energy model

In the physics literature, the GREM first appeared in Derrida (1985). For an introduction, the reader is referred to Chapter 10 in Bovier (2006).

Definition 0.3.11 (($\boldsymbol{\sigma}, \boldsymbol{\lambda}$)-GREM)**.** *Let N be the number of random variables in the field (the number of leaves in the tree) and let $M \in \mathbb{N}$ denote the number of levels in the tree. Also, define the following sets of parameters :*

$$\begin{aligned} \boldsymbol{\sigma} &\mathrel{\hat{=}} (\sigma_1, \sigma_2, \dots, \sigma_M) \in (0, \infty)^M, && \textit{(variance parameters)}, \\ \boldsymbol{\lambda} &\mathrel{\hat{=}} (\lambda_1, \lambda_2, \dots, \lambda_M) \in (0, 1]^M, && \textit{(scale parameters)}, \end{aligned} \tag{0.3.71}$$

where $0 \mathrel{\hat{=}} \lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_M \mathrel{\hat{=}} 1$. Define recursively a M-levels tree where, for every $i \leq M$, $N^{\lambda_i - \lambda_{i-1}}$ branches starting at the i-th level are attached to every vertices at the $(i-1)$-th level (see Figure 0.3.2a). The set of vertices at the M-th level is denoted by V_N. For all $v \in V_N$, denote by v_k the vertex at the k-th level on the shortest path from the origin of the tree to v. We assign i.i.d. Gaussian r.v.s $Z_{v_k} \sim \mathcal{N}(0, (\lambda_k - \lambda_{k-1})\sigma_k^2 \log N)$ to the k-th level branches of the tree (Z_{v_k} for the branch joining v_{k-1} and v_k). The field of interest is :

$$X_v \mathrel{\hat{=}} \sum_{k=1}^{M} Z_{v_k}, \quad v \in V_N. \tag{0.3.72}$$

The REM corresponds to the special case $M = 1$.

The parameters $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$ can be encoded simultaneously in the left-continuous step function

$$\sigma(s) \mathrel{\hat{=}} \sigma_1 \mathbf{1}_{\{0\}}(s) + \sum_{i=1}^{M} \sigma_i \mathbf{1}_{(\lambda_{i-1}, \lambda_i]}(s), \quad s \in [0,1]. \tag{0.3.73}$$

In the case of the GREM, it can be shown (see Section 1.2 of Arguin and Ouimet (2016) (Article 1) for an heuristic) that the answers to the questions of interest do not always take into account every single parameter in $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$ but only depend on the *effective parameters* that are encoded in the *concave hull* $\hat{\mathcal{J}}_{\sigma^2}$ of the *speed function* $\mathcal{J}_{\sigma^2}(s) \mathrel{\hat{=}} \int_0^s \sigma^2(r) dr$. More precisely, let $\bar{\sigma} : [0,1] \to \mathbb{R}$ be the unique left-continuous step function that satisfies $\mathcal{J}_{\bar{\sigma}^2}(s) = \hat{\mathcal{J}}_{\sigma^2}(s)$ for all $s \in (0,1]$. If the scales in $[0,1]$ where $\bar{\sigma}$ jumps are denoted by $0 \mathrel{\hat{=}} \lambda^0 < \lambda^1 < \cdots < \lambda^m \mathrel{\hat{=}} 1$, then we have the representation :

$$\bar{\sigma}(s) \mathrel{\hat{=}} \bar{\sigma}_1 \mathbf{1}_{\{0\}}(s) + \sum_{j=1}^{m} \bar{\sigma}_j \mathbf{1}_{(\lambda^{j-1}, \lambda^j]}(s), \quad s \in [0,1]. \tag{0.3.74}$$

where $\bar{\sigma}_j \mathrel{\hat{=}} \bar{\sigma}(\lambda^j)$ and $\{\lambda^j\}_{j=0}^m \subseteq \{\lambda_i\}_{i=0}^M$. See Figure 0.3.3 for an example of $\mathcal{J}_{\sigma^2}$ and $\hat{\mathcal{J}}_{\sigma^2}$.

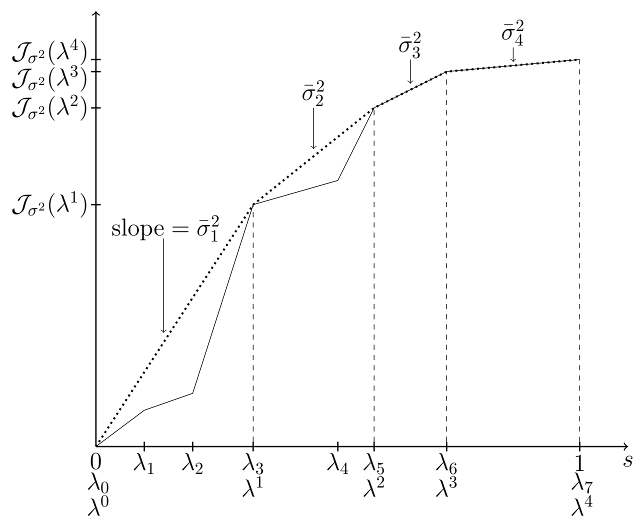


FIGURE 0.3.3. Example of $\mathcal{J}_{\sigma^2}$ (closed line) and $\hat{\mathcal{J}}_{\sigma^2}$ (dotted line) with 7 values for σ^2.

Below, we go through the same results that we proved for the REM (with the exception of $(Q4)$ and $(Q5)$) and show how the effective parameters influence the answers to the questions of interest. By using a second moment method and conditioning on the height of the maximal particle at every effective scale λ^j, we can answer $(Q1)$, $(Q2)$ and $(Q3)$.

Proposition 0.3.12 (First and second order of the maximum for the GREM)**.** *Let* $g \mathrel{\hat=} \sqrt{2}$*, then*

$$\lim_{N\to\infty} \frac{\max_{v\in V_N} X_v}{\log N} = g\int_0^1 \bar{\sigma}(s)ds \mathrel{\hat=} g\,\gamma^\star, \quad \text{in probability,} \tag{0.3.75}$$

for the first order, and

$$\lim_{N\to\infty} \frac{\max_{v\in V_N} X_v - g\gamma^\star \log N}{\log\log N} = -\sum_{j=1}^m \frac{1}{2}\cdot\frac{\bar{\sigma}_j}{g}, \quad \text{in probability,} \tag{0.3.76}$$

for the second order.

The number of γ-high points depends on *critical levels* defined by

$$\gamma^l \mathrel{\hat=} \int_0^1 \frac{\bar{\sigma}^2(s)}{\bar{\sigma}(s\wedge\lambda^l)}ds, \quad 1\leq l\leq m, \qquad \gamma^0 \mathrel{\hat=} 0. \tag{0.3.77}$$

For $\gamma \in (\gamma^{l-1}, \gamma^l]$, define

$$\mathcal{E}(\gamma) \mathrel{\hat=} (1-\lambda^{l-1}) - \frac{(\gamma - \int_0^{\lambda^{l-1}} \bar{\sigma}(s)ds)^2}{\int_{\lambda^{l-1}}^1 \bar{\sigma}^2(s)ds} \quad \text{and} \quad \mathcal{E}(0) \mathrel{\hat=} 1. \tag{0.3.78}$$

Proposition 0.3.13 (First order of the log-number of γ-high points for the GREM)**.** *Let* $g \mathrel{\hat=} \sqrt{2}$ *and* $\gamma \in (0,\gamma^\star)$*, then*

$$\lim_{N\to\infty} \frac{\log|\{v\in V_N : X_v \geq \gamma\, g\log N\}|}{\log N} = \mathcal{E}(\gamma), \quad \text{in probability.} \tag{0.3.79}$$

By applying Laplace's method using (0.3.75) and (0.3.79), we can find the limiting free energy and answer the first part of ($Q6$).

Proposition 0.3.14 (Limiting free energy of the GREM)**.** *For* $\beta > 0$*, we have*

$$\begin{aligned}\lim_{N\to\infty} f_N(\beta) &= \lim_{N\to\infty} \frac{1}{\log N}\log\sum_{v\in V_N} e^{\beta X_v} \\ &= \max_{\gamma\in[0,\gamma^\star]}\Big\{\beta\gamma + \mathcal{E}(\gamma)\Big\} = \sum_{j=1}^m f(\beta;\bar{\sigma}_j^2)\,(\lambda^j - \lambda^{j-1}),\end{aligned} \tag{0.3.80}$$

where the limit holds in probability and in L^1*, and where* $f(\beta;\sigma^2)$ *denotes the limiting free energy of the* σ*-REM from* (0.3.29)*.*

By perturbing the free energy between the scales $\alpha < \alpha'$, where $\lambda_{i-1} \leq \alpha < \alpha' \leq \lambda_i$ for some i, and by linking the derivative of the perturbed free energy with the two-overlap distribution using Gaussian integration by parts (as we did in the proof of Proposition 0.3.8), we can prove the following proposition, which answers the second part of ($Q6$).

Proposition 0.3.15 (Limiting two-overlap distribution of the GREM)**.** *Let $g \doteqdot \sqrt{2}$ and*

$$l_\beta \doteqdot \begin{cases} \min\{l \in \{1, \dots, m\} : \beta \leq \beta_c(\bar{\sigma}_l) \doteqdot \frac{g}{\bar{\sigma}_l}\}, & \text{if } \beta \leq \frac{g}{\bar{\sigma}_m}, \\ m+1, & \text{otherwise}, \end{cases} \tag{0.3.81}$$

and recall that $\rho(v, v') \doteqdot Corr(X_v, X_{v'})$. Then, for $\beta > 0$,

$$\lim_{N\to\infty} \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\Big[\mathbf{1}_{\{\rho(v,v')\leq q\}}\Big] = \begin{cases} 0, & \text{if } q < 0, \\ \frac{\beta_c(\bar{\sigma}_j)}{\beta}, & \text{if } q \in [x^{j-1}, x^j),\ j \leq l_\beta - 1, \\ 1, & \text{if } q \geq x^{l_\beta - 1}, \end{cases} \tag{0.3.82}$$

where $x^j \doteqdot \mathcal{J}_{\bar{\sigma}^2}(\lambda^j)/\mathcal{J}_{\bar{\sigma}^2}(1)$.

For $s \in \mathbb{N}$, let $(R_{\ell,\ell'}^N)_{1\leq \ell,\ell' \leq s} \doteqdot (\rho(v^\ell, v^{\ell'}))_{1\leq \ell,\ell'\leq s}$. (The variable N is there to remind us that the definition of the field $\{X_v\}_{v\in V_N}$ is N-dependent through the number of points and the variances/covariances.) By the representation theorem of Dovbysh and Sudakov (1982) for Gram-de Finetti matrices (for an accessible proof, see Panchenko (2010b)), we can show (since $[0,1]^{\mathbb{N}\times\mathbb{N}}$ being compact implies that the space of probability measures $\mathcal{M}_1([0,1]^{\mathbb{N}\times\mathbb{N}})$ is weakly compact) that there exists a subsequence $\{N_m\}_{m\in\mathbb{N}}$ converging to $+\infty$ such that for any $s \in \mathbb{N}$ and any continuous function $h : [0,1]^{s(s-1)/2} \to \mathbb{R}$, we have

$$\lim_{m\to\infty} \mathbb{E}\mathcal{G}_{\beta,N_m}^{\times\infty}\Big[h((R_{\ell,\ell'}^{N_m})_{1\leq\ell,\ell'\leq s})\Big] = E\mu_\beta^{\times\infty}\Big[h((R_{\ell,\ell'})_{1\leq\ell,\ell'\leq s})\Big], \tag{0.3.83}$$

where R is a random element of some probability space with measure P (and expectation E), generated by the random matrix of scalar products

$$(R_{\ell,\ell'})_{\ell,\ell'\in\mathbb{N}} = \Big((\rho_\ell, \rho_{\ell'})_{\mathcal{H}} + (1 - x^{l_\beta - 1})\mathbf{1}_{\{\ell=\ell'\}}\Big)_{\ell,\ell'\in\mathbb{N}}, \tag{0.3.84}$$

where $(\rho_\ell)_{\ell\in\mathbb{N}}$ is an i.i.d. sample from some random measure μ_β concentrated a.s. on the sphere of radius $\sqrt{x^{l_\beta - 1}}$ of a separable Hilbert space $\mathcal{H}$.

Here is the answer to the final part of $(Q6)$.

Proposition 0.3.16 (Extended Ghirlanda-Guerra identities in the limit for the GREM)**.** *Let $\beta > 0$ and let μ_β be a subsequential limit of $\{\mathcal{G}_{\beta,N}\}_{N\in\mathbb{N}}$ in the sense of (0.3.83). Then, for any $s \in \mathbb{N}$, any $k \in \{1,\dots,s\}$, and any functions $h : \{x^0, x^1, \dots, x^{l_\beta - 1}\}^{s(s-1)/2} \to \mathbb{R}$ and $g : \{x^0, x^1, \dots, x^{l_\beta - 1}\} \to \mathbb{R}$, we have*

$$\begin{aligned} E\mu_\beta^{\times(s+1)}\Big[g(R_{k,s+1})\, h((R_{i,i'})_{1\le i,i'\le s})\Big] &= \frac{1}{s} E\mu_\beta^{\times 2}\Big[g(R_{1,2})\Big] E\mu_\beta^{\times s}\Big[h((R_{i,i'})_{1\le i,i'\le s})\Big] \\ &\quad + \frac{1}{s}\sum_{\ell\neq k}^{s} E\mu_\beta^{\times s}\Big[g(R_{k,\ell})\, h((R_{i,i'})_{1\le i,i'\le s})\Big]. \end{aligned} \tag{0.3.85}$$

For a description of μ_β as a $(l_\beta - 1)$-levels tree of Poisson-Dirichlet processes (also called Ruelle probability cascade), see e.g. Corollary 7.2 in Ouimet (2017) (Article 2).

Here is the relevant literature for each question :

(Q1), (Q3), (Q4) and (Q5): See Theorem 1.1 in Bovier and Kurkova (2004a) for the convergence of the extremal process. The reader can also find a very good heuristic presentation in Section 2 of Kistler (2015).

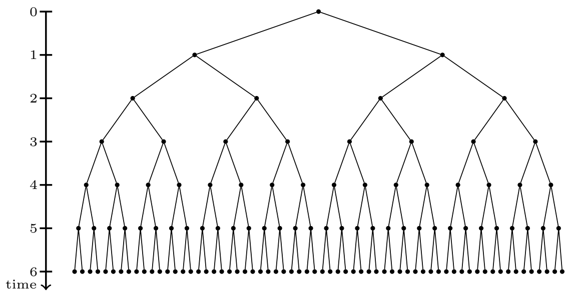
(Q2): I don't know any reference specific to the GREM, but it can easily be deduced from $(Q1)$ and the limiting free energy in $(Q6)$, since the problems are dual. Otherwise, see Arguin and Ouimet (2016) (Article 1) in the context of the IGFF and the solution of the related optimization problem in Appendix A of Ouimet (2014).

(Q6): For the limiting free energy, see Theorem 2.1 in Capocaccia et al. (1987). The rest (and more) is proved in Bovier and Kurkova (2004a) :

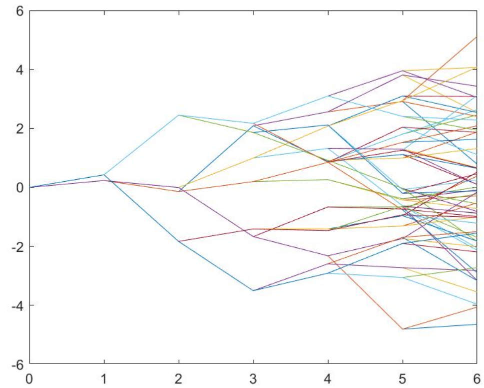
- Proposition 1.11 shows the limiting two-overlap distribution;
- Proposition 1.12 proves the Ghirlanda-Guerra identities in the limit;
- The joint distribution of the overlaps under the limiting mean Gibbs measure can be deduced from Theorem 1.9 and Theorem 1.13.

Remark 0.3.3. *Answers to $(Q1)$, $(Q2)$ and the limiting free energy (first part of $(Q6)$) for the 2-levels GREM with a random magnetic field (RMF) are provided in Persechino (2018). Deeper results can be found in Bovier and Klimovsky (2008) when the magnetic field is deterministic. For the REM with a RMF, the convergence of the extremal process and the convergence of the Gibbs weights are proved in Arguin and Kistler (2014).*

0.3.3.3. *The branching random walk (BRW)*



(A) BRW tree structure



(B) BRW simulation ($N = 2^6$)

FIGURE 0.3.4. The branching random walk

For an introduction to branching random walks, we refer the reader to Shi (2015a) and Zeitouni (2012). See also Athreya and Ney (1972) for a classic reference on branching processes. The branching random walk (BRW) is a limiting case of the GREM where the number of levels is proportional to the log-number of points in the field.

Definition 0.3.17 (BRW)**.** *The tree underlying the branching process can be described as follows. At time $k = 0$, there exists only one vertex o, called the origin, and we set $\mathbb{D}_0 \doteq \{o\}$. At time $k = 1$, there are 2 vertices and each of them is linked to o by an edge. Denote by $\mathbb{D}_1$ the set of vertices at time 1. At time $k = 2$, there are four vertices, two of which are linked to the first vertex in $\mathbb{D}_1$ and the other two are linked to the second vertex in $\mathbb{D}_1$. The set of vertices at time 2 is denoted by $\mathbb{D}_2$. The tree is defined iteratively in this manner up to time $k = n$, where $\mathbb{D}_k$ denotes the set of all vertices at time k and $|\mathbb{D}_k| = 2^k$. Figure 0.3.4a illustrates the tree structure. For all $v \in \mathbb{D}_n$, denote by v_k the ancestor of v at time k, namely the unique vertex in $\mathbb{D}_k$ that intersects the shortest path from o to v.*

Independent Gaussian random variables $Z_{v_k} \sim \mathcal{N}(0, \sigma^2)$ are assigned to each branch of the tree structure and the field of interest is $\{S_v(n)\}_{v\in\mathbb{D}_n}$, where $S_v(n)$ is the sum of the Gaussian r.v.s along the shortest path from o to v, namely

$$S_v(n) \doteq \sum_{k=1}^{n} Z_{v_k}, \quad v \in \mathbb{D}_n. \tag{0.3.86}$$

Remark 0.3.4. *There are $N = 2^n$ r.v.s in the field and they are log-correlated in the following sense. The branching time $\rho(u, v)$ is the latest time such that $u, v \in \mathbb{D}_n$ have the same ancestor. Formally,*

$$\rho(u, v) \mathrel{\hat{=}} \max\{k \in \{0, 1, \dots, n\} : u_k = v_k\}, \tag{0.3.87}$$

so that $d(u, v) \mathrel{\hat{=}} n - \rho(u, v)$ measures the proximity in time of u and v' latest common ancestor. The covariances of the field are then given by

$$Cov(S_u(n), S_v(n)) = \sigma^2 \rho(u, v) = \frac{\sigma^2}{\log 2} \cdot \frac{n - d(u, v)}{n} \log N. \tag{0.3.88}$$

As we proved in Section 0.3.3.1, by using a second-moment method, the second order correction for the maximum of the REM (N i.i.d. random variables with variance $\sigma^2 \log N$) is

$$-\frac{1}{2} \cdot \frac{\sigma}{g} \log \log N \tag{0.3.89}$$

for some model-specific constants σ and g (here : $g = \sqrt{2}$). When logarithmic correlations are introduced as in (0.3.88), then the second order correction becomes

$$-\frac{3}{2} \cdot \frac{\widetilde{\sigma}}{g} \log \log N \tag{0.3.90}$$

for some other model-specific constants $\widetilde{\sigma}$ and g. In the case of the BRW, we have

$$N = 2^n, \qquad g = \sqrt{2}, \qquad \widetilde{\sigma} = \frac{\sigma}{\sqrt{\log 2}}, \tag{0.3.91}$$

and the variances are

$$\sigma^2 n = \widetilde{\sigma}^2 \log N. \tag{0.3.92}$$

The additional $\log \log N$ factor between (0.3.89) and (0.3.90) comes from the fact that the naive version of the second-moment method (see the proof of Proposition 0.3.3) no longer works because of the correlations. Specifically, the second moment of the number of particles reaching above the level of the maximum is too large and thus the lower bound on the maximum is no longer “tight” when applying the Paley-Zygmund inequality.

When trying to guess the level of the maximum for a branching process at a given time, we have two factors to look at : the number of particles and how each particle fluctuates. It is a constant competition between the two. It seems at least intuitive that the level of the maximum should be achieved for a well chosen height $h(N)$ that satisfies

$$\mathbb{E}[\mathcal{N}] = O(1), \quad \text{with} \quad \mathcal{N} \mathrel{\hat{=}} \#\{v \in \mathbb{D}_n : S_v(n) \geq h(N)\}, \tag{0.3.93}$$

meaning that the number of particles reaching the level of the maximum at time n should be of the order of a constant (on average). If there was exponentially more particles than that, then there would be enough particles fluctuating near the maximal particle just before time n that we would be bound to find a particle that goes significantly higher than the anticipated level of the maximum. Now, in the case of the BRW, look at what happens if we naively evaluate $\mathbb{E}[\mathcal{N}]$ with $h(N) \mathrel{\hat{=}} \widetilde{\sigma} g \log N - \frac{3}{2} \cdot \frac{\widetilde{\sigma}}{g} \log\log N$. We get

$$\begin{aligned}
\mathbb{E}[\mathcal{N}] &\asymp \underbrace{N}_{\#\text{ of particles}} \times \underbrace{\frac{\sqrt{\widetilde{\sigma}^2 \log N}}{h(N)} \exp\left(-\frac{h(N)^2}{2\widetilde{\sigma}^2 \log N}\right)}_{\substack{\text{Gaussian tail estimate on each particle} \\ \text{(Lemma 11.1.1)}}} \\
&\asymp N \times (\log N)^{-1/2} N^{-1} (\log N)^{3/2} \\
&\asymp \log N,
\end{aligned} \tag{0.3.94}$$

which is not exactly what we want. However, if we modify $\mathcal{N}$ to

$$\mathcal{N}^\star \mathrel{\hat{=}} \#\left\{ v \in \mathbb{D}_n : \begin{array}{l} S_v(n) \geq h(N) \text{ and } t \mapsto S_v(t) \text{ stays below the barrier} \\ t \mapsto \underbrace{\tfrac{t}{n} \cdot h(N)}_{\substack{\text{linear path} \\ \text{leading to } h(N)}} + \underbrace{100 \tfrac{\widetilde{\sigma}}{g} \cdot \left\{1 + \log(1 + (t \wedge (n-t)))\right\}}_{\text{logarithmic barrier}} \end{array} \right\}$$

then, since the maximal particle behaves (approximately) as a discrete Brownian bridge around the linear path leading to the level of the maximum $h(N)$, we now have

$$\begin{aligned}
\mathbb{E}[\mathcal{N}^\star] &\asymp \underbrace{N}_{\#\text{ of particles}} \times \underbrace{\frac{\sqrt{\widetilde{\sigma}^2 \log N}}{h(N)} \exp\left(-\frac{h(N)^2}{2\widetilde{\sigma}^2 \log N}\right)}_{\substack{\text{Gaussian tail estimate on each particle} \\ \text{(Lemma 11.1.1)}}} \times \mathbb{P}\begin{pmatrix} \text{the particle} \\ \text{achieving the max} \\ \text{at time } n \text{ stays} \\ \text{below the barrier} \end{pmatrix} \\
&\asymp \log N \times \mathbb{P}\begin{pmatrix} \text{a discrete Brownian bridge stays} \\ \text{below a logarithmic barrier on} \\ \text{the time interval } [0,n] \end{pmatrix} \\
&\asymp 1.
\end{aligned} \tag{0.3.95}$$

The argument is illustrated in Figure 0.3.5. The estimate on the probability that a discrete Brownian bridge stays below a logarithmic barrier is part of a class of estimates known as *ballot theorems*, see the continuous analogue in Propositions 1 and 1' in Bramson (1978). In Lemmas 2.5 and 2.4 in Ouimet (2018c) (Article 3), Bramson's propositions are adapted to the discrete case using gambler's ruin estimates from Mogul'skiĭ (2009). For more info on ballot theorems, see e.g. Addario-Berry and Reed (2008) and Ford (2009).

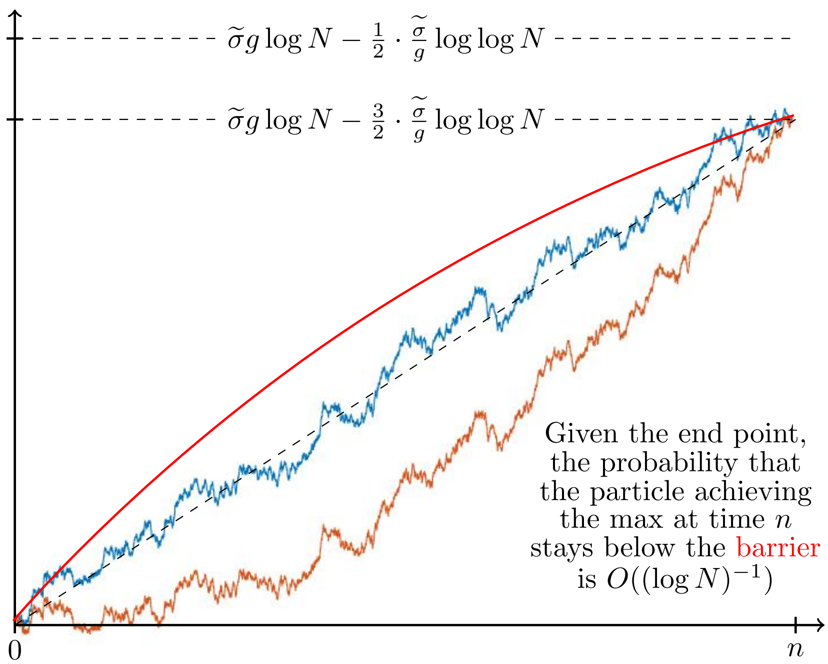


FIGURE 0.3.5. The maximal particle (blue) behaves as a Brownian bridge around the linear path $t \mapsto \frac{t}{n} h(N)$. If we impose that it stays below the red barrier, then there is a repulsion effect (orange).

The above heuristic is for the expectation of the number of particles reaching above the level of the maximum. When actually trying to bound the probability that the maximum of the field deviates from $h(N) = \widetilde{\sigma} g \log N - \frac{3}{2} \cdot \frac{\widetilde{\sigma}}{g} \log\log N$ by a factor $\varepsilon \log\log N$, then the idea is similar. Define

$$\begin{aligned} h_\varepsilon(N) &= \widetilde{\sigma} g \log N - (\tfrac{3}{2} - \varepsilon) \cdot \frac{\widetilde{\sigma}}{g} \log\log N, \\ b(t) &= 100 \frac{\widetilde{\sigma}}{g} \cdot \Big\{ 1 + \log(1 + (t \wedge (n - t))) \Big\}. \end{aligned} \tag{0.3.96}$$

If $S \circeq \{S(t)\}_{t=0}^n$ denotes a generic branch in the BRW, then a union bound yields

$$\mathbb{P}\begin{pmatrix} \max_{v \in \mathbb{D}_n} S_v(n) \\ \geq h_\varepsilon(N) \end{pmatrix} \leq \sum_{k=1}^n \underbrace{2^k}_{\substack{\#\text{ of particles} \\ \text{at time } k}} \times \underbrace{\mathbb{P}\begin{pmatrix} S \text{ crosses the barrier} \\ t \mapsto \frac{t}{n} \cdot h(N) + \frac{\varepsilon}{2} \cdot \frac{\widetilde{\sigma}}{g} \log\log N + b(t) \\ \text{for the first time at time } k \end{pmatrix}}_{(\star)}.$$

Using a Gaussian tail estimate (GTE) and the fact that $x \mapsto \frac{\log x}{x}$ is decreasing for $x \geq e$, we have

$$\begin{aligned}
(\star) &\ll \mathbb{P}\Big(S(k) \geq \tfrac{k}{n} \cdot h(N) + \frac{\varepsilon}{2} \cdot \frac{\widetilde{\sigma}}{g} \log\log N + b(k)\Big) \times \mathbb{P}\begin{pmatrix} \text{a Brownian bridge} \\ \text{on the time interval} \\ [0,k] \text{ stays below} \\ t \mapsto b(t) - \frac{t}{k} b(k) \end{pmatrix} \\
&\ll \underbrace{\frac{1}{\sqrt{k}}}_{\substack{\text{term in front of} \\ \text{the exponential} \\ \text{in the GTE}}} \cdot \underbrace{2^{-k}}_{\substack{\text{1st order} \\ \text{term squared} \\ \text{in the GTE}}} \cdot \underbrace{k^{3/2}}_{\substack{\text{cross between} \\ \text{1st order term and} \\ \text{2nd order term} \\ \text{in the GTE}}} \cdot \underbrace{n^{-\varepsilon/2}}_{\substack{\text{cross between} \\ \text{1st order term} \\ \text{and } \frac{\varepsilon}{2} \cdot \frac{\widetilde{\sigma}}{g} \log\log N \\ \text{in the GTE}}} \cdot \underbrace{\Big\{1 + (k \wedge (n-k))\Big\}^{-100}}_{\substack{\text{cross between} \\ \text{1st order term and} \\ \text{barrier term } b(k) \\ \text{in the GTE}}} \times \frac{1}{k}.
\end{aligned}$$

We deduce

$$\sum_{k=1}^n 2^k \times (\star) \ll n^{-\varepsilon/2} \sum_{k=1}^\infty \Big\{1 + (k \wedge (n-k))\Big\}^{-100} \ll n^{-\varepsilon/2}.$$

The argument is illustrated in Figure 0.3.6.

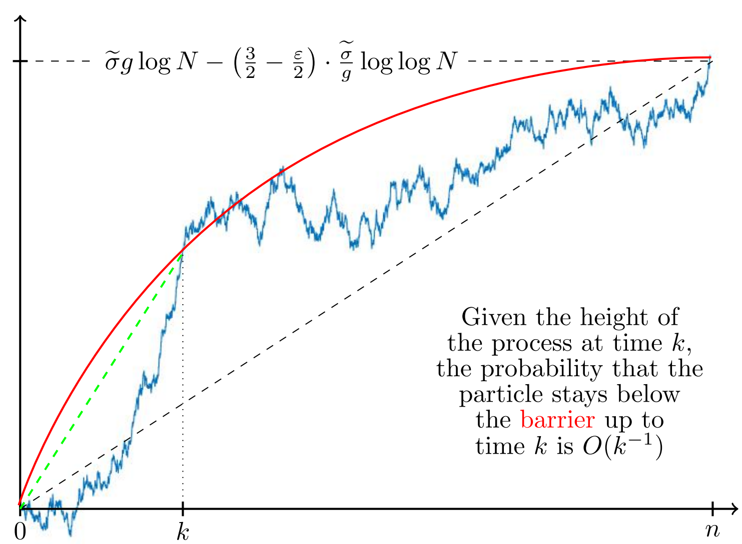


FIGURE 0.3.6. If there is a first time k at which the maximal particle crosses the red barrier, then given its height at time k, it behaves as a Brownian bridge around the green dashed line conditioned to stay below the red line.

For the lower bound, apply the Paley-Zygmund inequality (Lemma 11.1.2) by lowering the level to reach by $\varepsilon \log\log N$ and by adding barrier conditions in $\mathcal{N}$ (we need a lower barrier for the repulsion effect in Figure 0.3.5). The barriers will decrease the second moment enough that $\mathbb{E}[(\mathcal{N}_{\text{new}})^2] = (1 + o_\varepsilon(1))\, \mathbb{E}[\mathcal{N}_{\text{new}}]^2$ as $N \to \infty$, which will be sufficient to conclude. For the mathematical details of this heuristic, we refer to Kistler (2015).

Here are the answers to the questions of interest :

(Q1): See Theorem 4 in Biggins (1976). See also Hammersley (1974) and Kingman (1975) for earlier partial results.

(Q2): See Theorem 2 in Biggins (1977).

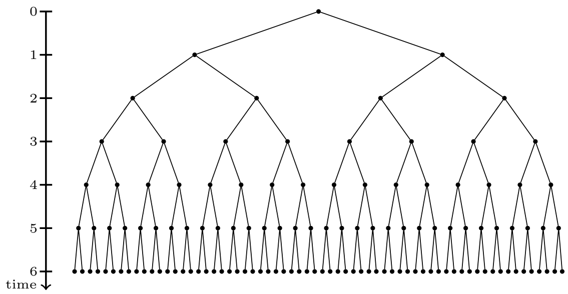
(Q3): See Theorem 1.2 in Hu and Shi (2009) for the second order of the maximum and almost-sure fluctuations (in the context of Galton-Watson trees). See also Theorem 2 in Roberts (2013) for a simple proof in the context of the BBM.

(Q4): See Theorem 3 in Addario-Berry and Reed (2009) for a proof under broader conditions (in the context of Galton-Watson trees). See also Mallein (2016) for a proof under near-optimal (even more general) integrability conditions.

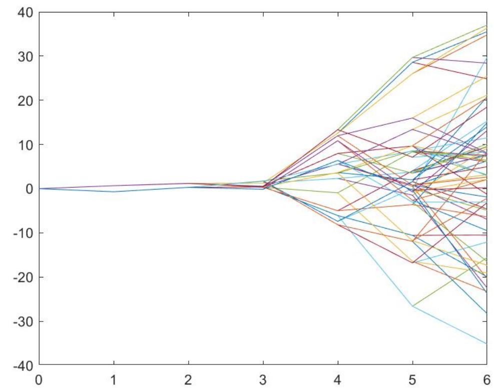
(Q5): For the convergence in law, see Aïdékon (2013), and Bramson et al. (2016b) for a simpler proof (under stronger assumptions). For the convergence of the extremal process, see Madaule (2017). The interpolation of the second order constant in the asymptotic expansion of the maximum between the REM and the BRW is explained in Kistler and Schmidt (2015).

(Q6): The proof for the limiting free energy originates from Theorem 1 in Chauvin and Rouault (1997). The proofs for the other results can easily be adapted from Arguin and Zindy (2014), which uses the Bovier-Kurkova technique : See Proposition 2.1 for the limiting free energy of the perturbed field, Theorem 1.4 for the limiting two-overlap distribution, and Theorem 1.5 for the joint distribution of the overlaps in the limit (Section 2.3 for the Ghirlanda-Guerra identities). A different proof is presented in Jagannath (2016), but it requires a much stronger control on the path of the maximal particle. A third approach is presented in Mallein (2018), where the weak limit of the (supercritical) Gibbs measure can be described as a consequence of the joint convergence of the extremal process with its genealogical information.

0.3.3.4. *The time-inhomogeneous branching random walk (IBRW)*



(A) IBRW tree structure



(B) IBRW simulation ($N = 2^6$) with 2 levels ($\lambda_1 = 1/2$ and $\sigma_1 < \sigma_2$)

FIGURE 0.3.7. The time-inhomogeneous branching random walk

The 2-levels version of this model was first introduced in Fang and Zeitouni (2012a).

Definition 0.3.18 (($\boldsymbol{\sigma}, \boldsymbol{\lambda}$)-IBRW)**.** *The tree structure of the* ($\boldsymbol{\sigma}, \boldsymbol{\lambda}$)*-IBRW is exactly the same as the one for the BRW. The only difference is that the variance of the branches in the tree changes macroscopically as time progresses. More precisely, let $M \in \mathbb{N}$ and define the sets of parameters :*

$$\begin{aligned} \boldsymbol{\sigma} &\mathrel{\hat{=}} (\sigma_1, \sigma_2, \dots, \sigma_M) \in (0, \infty)^M, && \textit{(variance parameters)}, \\ \boldsymbol{\lambda} &\mathrel{\hat{=}} (\lambda_1, \lambda_2, \dots, \lambda_M) \in (0, 1]^M, && \textit{(scale parameters)}, \end{aligned} \tag{0.3.97}$$

where $0 \mathrel{\hat{=}} \lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_M \mathrel{\hat{=}} 1$. Again, for all $v \in \mathbb{D}_n$, we denote by v_k the ancestor of v at time k, namely the unique vertex in $\mathbb{D}_k$ that intersects the shortest path from o to v. For $k \in (\lambda_{i-1} n, \lambda_i n]$, independent Gaussian r.v.s $Z_{v_k} \sim \mathcal{N}(0, \sigma_i^2)$ are assigned to each branch of the tree, and the field of interest is $\{S_v(n)\}_{v \in \mathbb{D}_n}$, where $S_v(n)$ is the sum of the Gaussian variables along the shortest path from o to v, namely

$$S_v(n) \mathrel{\hat{=}} \sum_{i=1}^{M} \sum_{k=\lambda_{i-1} n + 1}^{\lambda_i n} Z_{v_k}. \tag{0.3.98}$$

(For simplicity, we assume that $\lambda_i n \in \mathbb{N}$ for all i.)

Remark 0.3.5. *As for the GREM, the parameters $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$ can be encoded simultaneously in the left-continuous step function*

$$\sigma(s) \triangleq \sigma_1 \mathbf{1}_{\{0\}}(s) + \sum_{i=1}^{M} \sigma_i \mathbf{1}_{(\lambda_{i-1}, \lambda_i]}(s), \quad s \in (0,1]. \tag{0.3.99}$$

The covariances of the field are given by

$$Cov(S_u(n), S_v(n)) = \mathcal{J}_{\sigma^2}\Big(\frac{\rho(u,v)}{n}\Big) n = \frac{1}{\log 2} \cdot \mathcal{J}_{\sigma^2}\Big(\frac{n - d(u,v)}{n}\Big) \log N, \tag{0.3.100}$$

where $\rho(u,v)$ denotes the branching time as in (0.3.87), and $d(u,v) \triangleq n - \rho(u,v)$.

Here are the answers to the questions of interest :

(Q1), (Q3) and (Q4): The first proof appeared in Fang and Zeitouni (2012a) for the field with two levels (using the tightness of the maximum shifted by its median, from Theorem 1 in Fang (2012)). A result with much broader conditions was stated in Theorem 1.4 of Mallein (2015a), where the law of the increments isn't necessarily Gaussian. The proof uses a time-inhomogeneous version of the spinal decomposition for the BRW and thus rests crucially on the branching structure being *exact*. An alternate proof of the general case (but with Gaussian increments) was given independently in Ouimet (2014) and published later in Ouimet (2018c) (Article 3). The proof instead generalizes the argument of Fang and Zeitouni (2012a) and can be adapted to handle models with *approximate* branching structures.

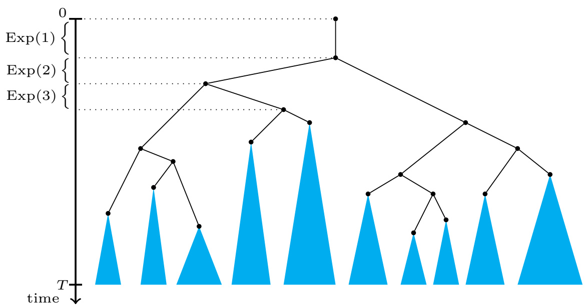
Answers are also provided for more general variance functions in Mallein (2015b) up to the second order of the maximum (and in some cases the tightness), under very general conditions on the increments. For similar results in the context of the variable speed BBM with (strictly) decreasing variances, see Fang and Zeitouni (2012b), Nolen et al. (2015) and Maillard and Zeitouni (2016).

(Q2): The statement and proof can be easily adapted from Theorem 1.3 in Arguin and Ouimet (2016) (Article 1).

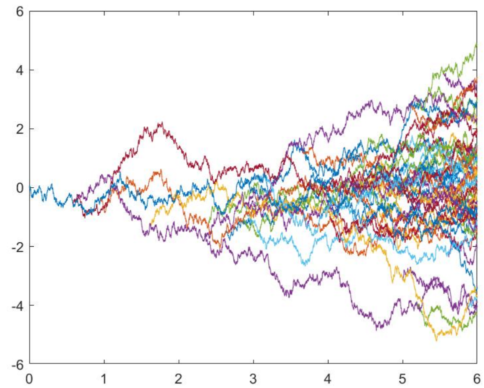
(Q5): The closest work on this question is done by Bovier and Hartung (2014, 2015, 2019) in the context of the BBM.

(Q6): The statements and proofs can be easily adapted from Ouimet (2017) (Article 2).

0.3.3.5. *The branching Brownian motion (BBM)*



(A) BBM tree structure



(B) BBM simulation ($n(T) = 2^6$)

FIGURE 0.3.8. The branching Brownian motion

For an introduction to the BBM, the reader is referred to Berestycki (2014), Bovier (2017) and Shi (2015b).

Definition 0.3.19 (BBM)**.** *The process is described as follows on the time interval $[0, T]$. At time $t = 0$, there is one particle, and it performs a Brownian motion until an exponential clock of mean 1 rings. When the clock rings, the particle splits into two particles, both of which have independent Brownian paths and both of which have independent exponential clocks with mean 1. After an exponential time of mean $1/2$, one of the two clocks rings and the corresponding particle splits into two. The now three particles have independent Brownian paths and independent exponential clocks of mean 1 (by the memoryless property), etc ... The particles continue to split and move in this manner up to time T.*

Let $n(t)$ be the number of particles in the underlying tree structure at time $t \in [0, T]$. It can be shown that

$$\mathbb{E}[n(t)] = e^t. \tag{0.3.101}$$

If we label the leaves of the tree at time t by $i_1(t)$, $i_2(t)$, ..., $i_{n(t)}(t)$, then the collection of correlated Brownian paths is $\{\{X_k^T(t)\}_{k=1}^{n(t)}\}_{t\in[0,T]}$ and the field of interest corresponds to the heights of the leaves at time $t = T$:

$$X_k^T(T), \quad k \in \{1, 2, \ldots, n(T)\}. \tag{0.3.102}$$

Remark 0.3.6. *The most recent common ancestor of $i_k(T)$ and $i_\ell(T)$ is given by the branching time :*

$$\rho(i_k(T), i_\ell(T)) \mathrel{\hat{=}} \sup\{t \leq T : i_k(t) = i_\ell(t)\}. \tag{0.3.103}$$

Thus, the covariances of the field are

$$Cov(X_k^T, X_\ell^T) = \rho(i_k(T), i_\ell(T)). \tag{0.3.104}$$

Here are the answers to the questions of interest :

(Q1): See Kolmogorov et al. (1937).

(Q2): See Theorem 1.1 in Aïdékon et al. (2017).

(Q3) and (Q4): The asymptotics of the median of the maximum are given up to an $O(1)$ error in Theorem 1 of Bramson (1978) (see Theorem 1 in Roberts (2013) for a simpler proof). The $O(1)$ term was improved to "Const. $+\, o(1)$" in Bramson (1983) using the Feynman-Kac formula.

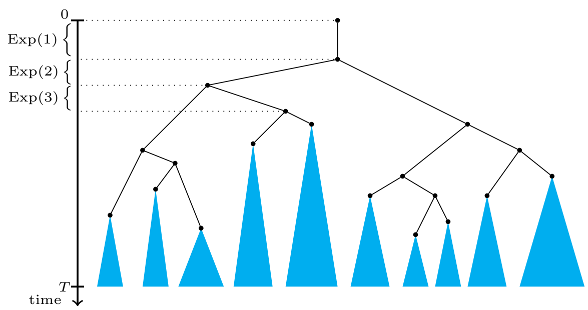
(Q5): For the convergence in law, see Theorem 1 in Lalley and Sellke (1987). For the convergence of the extremal process, see Theorem 2.1 in Arguin et al. (2013) (see also Arguin et al. (2011, 2012) for earlier relevant work). Around the same time, the convergence of the extremal process was also shown independently by Aïdékon et al. (2013), building on earlier results from Aïdékon (2013).

(Q6): The proofs can be adapted from Arguin and Zindy (2014). The proofs can also be adapted from Jagannath (2016), since we have a strong control on the path of the maximal particle in the case of the BBM (as in the case of the BRW).

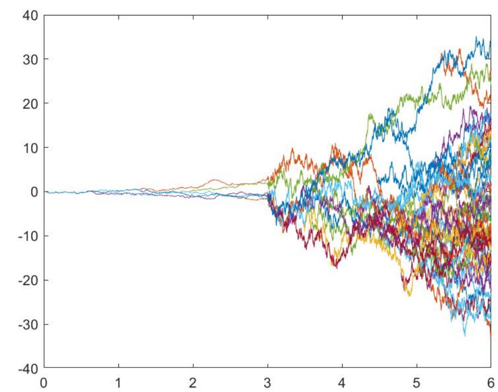
Remark 0.3.7. *In the case of the complex BBM (where we allow for arbitrary correlations between the real and imaginary parts of the energies), the fluctuations of the partition function and the phase diagram of the limiting free energy are described in Hartung and Klimovsky (2015, 2018). The phase diagram coincides with the one for the complex REM, previously studied in Kabluchko and Klimovsky (2014a). For analogous results in the context of the complex GREM, see Kabluchko and Klimovsky (2014b).*

Remark 0.3.8. *Answers to (Q1), (Q3) and (Q4) in the context of the d-dimensional BBM are provided in Theorem 1.1 of Mallein (2015c).*

0.3.3.6. *The variable speed branching Brownian motion (VSBBM)*



(A) VSBBM tree structure



(B) VSBBM simulation ($n(T) = 2^6$) with 2 levels ($\sigma_1 < \sigma_2$)

FIGURE 0.3.9. The variable speed branching Brownian motion

A similar model was originally considered in Derrida and Spohn (1988). For an introduction to the VSBBM, the reader is referred to Bovier and Hartung (2014, 2015, 2019) and Chapter 9 in Bovier (2017).

Definition 0.3.20 (VSBBM). *The tree structure is the same as the one described in Definition 0.3.19 for the BBM. The only difference is that the variance coefficient of the Brownian paths' increments are inhomogeneous in time. More specifically, take a "regular enough" variance function* $\sigma : [0,1] \to (0,\infty)$, *and consider the collection of time-changed Brownian paths* $\{\{X_k^T(t)\}_{k=1}^{n(t)}\}_{t\in[0,T]}$ *such that*

$$Var(X_k^T(t)) = \int_0^{t/T} \sigma^2(s)ds \cdot T \mathrel{\hat{=}} \mathcal{J}_{\sigma^2}\Big(\frac{t}{T}\Big) T, \quad k \leq n(t), \ t \in [0,T]. \tag{0.3.105}$$

Remark 0.3.9. *With the above definition, the covariances of the field are given by*

$$Cov(X_k^T, X_\ell^T) = \mathcal{J}_{\sigma^2}\Big(\frac{\rho(i_k(T), i_\ell(T))}{T}\Big) T, \tag{0.3.106}$$

where $\rho(i_k(T), i_\ell(T))$ *denotes the branching time as in* (0.3.103).

Here are the answers to the questions of interest :

(Q1), (Q3), (Q4) and (Q5): For the limiting law of the recentered maximum and the convergence of the extremal process when the VSBBM has two macroscopic levels, see Theorem 1.2 and Theorem 1.3 in Bovier and Hartung (2014). For the limiting law of the recentered maximum and the convergence of the extremal process when the VSBBM has one effective level and the speed function $\mathcal{J}_{\sigma^2}$ stays below the straight line $x \mapsto x$ (with only mild technical assumptions on the speed function), see Theorem 1.2 in Bovier and Hartung (2015). For the same result under stronger assumptions on the speed function, see Theorem 9.20 in Bovier (2017).

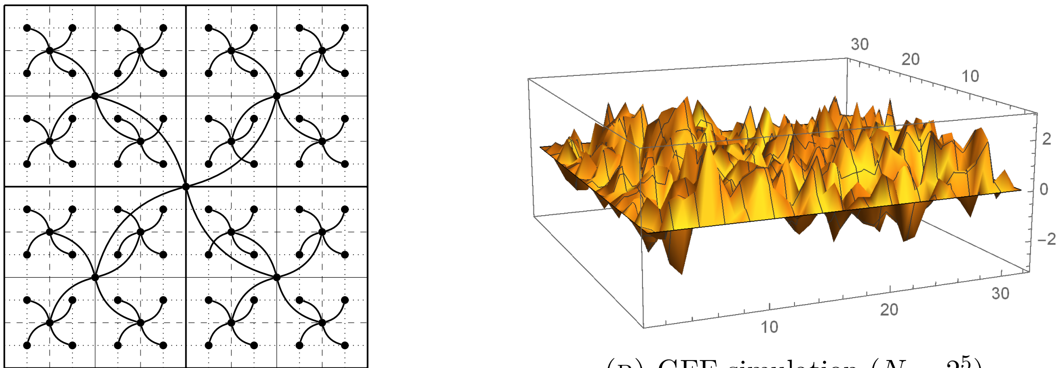
The interpolation of the second order constant in the asymptotic expansion of the maximum is explained in Bovier and Hartung (2019). (This is similar to the results of Kistler and Schmidt (2015) in the context of the REM.)

In the case of (strictly) decreasing variance functions, see Fang and Zeitouni (2012b) and Nolen et al. (2015) (the latter uses PDE techniques) for the second order of the maximum, and see Maillard and Zeitouni (2016) for the tightness and the convergence in law of the recentered maximum.

(Q2): For the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-VSBBM (meaning for the VSBBM where the variance function σ is the step function from (0.3.99)), the statement and proof can be adapted from Theorem 1.3 in Arguin and Ouimet (2016) (Article 1).

(Q6): For the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-VSBBM, the statements and proofs can be adapted from Ouimet (2017) (Article 2). For more general variance functions σ, see Bovier and Kurkova (2004b) in the context of the GREM model with a continuum of hierarchies.

0.3.3.7. *The Gaussian free field (GFF)*



(A) GFF tree structure

(B) GFF simulation ($N = 2^5$)

FIGURE 0.3.10. The Gaussian free field

For an introduction to the Gaussian free field, the reader is referred to Biskup (2018), Sheffield (2007), Berestycki (2016), Zeitouni (2017), and Section 10 in Chatterjee (2008).

Definition 0.3.21 (GFF). *Let $(W_k)_{k\geq 0}$ be a simple random walk (SRW) starting at $u \in \mathbb{Z}^2$ with law $\mathscr{P}_u$. For any finite box $B \subseteq \mathbb{Z}^2$, the (discrete) GFF on B is a centered Gaussian field $\phi \stackrel{\circ}{=} \{\phi_v\}_{v\in B}$ with covariance matrix*

$$G_B(u,v) \stackrel{\circ}{=} \frac{\pi}{2} \cdot \mathscr{E}_u\left[\sum_{k=0}^{\tau_{\partial B}-1} \mathbf{1}_{\{W_k=v\}}\right], \quad u,v \in B, \tag{0.3.107}$$

where $\tau_{\partial B}$ is the first hitting time of $(W_k)_{k\geq 0}$ on the boundary of B, $\partial B \stackrel{\circ}{=} \{v \in B \,|\, \exists z \notin B \text{ such that } \|v - z\|_2 \leq 1\}$, and $\|\cdot\|_2$ denotes the Euclidean distance in $\mathbb{Z}^2$. With this definition, B contains its boundary. Also, note that ϕ is identically zero on ∂B; this is the Dirichlet boundary condition. The GFF of interest to us is

$$\phi_v, \quad v \in V_N \stackrel{\circ}{=} \{0,1,\dots,N\}^2. \tag{0.3.108}$$

Remark 0.3.10. *The covariance function (0.3.107) is simply the (renormalized) Green function of the discrete Laplacian restricted to functions that are 0 outside $B\backslash\partial B$; it satisfies the following boundary value problem on B : for $x \in B\backslash\partial B$,*

$$\begin{aligned} \Delta G_B(x,y) &= \tfrac{\pi}{2} \cdot \mathbf{1}_{\{x=y\}}, & y \in B\backslash\partial B, \\ G_B(x,y) &= 0, & y \in \partial B. \end{aligned} \tag{0.3.109}$$

Remark 0.3.11. *For $\lambda \in (0,1)$ and $v = (v_1, v_2) \in V_N$, consider the closed neighborhood $[v]_\lambda$ in V_N consisting of the square box of width $N^{1-\lambda}$ centered at v that has been cut off by ∂V_N :*

$$[v]_\lambda \mathrel{\hat{=}} \Big((v_1, v_2) + \Big[-\frac{1}{2}N^{1-\lambda}, \frac{1}{2}N^{1-\lambda}\Big]^2\Big) \bigcap V_N.$$

By convention, we define $[v]_0 \mathrel{\hat{=}} V_N$ and $[v]_1 \mathrel{\hat{=}} \{v\}$. Let $\mathcal{F}_{\partial [v]_\lambda}$ be the σ-algebra generated by the r.v.s on the boundary of the box $[v]_\lambda$. It can be shown that the r.v.s $\mathbb{E}[\phi_v \,|\, \mathcal{F}_{\partial [v]_\lambda}]$, where the v's are the nodes at scale λ in Figure 0.3.10a (which we call the representatives at scale λ), play the same role as the leaves of the BRW at time λn.

To see the tree structure, define the branching scale between v and v' by

$$\rho(v, v') \mathrel{\hat{=}} \max\{\lambda \in [0,1] : [v]_\lambda \cap [v']_\lambda \neq \emptyset\}. \tag{0.3.110}$$

This is the largest λ for which the two neighborhoods $[v]_\lambda$ and $[v']_\lambda$ intersect. We always have by definition that $\|v - v'\|_2$ is of order $N^{1-\rho(v,v')}$. The branching scale plays the same role as the branching time (normalized to lie in $[0,1]$) in the BRW context. In fact, it can be shown that, for $v, v' \in V_N^\delta \mathrel{\hat{=}} \{v \in V_N : \min_{z \in \partial V_N} \|v - z\|_2 \geq \delta N\}$ with $\delta \in (0, \frac{1}{2})$,

$$Cov(\phi_v, \phi_{v'}) = G_{V_N}(v, v') = \log N^{1-\rho(v,v')} + O(1), \tag{0.3.111}$$

by using estimates on the potential kernel of the SRW.

An important property of the GFF is the Markov property (as a random field), which means that the value of the field inside a neighborhood is independent of the field outside given the boundary, see e.g. Dynkin (1980). It is a consequence of the strong Markov property of the SRW. In particular, for the neighborhood $[v]_\lambda$,

$$\phi_v(\lambda) \mathrel{\hat{=}} \mathbb{E}\Big[\phi_v \,|\, \mathcal{F}_{\partial [v]_\lambda \cup [v]_\lambda^c}\Big] = \mathbb{E}\Big[\phi_v \,|\, \mathcal{F}_{\partial [v]_\lambda}\Big]. \tag{0.3.112}$$

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

$$\phi_v(\lambda') - \phi_v(\lambda) \quad \text{is independent of} \quad \phi_{v'}(\mu') - \phi_{v'}(\mu). \tag{0.3.113}$$

This is because the shell $[v]_\lambda \cap [v]_{\lambda'}^c$ does not intersect the shell $[v']_\mu \cap [v']_{\mu'}^c$ in both cases. Thus, as in the BRW setting, the increments past the branching scale are independent.

The difference happens before the branching scale, where the increments are nearly perfectly correlated, and around the branching scale, where the decorrelation happens more smoothly.

Here are the answers to the questions of interest :

(Q1): See Theorem 2 in Bolthausen et al. (2001).

(Q2): See Theorem 1.3 in Daviaud (2006). For much stronger results, see Biskup and Louidor (2016b). In the continuous setting, see Theorem 1.1 and Theorem 1.2 in Hu et al. (2010) for results related to *thick points*.

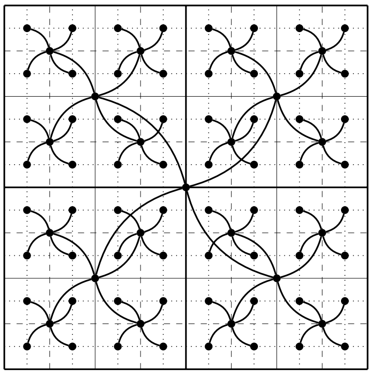
(Q3) and (Q4): A preliminary condition for the tightness was developed in Bolthausen et al. (2011) and proved in Bramson and Zeitouni (2012). Precise estimates for the tail probabilities of the recentered maximum were then found in Ding (2013) by bootstrapping estimates in smaller boxes using the FKG inequality (the sprinkling method). The estimates were later improved in Ding and Zeitouni (2014).

(Q5): See Theorem 1.1 and Theorem 2.5 in Bramson et al. (2016a); their strategy builds in part on earlier work from Ding and Zeitouni (2014). For a generalization of these two theorems to a large class of log-correlated Gaussian fiels, see Theorem 1.3 and Theorem 1.4 in Ding et al. (2017). For the convergence of the extremal process, see Theorem 1.1 in Biskup and Louidor (2016a) (local extrema) and Section 2 in Biskup and Louidor (2018).

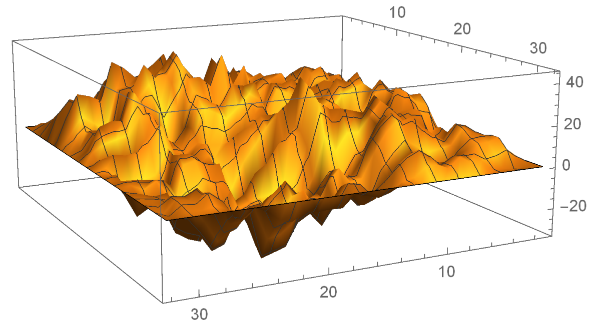
(Q6): In Arguin and Zindy (2015), see Theorem 2.1 for the limiting free energy of the perturbed field, Theorem 1.1 for the limiting two-overlap distribution, and Theorem 1.2 for the joint distribution of the overlaps in the limit. For the convergence in law of the Gibbs measure to a Poisson-Dirichlet process at low temperature, see also Corollary 2.7 in Biskup and Louidor (2018).

Remark 0.3.12. *The definition of the GFF extends naturally to other dimensions. Note however that $d = 2$ is the most interesting case (the critical dimension) from the point of view of extreme value theory (this is linked to the fact that the SRW is transient for $d \geq 3$). Results of interest for the GFF in dimensions $d \geq 3$ can be found for example in Cipriani and Hazra (2015, 2017), Chiarini et al. (2016, 2015), and Chen (2018a). See also the results on the closely related log-REM model in Carpentier and Le Doussal (2001), Fyodorov and Bouchaud (2008a,b), Fyodorov et al. (2009), and Cao et al. (2016).*

0.3.3.8. *The scale-inhomogeneous Gaussian free field (IGFF)*



(A) IGFF tree structure



(B) IGFF simulation ($N = 2^5$) with 2 levels ($\lambda_1 = 2/5$ and $\sigma_1 > \sigma_2$)

FIGURE 0.3.11. The scale-inhomogeneous Gaussian free field

The 2-levels IGFF first appeared in Arguin and Zindy (2015) and the general IGFF was introduced in Ouimet (2014) and Arguin and Ouimet (2016) (Article 1).

Definition 0.3.22 (($\boldsymbol{\sigma}, \boldsymbol{\lambda}$)-IGFF)**.** *Let $\phi \mathrel{\hat{=}} \{\phi_v\}_{v \in V_N}$ be the GFF on V_N. Take $M \in \mathbb{N}$ and $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$ as in (0.3.97). The $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-IGFF on V_N is a Gaussian field $\psi \mathrel{\hat{=}} \{\psi_v\}_{v \in V_N}$ defined by*

$$\psi_v \mathrel{\hat{=}} \sum_{i=1}^{M} \sigma_i \Big(\phi_v(\lambda_i) - \phi_v(\lambda_{i-1}) \Big). \tag{0.3.114}$$

Remark 0.3.13. *The field ψ can be seen as a martingale-transform of $(\phi_v(\lambda))_{\lambda \in [0,1]}$ applied simultaneously for every $v \in V_N$.*

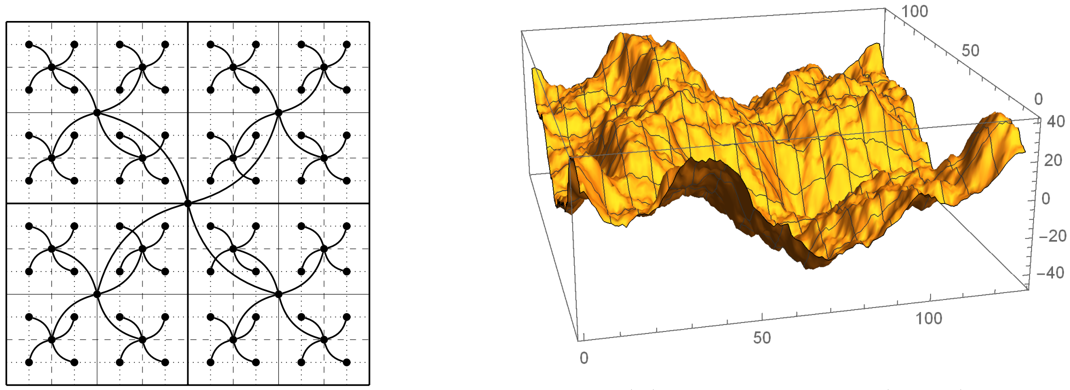
Here are the answers to the questions of interest :

(Q1) and (Q2): See Theorems 1.2 and 1.3 in Arguin and Ouimet (2016) (Article 1). In the continuous setting, see Theorem 7 in Chen (2018b) for a result related to the Hausdorff dimension of *f-steep points* under a sphere averaging regularization.

(Q3), (Q4) and (Q5): This is still open. Proving $(Q3)$ on V_N^δ should be feasible, but the difficulty of proving $(Q3)$ on V_N stems from the decay of the variances near ∂V_N. This is difficult to handle in the inhomogeneous case as each end of the increments of the field pulls the variances in opposite directions (see below (0.4.11)).

(Q6): In Ouimet (2017) (Article 2), see Theorem 6.1 for the limiting free energy, Theorem 6.3 for the limiting two-overlap distribution, and Corollary 7.2 for the identification of the limiting Gibbs measure as a cascade of Poisson-Dirichlet processes.

0.3.3.9. *The membrane model (MM)*



(A) MM tree structure ($d = 2$)

(B) MM simulation ($d = 2$) without the boundary condition

FIGURE 0.3.12. The membrane model

This model was first studied probabilistically by Sakagawa (2003). For an introduction, we refer the reader to Kurt (2009).

Definition 0.3.23 (MM)**.** *For any finite box $B \subseteq \mathbb{Z}^d$, the MM on B is a centered Gaussian field $\phi \circeq \{\phi_v\}_{v \in B}$ with covariance matrix given by the Green function of the discrete bi-Laplacian restricted to functions that are 0 outside $B \backslash \partial_2 B$, where $\partial_2 B \circeq \{v \in B \,|\, \exists z \notin B \text{ such that } \|v - z\|_2 \leq 2\}$. It satisfies the boundary value problem on B : for $x \in B \backslash \partial_2 B$,*

$$\begin{aligned} \Delta^2 G_B(x, y) &= \mathbf{1}_{\{x=y\}}, & y &\in B \backslash \partial_2 B, \\ G_B(x, y) &= 0, & y &\in \partial_2 B. \end{aligned} \tag{0.3.115}$$

The dimension of interest to us (*the critical dimension*) is $d = 4$. The covariance has no random walk representation, but it can be shown that, for $v, v' \in V_N^\delta$ with $\delta \in (0, \frac{1}{2})$,

$$\mathrm{Cov}(\phi_v, \phi_{v'}) = \frac{8}{\pi^2} \log N^{1 - \rho(v, v')} + O(1), \quad \text{similarly to the GFF in (0.3.111).}$$

Here are the answers to the questions of interest :

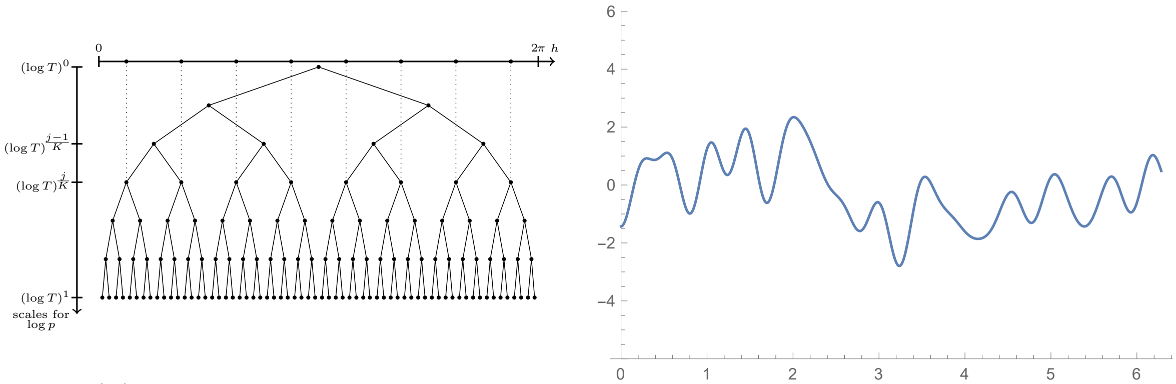
(Q1) and (Q2): See Theorem 1.2 in Kurt (2009), and Theorem 1.4 in Cipriani (2013).

(Q3) and (Q4): See Theorem 1.2 in Ding et al. (2017).

(Q5): See Theorem 1.1 in Schweiger (2019), which uses results from Ding et al. (2017).

(Q6): This is open, although see Sakagawa (2012, 2018) for results related to the free energy, and Cipriani et al. (2019) for the scaling limit of the membrane.

0.3.3.10. *The randomized log-modulus of the Riemann zeta function (RLM-RZF)*



(A) RLM-RZF tree structure ($K = \frac{1}{\log 2} \log\log T$ levels)

(B) RLM-RZF simulation ($T = 100\,000$-th prime)

Figure 0.3.13. The randomized log-modulus of the Riemann zeta function

This model was first defined by Harper (2013). For an introduction, we refer the reader to Arguin et al. (2017b) and Arguin (2017).

Definition 0.3.24 (RLM-RZF on the critical line)**.** *Let $(U_p,\ p$ primes$)$ be an i.i.d. sequence of uniform random variables on the unit circle in $\mathbb{C}$. The random field of interest is*

$$X_h \mathrel{\hat{=}} \sum_{p \le T} \frac{Re(U_p\, p^{-ih})}{p^{1/2}}, \quad h \in [0, 2\pi]. \tag{0.3.116}$$

(A sum over the variable p always denotes a sum over primes.)

Remark 0.3.14. *If $\tau \sim Uniform(T, 2T)$, the field $\{X_h\}_{h\in[0,2\pi]}$ is a good model to study the large values of $(\log|\zeta(\frac{1}{2} + i\tau + ih)|,\ h \in [0, 2\pi])$ for the following reason. Proposition 1 in Harper (2013) proves that, assuming the Riemann hypothesis, and for T large enough, there exists a set $B \subseteq [T, T + 2\pi]$, of Lebesgue measure at least 1.99π, such that*

$$\log|\zeta(\tfrac{1}{2} + it)| = Re\left(\sum_{p \le T} \frac{1}{p^{1/2+it}} \frac{\log(T/p)}{\log T}\right) + O(1), \quad t \in B. \tag{0.3.117}$$

If we ignore the smoothing term $\log(T/p)/\log T$ and note that the process $(p^{-i\tau},\ p$ primes$)$ converges, as $T \to \infty$ (in the sense of convergence of its finite-dimensional distributions), to a sequence of independent random variables distributed uniformly on the unit circle (by *computing the moments), then the model* (0.3.116) *follows. For more information, see Section 1.1 in* Arguin et al. (2017b)*.*

As for the BRW, we can show the logarithmic decay of the correlations for this model. Since $\mathrm{Re}(z) = (z + \overline{z})/2$, $\mathbb{E}[U_p^2] = \mathbb{E}[(\overline{U_p})^2] = 0$ and $\mathbb{E}[U_p \overline{U_p}] = 1$, it is easily shown from (0.3.116) that

$$\mathbb{E}[X_h X_{h'}] = \sum_{p \leq T} \frac{1}{2p} \cos(|h - h'| \log p), \quad h, h' \in [0,1]. \tag{0.3.118}$$

Using the prime number theorem (Montgomery and Vaughan, 2007, Theorem 6.9) which states that

$$\#\{p \text{ prime} : p \leq x\} = \int_2^x \frac{1}{\log u} du + R(x), \tag{0.3.119}$$

where $R(x) = O(x e^{-c\sqrt{\log x}})$, uniformly for $x \geq 2$, we can show (see, for example, page 20 of Appendix A in Harper (2013)) that

$$\mathrm{Corr}(X_h, X_{h'}) = \frac{\frac{1}{2}\log\Big((\log T) \wedge |h - h'|^{-1}\Big)}{\frac{1}{2}\log\log T} + O\Big((\log\log T)^{-1}\Big). \tag{0.3.120}$$

The analogy with BRWs is thus recovered when we break the sum (0.3.116) between the scales $\exp((\log T)^{\frac{j}{K}})$, $j \in \{0, 1, \dots, K\}$, with $K = \frac{1}{\log 2}\log\log T$ levels. This is illustrated in Figure 0.3.13a.

Here are the answers to the questions of interest :

(Q1): See Proposition 2 in Harper (2013).

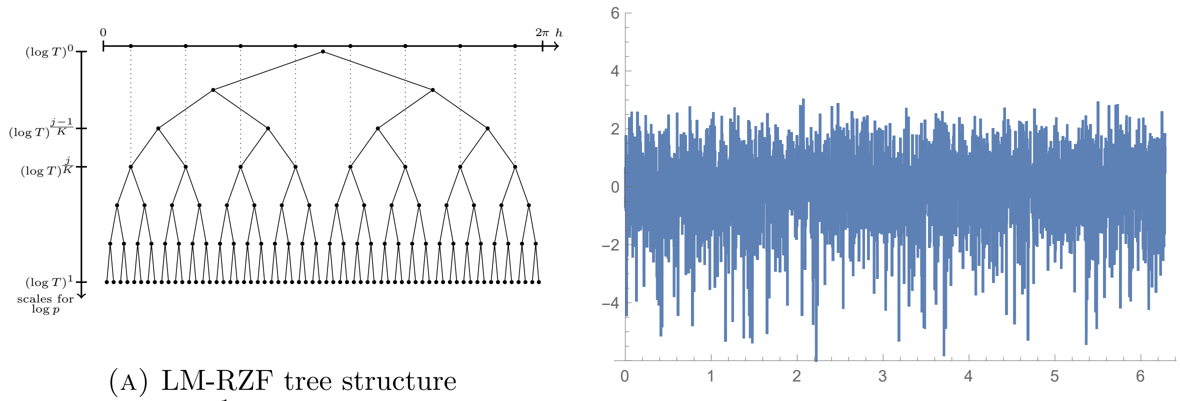
(Q2): See Lemma 12 in Arguin and Tai (2018) ($u = 0$), and see Arguin et al. (2019b) for the almost-sure convergence of the normalized Lebesgue measure of high points.

(Q3): See Theorem 1.2 in Arguin et al. (2017b).

(Q4) and (Q5): This is still open, although see Conjecture 10.1.4. A preliminary step for ($Q4$) has been made in Arguin and Ouimet (2019) (Article 5), where we show large deviations and continuity estimates for the derivative of the field.

(Q6): See Proposition 4 and Theorem 1 in Arguin and Tai (2018) for the limiting free energy of the perturbed field and the limiting two-overlap distribution, respectively. See Theorem 3.3 in Ouimet (2018b) (Article 4) for the joint distribution of the overlaps in the limit (Theorem 5.8 for the extended Ghirlanda-Guerra identities).

0.3.3.11. *The log-modulus of the Riemann zeta function (LM-RZF)*



(A) LM-RZF tree structure ($K = \frac{1}{\log 2} \log\log T$ levels)

(B) LM-RZF simulation ($T = 100\,000$)

FIGURE 0.3.14. The log-modulus of the Riemann zeta function

For good books on the properties of the Riemann zeta function, the reader is referred to Titchmarsh (1986), Laurinčikas (1996) and Ivić (2003).

Definition 0.3.25 (LM-RZF on the critical line)**.** *Let $s \in \mathbb{C}$. The Riemann zeta function is defined on $Re(s) > 1$ by the following absolutely convergent series and the related Euler product :*

$$\zeta(s) \triangleq \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ primes}} \left(1 - \frac{1}{p^s}\right)^{-1}. \tag{0.3.121}$$

It admits an analytic continuation to $\mathbb{C}\backslash\{1\}$, where $s = 1$ is a simple pole. For instance, the following representation is valid on the critical strip $Re(s) \in (0,1)$ (see e.g. Theorem 3.2 in De Koninck and Luca (2012)) :

$$\zeta(s) = \frac{s}{s-1} - s \int_1^{\infty} \frac{x - \lfloor x \rfloor}{x^{s+1}} dx. \tag{0.3.122}$$

If $\tau \sim Uniform(T, 2T)$, then the field of interest to us is

$$h \mapsto \log|\zeta(\tfrac{1}{2} + i\tau + ih)|, \quad h \in [0, 2\pi]. \tag{0.3.123}$$

Remark 0.3.15. *There are no obvious reasons for the tails of $\log|\zeta(\frac{1}{2} + i\tau)|$ to be Gaussian, but an influential theorem of Selberg (see Selberg (1946, 1992), see Radziwiłł and Soundararajan (2017) for a simpler proof, and see Bourgade (2010) for a multidimensional extension) shows, quite surprisingly, that the tails are approximately Gaussian when T is* *large. More precisely, for $V \in \mathbb{R}$ fixed and as $T \to \infty$,*

$$\begin{aligned}\mathbb{P}\Big(\log|\zeta(\tfrac{1}{2} + i\tau)| \geq V\sqrt{(1/2)\log\log T}\Big) \\ = (1 + o(1)) \int_V^\infty \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du.\end{aligned} \tag{0.3.124}$$

To push the analogy with the RLM-RZF and the Gaussian BRW, note that if we formally take the logarithm of the Euler product in (0.3.121) when $Re(s) = 1/2$, then, for an appropriate cutoff $X < T$, we have

$$\begin{aligned}\log|\zeta(\tfrac{1}{2} + i\tau + ih)| &= -Re \sum_p \log(1 - p^{-(\frac{1}{2} + i\tau + ih)}) \\ &\approx \sum_{p \leq X} \frac{Re(p^{-i\tau} p^{-ih})}{p^{1/2}},\end{aligned} \tag{0.3.125}$$

where the $p^{-i\tau}$'s play the same role as the U_p's in Definition 0.3.24. This approximation can be made rigorous by the main proposition upper bound in Soundararajan (2009) (conditional on RH) and the mollification argument (near the critical line) used to prove Selberg's theorem in Radziwiłł and Soundararajan (2017) and used to prove the lower bound of the maximum in Arguin et al. (2019a). Therefore, the tree structure in Figure 0.3.14a and the logarithmic decay of the correlations hold (approximately) as it does for the RLM-RZF.

Let $\theta > -1$ and let $I \mathrel{\hat{=}} [-\log^\theta T, \log^\theta T]$. We present below a rigorous comparison between the correlations of $\{\log|\zeta(\frac{1}{2} + i\tau + ih)|\}_{h \in I}$ and the correlations of the real-part of the Dirichlet polynomial

$$P(s) \mathrel{\hat{=}} \sum_{p \leq T^\varepsilon} p^{-s}, \quad \varepsilon \in (0, 1), \tag{0.3.126}$$

on the critical line when $s = \frac{1}{2} + i\tau + ih$. The argument is due to Maksym Radziwiłł (in a private email exchange). For $h, h' \in I$, define the overlaps of $\{\log|\zeta(\frac{1}{2} + i\tau + ih)|\}_{h \in I}$ by

$$\rho(h, h') \mathrel{\hat{=}} \frac{\mathbb{E}\Big[\log|\zeta(\frac{1}{2} + i\tau + ih)| \, \log|\zeta(\frac{1}{2} + i\tau + ih')|\Big]}{\sqrt{\mathbb{E}\Big[(\log|\zeta(\frac{1}{2} + i\tau + ih)|)^2\Big]}\sqrt{\mathbb{E}\Big[(\log|\zeta(\frac{1}{2} + i\tau + ih')|)^2\Big]}}. \tag{0.3.127}$$

Then, we have the following result.

Proposition 0.3.26 (Comparison between the overlaps of $\log|\zeta|$ and $\mathrm{Re}P$)**.** *Let* $\theta > -1$*. For all* $h, h' \in I$*, we have*

$$\rho(h,h') = \frac{\frac{1}{2}\log((\log T)\wedge|h-h'|^{-1})}{\frac{1}{2}\log\log T} + O\Big((\log\log T)^{-1/2}\Big). \tag{0.3.128}$$

Proof. For a lighter notation, set

$$f(h) \mathrel{\hat{=}} \log|\zeta(\tfrac{1}{2}+i\tau+ih)| \quad \text{and} \quad g(h) \mathrel{\hat{=}} \mathrm{Re}\sum_{p\leq T^{\varepsilon}} p^{-(\frac{1}{2}+i\tau+ih)}, \tag{0.3.129}$$

where $0 < \varepsilon < \frac{1}{200}$ is fixed arbitrarily. We have

$$\begin{aligned}\mathbb{E}\Big[f(h)f(h')\Big] = \mathbb{E}\Big[g(h)g(h')\Big] &+ O\Big(\mathbb{E}\Big[|f(h)-g(h)||f(h')|\Big]\Big)\\ &+ O\Big(\mathbb{E}\Big[|g(h)||f(h')-g(h')|\Big]\Big).\end{aligned} \tag{0.3.130}$$

We need to control the two error terms.

Based on the work of Selberg (1946), Theorem 5.1 in Tsang (1984) shows that, for any given $0 < \varepsilon < \frac{1}{200}$,

$$\max_{h\in I}\mathbb{E}\Big[|f(h)-g(h)|^2\Big] \ll 1. \tag{0.3.131}$$

By a standard second moment estimate on g, this implies

$$\begin{aligned}\max_{h\in I}\mathbb{E}\Big[|f(h)|^2\Big] &\ll \max_{h\in I}\mathbb{E}\Big[|g(h)|^2\Big] + \max_{h\in I}\mathbb{E}\Big[|f(h)-g(h)|^2\Big]\\ &\ll \log\log T.\end{aligned} \tag{0.3.132}$$

By applying the Cauchy-Schwarz inequality for both error terms in (0.3.130), and then using (0.3.131) and (0.3.132), we find that, for all $h, h' \in I$,

$$\mathbb{E}\Big[f(h)f(h')\Big] = \mathbb{E}\Big[g(h)g(h')\Big] + O\Big((\log\log T)^{1/2}\Big). \tag{0.3.133}$$

In particular, for all $h \in I$,

$$\mathbb{E}\Big[f(h)^2\Big] = \mathbb{E}\Big[g(h)^2\Big] + O\Big((\log\log T)^{1/2}\Big). \tag{0.3.134}$$

Since $\mathbb{E}\Big[g(h)g(h')\Big] = \frac{1}{2}\log((\log T)\wedge|h-h'|^{-1}) + O(1)$ for all $h, h' \in I$ by an argument very similar to page 20 in Harper (2013), we get the conclusion from (0.3.133) and (0.3.134). This ends the proof. $\square$

Here are the answers to the questions of interest :

(Q1): When the length of the interval is $O(1)$ (i.e. $\theta = 0$), see Theorem 1.1 in Arguin et al. (2019a). For a generalization to intervals of length $O(\log^\theta T)$, $\theta > -1$, see Theorem 1.2 in Arguin et al. (2019c) (Article 6).

(Q2) and (Q6): For the asymptotics of the log-number of high points and the free energy on intervals of length $O(\log^\theta T)$, $\theta > -1$, see Theorem 1.1 in Arguin et al. (2019c) (Article 6) and its proof. For moments and high points conjectures when $\theta = 0$, see Keating and Snaith (2000) and Fyodorov and Keating (2014). For the limiting two-overlap distribution and the extended Ghirlanda-Guerra identities, see Conjectures 10.1.1 and 10.1.2.

(Q3), (Q4) and (Q5): This is open, although see Conjecture 10.1.3 (due to L.-P. Arguin) for the second order of the maximum on intervals of length $O(\log^\theta T)$, $\theta > -1$. When $\theta = 0$, see Theorem 2 in Harper (2019) for an upper bound on the second order of the maximum. Also, see Fyodorov and Keating (2014) for a conjecture on the convergence in law of the recentered maximum. In Saksman and Webb (2018), it is proved that the field converges to a non-trivial random generalized function with a Gaussian multiplicative chaos factor.

Remark 0.3.16. *The general question of understanding the moments of zeta* (*Q*6) *is tied to the Lindelöf hypothesis, which conjectures that*

$$\zeta(\tfrac{1}{2} + it) \ll t^\varepsilon, \quad \forall \varepsilon > 0, \quad \text{as } t \to \infty. \tag{0.3.135}$$

It is equivalent to

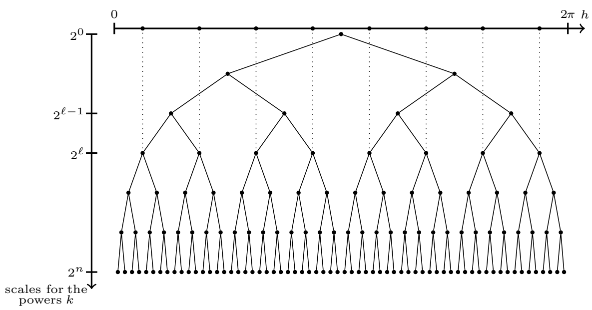
$$\frac{1}{T}\int_1^T (\zeta(\tfrac{1}{2} + it))^{2k} dt \ll T^\varepsilon, \quad \forall k \in \mathbb{N},\ \forall \varepsilon > 0, \tag{0.3.136}$$

see e.g. Theorem 13.2 in Titchmarsh (1986). It is a weak form of the Riemann hypothesis as it can be restated in terms of the concentration of non-trivial zeros around the critical line : for every $\varepsilon > 0$,

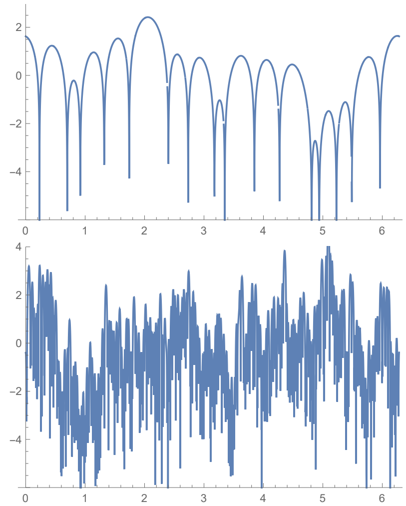
$$\#\Big\{z \in \mathbb{C} : \zeta(z) = 0,\ \tfrac{1}{2} + \varepsilon \leq \operatorname{Re} z < 1,\ T \leq \operatorname{Im} z \leq T + 1\Big\} = o(\log T), \tag{0.3.137}$$

see Backlund (1918–1919).

0.3.3.12. *The log-characteristic polynomials of the CUE field (LCP-CUE)*



(A) LCP-CUE tree structure ($n = \frac{1}{\log 2} \log N$ levels)



(B) LCP-CUE simulation ($N = 2^4$; $N = 2^8$)

FIGURE 0.3.15. The log-characteristic polynomials of the CUE field

For an introduction to random matrix theory, the reader is referred to Anderson et al. (2010) and Tao (2012).

Definition 0.3.27 (LCP-CUE)**.** *For $N \in \mathbb{N}$, let U_N be a random matrix sampled from the group of $N \times N$ unitary matrices under the Haar measure. The field of interest is the real-part of the log-characteristic polynomial of U_N, namely*

$$X_h \mathrel{\hat=} \log|\det(e^{ih} I_N - U_N)| = \sum_{j=1}^{N} \log|1 - e^{i(\lambda_j - h)}|, \quad h \in [0, 2\pi], \tag{0.3.138}$$

where I_N denotes the identity matrix of order N and $e^{i\lambda_j}$ is the j-th eigenvalue of U_N. By expanding the logarithm, we have

$$X_h = \sum_{j=1}^{N} \sum_{k=1}^{\infty} \frac{-Re(e^{ik(\lambda_j - h)})}{k} \approx \sum_{k=1}^{N} \frac{-Re(e^{-ikh}\, Tr\, U_N^k)}{k}. \tag{0.3.139}$$

Remark 0.3.17. *Traces of powers of U_N play a role analogous to the U_p's in the RLM-RZF model of Definition 0.3.24. If $N = 2^n$, then we can recover the BRW analogy by breaking the sum (0.3.139) into $\sum_{\ell=1}^{n} \sum_{2^{\ell-1} < k \leq 2^\ell} \frac{-Re(e^{-ikh}\, Tr\, U_N^k)}{k}$; see Figure 0.3.15a for an illustration of the branching structure.*

Remark 0.3.18. *To compute the correlations, note that* $\mathbb{E}\Big[Tr U_N^k \, Tr U_N^{k'}\Big] = 0$ *by rotation invariance of the Haar measure, and also that* $\mathbb{E}\Big[Tr U_N^k \, \overline{Tr U_N^{k'}}\Big] = \mathbf{1}_{\{k=k'\}} \min\{k', N\}$ *from Theorem 2.1 in Diaconis and Evans (2001) (see also Diaconis and Shahshahani (1994)). It follows from* $Re(z) = (z+\overline{z})/2$ *and* (0.3.139) *that*

$$Cov(X_h, X_{h'}) \approx \sum_{k=1}^{N} \frac{e^{ik(h-h')} + e^{-ik(h-h')}}{4k} = \frac{1}{2} \sum_{k=1}^{N} \frac{\cos(k\,|h-h'|)}{k}, \tag{0.3.140}$$

so that

$$Corr(X_h, X_{h'}) \approx \frac{\frac{1}{2}\log(N \wedge |h-h'|^{-1})}{\frac{1}{2}\log N} + O((\log N)^{-1}), \tag{0.3.141}$$

showing again the logarithmic decay of the correlations with the distance.

Here are the answers to the questions of interest :

(Q1) and (Q2): See Theorem 1.2 and Theorem 1.3 in Arguin et al. (2017a).

(Q3): See Theorem 1.2 in Paquette and Zeitouni (2018).

(Q4): See Theorem 1.2 in Chhaibi et al. (2018). In fact, their article proves the tightness of the recentered maximum for the log-characteristic polynomials of the CβE field for all $\beta > 0$ (the CUE field corresponds to the special case $\beta = 2$).

(Q5): This is open, although see Fyodorov and Keating (2014) and Fyodorov et al. (2018) for some conjectures. In Chhaibi et al. (2017), the field $\{\det(e^{ih} I_N - U_N)\}_{h \in [0,2\pi]}$ is shown to converge, after proper scaling at the microscopic level, to a random analytic function whose zeros form a determinantal point process with sine kernel.

(Q6): This is open in principle, but it should be possible to write it down. The limiting free energy is shown in Corollary 1.4 of Arguin et al. (2017a). For the rest, one can take inspiration from Arguin and Tai (2018) to find the limiting two-overlap distribution, and from Ouimet (2018b) (Article 4) to prove the Ghirlanda-Guerra identities. The reader can find conjectures on moments and high points in Keating and Snaith (2000), Fyodorov and Keating (2014) and Fyodorov et al. (2018).

Remark 0.3.19. *In Fyodorov and Keating (2014), it was conjectured that, as* $N \to \infty$, *the Radon measure* $\frac{e^{\gamma X_h}}{\mathbb{E}[e^{\gamma X_h}]} dh$ *converges to a GMC measure. This result was proved for the* L^2*-phase in Webb (2015), and for the* L^1*-phase in Nikula et al. (2018).*

0.4. A summary of the new results and ideas

The thesis contains 9 articles of which 7 are already published in peer-reviewed journals. The reader can find the articles listed on my personal website :

https://sites.google.com/site/fouimet26/research.

I present a brief summary of the new results and ideas in the three subsections below. In order to not repeat too much material, it is not a comprehensive review. If the reader wants to know more, I invite him/her to read the beginning sections of each article.

0.4.1. In Part 1

Part 1 of the thesis contains three articles on log-correlated Gaussian fields :

Article 1

Arguin and Ouimet (2016) (Article 1) shows the first order of the maximum ($Q1$) and the first order of the log-number of γ-high points ($Q2$) for the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-IGFF on its full domain. Specifically, we have the following theorem.

Theorem 0.4.1. *Let $\{\psi_v\}_{v \in V_N}$ be the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-IGFF of Definition 0.3.22. Then,*

$$\lim_{N \to \infty} \frac{\max_{v \in V_N} \psi_v}{\log N^2} = \int_0^1 \bar{\sigma}(s) ds \doteqdot \gamma^\star, \quad \textit{in probability.} \tag{0.4.1}$$

The log-number of γ-high points depends on critical levels defined by

$$\gamma^l \doteqdot \int_0^1 \frac{\bar{\sigma}^2(s)}{\bar{\sigma}(s \wedge \lambda^l)} ds, \quad 1 \leq l \leq m, \qquad \gamma^0 \doteqdot 0. \tag{0.4.2}$$

For $\gamma \in (\gamma^{l-1}, \gamma^l]$, define

$$\mathcal{E}(\gamma) \doteqdot (1 - \lambda^{l-1}) - \frac{(\gamma - \int_0^{\lambda^{l-1}} \bar{\sigma}(s) ds)^2}{\int_{\lambda^{l-1}}^1 \bar{\sigma}^2(s) ds} \quad \textit{and} \quad \mathcal{E}(0) \doteqdot 1. \tag{0.4.3}$$

Let $\gamma \in [0, \gamma^\star)$, then

$$\lim_{N \to \infty} \frac{\log |\{v \in V_N : \psi_v \geq \gamma \log N^2\}|}{\log N^2} = \mathcal{E}(\gamma), \quad \textit{in probability.} \tag{0.4.4}$$

Here is an heuristic. Let $\nabla_i f \mathrel{\hat{=}} f(i) - f(i-1)$ be the difference operator (we omit the subscript when the variable is obvious from the context). Consider the set of v's that are representatives at scale λ_k (recall Remark 0.3.11 and Figure 0.3.10a) and for which the increments of the field ψ reach level $\nabla\gamma_i$ in the scale interval $[\lambda_{i-1}, \lambda_i]$, for every $i \leq k$:

$$\Lambda_{N,k} \mathrel{\hat{=}} \{v \in R_{\lambda_k} : \nabla\psi_v(\lambda_i) \geq \nabla\gamma_i \log N^2 \text{ for all } i \in \{1, 2, \dots, k\}\}. \tag{0.4.5}$$

By identifying each point of the field $\psi_v(\lambda_i)$ at each scale λ_i with his closest representative $\psi_{v_{\lambda_i}}(\lambda_i)$, then Gaussian tail estimates (Lemma 11.1.1) and the independence between the increments $\nabla\psi_v(\lambda_i)$ (coming from the Markov property of the GFF) yield

$$\mathbb{E}[|\Lambda_{N,k}|] \asymp N^{2\lambda_k} \prod_{i=1}^{k} \mathbb{P}(\nabla\psi_v(\lambda_i) \geq 2\nabla\gamma_i \log N) \asymp \frac{N^{2\lambda_k} N^{-2\sum_{i=1}^{k} \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i}}}{(\log N)^{k/2}}, \tag{0.4.6}$$

since there are $\asymp N^{2\lambda_k}$ representatives at scale λ_k and the variance of the increments is $\mathbb{V}(\nabla\psi_v(\lambda_i)) = \sigma_i^2 \nabla\lambda_i \log N + O(1)$ if we ignore the boundary effect. In other words,

$$\lim_{N\to\infty} \frac{\log(\mathbb{E}[|\Lambda_{N,k}|])}{\log N^2} = \sum_{i=1}^{k} \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i}\right). \tag{0.4.7}$$

Since there should be representatives at each scale λ_k that ultimately yield a high value at scale λ_M, it is intuitive that the level of the maximum should be found by maximizing

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

If the end level is fixed to $\gamma_M = \gamma$, then the same argument suggests that the log-number of γ-high points (for $0 < \gamma < \gamma^\star$) should be found by maximizing

$$\lim_{N\to\infty} \frac{\log(\mathbb{E}[|\Lambda_{N,M}|])}{\log N^2} = \sum_{i=1}^{M-1} \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i}\right) + \left(\nabla\lambda_M - \frac{(\gamma - \gamma_{M-1})^2}{\sigma_M^2 \nabla\lambda_M}\right) \tag{0.4.8}$$

under the constraints

$$\sum_{i=1}^{k} \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i}\right) \geq 0, \quad 1 \leq k \leq M-1. \tag{0.4.9}$$

The unique solution to both optimization problems can be found explicitly by using the Karush-Kuhn-Tucker conditions, see Appendix A in Ouimet (2014).

An analysis of the solutions tells us where the maximal particle (respectively, most particles reaching above $\gamma \log N^2$) should be at every effective scale λ^j, with high probability. Therefore, if the difference of contribution between each point of the field and its closest representative can be shown to be negligible at every scale, then we can work on the underlying branching structure of the field (recall Figure 0.3.11a) and adapt a second method (by being careful around the branching point in the lower bound) to prove Theorem 0.4.1. Away from the boundary, i.e. on $V_N^\delta = \{v \in V_N : \min_{z\in\partial V_N} \|z - v\|_2 \geq \delta N\}$ with $\delta \in (0, \frac{1}{2})$, this is not so difficult because $\mathbb{V}(\nabla\psi_v(\lambda_i)) = \sigma_i^2 \nabla\lambda_i \log N + O(1)$ holds true. It was proved (with some mistakes and different notation) in Ouimet (2014).

The main innovation of the article is in proving that Theorem 0.4.1 not only holds on V_N^δ, but also on V_N, even though the variance of the increments decreases to zero as we approach ∂V_N. The two ingredients necessary to extend the proof from V_N^δ to V_N are essentially (see the appendix in Arguin and Ouimet (2016) (Article 1)) :

(a) For any $i \in \{1, 2, \ldots, M\}$, we have

$$\max_{v\in V_N} \mathrm{Var}(\psi_v(\lambda_i) - \psi_v(\lambda_{i-1})) \leq \sigma_i^2 \nabla\lambda_i \log N + \mathrm{Const.}(\sigma_i), \tag{0.4.10}$$

meaning that the variance of the increments of the field is uniformly bounded everywhere by the variance of the increments in the center of V_N, up to a constant.

(b) For any $i \in \{1, 2, \ldots, M\}$, then

$$\max_{v\in V_N} \mathrm{Var}(\psi_v(\lambda_i) - \psi_{v_{\lambda_i}}(\lambda_i)) \leq \mathrm{Const.}(\sigma_1, \ldots, \sigma_i), \tag{0.4.11}$$

meaning that the difference of contributions between each point of the field and its closest representative is in fact negligible at any given scale, uniformly on V_N.

Both results are consequences of careful estimates on Green functions using the potential kernel of the simple random walk underlying the definition of the variances of the field. In the case of the GFF (one scale), only an upper bound on $\max_{v\in V_N} \mathrm{Var}(\psi_v)$ is needed. Such a bound is trivial since the Green function (0.3.107) maximizes in the center of any box. What makes (a) and (b) particularly non-trivial in the scale-inhomogeneous case is that we have to estimate <u>**differences**</u> of Green functions at different scales, where both estimates pull in opposite directions. These computations also have to be done near the boundary ∂V_N, where a shell $[v]_{\lambda_{i-1}} \cap [v]_{\lambda_i}^c$ might be cut off in various ways. Because of this, (a) is no longer obvious. To obtain (b), we can compare successively $\psi_v(\lambda_i)$, $\mathbb{E}[\psi_v \,|\, \mathcal{F}_{\partial B}]$, $\mathbb{E}[\psi_{v_{\lambda_i}} \,|\, \mathcal{F}_{\partial B}]$ and $\psi_{v_{\lambda_i}}(\lambda_i)$, for a large enough box $B \supseteq [v]_{\lambda_i} \cup [v_{\lambda_i}]_{\lambda_i}$ of width $O(N^{1-\lambda_i})$, where the boundary effects are again an issue.

Article 2

Ouimet (2017) (Article 2) uses the results from Arguin and Ouimet (2016) (Article 1) to show that the limiting law of the Gibbs measure for the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-IGFF is a tree of Poisson-Dirichlet processes with a number of levels that depends on the inverse temperature parameter β. It is a consequence of the following theorem, which answers $(Q6)$.

Theorem 0.4.2. *Let $\{\psi_v\}_{v \in V_N}$ be the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-IGFF on V_N of Definition 0.3.22, let*

$$q^N(v, v') \doteq \frac{\mathbb{E}[\psi_v \psi_{v'}]}{\max_{v \in V_N} Var(\psi_v)}, \quad v, v' \in V_N, \tag{0.4.12}$$

denote the overlaps of the field, and let

$$l_\beta \doteq \begin{cases} \min\{l \in \{1, \dots, m\} : \beta \leq \beta_c(\bar{\sigma}_l) \doteq 2/\bar{\sigma}_l\}, & \text{if } \beta \leq 2/\bar{\sigma}_m, \\ m + 1, & \text{otherwise.} \end{cases} \tag{0.4.13}$$

Then, for any $\beta > 0$, the limiting two-overlap distribution of the field ψ is

$$\lim_{N \to \infty} \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\Big[\mathbf{1}_{\{q^N(v,v') \leq r\}}\Big] = \begin{cases} 0, & \text{if } r < 0, \\ \beta_c(\bar{\sigma}_j)/\beta, & \text{if } r \in [x^{j-1}, x^j), \ j \leq l_\beta - 1, \\ 1, & \text{if } r \geq x^{l_\beta - 1}, \end{cases} \tag{0.4.14}$$

where $x^j \doteq \mathcal{J}_{\bar{\sigma}^2}(\lambda^j)/\mathcal{J}_{\bar{\sigma}^2}(1)$ and $\mathcal{G}_{\beta,N}(\{v\}) \doteq e^{\beta \psi_v} / \sum_{v' \in V_N} e^{\beta \psi_{v'}}$, $v \in V_N$.

Also, let $\beta > 0$ and let μ_β be a subsequential limit of $\{\mathcal{G}_{\beta,N}\}_{N \in \mathbb{N}}$ as in (0.3.83). Then, for any $s \in \mathbb{N}$, any $k \in \{1, \dots, s\}$, and any functions $h : \{x^0, x^1, \dots, x^{l_\beta - 1}\}^{s(s-1)/2} \to \mathbb{R}$ and $g : \{x^0, x^1, \dots, x^{l_\beta - 1}\} \to \mathbb{R}$, we have the extended Ghirlanda-Guerra identities :

$$\begin{aligned} E\mu_\beta^{\times(s+1)}\Big[g(R_{k,s+1})\, h((R_{i,i'})_{1 \leq i,i' \leq s})\Big] &= \frac{1}{s} E\mu_\beta^{\times 2}\Big[g(R_{1,2})\Big] E\mu_\beta^{\times s}\Big[h((R_{i,i'})_{1 \leq i,i' \leq s})\Big] \\ &\quad + \frac{1}{s} \sum_{\ell \neq k}^{s} E\mu_\beta^{\times s}\Big[g(R_{k,\ell})\, h((R_{i,i'})_{1 \leq i,i' \leq s})\Big]. \end{aligned} \tag{0.4.15}$$

The first part of the theorem says that, in the limit $N \to \infty$, the particles are sampled under the Gibbs measure in such a way that their overlaps can only take finitely many values (on average) : $x^0 < x^1 < \cdots < x^{l_\beta - 1}$, where x^j has probability $\nabla_j(1 \wedge (\beta_c(\bar{\sigma}_{j+1})/\beta))$ with the convention $\bar{\sigma}_0 \triangleq \infty$ and $\bar{\sigma}_{m+1} \triangleq 0$. As β gets larger (i.e. as temperature decreases) and passes certain critical thresholds

$$\beta_c(\bar{\sigma}_1) < \beta_c(\bar{\sigma}_2) < \cdots < \beta_c(\bar{\sigma}_m), \tag{0.4.16}$$

the number of possible values x^j for the overlaps, namely l_β, also increases. There are m critical thresholds in general because there are m effective scales (in the case of the GFF, there is only one threshold because there is only one effective scale). As we already know from Panchenko's work, the extended Ghirlanda-Guerra identities in (0.4.15) together with the limiting two-overlap distribution in (0.4.14) imply that the joint distribution of the overlaps in the limit is completely determined under $E\mu_\beta^{\times\infty}$ and the underlying limiting Gibbs measure μ_β can be identified (see Section 7 in Ouimet (2017) (Article 2)) as a $(l_\beta - 1)$-levels tree of Poisson-Dirichlet processes (also called Ruelle probability cascade). These results coincide with the ones for the GREM in Bovier and Kurkova (2004a). The surprising part is that our results coincide **despite** the branching structure of the IGFF being only approximate when N is finite and despite the decay of the variances near the boundary of the domain V_N. This is one additional piece of evidence (among many) for the universality of the extended Ghirlanda-Guerra identities, the extent of which is discussed and quantified in Jagannath (2017).

The proof of Theorem 0.4.2 follows the same steps (essentially) as in the proof of the extended Ghirlanda-Guerra identities for the REM in Section 0.3.3.1. The two major differences are that we have to apply the arguments (the Bovier-Kurkova technique, Gaussian integration by parts, Panchenko's concentration argument, etc.) on the increments at every scales and we also have to be careful about the boundary effect of the variances on the Gibbs measure.

The article generalizes Arguin and Zindy (2015) where the same techniques were used in the simpler case of the GFF (one effective scale). By applying Laplace's method with the knowledge of the first order of the maximum and log-number of γ-high points from Arguin and Ouimet (2016) (Article 1), we can easily derive the limit (in probability and in L^p, $p \geq 1$) of the free energy on V_N. The same expression also holds for the limiting free energy away from the boundary, i.e. on the set

$$A_{N,\rho} \mathrel{\hat{=}} \left\{ v \in V_N : \min_{z \in \mathbb{Z}^2 \setminus V_N} \|v - z\|_2 \geq N^{1-\rho} \right\}, \quad \text{with } \rho \in (0,1). \tag{0.4.17}$$

In Arguin and Zindy (2015), they apply Slepian's lemma to compare the GFF with the REM (which has no correlations) in order to show that the Gibbs measure of the set $A_{N,\rho}^c$ is negligible in the limit (in probability), meaning that the contribution of the boundary doesn't impact the expression of the limiting two-overlap distribution. It also means that the Ghirlanda-Guerra identities need only to hold on $A_{N,\rho}$ as $N \to \infty$. In my article, such a comparison is far from obvious as we would have to compare the increments of the field at every effective scale with the increments of a GREM and somehow put everything together to make a global comparison. The fact that the set of representatives changes with each scale make us lose the independence property of the increments $\nabla \psi_v(\lambda_i)$ that we have when v is fixed, so it is not clear how to connect the comparisons to make it global.

The approach that I use is completely different and instead compares the related optimization problems on V_N and $A_{N,\rho}^c$ respectively. A monotony argument shows that the log-number of γ-high points on $A_{N,\rho}^c$ and the limiting free energy on $A_{N,\rho}^c$ are both strictly smaller than their counterparts on V_N, from which it follows that, for ρ small enough but fixed, we have $\mathcal{G}_{\beta,N}(A_{N,\rho}^c) \to 0$ as $N \to \infty$, in probability.

Once the problem is reduced to $A_{N,\rho}$, the structure of the argument is the same as in Section 0.3.3.1; each argument is just more laborious as we have to prove everything on the increments $\psi_v(\alpha') - \psi_v(\alpha)$, $\lambda_{i-1} \leq \alpha < \alpha' \leq \lambda_i$. The Bovier-Kurkova technique for instance is the idea that relates the limiting two-overlap distribution and the derivative of the limiting free energy with respect to a perturbation parameter u, via Gaussian integration by parts and estimates on the increments of overlaps. In the scale-inhomogeneous case, the details needed to find the expression of the derivative at $u = 0$ are much more involved and probably the most technical part of the article.

Article 3

Ouimet (2018c) (Article 3) shows the tightness of the recentered maximum ($Q4$) for the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-IBRW of Definition 0.3.18.

Theorem 0.4.3. *For all $\varepsilon > 0$, there exists $K_\varepsilon > 0$ large enough that for all $n \in \mathbb{N}$,*

$$\mathbb{P}\left(\left|\max_{v \in \mathbb{D}_n} S_v(n) - \sum_{j=1}^{m} \left[\sqrt{2\log 2}\,\bar{\sigma}_j \nabla \lambda^j n - \frac{(1 + \delta_j^{\text{left}} + \delta_j^{\text{right}})\bar{\sigma}_j}{2\sqrt{2\log 2}} \log n\right]\right| \geq K_\varepsilon\right) < \varepsilon,$$

where $\delta_j^{\text{left}} \mathrel{\hat{=}} 1$ when $\mathcal{J}_{\sigma^2}$ and $\mathcal{J}_{\bar{\sigma}^2}$ coincide on $[\lambda^{j-1}, a]$ for some $a > \lambda^{j-1}$, and $\delta_j^{\text{left}} \mathrel{\hat{=}} 0$ otherwise. Similarly, $\delta_j^{\text{right}} \mathrel{\hat{=}} 1$ when $\mathcal{J}_{\sigma^2}$ and $\mathcal{J}_{\bar{\sigma}^2}$ coincide on $[b, \lambda^j]$ for some $b < \lambda^j$, and $\delta_j^{\text{right}} \mathrel{\hat{=}} 0$ otherwise.

The proof uses discrete Brownian bridge estimates adapted from Bramson (1978) and a refined second moment method between each effective scale λ^j, $1 \leq j \leq m$, that generalizes the one used in Fang and Zeitouni (2012a) (slightly more precise than the heuristic below (0.3.96)). The article describes how the variance and scale parameters influence the first and second order of the maximum. In particular, everytime the speed function $\mathcal{J}_{\sigma^2}$ and its concave hull $\mathcal{J}_{\bar{\sigma}^2}$ coincide to the immediate left or right of an effective scale λ^j (in the homogeneous case of the BRW, note that they coincide on both sides), then halves of Brownian bridge estimates need to be added for the second moment method to work, as we already explained with an heuristic in Section 0.3.3.3.

In the homogeneous case of the BRW, both halves of the Brownian bridge estimate are needed, so that $\delta_1^{\text{left}} = \delta_1^{\text{right}} = 1$ and we recover the $-\frac{3}{2} \cdot \frac{\bar{\sigma}_1}{\sqrt{2\log 2}} \log n$ correction from (0.3.90). By applying the heuristic argument between each effective scale λ^j, the result of Theorem 0.4.3 follows from the exponential decay of the upper bound probability together with a second order lower bound and the tightness of the maximum around its median (from Theorem 1 in Fang (2012)).

A stronger result was independently obtained by Mallein (2015a) when the law of the increments is not necessarily Gaussian. However, our proof (which doesn't use the spinal decomposition) is more robust in the presence of an approximate branching structure.

0.4.2. In Part 2

Part 2 of the thesis contains three articles on the Riemann zeta function :

Article 4

Ouimet (2018b) (Article 4) finds the joint distribution of the overlaps in the limit for the RLM-RZF model of Definition 0.3.24 on the critical line. As for the REM, we identify the weak limit of the Gibbs measure

$$G_{\beta,T}(A) = \int_A \frac{e^{\beta X_h}}{\int_{[0,1]} e^{\beta X_{h'}} dh'} dh, \quad A \in \mathcal{B}([0,1]), \tag{0.4.18}$$

as a Poisson-Dirichlet process (at low temperature) in the following sense.

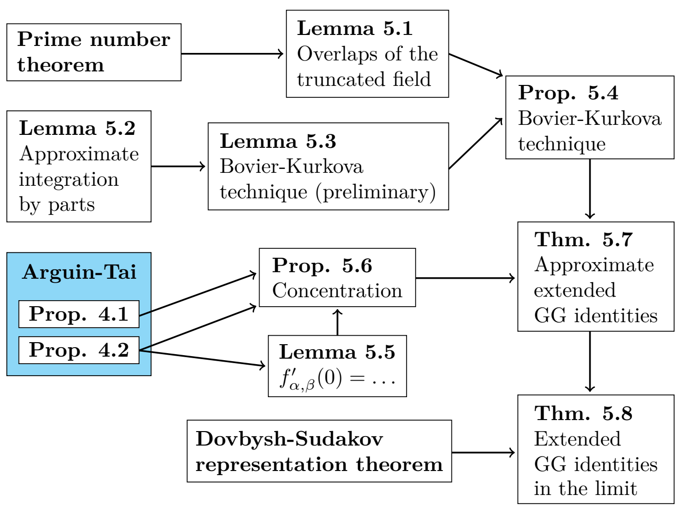
Theorem 0.4.4 (Answer to (Q6))**.** *Let $\beta > \beta_c \mathrel{\hat{=}} 2$ and let $\xi = (\xi_k)_{k\in\mathbb{N}}$ be a Poisson-Dirichlet variable of parameter β_c/β. Denote by E the expectation with respect to ξ. For any continuous function $\phi : [0,1]^{s(s-1)/2} \to \mathbb{R}$ of the overlaps of s points,*

$$\begin{aligned} &\lim_{T\to\infty} \mathbb{E}G_{\beta,T}^{\times s}\left[\phi\Big(\big(\rho(h_\ell, h_{\ell'})\big)_{1\leq \ell,\ell' \leq s}\Big)\right] \\ &\qquad = E\left[\sum_{k_1,\dots,k_s\in\mathbb{N}} \xi_{k_1}\cdots\xi_{k_s}\phi\Big(\big(\mathbf{1}_{\{k_\ell = k_{\ell'}\}}\big)_{1\leq \ell,\ell'\leq s}\Big)\right]. \end{aligned} \tag{0.4.19}$$

The limiting free energy of the perturbed field and the limiting two-overlap distribution were previously found in Arguin and Tai (2018). To obtain the limiting free energy of the perturbed field, Laplace's method was applied. The estimates on the log-number of high points for the increments of the field come from estimates on the joint Laplace transform using an asymptotic formula for the modified Bessel function of the first kind and prime number theorem estimates. The limiting two-overlap distribution was found by adapting the Bovier-Kurkova technique for each prime separately and by approximating the integration by parts (recall Lemma 1.1 in Panchenko (2013b)) to the case of non-Gaussian fields that still have some specific properties on their first three moments. The field in question there is (U_p, p primes) and they use the fact that $\mathbb{E}[U_p] = \mathbb{E}[U_p^2] = 0$, $\mathbb{E}[U_p\overline{U_p}] = 1$ and $\mathbb{E}[|U_p|^3] < \infty$ to obtain $\mathbb{E}[\xi_p F(\xi_p, \overline{\xi}_p)] = \mathbb{E}[\xi_p\overline{\xi}_p]\mathbb{E}[\partial_{\overline{z}} F(\xi_p, \overline{\xi}_p)] + O(p^{-3/2})$

for an appropriate r.v. ξ_p depending on U_p and a function F which is related to the prime p summand of expectations under the Gibbs measure. Hence, the analogue of (0.3.55) can be proved for each prime p separately; we get the full relation simply by summing over the primes. The estimates on the increments of overlaps needed to complete the Bovier-Kurkova technique are straightforward consequences of prime number theorem estimates.

Our article generalizes Arguin and Tai' arguments to the multidimensional case $\boldsymbol{h} \mathrel{\hat{=}} (h_1, h_2, \dots, h_s)$ (including the approximate integration by parts and the Bovier-Kurkova technique) and adapts a concentration result from Panchenko (2010a) for the mixed p-spin model to show the extended Ghirlanda-Guerra identities in the limit. The following figure summarizes the proof structure better than words.



Since the sequence $(p^{-i\tau},\, p \text{ primes})$, for $\tau \sim \text{Uniform}(T, 2T)$ and T large, behaves very closely to the sequence of uniform random variables $(U_p,\, p \text{ primes})$ on the unit circle in $\mathbb{C}$, we can actually extend the arguments above and prove the extended Ghirlanda-Guerra identities for the real-part of the Dirichlet polynomials

$$\sum_{p \leq T^{\varepsilon}} p^{-(\frac{1}{2} + i\tau + ih)}, \quad h \in [0, 1], \tag{0.4.20}$$

for which we already compared the correlations with the ones for the log-modulus of the Riemann zeta function on the critical line in Section 0.3.3.11. This leads us to believe that Theorem 0.4.4 could be true for $\log|\zeta|$ itself, see Conjectures 10.1.1 and 10.1.2.

Article 5

Arguin and Ouimet (2019) (Article 5) deals with the open problem of the tightness of the recentered maximum ($Q4$) for the RLM-RZF of Definition 0.3.24 on the critical line. We simplify the problem by showing the following.

Theorem 0.4.5. *With arbitrarily high probability, we have*

$$\max_{h\in[0,1]} X_h = \max_{h\in\mathcal{S}} X_h + O(1), \quad \text{as } T \to \infty, \tag{0.4.21}$$

for some set $\mathcal{S} \subseteq [0,1]$ that contains $O(\log T\sqrt{\log\log T})$ equidistant points.

The main idea is to apply, for every $\omega \in \Omega$, a mean-value theorem for the maximum $e^{X_{h^\star(\omega)}(\omega)} = e^{\max_{h\in[0,1]} X_h(\omega)}$ around its closest neighbor $h(\omega) \in \mathcal{S}$ (assuming that $\mathcal{S}$ is chosen such that we always have $|h(\omega) - h^\star(\omega)| \leq (C \cdot \log T\sqrt{\log\log T})^{-1}$ for some $C > 0$) :

$$e^{X_{h^\star(\omega)}(\omega)} - e^{X_{h(\omega)}} = \frac{d}{dh}X_h(\omega)\bigg|_{h=\xi(\omega)} e^{X_{\xi(\omega)}(\omega)}(h^\star(\omega) - h(\omega)), \tag{0.4.22}$$

where $\xi(\omega)$ is lying between $h(\omega)$ and $h^\star(\omega)$. By obtaining Laplace transform estimates for the derivative of the field (using an asymptotic formula for the modified Bessel function of the first kind and prime number theorem estimates), we can obtain continuity estimates on $h \mapsto \frac{d}{dh}X_h$ by adapting a chaining argument from Arguin et al. (2017b) and deduce large deviation estimates on $h \mapsto \frac{d}{dh}X_h$, which show that, uniformly on any fixed interval $J \subseteq [0,1]$ of length $\ll (\log T)^{-3}$, we have $\mathbb{P}(\max_{h\in J} \frac{d}{dh}X_h \geq x) \ll \exp(-2\cdot(x/\log T)^2)$. A union bound then yields

$$\max_{h\in[0,1]} \frac{d}{dh}X_h \ll \log T\sqrt{\log\log T}, \tag{0.4.23}$$

with high probability. Using this fact in (0.4.22) together with the fact that

$$e^{X_{\xi(\omega)}(\omega)} \leq e^{X_{h^\star(\omega)}(\omega)}, \tag{0.4.24}$$

the theorem follows if we choose $C > 0$ large enough.

Theorem 0.4.5 improves on the application of a Bernstein type inequality for trigonometric polynomials, which would show that (0.4.21) holds true with $O(\log T \cdot \log\log T)$ points instead.

Article 6

Arguin et al. (2019c) (Article 6) shows the limiting free energy (asymptotics of the moments) and the first order of the maximum for the LM-RZF of Definition 0.3.25 on short intervals of length $O(\log^\theta T)$, $\theta > -1$, on the critical line.

Theorem 0.4.6 (Moments; answers the first part of (*Q*6))**.** *Let $\theta > -1$, $\beta > 0$ and $\varepsilon > 0$ be given. Let τ be a random variable uniformly distributed on $[T, 2T]$ under the probability measure $\mathbb{P}$. Then, for some explicit exponent $f_\theta(\beta)$, as $T \to \infty$, we have*

$$\mathbb{P}\Big(\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)|^\beta dh < (\log T)^{f_\theta(\beta)-\varepsilon} \Big) = o(1). \tag{0.4.25}$$

Moreover, if $\theta \le 3$ or if the Riemann hypothesis holds, then as $T \to \infty$,

$$\mathbb{P}\Big(\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)|^\beta dh > (\log T)^{f_\theta(\beta)+\varepsilon} \Big) = o(1). \tag{0.4.26}$$

This proves an extended version of a conjecture of Fyodorov and Keating (2014). The form of the exponent f_θ differs between mesoscopic intervals ($-1 < \theta < 0$) and macroscopic intervals ($\theta > 0$), a phenomenon that stems from an approximate tree structure for the correlations of $\log|\zeta|$ (recall Section 0.3.3.11). We also have the following.

Theorem 0.4.7 (Local maximum; answers (*Q*1))**.** *Let $\theta > -1$ and $\varepsilon > 0$ be given. Let τ be a random variable uniformly distributed on $[T, 2T]$ under the probability measure $\mathbb{P}$. Then, for some explicit $m(\theta)$, as $T \to \infty$, we have*

$$\mathbb{P}\Big(\max_{|h| \le \log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)| < (\log T)^{m(\theta)-\varepsilon} \Big) = o(1). \tag{0.4.27}$$

Moreover, if $\theta \le 3$ or if the Riemann hypothesis holds, then as $T \to \infty$,

$$\mathbb{P}\Big(\max_{|h| \le \log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)| > (\log T)^{m(\theta)+\varepsilon} \Big) = o(1). \tag{0.4.28}$$

This generalizes earlier results of Najnudel (2018) and Arguin et al. (2019a) for $\theta = 0$. The proofs of both theorems are unconditional, except for the upper bounds when $\theta > 3$, where the Riemann hypothesis is assumed.

For $\theta > 0$, the upper bound part of Theorem 0.4.6 and Theorem 0.4.7 follows from the moment estimates

$$\mathbb{E}\Big[|\zeta(\tfrac{1}{2}+i\tau)|^{\beta}\Big] \ll (\log T)^{\beta^2/4+\varepsilon}, \tag{0.4.29}$$

(see Heap et al. (2019) when $\beta \leq 4$, and see Soundararajan (2009) for all $\beta > 0$ when assuming the Riemann hypothesis) and from a discretization result which shows that the process $(\zeta(\frac{1}{2}+i\tau+ih),\ |h| \leq \log^{\theta} T)$ varies on a $(\log T)^{-1}$ scale, so that the maximum and moments on an interval of length $O(\log^{\theta} T)$ behave (approximately) as those of

$$\boxed{O(\log^{1+\theta} T) \text{ i.i.d. Gaussian r.v.s of variance } \tfrac{1}{2}\log\log T.}$$

The limitation to $\theta \leq 3$ comes from the fact that the upper bounds (0.4.29) are not known unconditionally for $\beta > 4$.

When $\theta < 0$, the upper bounds in Theorem 0.4.6 and Theorem 0.4.7 follow the same strategy, but with the function

$$(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau), \quad \text{where } \mathcal{P}_{\alpha}(s) = \sum_{\log p \leq \log^{\alpha} T} \frac{1}{p^s}, \tag{0.4.30}$$

instead of $\zeta(\frac{1}{2}+i\tau)$. This is because, when $\theta < 0$, there is only one branch in the underlying tree structure up to scale $|\theta|$. Indeed, by a mean value formula (see e.g. (Tenenbaum, 2015, Theorem 2.10)), we have, for all $h, h' \in [-\log^{|\theta|} T, \log^{|\theta|} T]$,

$$\begin{aligned} &\mathbb{E}\Big[\Big|\mathcal{P}_{|\theta|}(\tfrac{1}{2}+i\tau+ih) - \mathcal{P}_{|\theta|}(\tfrac{1}{2}+i\tau+ih')\Big|^2\Big] \\ &\qquad \asymp \sum_{\log p \leq \log^{|\theta|} T} \frac{2-2\cos(|h-h'|\log p)}{p} = O(1). \end{aligned} \tag{0.4.31}$$

For a specific event $\mathcal{A}(T)$ on which the process $((\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\frac{1}{2}+i\tau+ih),\ |h| \leq \log^{\theta} T)$ can be discretized, and such that $\mathbb{P}(\mathcal{A}(T)) = 1 - o(1)$, we can show that, for $\beta \leq 2$,

$$\mathbb{E}\Big[\Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau)\Big|^{\beta}\mathbf{1}_{\mathcal{A}(T)}\Big] \ll (\log T)^{(\beta^2/4)\cdot(1+\theta)+\varepsilon}. \tag{0.4.32}$$

In this case, the maximum and moments will behave (approximately) as

$$\boxed{O(\log^{1+\theta} T) \text{ i.i.d. Gaussian r.v.s of variance } \tfrac{1+\theta}{2}\log\log T.}$$

Now, for the lower bounds, note that $f_\theta(\beta) = \beta m(\theta) - 1$ when $\beta > \beta_c(\theta)$, so the lower bound in Theorem 0.4.6 implies that for β large enough with respect to ε and θ, we have, with probability $1 - o(1)$,

$$\max_{|h| \le \log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)| \ge \left(\frac{1}{2\log^\theta T} \int_{-\log^\theta T}^{\log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)|^\beta dh \right)^{1/\beta}$$
$$\gg (\log T)^{m(\theta) - \frac{(1+\varepsilon+\theta)}{\beta}} \ge (\log T)^{m(\theta)-\varepsilon}, \tag{0.4.33}$$

which is the lower bound in Theorem 0.4.7.

It remains to consider the lower bound in Theorem 0.4.6. The problem is first reduced to obtaining lower bounds for moments off the critical line using a localization on the scale of $[-\log^\theta T, \log^\theta T]$. It is shown, uniformly in $\frac{1}{2} \le \sigma \le \frac{1}{2} + (\log T)^{\theta - 3\varepsilon}$ and for D a Dirichlet polynomial that approximates zeta, that with probability $1 - o(1)$,

$$\begin{aligned} \int_{-\log^\theta T}^{\log^\theta T} |D(\sigma + i\tau + iu)|^\beta du &\ll \int_{\mathbb{R}} |D(\sigma + i\tau + iu)|^\beta \cdot |\Phi(\sigma + iu)|^\beta du \\ &\ll \int_{\mathbb{R}} |D(\tfrac{1}{2} + i\tau + iu)|^\beta \cdot |\Phi(\tfrac{1}{2} + iu)|^\beta du \\ &\ll \int_{-2\log^\theta T}^{2\log^\theta T} |D(\tfrac{1}{2} + i\tau + iu)|^\beta du + \frac{1}{(\log T)^7}. \end{aligned} \tag{0.4.34}$$

The decentering follows from a result of Gabriel (1927) for subharmonic functions and the construction of an explicit entire function $\Phi(\sigma + iu)$ which is a good approximation to the indicator function of the rectangle $\mathcal{R} = \{\sigma + iu : |u| \le (\log T)^\theta, \frac{1}{2} \le \sigma \le \frac{1}{2} + (\log T)^{\theta - 3\varepsilon}\}$ in the whole strip $\frac{1}{2} \le \operatorname{Re} s$. The problem is reduced to obtaining a lower bound for

$$\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\sigma_0 + i\tau + ih)|^\beta dh, \quad \text{with } \sigma_0 = \frac{1}{2} + \frac{1}{(\log T)^{1-\delta}}, \tag{0.4.35}$$

for an appropriate $\delta > 0$. We adapt mollification results from Arguin et al. (2019a) to show that, outside of an event of probability $o(1)$, the problem can be reduced to understanding

$$\int_{-\log^\theta T}^{\log^\theta T} \exp\Big(\beta \operatorname{Re} \mathcal{P}_{1-\delta}(\sigma_0 + i\tau + ih)\Big) dh. \tag{0.4.36}$$

Since the process $\Big(\operatorname{Re} \mathcal{P}_{1-\delta}(\sigma_0 + i\tau + ih),\ |h| \le \log^\theta T\Big)$ behaves approximatively like a log-correlated Gaussian field (recall the correlations in Proposition 0.3.26, and see Lemma 6.4.3 in Arguin et al. (2019c) (Article 6) for the approximate Gaussian moments), so the remaining part of the argument follows from a second moment method.

0.4.3. In Part 3

Part 3 of the thesis contains three articles on asymptotic statistics :

Article 7

Ouimet (2018a) (Article 7) proves the complete monotonicity of multinomial probabilities by generalizing the computations first made in Alzer (2018) for the binomial distribution. Denote the d-dimensional simplex by $\mathcal{S} \mathrel{\hat{=}} \left\{\boldsymbol{x} \in [0,1]^d : \|\boldsymbol{x}\|_1 \mathrel{\hat{=}} \sum_{i=1}^d |x_i| \leq 1\right\}$. We have the following theorem.

Theorem 0.4.8. *For any $d \in \mathbb{N}$, $M \in (0,\infty)$, $\boldsymbol{x} \in Int(\mathcal{S})$, $x_{d+1} \mathrel{\hat{=}} 1 - \|\boldsymbol{x}\|_1 > 0$, and any $\boldsymbol{\gamma} \in [0,\infty)^d$ such that $\|\boldsymbol{\gamma}\|_1 \leq M$ and $\gamma_{d+1} \mathrel{\hat{=}} M - \|\boldsymbol{\gamma}\|_1 \geq 0$, the function*

$$g(a) \mathrel{\hat{=}} \frac{\Gamma(aM+1)}{\prod_{i=1}^{d+1} \Gamma(a\gamma_i + 1)} \prod_{i=1}^{d+1} x_i^{a\gamma_i} \tag{0.4.37}$$

is completely monotonic on $(0,\infty)$, meaning that g has derivatives of all orders and satisfies

$$(-1)^n g^{(n)}(a) \geq 0, \quad \text{for all } n \in \mathbb{N}_0,\ a \in (0,\infty). \tag{0.4.38}$$

In fact, we prove a stronger result by showing that $(-\log g)'$ is completely monotonic. The proof follows from an integral representation for the trigamma function, the convexity of the function $c \mapsto 1/(y^{1/c} - 1) + 1/(y^{1/(1-c)} - 1)$ on $(0,1)$, some identity and asymptotic formula for the digamma function, and finally, the fact that the Kullback-Leibler divergence is non-negative, which is a consequence of Jensen's inequality.

The most interesting part is not the theorem itself, but the consequences. Non-constant completely monotonic functions are strictly convex, (strictly) decreasing and positive. Since the proof shows that $(-\log g)'$ is completely monotonic, it means that g is strictly log-convex, which automatically implies the following Holder-type inequality for multinomial coefficients : For all $\lambda_j \in (0,1)$, $j \in \{1,2,\ldots,k\}$, such that $\sum_{j=1}^k \lambda_j = 1$, we have

$$\frac{\Gamma((\sum_{j=1}^k a_j \lambda_j) M + 1)}{\prod_{i=1}^{d+1} \Gamma((\sum_{j=1}^k a_j \lambda_j)\gamma_i + 1)} \leq \prod_{j=1}^k \left(\frac{\Gamma(a_j M + 1)}{\prod_{i=1}^{d+1} \Gamma(a_j \gamma_i + 1)}\right)^{\lambda_j} \tag{0.4.39}$$

where equality holds if and only if all the a_j's are the same. Other inequalities are listed in Corollary 7.3.1 of Ouimet (2018a) (Article 7).

By generalizing arguments from Leblanc and Johnson (2007) and Leblanc (2010), the article also shows how the monotonicity result from Theorem 0.4.8 can be used to study the asymptotic properties of Bernstein estimators on the simplex. More specifically, in Leblanc (2010), the following family of polynomials,

$$S_{r,s,m}(\boldsymbol{x}) \mathrel{\hat{=}} \sum_{\boldsymbol{k} \in \mathbb{N}_0^d : \|\boldsymbol{k}\|_1 \leq m} P_{r\boldsymbol{k}, rm}(\boldsymbol{x}) P_{s\boldsymbol{k}, sm}(\boldsymbol{x}), \quad r, s, m \in \mathbb{N}, \tag{0.4.40}$$

where

$$P_{\boldsymbol{k},m}(\boldsymbol{x}) \mathrel{\hat{=}} \frac{m!}{\prod_{i=1}^{d+1} k_i!} \prod_{i=1}^{d+1} x_i^{k_i}, \quad \boldsymbol{x} \in \mathcal{S}, \tag{0.4.41}$$

was studied, when $d = 1$, in order to construct a density estimator on $[0,1]$, using Bernstein polynomials, with a lower bias and a higher rate of convergence than the base Bernstein density estimator. Our article proves certain convergence results for the above family of polynomials (for all $d \in \mathbb{N}$), namely :

Proposition 0.4.9. *Let $r, s, m \in \mathbb{N}$ and let $h : \mathcal{S} \to \mathbb{R}$ be any bounded measurable function. As $m \to \infty$,*

(a) $m^{d/2} \int_{\mathcal{S}} S_{r,s,m}(\boldsymbol{x}) d\boldsymbol{x} = \frac{2^{-d}\sqrt{\pi}}{\Gamma(d/2+1/2)} + O(m^{-1}) = \int_{\mathcal{S}} \phi_{r,s}(\boldsymbol{x}) d\boldsymbol{x} + O(m^{-1})$,

(b) $\int_{\mathcal{S}} h(\boldsymbol{x})(m^{d/2} S_{r,s,m}(\boldsymbol{x}) - \phi_{r,s}(\boldsymbol{x})) d\boldsymbol{x} = o(1)$,

where

$$\phi_{r,s}(\boldsymbol{x}) \mathrel{\hat{=}} \frac{(\gcd(r,s))^d}{(2\pi)^{d/2} (\det(rs(r+s)(\mathrm{diag}(\boldsymbol{x}) - \boldsymbol{x}\boldsymbol{x}^T)))^{1/2}}. \tag{0.4.42}$$

The first step to prove this proposition is to show that, for any fixed $\boldsymbol{x} \in \mathrm{Int}(\mathcal{S})$,

$$m^{d/2} S_{r,s,m}(\boldsymbol{x}) = \phi_{r,s}(\boldsymbol{x}) + o_{\boldsymbol{x}}(1), \quad \text{as } m \to \infty, \tag{0.4.43}$$

by expressing $S_{r,s,m}(\boldsymbol{x})$ as the probability that a difference of linear combinations of independent multinomial r.v.s equals 0 (this generalizes a trick from Theorem 3 of Section XV.5 in Feller (1971)) and then using a local central limit theorem for random vectors with lattice distributions. The second step is to prove the convergence in (a) in the special case $r = s = 1$ using the Chu-Vandermonde convolution formula, Legendre's duplication formula, the Cholesky decomposition of covariance matrices for the multinomial distribution, etc. Combining (0.4.43) and (a) yields that $\{S_{1,1,m}(\cdot)\}_{m\in\mathbb{N}}$ is uniformly integrable. By Theorem 0.4.8, $a \mapsto P_{a\boldsymbol{k},am}$ is decreasing on $(0,\infty)$, so

$$S_{r,s,m}(\boldsymbol{x}) \leq \sum_{\|\boldsymbol{k}\|_1 \leq m} (P_{\boldsymbol{k},m}(\boldsymbol{x}))^2 = S_{1,1,m}(\boldsymbol{x}), \tag{0.4.44}$$

which implies that $\{S_{r,s,m}(\cdot)\}_{m\in\mathbb{N}}$ is also uniformly integrable. Hence, by (0.4.43), we must have (a) in the general case $r, s \in \mathbb{N}$. Finally, the almost-everywhere convergence and the uniform integrability imply the L^1 convergence, so (b) follows immediately from Jensen's inequality and the fact that h is bounded.

Following Leblanc (2010) ($d = 1$), Proposition 0.4.9 opens to the door to a similar application of bias-reduction for the Bernstein density estimator on the d-dimensional simplex. To the best of my knowledge, bias reduction in this context is a completely untouched subject and there is only one article (Tenbusch (1994)) that even discusses (some) asymptotic properties of the base estimator, and only in the special case $d = 2$. For more details, see (3) from the list of open problems in Section 10.2.

This subject is worth investigating because there are instances in practice where the distribution that we would like to estimate lives naturally on the d-dimensional simplex. One such example is the Dirichlet distribution, which is the conjugate prior of the multinomial distribution in Bayesian estimation, see e.g. Lange (1995) for an application in the context of allele frequency estimation in genetics. In those instances, we would expect that Bernstein estimators defined on the simplex perform better than Bernstein estimators defined on the unit hypercube, especially near the boundary $\|\boldsymbol{x}\|_1 = 1$.

Article 8

Lafaye de Micheaux and Ouimet (2018) (Article 8) proves a new uniform law of large numbers for summands that "blow up". The article was motivated by an open question originally raised by Pierre Lafaye de Micheaux about the convergence in probability of $\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \neq \hat{\mu}_n\}} \log |X_i - \hat{\mu}_n|$, where X_i are i.i.d. with Laplace(μ) distribution and $\hat{\mu}_n$ is a consistent estimator (such as the median, which is the maximum likelihood estimator in this case). Under technical conditions (roughly, good integrability properties for $X_i - \mu$ and $X_i - \mu_n$ around the blow-up point and a good control on the tail distribution of the X_i's and μ_n, uniformly for μ_n between μ and $\hat{\mu}_n$), the article shows that, for $h : \mathbb{R}\backslash\{0\} \longrightarrow \mathbb{R}$ a function that blows up at 0 but otherwise is integrable and has "good" properties elsewhere, we have

$$\lim_{n\to\infty} \sup_{v\in[0,1]} \mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \neq \mu + v(\hat{\mu}_n - \mu)\}} h(X_i - (\mu + v(\hat{\mu}_n - \mu))) - \mathbb{E}\Big[h(X_1 - \mu)\Big]\right| = 0,$$

so that $\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \neq \hat{\mu}_n\}} \log |X_i - \hat{\mu}_n|$ indeed converges.

The appeal of such results is that entropy conditions for uniform laws of large numbers (see e.g. van der Vaart and Wellner (1996) and van de Geer (2000)) cannot be applied here because the envelope function of the class of functions $\{\mathbf{1}_{\{X_i \neq t\}} \log |\cdot - t|\}_{t:|t-\mu|<\delta}$ is infinite in any small neighborhood of μ.

Article 9

Desgagné et al. (2018) (Article 9) finds the limiting law of a modified score statistic (under H_0 and under local alternatives) when we test a given exponential power distribution (H_0) against the family of asymmetric power distributions (H_1). The asymmetric power distribution, introduced in Komunjer (2007), is a reparametrization of the skewness exponential power distribution from Fernández et al. (1995). The score is modified in the sense that we assume the location and scale parameters of the exponential power distribution to be unknown and we replace them by their maximum likelihood estimators. Using a first order Taylor expansion, we expand the modified score and prove the convergence of the derivative component using a standard uniform law of large numbers from

Lucien LeCam (see e.g. Chapter 16 in Ferguson (1996)). This is possible except in the case where the X_i's are Laplace distributed under H_0, which was the ultimate motivation for the question raised in the previous paragraph and answered in Lafaye de Micheaux and Ouimet (2018) (Article 8).

Bibliography

L. Addario-Berry and B. Reed. Minima in branching random walks. *Ann. Probab.*, 37(3): 1044–1079, 2009. MR2537549.

L. Addario-Berry and B. A. Reed. Ballot theorems, old and new. In *Horizons of combinatorics*, volume 17 of *Bolyai Soc. Math. Stud.*, pages 9–35. Springer, Berlin, 2008. MR2432525.

R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Monographs in Mathematics. Springer, New York, 2007. MR2319516.

E. Aïdékon. Convergence in law of the minimum of a branching random walk. *Ann. Probab.*, 41(3A):1362–1426, 2013. MR3098680.

E. Aïdékon, J. Berestycki, É. Brunet, and Z. Shi. Branching Brownian motion seen from its tip. *Probab. Theory Related Fields*, 157(1-2):405–451, 2013. MR3101852.

E. Aïdékon, Y. Hu, and Z. Shi. Large deviations for level sets of branching Brownian motion and Gaussian free fields. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 457(Veroyatnost' i Statistika. 25):12–36, 2017. MR3723574.

H. Alzer. Complete monotonicity of a function related to the binomial probability. *J. Math. Anal. Appl.*, 459(1):10–15, 2018. MR3730425.

G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. MR2760897.

L.-P. Arguin. Extrema of log-correlated random variables principles and examples. In *Advances in disordered systems, random processes and some applications*, pages 166–204. Cambridge Univ. Press, Cambridge, 2017. MR3644278.

L.-P. Arguin and M. Aizenman. On the structure of quasi-stationary competing particle systems. *Ann. Probab.*, 37(3):1080–1113, 2009. MR2537550.

L.-P. Arguin and S. Chatterjee. Random overlap structures: properties and applications to spin glasses. *Probab. Theory Related Fields*, 156(1-2):375–413, 2013. MR3055263.

L.-P. Arguin and N. Kistler. Microcanonical analysis of the random energy model in a random magnetic field. *J. Stat. Phys.*, 157(1):1–16, 2014. MR3249901.

L.-P. Arguin and F. Ouimet. Extremes of the two-dimensional Gaussian free field with scale-dependent variance. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(2):779–808, 2016. MR3541850.

L.-P. Arguin and F. Ouimet. Large deviations and continuity estimates for the derivative of a random model of $\log|\zeta|$ on the critical line. *J. Math. Anal. Appl.*, 472(1):687–695, 2019. MR3906393.

L.-P. Arguin and W. Tai. Is the Riemann zeta function in a short interval a 1-RSB spin glass ? *Preprint. To appear in Sojourns in probability and statistical physics, in honor of the 70th birthday of C. Newman*, pages 1–20, 2018. arXiv:1706.08462.

L.-P. Arguin and O. Zindy. Poisson-Dirichlet statistics for the extremes of a log-correlated Gaussian field. *Ann. Appl. Probab.*, 24(4):1446–1481, 2014. MR3211001.

L.-P. Arguin and O. Zindy. Poisson-Dirichlet statistics for the extremes of the two-dimensional discrete Gaussian free field. *Electron. J. Probab.*, 20:no. 59, 19, 2015. MR3354619.

L.-P. Arguin, A. Bovier, and N. Kistler. Genealogy of extremal particles of branching Brownian motion. *Comm. Pure Appl. Math.*, 64(12):1647–1676, 2011. MR2838339.

L.-P. Arguin, A. Bovier, and N. Kistler. Poissonian statistics in the extremal process of branching Brownian motion. *Ann. Appl. Probab.*, 22(4):1693–1711, 2012. MR2985174.

L.-P. Arguin, A. Bovier, and N. Kistler. The extremal process of branching Brownian motion. *Probab. Theory Related Fields*, 157(3-4):535–574, 2013. MR3129797.

L.-P. Arguin, D. Belius, and P. Bourgade. Maximum of the characteristic polynomial of random unitary matrices. *Comm. Math. Phys.*, 349(2):703–751, 2017a. MR3594368.

L.-P. Arguin, D. Belius, and A. J. Harper. Maxima of a randomized Riemann zeta function, and branching random walks. *Ann. Appl. Probab.*, 27(1):178–215, 2017b. MR3619786.

L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł, and K. Soundararajan. Maximum of the Riemann zeta function on a short interval of the critical line. *Comm. Pure Appl.*

Math., 72(3):500–535, 2019a. MR3911893.

L.-P. Arguin, L. Hartung, and N. Kistler. High points of a random model of the Riemann-zeta function and Gaussian multiplicative chaos. *Preprint*, pages 1–13, 2019b. arXiv:1906.08573.

L.-P. Arguin, F. Ouimet, and M. Radziwiłł. Moments of the Riemann zeta function on short intervals of the critical line. *Preprint*, pages 1–34, 2019c. arXiv:1901.04061.

K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196. MR0373040.

R. Backlund. über die Beziehung zwischen Anwachsen und Nullstellen der Zeta-funktion. *Ofversigt Finska Vetensk. Soc.*, 61(9), 1918–1919.

J. Barral, X. Jin, R. Rhodes, and V. Vargas. Gaussian multiplicative chaos and KPZ duality. *Comm. Math. Phys.*, 323(2):451–485, 2013. MR3096527.

M. Belalia. On the asymptotic properties of the Bernstein estimator of the multivariate distribution function. *Statist. Probab. Lett.*, 110:249–256, 2016. MR3474765.

J. Berestycki. Topics on branching Brownian motion. *Lecture notes*, 2014. URL `http://www.stats.ox.ac.uk/~berestyc/Articles/EBP18_v2.pdf`.

N. Berestycki. Introduction to the Gaussian Free Field and Liouville Quantum Gravity. *Lecture notes*, 2016. URL `http://www.statslab.cam.ac.uk/~beresty/Articles/oxford4.pdf`.

N. Berestycki. An elementary approach to Gaussian multiplicative chaos. *Electron. Commun. Probab.*, 22:Paper No. 27, 12, 2017. MR3652040.

J. D. Biggins. The first- and last-birth problems for a multitype age-dependent branching process. *Advances in Appl. Probability*, 8(3):446–459, 1976. MR0420890.

J. D. Biggins. Chernoff's theorem in the branching random walk. *J. Appl. Probability*, 14 (3):630–636, 1977. MR0464415.

M. Biskup. Extrema of the two-dimensional Discrete Gaussian Free Field. *Lecture notes*, 2018. URL `https://arxiv.org/pdf/1712.09972.pdf`.

M. Biskup and O. Louidor. Extreme local extrema of two-dimensional discrete Gaussian free field. *Comm. Math. Phys.*, 345(1):271–304, 2016a. MR3509015.

M. Biskup and O. Louidor. On intermediate level sets of two-dimensional discrete Gaussian free field. *Preprint. To appear in Annales de l'Institut Henri Poincare B*, pages 1–41, 2016b. arXiv:1612.01424.

M. Biskup and O. Louidor. Full extremal process, cluster law and freezing for the two-dimensional discrete Gaussian free field. *Adv. Math.*, 330:589–687, 2018. MR3787554.

E. Bolthausen and A.-S. Sznitman. *Ten lectures on random media*, volume 32 of *DMV Seminar*. Birkhäuser Verlag, Basel, 2002. MR1890289.

E. Bolthausen, J.-D. Deuschel, and G. Giacomin. Entropic repulsion and the maximum of the two-dimensional harmonic crystal. *Ann. Probab.*, 29(4):1670–1692, 2001. MR1880237.

E. Bolthausen, J. D. Deuschel, and O. Zeitouni. Recursions and tightness for the maximum of the discrete, two dimensional Gaussian free field. *Electron. Commun. Probab.*, 16: 114–119, 2011. MR2772390.

P. Bourgade. Mesoscopic fluctuations of the zeta zeros. *Probab. Theory Related Fields*, 148(3-4):479–500, 2010. MR2678896.

A. Bovier. *Statistical mechanics of disordered systems*, volume 18 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2006. A mathematical perspective. MR2252929.

A. Bovier. *Gaussian processes on trees*, volume 163 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2017. From spin glasses to branching Brownian motion. MR3618123.

A. Bovier and L. Hartung. The extremal process of two-speed branching Brownian motion. *Electron. J. Probab.*, 19:no. 18, 28, 2014. MR3164771.

A. Bovier and L. Hartung. Variable speed branching Brownian motion 1. Extremal processes in the weak correlation regime. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12(1): 261–291, 2015. MR3351476.

A. Bovier and L. Hartung. From 1 to 6: a finer analysis of perturbed branching Brownian motion. *Preprint*, pages 1–28, 2019. arXiv:1808.05445.

A. Bovier and A. Klimovsky. Fluctuations of the partition function in the generalized random energy model with external field. *J. Math. Phys.*, 49(12):125202, 27, 2008. MR2484333.

A. Bovier and I. Kurkova. Derrida's generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):439–480, 2004a. MR2070334.

A. Bovier and I. Kurkova. Derrida's generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):481–495, 2004b. MR2070335.

M. Bramson. Maximal displacement of branching Brownian motion. *Comm. Pure Appl. Math.*, 31(5):531–581, 1978. MR0494541.

M. Bramson. Convergence of solutions of the Kolmogorov equation to travelling waves. *Mem. Amer. Math. Soc.*, 44(285):iv+190, 1983. MR705746.

M. Bramson and O. Zeitouni. Tightness of the recentered maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.*, 65(1):1–20, 2012. MR2846636.

M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.*, 69(1):62–123, 2016a. MR3433630.

M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of nonlattice branching random walk. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1897–1924, 2016b. MR3573300.

X. Cao, Y. V. Fyodorov, and P. Le Doussal. One step replica symmetry breaking and extreme order statistics of logarithmic REMs. *SciPost Phys.*, 1(2):011, 1–58, 2016. doi:10.21468/SciPostPhys.1.2.011.

D. Capocaccia, M. Cassandro, and P. Picco. On the existence of thermodynamics for the generalized random energy model. *J. Statist. Phys.*, 46(3-4):493–505, 1987. MR883541.

D. Carpentier and P. Le Doussal. Glass transition of a particle in a random potential, front selection in nonlinear renormalization group, and entropic phenomena in Liouville and sinh-Gordon models. *Phys. Rev. E*, 63(026110):1–33, 2001. doi:10.1103/PhysRevE.63.026110.

S. Chatterjee. Chaos, concentration, and multiple valleys. *Preprint*, pages 1–63, 2008. arXiv:0810.4221.

B. Chauvin and A. Rouault. Boltzmann-Gibbs weights in the branching random walk. In

Classical and modern branching processes (Minneapolis, MN, 1994), volume 84 of *IMA Vol. Math. Appl.*, pages 41–50. Springer, New York, 1997. MR1601693.

L. Chen. Thick points of high-dimensional Gaussian free fields. *Ann. Inst. Henri Poincaré Probab. Stat.*, 54(3):1492–1526, 2018a. MR3825889.

L. Chen. Steep points of Gaussian free fields in any dimension. *Preprint*, pages 1–43, 2018b. arXiv:1802.03034.

L. Chen and D. Jakobson. Gaussian free fields and KPZ relation in $\mathbb{R}^4$. *Ann. Henri Poincaré*, 15(7):1245–1283, 2014. MR3225731.

R. Chhaibi, J. Najnudel, and A. Nikeghbali. The circular unitary ensemble and the Riemann zeta function: the microscopic landscape and a new approach to ratios. *Invent. Math.*, 207(1):23–113, 2017. MR3592756.

R. Chhaibi, T. Madaule, and J. Najnudel. On the maximum of the CβE field. *Duke Math. J.*, 167(12):2243–2345, 2018. MR3848391.

A. Chiarini, A. Cipriani, and R. S. Hazra. A note on the extremal process of the supercritical Gaussian free field. *Electron. Commun. Probab.*, 20:no. 74, 10, 2015. MR3417446.

A. Chiarini, A. Cipriani, and R. S. Hazra. Extremes of the supercritical Gaussian free field. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(2):711–724, 2016. MR3531397.

A. Cipriani. High points for the membrane model in the critical dimension. *Electron. J. Probab.*, 18:no. 86, 17, 2013. MR3109625.

A. Cipriani and R. S. Hazra. Thick points for a Gaussian free field in 4 dimensions. *Stochastic Process. Appl.*, 125(6):2383–2404, 2015. MR3322868.

A. Cipriani and R. S. Hazra. Thick points for Gaussian free fields with different cut-offs. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(1):79–97, 2017. MR3606735.

A. Cipriani, B. Dan, and R. S. Hazra. The scaling limit of the membrane model. *Preprint. To appear in Annals of Probability*, pages 1–35, 2019. arXiv:1801.05663.

N. Cressie. A finely tuned continuity correction. *Ann. Inst. Statist. Math.*, 30(3):435–442, 1978. MR538319.

O. Daviaud. Extremes of the discrete two-dimensional Gaussian free field. *Ann. Probab.*, 34(3):962–986, 2006. MR2243875.

F. David, A. Kupiainen, R. Rhodes, and V. Vargas. Liouville quantum gravity on the

Riemann sphere. *Comm. Math. Phys.*, 342(3):869–907, 2016. MR3465434.

J.-M. De Koninck and F. Luca. *Analytic number theory*, volume 134 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012. MR2919246.

B. Derrida. Random-energy model: limit of a family of disordered models. *Phys. Rev. Lett.*, 45(2):79–82, 1980. MR575260.

B. Derrida. Random-energy model: an exactly solvable model of disordered systems. *Phys. Rev. B (3)*, 24(5):2613–2626, 1981. MR627810.

B. Derrida. A generalization of the Random Energy Model which includes correlations between energies. *J. Physique Lett.*, 46(9):401–407, 1985. doi:10.1051/jphyslet:01985004609040100.

B. Derrida and H. Spohn. Polymers on disordered trees, spin glasses, and traveling waves. *J. Statist. Phys.*, 51(5-6):817–840, 1988. MR971033.

A. Desgagné, P. Lafaye de Micheaux, and F. Ouimet. Asymptotic law of a modified score statistic for the asymmetric power distribution with unknown location and scale parameter. *Preprint*, pages 1–15, 2018. arXiv:1809.02852.

P. Diaconis and S. N. Evans. Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.*, 353(7):2615–2633, 2001. MR1828463.

P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. *J. Appl. Probab.*, 31A:49–62, 1994. MR1274717.

J. Ding. Exponential and double exponential tails for maximum of two-dimensional discrete Gaussian free field. *Probab. Theory Related Fields*, 157(1-2):285–299, 2013. MR3101848.

J. Ding and O. Zeitouni. Extreme values for two-dimensional discrete Gaussian free field. *Ann. Probab.*, 42(4):1480–1515, 2014. MR3262484.

J. Ding, R. Roy, and O. Zeitouni. Convergence of the centered maximum of log-correlated Gaussian fields. *Ann. Probab.*, 45(6A):3886–3928, 2017. MR3729618.

L. N. Dovbysh and V. N. Sudakov. Gram-de Finetti matrices. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 119:77–86, 1982. MR666087.

B. Duplantier and S. Sheffield. Liouville quantum gravity and KPZ. *Invent. Math.*, 185 (2):333–393, 2011. MR2819163.

B. Duplantier, R. Rhodes, S. Sheffield, and V. Vargas. Renormalization of critical Gaussian multiplicative chaos and KPZ relation. *Comm. Math. Phys.*, 330(1):283–330, 2014a. MR3215583.

B. Duplantier, R. Rhodes, S. Sheffield, and V. Vargas. Critical Gaussian multiplicative chaos: convergence of the derivative martingale. *Ann. Probab.*, 42(5):1769–1808, 2014b. MR3262492.

E. B. Dynkin. Markov processes and random fields. *Bull. Amer. Math. Soc. (N.S.)*, 3(3): 975–999, 1980. MR585179.

M. Fang. Tightness for maxima of generalized branching random walks. *J. Appl. Probab.*, 49(3):652–670, 2012. MR3012090.

M. Fang and O. Zeitouni. Branching random walks in time inhomogeneous environments. *Electron. J. Probab.*, 17:no. 67, 18, 2012a. MR2968674.

M. Fang and O. Zeitouni. Slowdown for time inhomogeneous branching Brownian motion. *J. Stat. Phys.*, 149(1):1–9, 2012b. MR2981635.

W. Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971. MR0270403.

S.-J. Feng. On Karatsuba conjecture and the Lindelöf hypothesis. *Acta Arith.*, 114(3): 295–300, 2004. MR2071084.

T. S. Ferguson. *A course in large sample theory.* Texts in Statistical Science Series. Chapman & Hall, London, 1996. MR1699953.

C. Fernández, J. Osiewalski, and M. F. J. Steel. Modeling and inference with v-spherical distributions. *J. Amer. Statist. Assoc.*, 90(432):1331–1340, 1995. MR1379475.

K. Ford. Sharp probability estimates for random walks with barriers. *Probab. Theory Related Fields*, 145(1-2):269–283, 2009. MR2520128.

Y. V. Fyodorov and J.-P. Bouchaud. Statistical mechanics of a single particle in a multiscale random potential: Parisi landscapes in finite-dimensional Euclidean spaces. *J. Phys. A*, 41(32):324009, 25, 2008a. MR2425780.

Y. V. Fyodorov and J.-P. Bouchaud. Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential. *J. Phys. A*, 41(37):372001, 12, 2008b. MR2430565.

Y. V. Fyodorov and J. P. Keating. Freezing transitions and extreme values: random matrix theory, $\zeta(\frac{1}{2}+it)$ and disordered landscapes. *Philos. Trans. R. Soc. A*, 372(20120503): 1–32, 2014. MR3151088.

Y. V. Fyodorov, P. Le Doussal, and A. Rosso. Statistical mechanics of logarithmic REM: duality, freezing and extreme value statistics of $1/f$ noises generated by Gaussian free fields. *J. Stat. Mech. Theory Exp.*, 2009(10):P10005, 32, 2009. MR2882779.

Y. V. Fyodorov, S. Gnutzmann, and J. P. Keating. Extreme values of CUE characteristic polynomials: a numerical study. *J. Phys. A*, 51(46):464001, 22, 2018. MR3876698.

R. M. Gabriel. Some Results Concerning the Integrals of Moduli of Regular Functions Along Certain Curves. *J. London Math. Soc.*, 2(2):112–117, 1927. MR1574583.

C. Garban. Quantum gravity and the KPZ formula [after Duplantier-Sheffield]. *Astérisque*, pages Exp. No. 1052, ix, 315–354, 2013. MR3087350.

C. Garban, R. Rhodes, and V. Vargas. Liouville Brownian motion. *Ann. Probab.*, 44(4): 3076–3110, 2016. MR3531686.

S. Ghirlanda and F. Guerra. General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A*, 31(46):9149–9155, 1998. MR1662161.

J. M. Hammersley. Postulates for subadditive processes. *Ann. Probability*, 2:652–680, 1974. MR0370721.

A. J. Harper. A note on the maximum of the Riemann zeta function, and log-correlated random variables. *Preprint*, pages 1–26, 2013. arXiv:1304.0677.

A. J. Harper. On the partition function of the Riemann zeta function, and the Fyodorov–Hiary–Keating conjecture. *Preprint*, pages 1–40, 2019. arXiv:1906.05783.

L. Hartung and A. Klimovsky. The glassy phase of the complex branching Brownian motion energy model. *Electron. Commun. Probab.*, 20:no. 78, 15, 2015. MR3417450.

L. Hartung and A. Klimovsky. The phase diagram of the complex branching Brownian motion energy model. *Electron. J. Probab.*, 23:Paper No. 127, 27, 2018. MR3896864.

W. Heap, M. Radziwiłł, and K. Soundararajan. Sharp upper bounds for fractional moments of the Riemann zeta function. *Preprint*, pages 1–10, 2019. arXiv:1901.08423.

X. Hu, J. Miller, and Y. Peres. Thick points of the Gaussian free field. *Ann. Probab.*, 38

(2):896–926, 2010. MR2642894.

Y. Hu and Z. Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.*, 37(2):742–789, 2009. MR2510023.

Aleksandar Ivić. *The Riemann zeta-function*. Dover Publications, Inc., Mineola, NY, 2003. Reprint of the 1985 original. MR1994094.

A. Jagannath. On the overlap distribution of branching random walks. *Electron. J. Probab.*, 21:Paper No. 50 16, 2016. MR3539644.

A. Jagannath. Approximate ultrametricity for random measures and applications to spin glasses. *Comm. Pure Appl. Math.*, 70(4):611–664, 2017. MR3628881.

J. Junnila and E. Saksman. Uniqueness of critical Gaussian chaos. *Electron. J. Probab.*, 22:Paper No. 11, 31, 2017. MR3613704.

Z. Kabluchko and A. Klimovsky. Complex random energy model: zeros and fluctuations. *Probab. Theory Related Fields*, 158(1-2):159–196, 2014a. MR3152783.

Z. Kabluchko and A. Klimovsky. Generalized Random Energy Model at complex temperatures. *Preprint*, pages 1–109, 2014b. arXiv:1402.2142.

J.-P. Kahane. Sur le chaos multiplicatif. *Ann. Sci. Math. Québec*, 9(2):105–150, 1985. MR829798.

J.-P. Kahane. Comments and corrections: "Multiplicative chaos" [Ann. Sci. Math. Québec **9** (1985), no. 2, 105–150]. *Ann. Sci. Math. Québec*, 10(2):117–118, 1986. MR877388.

J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2 + it)$. *Comm. Math. Phys.*, 214(1):57–89, 2000. MR1794265.

J. F. C. Kingman. The first birth problem for an age-dependent branching process. *Ann. Probability*, 3(5):790–801, 1975. MR0400438.

N. Kistler. Derrida's random energy models. From spin glasses to the extremes of correlated random fields. In *Correlated Random Systems: Five Different Methods*, volume 2143 of *Lecture Notes in Mathematics*. Springer, 2015. MR3380419.

N. Kistler and M. A. Schmidt. From Derrida's random energy model to branching random walks: from 1 to 3. *Elec. Comm. Prob.*, 20:no. 47, 12, 2015. MR3358969.

A. N. Kolmogorov, I. Petrovski, and N. Piscounov. Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Moscow Univ. Math. Bull.*, 1:1–25, 1937.

I. Komunjer. Asymmetric power distribution: theory and applications to risk measurement. *J. Appl. Econometrics*, 22(5):891–921, 2007. MR2395888.

I. Kurkova. Temperature dependence of the Gibbs state in the random energy model. *J. Statist. Phys.*, 111(1-2):35–56, 2003. MR1964264.

N. Kurt. Maximum and entropic repulsion for a Gaussian membrane model in the critical dimension. *Ann. Probab.*, 37(2):687–725, 2009. MR2510021.

P. Lafaye de Micheaux and F. Ouimet. A uniform L^1 law of large numbers for functions of i.i.d. random variables that are translated by a consistent estimator. *Statist. Probab. Lett.*, 142:109–117, 2018. MR3842623.

S. P. Lalley and T. Sellke. A conditional limit theorem for the frontier of a branching Brownian motion. *Ann. Probab.*, 15(3):1052–1061, 1987. MR893913.

K. Lange. Applications of the Dirichlet distribution to forensic match probabilities. *Genetica*, 96(1-2):107–117, 1995. doi:10.1007/BF01441156.

A. Laurinčikas. *Limit theorems for the Riemann zeta-function*, volume 352 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996. MR1376140.

A. Leblanc. A bias-reduced approach to density estimation using Bernstein polynomials. *J. Nonparametr. Stat.*, 22(3-4):459–475, 2010. MR2662607.

A. Leblanc. On estimating distribution functions using Bernstein polynomials. *Ann. Inst. Statist. Math.*, 64(5):919–943, 2012. MR2960952.

A. Leblanc and B. C. Johnson. On a uniformly integrable family of polynomials defined on the unit interval. *JIPAM. J. Inequal. Pure Appl. Math.*, 8(3): Article 67, 5pp., 2007. MR2345922.

T. Madaule. Convergence in law for the branching random walk seen from its tip. *J. Theoret. Probab.*, 30(1):27–63, 2017. MR3615081.

P. Maillard and O. Zeitouni. Slowdown in branching Brownian motion with inhomogeneous variance. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(3):1144–1160, 2016. MR3531703.

B. Mallein. Maximal displacement in a branching random walk through interfaces. *Electron. J. Probab.*, 20:no. 68, 40, 2015a. MR3361256.

B. Mallein. Maximal displacement of a branching random walk in time-inhomogeneous environment. *Stoch. Proc. Appl.*, 125(10):3958–4019, 2015b. MR3373310.

B. Mallein. Maximal displacement of d-dimensional branching Brownian motion. *Electron. Commun. Probab.*, 20:no. 76, 12, 2015c. MR3417448.

B. Mallein. Asymptotic of the maximal displacement in a branching random walk. *Grad. J. Math.*, 1(2):92–104, 2016. MR3850767.

B. Mallein. Genealogy of the extremal process of the branching random walk. *ALEA Lat. Am. J. Probab. Math. Stat.*, 15(2):1065–1087, 2018. MR3852245.

M. Mézard, G. Parisi, and M. A. Virasoro. *Spin glass theory and beyond*, volume 9 of *World Scientific Lecture Notes in Physics*. World Scientific Publishing Co., Inc., Teaneck, NJ, 1987. MR1026102.

A. A. Mogul'skiĭ. A local limit theorem for the first hitting time of a fixed level by a random walk. *Mat. Tr.*, 12(2):126–138, 2009. MR2599428.

H. L. Montgomery and R. C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007. MR2378655.

J. Najnudel. On the extreme values of the Riemann zeta function on random intervals of the critical line. *Probab. Theory Related Fields*, 172(1-2):387–452, 2018. MR3851835.

M. Nikula, E. Saksman, and C. Webb. Multiplicative chaos and the characteristic polynomial of the CUE: the L^1-phase. *Preprint*, pages 1–49, 2018. arXiv:1806.01831.

J. Nolen, J.-M. Roquejoffre, and L. Ryzhik. Power-like delay in time inhomogeneous Fisher-KPP equations. *Comm. Partial Differential Equations*, 40(3):475–505, 2015. MR3285242.

F. Ouimet. *Étude du maximum et des hauts points de la marche aléatoire branchante inhomogène et du champ libre gaussien inhomogène*. Master's thesis, Université de Montréal, 2014. URL `https://papyrus.bib.umontreal.ca/xmlui/handle/1866/11510`.

F. Ouimet. Geometry of the Gibbs measure for the discrete 2D Gaussian free field with scale-dependent variance. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(2):851–902, 2017. MR3731796.

F. Ouimet. Complete monotonicity of multinomial probabilities and its application to

Bernstein estimators on the simplex. *J. Math. Anal. Appl.*, 466(2):1609–1617, 2018a. MR3825458.

F. Ouimet. Poisson-Dirichlet statistics for the extremes of a randomized Riemann zeta function. *Electron. Commun. Probab.*, 23:Paper No. 46, 15, 2018b. MR3841407.

F. Ouimet. Maxima of branching random walks with piecewise constant variance. *Braz. J. Probab. Stat.*, 32(4):679–706, 2018c. MR3845025.

M. Pain and O. Zindy. Absence of temperature chaos for the 2D discrete Gaussian free field: an overlap distribution different from the random energy model. *Preprint*, pages 1–24, 2018. arXiv:1811.08313.

R.E.A.C. Paley and A. Zygmund. A note on analytic functions in the unit circle. *Math. Proc. Cambridge Philos. Soc.*, 28(3):266–272, 1932. doi:10.1017/S0305004100010112.

D. Panchenko. The Ghirlanda-Guerra identities for mixed p-spin model. *C. R. Math. Acad. Sci. Paris*, 348(3-4):189–192, 2010a. MR2600075.

D. Panchenko. On the Dovbysh-Sudakov representation result. *Electron. Commun. Probab.*, 15:330–338, 2010b. MR2679002.

D. Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013a. MR2999044.

D. Panchenko. *The Sherrington-Kirkpatrick Model*. Springer Monographs in Mathematics. Springer, New York, 2013b. MR3052333.

E. Paquette and O. Zeitouni. The maximum of the CUE field. *Int. Math. Res. Not. IMRN*, 2018(16):5028–5119, 2018. MR3848227.

R. Persechino. *Le modèle GREM à deux niveaux en présence d'un champ magnétique aléatoire*. PhD thesis, Université de Montréal, 2018. URL `https://papyrus.bib.umontreal.ca/xmlui/handle/1866/21150`.

M. Radziwiłł and K. Soundararajan. Selberg's central limit theorem for $\log|\zeta(1/2+it)|$. *Enseign. Math.*, 63(1-2):1–19, 2017. MR3832861.

S. I. Resnick. *Extreme values, regular variation and point processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original. MR2364939.

R. Rhodes and V. Vargas. KPZ formula for log-infinitely divisible multifractal random measures. *ESAIM Probab. Stat.*, 15:358–371, 2011. MR2870520.

R. Rhodes and V. Vargas. Gaussian multiplicative chaos and applications: a review. *Probab. Surv.*, 11:315–392, 2014. MR3274356.

R. Rhodes and V. Vargas. Gaussian multiplicative chaos and Liouville quantum gravity. In *Stochastic processes and random matrices*, pages 548–577. Oxford Univ. Press, Oxford, 2017. MR3728724.

R. Robert and V. Vargas. Gaussian multiplicative chaos revisited. *Ann. Probab.*, 38(2): 605–631, 2010. MR2642887.

M. I. Roberts. A simple path to asymptotics for the frontier of a branching Brownian motion. *Ann. Probab.*, 41(5):3518–3541, 2013. MR3127890.

R. T. Rockafellar. *Convex Analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970. MR0274683.

D. Ruelle. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.*, 108(2):225–239, 1987. MR875300.

H. Sakagawa. Entropic repulsion for a Gaussian lattice field with certain finite range interaction. *J. Math. Phys.*, 44(7):2939–2951, 2003. MR1982781.

H. Sakagawa. On the free energy of a Gaussian membrane model with external potentials. *J. Stat. Phys.*, 147(1):18–34, 2012. MR2922757.

H. Sakagawa. Localization of a Gaussian membrane model with weak pinning potentials. *ALEA Lat. Am. J. Probab. Math. Stat.*, 15(2):1123–1140, 2018. URL `http://alea.impa.br/articles/v15/15-41.pdf`.

E. Saksman and C. Webb. The Riemann zeta function and Gaussian multiplicative chaos: statistics on the critical line. *Preprint*, pages 1–85, 2018. arXiv:1609.00027.

F. Schweiger. The maximum of the four-dimensional membrane model. *Preprint*, pages 1–25, 2019. arXiv:1903.02522.

A. Selberg. Contributions to the theory of the Riemann zeta-function. *Arch. Math. Naturvid.*, 48(5):89–155, 1946. MR0020594.

A. Selberg. Old and new conjectures and results about a class of Dirichlet series. In *Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989)*, pages 367–385. Univ. Salerno, Salerno, 1992. MR1220477.

A. Shamov. On Gaussian multiplicative chaos. *J. Funct. Anal.*, 270(9):3224–3261, 2016. MR3475456.

S. Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields*, 139 (3-4):521–541, 2007. MR2322706.

Z. Shi. *Branching random walks*, volume 2151 of *Lecture Notes in Mathematics*. Springer, Cham, 2015a. Lecture notes from the 42nd Probability Summer School held in Saint Flour, 2012. MR0274683.

Z. Shi. Branching Brownian motion and the spinal decomposition. *Lecture notes*, 2015b. URL `https://www.lpsm.paris//pageperso/zhan/pdffile/DarmstadtBBM.pdf`.

K. Soundararajan. Moments of the Riemann zeta function. *Ann. of Math. (2)*, 170(2): 981–993, 2009. MR2552116.

M. Talagrand. *Spin Glasses: A Challenge for Mathematicians*, volume 46 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*. Springer-Verlag, Berlin, 2003. MR1993891.

M. Talagrand. *Mean Field Models for Spin Glasses. Volume I: Basic Examples*, volume 54. Springer-Verlag, Berlin, 2011a. MR2731561.

M. Talagrand. *Mean Field Models for Spin Glasses. Volume II: Advanced Replica-Symmetry and Low Temperature*, volume 55. Springer, Heidelberg, 2011b. MR3024566.

T. Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012. MR2906465.

T. Tao. The Poisson-Dirichlet process, and large prime factors of a random number. 2013. URL `https://terrytao.wordpress.com/2013/09/21/the-poisson-dirichlet-process-and-large-prime-factors-of-a-random-number/#more-7040`.

A. Tenbusch. Two-dimensional Bernstein polynomial density estimators. *Metrika*, 41(3-4): 233–253, 1994. MR1293514.

G. Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 163 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, third edition, 2015. MR3363366.

E. C. Titchmarsh. *The theory of the Riemann zeta-function*. The Clarendon Press, Oxford University Press, New York, second edition, 1986. MR882550.

K.-M. Tsang. *The distribution of the values of the Riemann zeta-function.* PhD thesis, Princeton University, 1984. URL `http://www.math.sjsu.edu/~goldston/TsangThesis.htm`.

S. A. van de Geer. *Applications of empirical process theory*, volume 6 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2000. MR1739079.

A. W. van der Vaart and J. A. Wellner. *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York, 1996. MR1385671.

C. Webb. The characteristic polynomial of a random unitary matrix and Gaussian multiplicative chaos—the L^2-phase. *Electron. J. Probab.*, 20:no. 104, 21, 2015. MR3407221.

O. Zeitouni. Branching random walks and Gaussian fields. *Notes for Lectures*, 2012. URL `http://www.wisdom.weizmann.ac.il/~zeitouni/pdf/notesBRW.pdf`.

O. Zeitouni. Gaussian fields. *Notes for Lectures*, 2017. URL `http://www.cims.nyu.edu/~zeitouni/notesGauss.pdf`.

Part 1 : Log-correlated Gaussian fields

Article 1

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

by

Louis-Pierre Arguin[1] and Frédéric Ouimet[2]

([1]) Baruch College and Graduate Center (CUNY), New York, NY 10010, USA.

([2]) Université de Montréal, Montréal, QC H3T 1J4, Canada.

Here is the complete citation :

Arguin, L.-P. and Ouimet, F. (2016). Extremes of the two-dimensional Gaussian free field with scale-dependent variance. ALEA, Lat. Am. J. Probab. Math. Stat. 13, no. 2, 779–808. http://alea.impa.br/articles/v13/13-31.pdf.

Contributions

I wrote the majority of the proofs and Louis-Pierre wrote the majority of the introduction. Louis-Pierre also contributed by introducing a better notation and by refining the presentation in certain places.

Innovations with respect to my master's thesis

This article extends the asymptotic results of Chapter 4 of my master's thesis from V_N^δ to V_N. It was the hardest part. The biggest innovations are Lemma 1.3.2 and Section 1.4.1, for which I am the biggest contributor.

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

1.1. Introduction

1.1.1. The model

Let $(W_k)_{k\geq 0}$ be a simple random walk starting at $u \in \mathbb{Z}^2$ with law $\mathscr{P}_u$. For every finite box $B \subseteq \mathbb{Z}^2$, the Gaussian free field (GFF) on B is a centered Gaussian field $\phi \mathrel{\hat=} \{\phi_v\}_{v\in B}$ with covariance matrix

$$G_B(u,v) \mathrel{\hat=} \frac{\pi}{2} \cdot \mathscr{E}_u \left[\sum_{k=0}^{\tau_{\partial B}-1} \mathbf{1}_{\{W_k = v\}} \right], \quad u,v \in B, \tag{1.1.1}$$

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

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

and $\|\cdot\|_2$ denotes the Euclidean distance in $\mathbb{Z}^2$. With this definition, B contains its boundary. We let $B^o \mathrel{\hat=} B \backslash \partial B$. By convention, summations are zero when there are no indices, so ϕ is identically zero on ∂B. This is the *Dirichlet boundary condition*. The constant $\pi/2$ in (1.1.1) is a convenient normalization for the variance.

In this paper, we consider a family of Gaussian fields constructed from the GFF $\{\phi_v\}_{v\in V_N}$ on the square box $V_N \mathrel{\hat=} \{0,1,...,N\}^2$. These Gaussian fields are the analogues, in the context of the GFF, of the time-inhomogeneous branching random walks studied in Bovier and Kurkova (2004); Fang and Zeitouni (2012a); Bovier and Hartung (2014); Ouimet (2018). We study the maxima and the number of high points of this family of Gaussian fields as $N \to \infty$.

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

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

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

$$\phi_v(\lambda) \mathrel{\hat=} \mathbb{E}\left[\phi_v \,|\, \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}\right],$$

then

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

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

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

$$\boldsymbol{\sigma} \mathrel{\hat=} (\sigma_1, \sigma_2, ..., \sigma_M) \in (0,\infty)^M, \qquad \text{(variance parameters)}$$

$$\boldsymbol{\lambda} \mathrel{\hat=} (\lambda_1, \lambda_2, ..., \lambda_M) \in (0,1]^M, \qquad \text{(scale parameters)}$$

where $0 \mathrel{\hat=} \lambda_0 < \lambda_1 < ... < \lambda_M \mathrel{\hat=} 1$. The parameters $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$ can be encoded simultaneously in the left-continuous step function

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

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

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

Definition 1.1.1 (Scale-inhomogeneous Gaussian free field)**.** *Let $\phi \doteq \{\phi_v\}_{v\in V_N}$ be the GFF on V_N. The $(\boldsymbol{\sigma},\boldsymbol{\lambda})$-GFF on V_N is a Gaussian field $\psi \doteq \{\psi_v\}_{v\in V_N}$ defined by*

$$\psi_v \doteq \sum_{i=1}^{M} \sigma_i \nabla\phi_v(\lambda_i) = \sum_{i=1}^{M} \sigma_i\Big(\phi_v(\lambda_i) - \phi_v(\lambda_{i-1})\Big). \tag{1.1.2}$$

Similarly to the GFF, we define

$$\psi_v(\lambda) \doteq \mathbb{E}\Big[\psi_v \,|\, \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}\Big].$$

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

1.1.2. Main results

The main results of this paper are the derivation of the first order of the maximum and the log-number of high points for the scale-inhomogeneous Gaussian free field of Definition 1.1.1. The methods of proof are general and directly applicable to time-inhomogeneous branching random walks and to other log-correlated Gaussian fields.

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

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

It turns out that the first order of the maximum and the log-number of high points are controlled by the *concavification* of $\mathcal{J}_{\sigma^2}(\cdot)$. Let $\hat{\mathcal{J}}_{\sigma^2}$ be the function whose graph is the concave hull of $\mathcal{J}_{\sigma^2}$. Its graph is an increasing and concave polygonal line, see Figure 1.1.1 for an example. There exists a unique non-increasing left-continuous step function $s \mapsto \bar{\sigma}(s)$ such that

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

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

$$0 \mathrel{\hat{=}} \lambda^0 < \lambda^1 < \ldots < \lambda^m \mathrel{\hat{=}} 1, \tag{1.1.3}$$

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

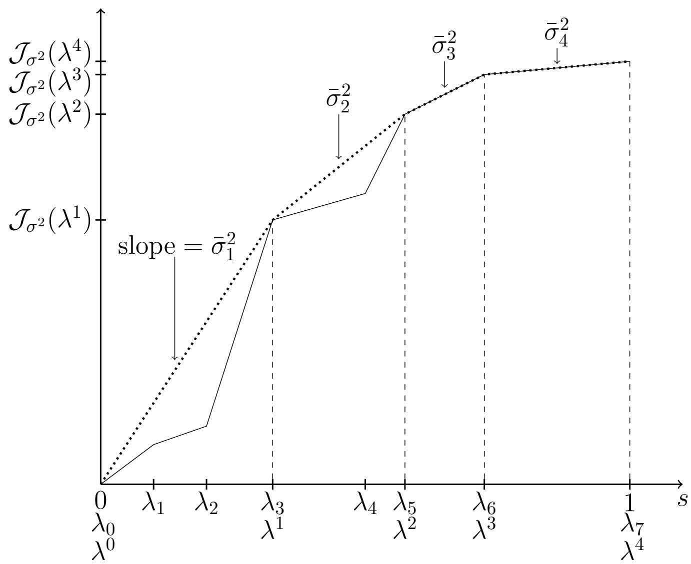


FIGURE 1.1.1. Example of $\mathcal{J}_{\sigma^2}$ (closed line) and $\hat{\mathcal{J}}_{\sigma^2}$ (dotted line) with 7 values for σ^2.

Theorem 1.1.2 (First order of the maximum). *Let $\{\psi_v\}_{v\in V_N}$ be the $(\boldsymbol{\sigma},\boldsymbol{\lambda})$-GFF on V_N of Definition 1.1.1, then*

$$\lim_{N\to\infty} \frac{\max_{v\in V_N}\psi_v}{\log N^2} = \mathcal{J}_{\sigma^2/\bar{\sigma}}(1) \mathrel{\hat{=}} \gamma^\star \quad \textit{in probability.}$$

In the homogeneous case where $M = 1$ and $\sigma_1 = 1$, the result reduces to $\gamma^\star = 1$, as proved in Bolthausen et al. (2001), which corresponds to the first order of the maximum of N^2 i.i.d. Gaussian variables of mean 0 and variance $\log N$. Note that the result of Theorem 1.1.2 can be written as follows :

$$\gamma^\star = \mathcal{J}_{\sigma^2/\bar{\sigma}}(1) = \sum_{l=1}^{m} \int_{\lambda^{l-1}}^{\lambda^l} \frac{\sigma^2(s)}{\bar{\sigma}(s)} ds = \int_0^1 \bar{\sigma}(s) ds. \tag{1.1.4}$$

This is simply a weighted average of homogeneous cases on the intervals $[\lambda^{l-1},\lambda^l]$ with variance parameter $\bar{\sigma}_l$. We say that $s \mapsto \bar{\sigma}^2(s)$ act as the *effective variance* of the field.

We stress that $\gamma^\star$ is strictly smaller than $\bar{\sigma}_1$ in cases where the concave hull is not a straight line. In particular, the upper bound on the level of the maximum cannot be proved by a simple union bound as in the homogeneous case.

The set of γ-high points of the field ψ is defined as

$$\mathcal{H}_N^\gamma \mathrel{\hat{=}} \{v \in V_N \,|\, \psi_v \geq \gamma \log N^2\}, \quad \text{for all } 0 \leq \gamma < \gamma^\star.$$

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

$$\gamma^l \mathrel{\hat{=}} \int_0^1 \frac{\sigma^2(s)}{\bar{\sigma}(s \wedge \lambda^l)} ds = \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^l) + \frac{\mathcal{J}_{\sigma^2}(\lambda^l, 1)}{\bar{\sigma}_l}, \quad 1 \leq l \leq m, \quad \gamma^0 \mathrel{\hat{=}} 0. \tag{1.1.5}$$

Theorem 1.1.3 (Log-number of high points or Entropy)**.** *Let $\{\psi_v\}_{v \in V_N}$ be the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF on V_N of Definition 1.1.1 and let $\gamma^{l-1} \leq \gamma < \gamma^l$ for some $l \in \{1, ..., m\}$, then*

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

The homogeneous case where $M = 1$ and $\sigma_1 = 1$ was proved in Daviaud (2006). In that case, we have $\mathcal{E}_\gamma = 1 - \gamma^2$ as for N^2 i.i.d. Gaussian variables of mean 0 and variance $\log N$. The proofs of Theorems 1.1.2 and 1.1.3 are deferred to Section 1.3. The method of proof is explained in Section 1.2. It is a refinement of the second moment method based on the control of the increments of high points at every scale. The method was used in Kistler (2015) to obtain a new proof of the first order of the maximum in the homogeneous case. Here we extend this method to the log-number of high points in all settings and to the first order of the maximum in the inhomogeneous setting. In the scale-dependent case, as opposed to the homogeneous case, it is necessary to truncate the first moment using the information at every scale λ^l to get the correct upper bound.

1.1.3. Related works and conjectures

The scale-inhomogeneous GFF is the equivalent of the time-inhomogeneous branching random walk (IBRW) where the variance of the random walk is a function of time. In particular, Theorems 1.1.2 and 1.1.3 can be proved for branching random walks using the same technique, see Section 2 of Ouimet (2014). In fact, much more precise information is known about the maxima of these models. In Bovier and Kurkova (2004), the authors introduce a continuous version of Derrida's *Generalized Random Energy Model* (GREM) Derrida (1985), which is akin to a time-inhomogeneous branching random walk, for which they obtain the first order of the maximum and the free energy. In particular, they noticed the concavification phenomenon for the first order. This observation also appears in Capocaccia et al. (1987) for the GREM. A model interpolating between the GREM and the branching random walk was introduced in Kistler and Schmidt (2015) where Poisson statistics of the extremes are proved. For Gaussian IBRWs with two values of the variance ($M = 2$), the lower order corrections for the maximum and tightness of the law were proved in Fang and Zeitouni (2012a). In this case, convergence of the extremal processes and of the law of the recentered maximum have been shown in Bovier and Hartung (2014). This is also proved in the case where the integral of the variance remains strictly below its concave hull (for example, in the case of increasing variances), see Bovier and Hartung (2015). For strictly decreasing variances, the lower order corrections for IBBMs exhibit a slowdown of the order $t^{1/3}$ as proved in Fang and Zeitouni (2012b); Maillard and Zeitouni (2016). Similar results for non-Gaussian IBRWs and more general variances are proved in Mallein (2015), though not at the level of convergence of the law. In Ouimet (2018), the second order of the maximum for the Gaussian IBRW with a finite number of variances is shown by generalizing the approach of Fang and Zeitouni (2012a) and the tightness follows from Fang (2012).

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

One interest of Definition 1.1.1 for the scale-inhomogeneous GFF is that it can be extended to a piecewise smooth variance function $\sigma : [0,1] \to [a,b]$ where $a > 0$. Consider the two-dimensional continuous Gaussian free field $\phi = \{\phi_v\}_{v\in[0,1]^2}$ on the unit square $[0,1]^2$, see e.g. Sheffield (2007) for a definition. The field ϕ cannot be defined as a random function. However, averages over sets make sense as random variables. In particular, for every $v \in [0,1]^2$ and $\lambda \in [0,1]$, one can define $\phi_v^r(\lambda)$ as the average of the field over a circle of radius r^λ :

$$\phi_v^r(\lambda) \mathrel{\hat{=}} \frac{1}{2\pi r^\lambda} \int_0^{2\pi} \phi_{v+r^\lambda e^{i\theta}} \, d\theta. \tag{1.1.6}$$

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

$$\psi_v^r(1) \mathrel{\hat{=}} \int_0^1 \sigma(\lambda) \, d\phi_v^r(\lambda), \quad v \in [0,1]^2.$$

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

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

where it is assumed that the continuous Green function on $[0,1]^2$ associated to ϕ has been normalized as in (1.1.1). This is the notion analogous to γ-high points. It was shown in Hu et al. (2010) that the Hausdorff dimension of the set of γ-thick points is $2(1-\gamma^2)$ when $\sigma \equiv 1$. In view of Theorem 1.1.3, it is reasonable to conjecture that the Hausdorff dimension of the set of γ-thick points of ψ is

$$2\left((1-\lambda_\star) - \frac{(\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda_\star))^2}{\mathcal{J}_{\sigma^2}(\lambda_\star, 1)}\right), \tag{1.1.7}$$

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

1.2. Outline of Proof

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

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

$$\rho(v, v') \mathrel{\mathring{=}} \max\{\lambda \in [0,1] : [v]_\lambda \cap [v']_\lambda \neq \emptyset\}. \tag{1.2.1}$$

This is the largest λ for which the two neighborhoods $[v]_\lambda$ and $[v']_\lambda$ intersect. We always have by definition that $\|v - v'\|_2$ is of order $N^{1-\rho(v,v')}$. The branching scale plays the same role as the branching time (normalized to lie in $[0,1]$) in branching random walk. More precisely, let $\{\phi_v\}_{v \in V_N}$ be a homogeneous GFF and consider the increments $\phi_v(\lambda') - \phi_v(\lambda)$ and $\phi_{v'}(\mu') - \phi_{v'}(\mu)$ for some choice of $\lambda < \lambda'$ and $\mu < \mu'$. The Markov property of the Gaussian free field (see Section 1.4.1) implies that for $\lambda, \mu > \rho(v, v')$,

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

because the neighborhoods $[v]_\lambda$ and $[v']_\mu$ are disjoint, see Figure 1.2.2. This means that the increments after the branching scale are independent.

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

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

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

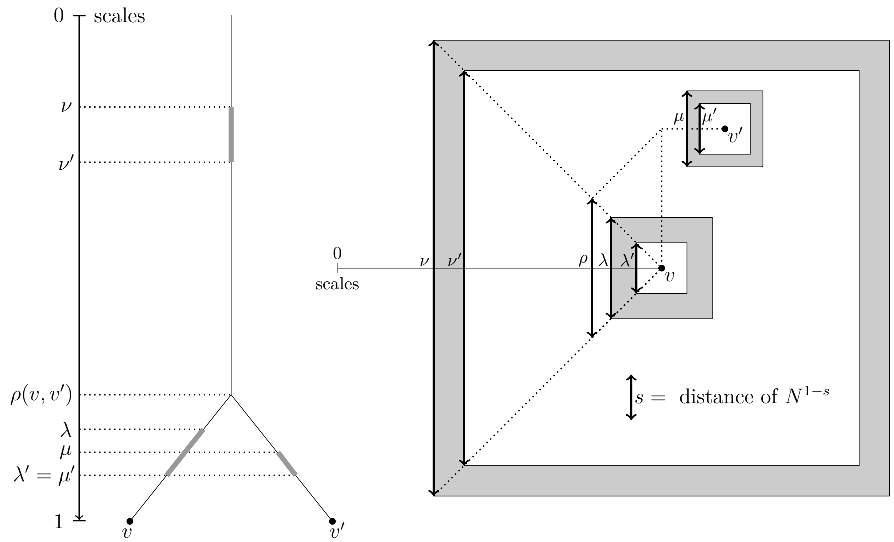


FIGURE 1.2.2. The branching structure of the GFF.

Of course, the branching structure here is not exact as in branching random walk. In particular, nothing precise can be said on the increments $\phi_v(\lambda') - \phi_v(\lambda)$ and $\phi_{v'}(\lambda') - \phi_{v'}(\lambda)$ in the case where $\lambda < \rho < \lambda'$. However, the contribution of such increments can be made negligible by considering a large number of increments, as we shall do. This branching structure holds also for the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF, since it is defined in terms of the increments of ϕ, see (1.1.2) and Lemma 1.4.1.

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

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

$$\Lambda_{N,M} \mathrel{\mathring{=}} \{v \in V_N \,|\, \nabla\psi_v(\lambda_i) \geq \nabla\gamma_i \log N^2 \text{ for all } i \in \{1, 2, \ldots, M\}\},$$

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

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

There are $O(N^{2\lambda_k})$ representatives at scale λ_k and the variance of the increments is $\mathbb{V}(\nabla\psi_v(\lambda_i)) = \sigma_i^2 \nabla\lambda_i \log N + O(1)$ if we ignore the boundary effect. Therefore, using the independence between the increments and standard Gaussian estimates (see Lemma 1.4.7, it will be used repeatedly) :

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

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

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

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

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

This optimization problem can be solved using the Karush-Kuhn-Tucker theorem (see Lemma 1.4.10). We write $(\gamma_1^\star, \gamma_2^\star, ..., \gamma_M^\star)$ for the unique solution. We will make extensive use of the polygonal line $L_N^\star(\cdot)$ linking the points $(0,0)$, $(\lambda_1, \gamma_1^\star \log N^2)$, $(\lambda_2, \gamma_2^\star \log N^2)$, $\dots$, $(\lambda_M, \gamma_M^\star \log N^2)$ to prove Theorem 1.1.2 and 1.1.3 :

$$L_N^\star(s) \doteq \int_0^s \frac{\sigma^2(r)}{\bar{\sigma}(r)} dr \log N^2 = \mathcal{J}_{\sigma^2/\bar{\sigma}}(s) \log N^2, \quad s \in [0,1]. \tag{1.2.2}$$

This is the *optimal path for the maximum*. Figure 1.2.3 shows an example of such a path. In particular, it is important to note that the optimal path coincides with its concave hull at each scale λ^l, namely

$$L_N^\star(\lambda^l) = \hat{L}_N^\star(\lambda^l) = \hat{\mathcal{J}}_{\sigma^2/\bar{\sigma}}(\lambda^l) \log N^2 = \mathcal{J}_{\bar{\sigma}}(\lambda^l) \log N^2, \quad 1 \leq l \leq m. \tag{1.2.3}$$

The same heuristic can be used to determine the optimal path $L_N^\gamma(\cdot)$ for γ-high points,

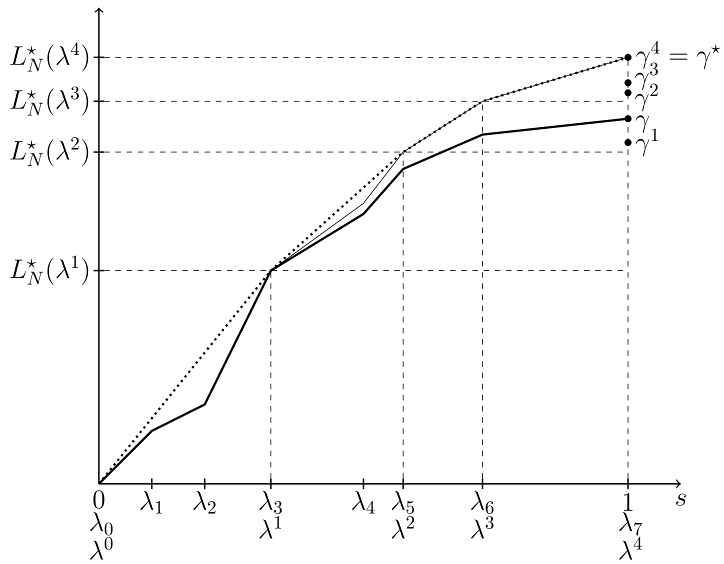


FIGURE 1.2.3. Example of L_N^γ (bold line), $L_N^\star$ (thin line) and its concavified version $\hat{L}_N^\star$ (dotted line), with 7 values for σ^2 and $\gamma^1 < \gamma < \gamma^2$.

$0 < \gamma < \gamma^\star$. Setting now $\gamma_M = \gamma$, we get

$$\lim_{N\to\infty} \frac{\log(\mathbb{E}[|\Lambda_{N,M}|])}{\log N^2} = \sum_{i=1}^{M-1} \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i} \right) + \left(\nabla\lambda_M - \frac{(\gamma - \gamma_{M-1})^2}{\sigma_M^2 \nabla\lambda_M} \right). \tag{1.2.4}$$

A lower bound for the log-number of γ-high points can be found by maximizing (1.2.4) with respect to $\gamma_1, \gamma_2, \ldots, \gamma_{M-1}$ and under the constraints

$$\sum_{i=1}^{k} \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i} \right) \geq 0, \quad 1 \leq k \leq M-1. \tag{1.2.5}$$

The unique solution to this problem is found in Lemma 1.4.11 using again the Karush-Kuhn-Tucker theorem. The form of the path will always depend on the critical levels defined in (1.1.5). Whenever $\gamma^{l-1} \leq \gamma < \gamma^l$, the *optimal path for γ-high points* is :

$$L_N^\gamma(s) \mathrel{\hat{=}} \begin{cases} \mathcal{J}_{\sigma^2/\bar{\sigma}}(s) \log N^2, & 0 \leq s \leq \lambda^{l-1}, \\ \left(\mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}) + \frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, s)}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} (\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1})) \right) \log N^2, & \lambda^{l-1} \leq s \leq 1. \end{cases} \tag{1.2.6}$$

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

1.3. Proofs of the main results

1.3.1. Preliminaries

For all $\lambda \in [0,1]$, recall that $\psi_v(\lambda) \mathrel{\hat=} \mathbb{E}[\psi_v \,|\, \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}]$. By the Markov property of the GFF (see Lemma 1.4.1), it is not hard to show that for any partition $0 \mathrel{\hat=} s_0 < s_1 < \ldots < s_K \mathrel{\hat=} 1$ of $[0,1]$ such that $\{\lambda_i\}_{i=0}^M \subseteq \{s_j\}_{j=0}^K$, we have for all $1 \le k \le l \le K$:

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

In particular, the independence of the increments of ψ follows directly from the one for ϕ. Moreover, using standard estimates on Green functions, Lemma 1.4.2 shows that

$$-C_1(\delta) \le \mathbb{V}(\psi_v(s_l) - \psi_v(s_{k-1})) - \mathcal{J}_{\sigma^2}(s_{k-1}, s_l) \log N \le C_2 \tag{1.3.1}$$

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

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

The set V_N^δ contains the points that are at a distance at least δN from the boundary of V_N. Lemma 1.4.3 proves that the upper bound in (1.3.1) holds on V_N, that is

$$\max_{v \in V_N} \mathbb{V}(\psi_v(s_l) - \psi_v(s_{k-1})) \le \mathcal{J}_{\sigma^2}(s_{k-1}, s_l) \log N \; + \; C \tag{1.3.2}$$

for N large enough.

Remark 1.3.1. *Throughout the proofs, c and C will denote positive constants whose value can change at different occurrences and might depend on the parameters $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$. For simplicity, equations in the proofs are implicitly stated to hold for N large enough where it is needed.*

1.3.2. First order of the maximum

Theorem 1.1.2 is a direct consequence of Lemma 1.3.1, which proves that $\gamma^\star \log N^2$ is an upper bound on the first order of the maximum, and Lemma 1.3.3 which shows the corresponding lower bound.

Lemma 1.3.1 (Upper bound on the first order of the maximum). *Let $\{\psi_v\}_{v\in V_N}$ be the $(\boldsymbol{\sigma},\boldsymbol{\lambda})$-GFF on V_N of Definition 1.1.1 and $\gamma^\star$ as in Theorem 1.1.2. For all $\varepsilon > 0$, there exists a constant $c = c(\varepsilon,\boldsymbol{\sigma},\boldsymbol{\lambda}) > 0$ such that*

$$\mathbb{P}\Big(\max_{v\in V_N}\psi_v \geq (\gamma^\star + m\varepsilon)\log N^2\Big) \leq N^{-c} \tag{1.3.3}$$

for N large enough.

Proof. Recall the definition of the optimal path $L_N^\star$ from (1.2.2) and define

$$L_N^{\star,z}(s) \mathrel{\hat=} L_N^\star(s) + z\log N^2, \quad s\in[0,1].$$

Recall the definition of λ^j in (1.1.3) and the notation R_{λ^j} for the set of representatives at scale λ^j. Consider the set of representatives whose value reached just over the optimal level at λ^j :

$$\mathcal{H}_{N,j}^{\star,\varepsilon} \mathrel{\hat=} \Big\{v\in R_{\lambda^j}\,|\,\psi_v(\lambda^j) \geq L_N^{\star,j\varepsilon}(\lambda^j)\Big\}, \quad 1\leq j\leq m.$$

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

$$A_\varepsilon \mathrel{\hat=} \Big\{|\psi_v(\lambda^j) - \psi_{v_{\lambda^j}}(\lambda^j)| \leq \eta_\varepsilon \log N^2 \ \text{ for all } j\in\{1,...,m\} \text{ and all } v\in V_N\Big\}$$

where v_λ denotes any representative in R_λ that is closest to v. Here we introduced the event A_ε to approximate the branching structure of the field ψ. Since $R_{\lambda^m} = V_N$ by definition and $L_N^{\star,m\varepsilon}(\lambda^m) = (\gamma^\star + m\varepsilon)\log N^2$, a union bound gives the following upper bound on the probability in (1.3.3) :

$$\begin{aligned}
&\mathbb{P}\Big(|\mathcal{H}_{N,m}^{\star,\varepsilon}|\geq 1\Big) \leq \mathbb{P}(A_\varepsilon^c) + \sum_{l=1}^m \mathbb{P}\Big(\Big\{|\mathcal{H}_{N,1}^{\star,\varepsilon}| = ... = |\mathcal{H}_{N,l-1}^{\star,\varepsilon}| = 0, |\mathcal{H}_{N,l}^{\star,\varepsilon}|\geq 1\Big\}\cap A_\varepsilon\Big)\\
&\leq \mathbb{P}(A_\varepsilon^c) + \sum_{l=1}^m \mathbb{P}\left(\left\{\begin{array}{l}\exists v\in R_{\lambda^l}\cap V_N^o \ \text{ s.t. } \psi_v(\lambda^l)\geq L_N^{\star,l\varepsilon}(\lambda^l) \text{ and}\\ \psi_{v_{\lambda^j}}(\lambda^j) < L_N^{\star,j\varepsilon}(\lambda^j) \ \text{ for all } 1\leq j\leq l-1\end{array}\right\}\cap A_\varepsilon\right)\\
&\leq Ce^{-c(\eta_\varepsilon)(\log N)^2} + \sum_{l=1}^m N^{2\lambda^l}\max_{v\in R_{\lambda^l}\cap V_N^o}\mathbb{P}\left(\left\{\begin{array}{l}\psi_v(\lambda^l)\geq L_N^{\star,l\varepsilon-\eta_\varepsilon}(\lambda^l) \text{ and}\\ \psi_v(\lambda^j) < L_N^{\star,j(\varepsilon+\eta_\varepsilon)}(\lambda^j)\\ \text{for all } 1\leq j\leq l-1\end{array}\right\}\right)
\end{aligned} \tag{1.3.4}$$

The bound on $\mathbb{P}(A_\varepsilon^c)$ follows easily from a union bound (with $m\cdot(N+1)^2$ terms), Gaussian estimates (Lemma 1.4.7) and the variance estimates of Lemma 1.4.6.

It remains to consider the terms in the sum in (1.3.4). We look at the case $l=1$. Since $\max_{v\in V_N}\mathbb{V}(\psi_v(\lambda^1))\leq \lambda^1\bar{\sigma}_1^2\log N + C$ from (1.3.2) and $L_N^\star(\lambda^1)=\lambda^1\bar{\sigma}_1\log N^2$, a Gaussian estimate shows that

$$\begin{aligned}\mathbb{P}\Big(\psi_v(\lambda^1)\geq L_N^{\star,\varepsilon-\eta_\varepsilon}(\lambda^1)\Big) &\leq \frac{\sqrt{\mathbb{V}(\psi_v(\lambda^1))}}{L_N^{\star,\varepsilon-\eta_\varepsilon}(\lambda^1)}\exp\left(-\frac{(L_N^{\star,\varepsilon-\eta_\varepsilon}(\lambda^1))^2}{2\mathbb{V}(\psi_v(\lambda^1))}\right)\\ &\leq \frac{C}{\sqrt{\log N}}N^{-2\lambda^1}N^{-4\frac{(\varepsilon-\eta_\varepsilon)}{\bar{\sigma}_1}}.\end{aligned}$$

After multiplying by $N^{2\lambda^1}$, we conclude that the $l=1$ term in (1.3.4) goes to 0 like $N^{-c(\varepsilon)}$. We now show a similar estimate for a fixed $l\in\{2,...,m\}$. To simplify the notation, denote $(X_v^1,...,X_v^l)\circeq(\psi_v(\lambda^1),...,\psi_v(\lambda^l))$. By conditioning on the value of the vector $\boldsymbol{X}\circeq(X_v^1,...,X_v^{l-1})$, the probability in (1.3.4) is equal to

$$\int_{-\infty}^{L_N^{\star,1(\varepsilon+\eta_\varepsilon)}(\lambda^1)}\dots\int_{-\infty}^{L_N^{\star,(l-1)(\varepsilon+\eta_\varepsilon)}(\lambda^{l-1})}\mathbb{P}\Big(X_v^l\geq L_N^{\star,l\varepsilon-\eta_\varepsilon}(\lambda^l)\,|\,\boldsymbol{X}=\boldsymbol{x}\Big)\,f_v(\boldsymbol{x})\,d\boldsymbol{x}$$

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

$$\int_{-\infty}^{L_N^{\star,1(\varepsilon+\eta_\varepsilon)}(\lambda^1)}\dots\int_{-\infty}^{L_N^{\star,(l-1)(\varepsilon+\eta_\varepsilon)}(\lambda^{l-1})}\mathbb{P}\Big(\nabla X_v^l\geq L_N^{\star,l\varepsilon-\eta_\varepsilon}(\lambda^l)-x_{l-1}\Big)\,f_v(\boldsymbol{x})\,d\boldsymbol{x}. \tag{1.3.5}$$

Since $l\varepsilon-\eta_\varepsilon=(\varepsilon-l\eta_\varepsilon)+(l-1)(\varepsilon+\eta_\varepsilon)$, a Gaussian estimate and the bound $\max_{v\in V_N}\mathbb{V}\Big(\nabla X_v^l\Big)\leq\bar{\sigma}_l^2\nabla\lambda^l\log N+C$ from (1.3.2) give

$$\begin{aligned}&\mathbb{P}\Big(\nabla X_v^l\geq L_N^{\star,l\varepsilon-\eta_\varepsilon}(\lambda^l)-x_{l-1}\Big)\\ &\quad\leq\frac{\sqrt{\mathbb{V}(\nabla X_v^l)}}{L_N^{\star,l\varepsilon-\eta_\varepsilon}(\lambda^l)-x_{l-1}}\exp\left(\frac{-(\nabla L_N^\star(\lambda^l)+L_N^{\star,l\varepsilon-\eta_\varepsilon}(\lambda^{l-1})-x_{l-1})^2}{2\mathbb{V}(\nabla X_v^l)}\right)\\ &\quad\leq\frac{C}{\sqrt{\log N}}N^{-2\nabla\lambda^l}\exp\left(-2\frac{(L_N^{\star,l\varepsilon-\eta_\varepsilon}(\lambda^{l-1})-x_{l-1})}{\bar{\sigma}_l}\right)\\ &\quad=\frac{C}{\sqrt{\log N}}N^{-2\nabla\lambda^l}N^{-4\frac{\varepsilon-l\eta_\varepsilon}{\bar{\sigma}_l}}\exp\left(-2\frac{(L_N^{\star,(l-1)(\varepsilon+\eta_\varepsilon)}(\lambda^{l-1})-x_{l-1})}{\bar{\sigma}_l}\right).\end{aligned} \tag{1.3.6}$$

To get the second inequality, we bounded the ratio using

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

from the integration limits of x_{l-1} in (1.3.5). It is convenient to do the change of variables $Y_{v,j} \doteq (\varepsilon+\eta_\varepsilon)\log N^2 + \nabla L_N^\star(\lambda^j) - \nabla X_v^j$ for all $j \in \{1, ..., l-1\}$. Equation (1.3.5) is then bounded, using (1.3.6), by

$$\frac{CN^{-4\frac{\varepsilon-l\eta_\varepsilon}{\bar{\sigma}_l}}}{N^{2\lambda^l}\sqrt{\log N}}\; N^{2\lambda^{l-1}} \int_0^\infty \int_{-y_1}^\infty ... \int_{-\sum_{j=1}^{l-2} y_j}^\infty \prod_{j=1}^{l-1} e^{-2\frac{y_j}{\bar{\sigma}_l}}\; \frac{e^{-\frac{\left((y_j-(\varepsilon+\eta_\varepsilon)\log N^2)-\nabla L_N^\star(\lambda^j)\right)^2}{2\mathbb{V}\left(Y_{v,j}\right)}}}{\sqrt{2\pi\mathbb{V}(Y_{v,j})}} d\boldsymbol{y}. \tag{1.3.7}$$

After multiplying by $N^{2\lambda^l}$, the l-th term of the sum in (1.3.4) has the right decay if we show that the integral in (1.3.7) is bounded by $\tilde{C}N^{-2\lambda^{l-1}}$. From (1.3.2), we have

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

for all $v \in V_N^o$. If the variances were all equal to $\bar{\sigma}_j^2\nabla\lambda^j \log N + C$, the argument would be simpler. Extra work is needed to take care of the boundary effect of the GFF. We gather the result into a lemma for later use in the proof of Lemma 1.3.4. □

Lemma 1.3.2. *Let $2 \leq l \leq m$ and $\boldsymbol{z} \doteq (z_j)_{j=1}^{l-1}$ be such that $0 < z_j \leq \bar{\sigma}_j^2\nabla\lambda^j \log N + C$. For all $\tilde{\varepsilon} > 0$, consider the integral*

$$I_{\tilde{\varepsilon}}(\boldsymbol{z}) \doteq \int_0^\infty g_1(y_1) ... \int_{-\sum_{j=1}^{l-3} y_j}^\infty g_{l-2}(y_{l-2}) \int_{-\sum_{j=1}^{l-2} y_j}^\infty e^{-2a_l \sum_{j=1}^{l-1} y_j} g_{l-1}(y_{l-1})\; d\boldsymbol{y}$$

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

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

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

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

$$e^{-2a_l \sum_{j=1}^{l-2} y_j} \int_{-\sum_{j=1}^{l-2} y_j}^\infty e^{-2a_l y_{l-1}} \frac{1}{\sqrt{2\pi z_{l-1}}} e^{-\frac{1}{2z_{l-1}}\left((y_{l-1}-\tilde{\varepsilon}\log N^2)-\nabla L_N^\star(\lambda^{l-1})\right)^2} dy_{l-1}$$

$$\le e^{-2a_l \sum_{j=1}^{l-2} y_j} \frac{1}{\sqrt{z_{l-1}}} e^{-\frac{1}{2z_{l-1}}(\nabla L_N^\star(\lambda^{l-1}))^2} \int_{-\sum_{j=1}^{l-2} y_j}^{\infty} e^{-2(a_l-\beta_{l-1})y_{l-1}} dy_{l-1}$$

$$= e^{-2a_l \sum_{j=1}^{l-2} y_j} \frac{1}{\sqrt{z_{l-1}}} e^{-\frac{1}{2z_{l-1}}(\nabla L_N^\star(\lambda^{l-1}))^2} \frac{1}{2(a_l-\beta_{l-1})} e^{2(a_l-\beta_{l-1})\sum_{j=1}^{l-2} y_j}.$$

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

$$C e^{-2\beta_{l-1} \sum_{j=1}^{l-2} y_j} N^{-2\nabla \lambda^{l-1}}. \tag{1.3.8}$$

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

$$e^{-2a_l \sum_{j=1}^{l-2} y_j} \int_{-\sum_{j=1}^{l-2} y_j}^{\infty} \frac{e^{-2a_l y_{l-1}}}{\sqrt{2\pi z_{l-1}}} e^{-\frac{\left((y_{l-1}-\tilde{\varepsilon}\log N^2)-\nabla L_N^\star(\lambda^{l-1})\right)^2}{2z_{l-1}}} dy_{l-1}$$

$$\le e^{-2a_l \sum_{j=1}^{l-2} y_j} \int_{-\sum_{j=1}^{l-2} y_j}^{\infty} \frac{e^{-2a_l (y_{l-1}-\tilde{\varepsilon}\log N^2)}}{\sqrt{2\pi z_{l-1}}} e^{-\frac{\left((y_{l-1}-\tilde{\varepsilon}\log N^2)-\nabla L_N^\star(\lambda^{l-1})\right)^2}{2z_{l-1}}} dy_{l-1}$$

$$= e^{-2a_l \sum_{j=1}^{l-2} y_j} e^{-2a_l \nabla L_N^\star(\lambda^{l-1})} e^{2a_l^2 z_{l-1}} \cdot \int_{-\sum_{j=1}^{l-2} y_j}^{\infty} \frac{1}{\sqrt{2\pi z_{l-1}}} e^{-\frac{\left((y_{l-1}-\tilde{\varepsilon}\log N^2)-\left(\nabla L_N^\star(\lambda^{l-1})-2a_l z_{l-1}\right)\right)^2}{2z_{l-1}}} dy_{l-1}$$

$$\le \exp\Big(-2a_l \sum_{j=1}^{l-2} y_j - 2a_l \nabla L_N^\star(\lambda^{l-1}) + 2a_l^2 z_{l-1}\Big).$$

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

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

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

$$e^{-2a_l \sum_{j=1}^{l-2} y_j} N^{-2\nabla\lambda^{l-1}}. \tag{1.3.9}$$

The second integral with respect to y_{l-2} in $I_{\tilde{\varepsilon}}(\boldsymbol{z})$ is evaluated similarly using (1.3.8) and (1.3.9) by taking $a_{l-1} \mathrel{\hat{=}} \min\{a_l, \beta_{l-1}\}$ and considering whether $a_{l-1} - \beta_{l-2} \ge 1/\sqrt{z_{l-2}}$ or not. Note that $a_{l-1} > 1/\bar{\sigma}_{l-2}$ holds since $a_l > 1/\bar{\sigma}_{l-1} > 1/\bar{\sigma}_{l-2}$ (because the steps of $\bar{\sigma}$ are decreasing in height) and $\beta_{l-1} \ge 1/\bar{\sigma}_{l-1} - O((\log N)^{-1}) > 1/\bar{\sigma}_{l-2}$ from the bounds on z_{l-1}. This recursive reasoning shows that $I_{\tilde{\varepsilon}}(\boldsymbol{z})$ is smaller than $\tilde{C} N^{-2\lambda^{l-1}}$. $\square$

Lemma 1.3.3 (Lower bound on the first order of the maximum)**.** *Let* $\{\psi_v\}_{v\in V_N}$ *be the* $(\boldsymbol{\sigma},\boldsymbol{\lambda})$*-GFF on* V_N *of Definition 1.1.1 and* $\gamma^\star$ *as in Theorem 1.1.2. For all* $0<\varepsilon<1$*, there exists a constant* $c=c(\varepsilon,\boldsymbol{\sigma},\boldsymbol{\lambda})>0$ *such that*

$$\mathbb{P}\Big(\max_{v\in V_N}\psi_v \le (1-\varepsilon)\gamma^\star \log N^2\Big)\le N^{-c} \tag{1.3.10}$$

for N *large enough.*

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

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

As a particular case of Lemma 1.4.3, note that

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

If we can show Lemma 1.3.3 when the λ_i's are rational numbers, then a union bound and a Gaussian estimate yield

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

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

The proof of Lemma 1.3.3 is based on a coarse-graining of the scales introduced in Kistler (2015). Consider $\alpha_k\mathrel{\hat{=}}\frac{k}{K}$, $0\le k\le K$. The parameter $K\in\mathbb{N}$ will be chosen large enough depending on ε during the proof. By the argument above, we can assume that $\lambda_i K\in\mathbb{N}_0$ for all $i\in\{0,...,M\}$, so that the α_k's form a finer partition of $[0,1]$ than the

λ_i's. The bounds in (1.3.1) imply that for all $k \in \{1, ..., K\}$ and for all $v \in V_N^\delta$:

$$|\mathbb{V}(\nabla\psi_v(\alpha_k)) - \sigma^2(\alpha_k)\nabla\alpha_k \log N| \leq C(\delta). \tag{1.3.11}$$

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

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

The proof relies on the Paley-Zygmund inequality (see Lemma 1.4.8) applied to a *modified number of exceedances*. In fact, we consider only points in V_N^δ whose increments are almost optimal. Moreover, and crucially, we drop the first r increments. We will choose r during the proof. This allows more independence between the variables of the field, which is needed to find a tight lower bound using the Paley-Zygmund inequality. More precisely, define

$$\mathcal{N}^\star_\varepsilon \mathrel{\hat{=}} \sum_{v \in V_N^\delta} \mathbf{1}_{A_v} \quad \text{where} \quad A_v \mathrel{\hat{=}} \{\nabla\psi_v(\alpha_j) \geq \nabla L^\star_{N,\varepsilon}(\alpha_j) \;\; \forall j \in \{r+1, ..., K\}\}.$$

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

$$\mathbb{P}\left(\max_{v \in V_N} \psi_v \geq (1-3\varepsilon)\gamma^\star \log N^2\right) \geq \mathbb{P}(\mathcal{N}^\star_\varepsilon \geq 1) - O(N^{-c}). \tag{1.3.12}$$

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

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

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

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

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

$$\mathbb{P}\left(\max_{v\in V_N^\delta}\psi_v(\alpha_r)\geq \varepsilon\gamma^\star \log N^2\right)\leq N^2\exp\left(-\frac{(\varepsilon\gamma^\star)^2\log N^2}{\max_i \sigma_i^2\alpha_r}\right) \tag{1.3.13}$$

where we used a union bound, a Gaussian estimate and (1.3.2) to get the inequality. This is $O(N^{-c})$ by choosing K large enough for a fixed ε and r. On the other hand, the first term is smaller than $\mathbb{P}(\max_{v\in V_N^\delta}\psi_v\geq(1-3\varepsilon)\gamma^\star\log N^2)$. Since $V_N\supseteq V_N^\delta$, this implies (1.3.12) as claimed.

Proof of Lemma 1.3.3. In view of (1.3.12), it suffices to show $\mathbb{P}(\mathcal{N}_\varepsilon^\star\geq 1)=1-O(N^{-c})$. The Paley-Zygmund inequality implies

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

We show

$$\mathbb{E}\left[(\mathcal{N}_\varepsilon^\star)^2\right]\leq(1+O(N^{-\frac{1}{2K}(1-(1-\varepsilon)^2)}))\ (\mathbb{E}[\mathcal{N}_\varepsilon^\star])^2, \tag{1.3.14}$$

which proves the claim.

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

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

Using Gaussian estimates and the variance estimates in (1.3.11), the probabilities are for every j and $v\in V_N^\delta$:

$$\begin{aligned}p_{v,j} &\doteq \mathbb{P}\Big(\nabla\psi_v(\alpha_j)\geq\nabla L_{N,\varepsilon}^\star(\alpha_j)\Big)\\ &\asymp\frac{1}{\sqrt{\log N}}\exp\left(-(1-\varepsilon)^2\frac{\sigma^2(\alpha_j)}{\bar\sigma^2(\alpha_j)}\nabla\alpha_j\log N^2\right).\end{aligned} \tag{1.3.15}$$

Write e_j for the exponential term on the right-hand side of (1.3.15). The first moment satisfies

$$\mathbb{E}[\mathcal{N}_\varepsilon^\star]=\sum_{v\in V_N^\delta}\mathbb{P}(A_v)\geq\frac{c(\varepsilon,\delta)}{(\log N)^{\frac12(K-r)}}\times|V_N^\delta|\times\prod_{j=r+1}^{K}e_j. \tag{1.3.16}$$

Now, we compare this with the second moment :

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

We divide the sum depending on the correlations between ψ_v and $\psi_{v'}$. More precisely, recall the definition of the branching scale in (1.2.1) :

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

Write the second moment as

$$\sum_{\substack{v,v' \in V_N^\delta \\ \rho(v,v')<\alpha_r}} \mathbb{P}(A_v \cap A_{v'}) \;+ \sum_{k=r+1}^{K-1} \sum_{\substack{v,v' \in V_N^\delta \\ \alpha_{k-1} \le \rho(v,v') < \alpha_k}} \mathbb{P}(A_v \cap A_{v'}) \;+ \sum_{\substack{v,v' \in V_N^\delta \\ \rho(v,v') \ge \alpha_{K-1}}} \mathbb{P}(A_v \cap A_{v'})\,. \tag{1.3.17}$$

In particular, the first term in (1.3.17) is equal to

$$\sum_{\substack{v,v' \in V_N^\delta \\ \rho(v,v')<\alpha_r}} \mathbb{P}(A_v)\,\mathbb{P}(A_{v'}) \le \sum_{v,v' \in V_N^\delta} \mathbb{P}(A_v)\,\mathbb{P}(A_{v'}) = (\mathbb{E}[\mathcal{N}_\varepsilon^\star])^2. \tag{1.3.18}$$

It remains to show that the second and third term in (1.3.17) are negligible compared to $(\mathbb{E}[\mathcal{N}_\varepsilon^\star])^2$. We write the details for the second term since the last term is done similarly and is easier. By Lemma 1.4.1 (following the Markov property of the GFF), note that if $\alpha_{k-1} \le \rho(v,v') < \alpha_k$ for some $k \ge r+1$, then $\nabla\psi_{v'}(\alpha_{j'})$, $j' \ge k+1$, is independent of $\nabla\psi_v(\alpha_j)$ for $j \le k-2$ and $j \ge k+1$. Therefore, for $v,v' \in V_N^\delta$ such that $\alpha_{k-1} \le \rho(v,v') < \alpha_k$, we have

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

where we dropped the conditions on $j \in \{k-1,k\}$ for v as well as the conditions on $j \le k$ for v' in the first inequality. We simply rearranged the probabilities and eliminated the log terms to get the last inequality. The number of pairs $v,v' \in V_N^\delta$ such that $\alpha_{k-1} \le \rho(v,v') < \alpha_k$ is at most $|V_N^\delta| \times N^{2(1-\alpha_{k-1})}$. Therefore, by (1.3.16),

$$\sum_{\substack{v,v' \in V_N^\delta \\ \alpha_{k-1} \le \rho(v,v') < \alpha_k}} \mathbb{P}(A_v \cap A_{v'}) \le \frac{(\mathbb{E}[\mathcal{N}_\varepsilon^\star])^2}{(\log N)^{-K}} \times N^{-2\alpha_{k-1}(1-(1-\varepsilon)^2)}\left(\frac{\prod_{j=1}^r e_j}{e_{k-1}e_k^2}\right) \times \frac{N^{-2\alpha_{k-1}(1-\varepsilon)^2}}{\prod_{j=1}^{k-1} e_j} \tag{1.3.19}$$

The right-hand side of (1.3.19) is separated in three factors by $\times$. The third factor is bounded by 1 because

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

by definition of $\bar{\sigma}$. To bound the second factor, set $r \geq 3$ independently of any other variable. Note that if r depended on K, the bound in (1.3.13) would not necessarily tend to 0. There are two cases to consider : $\alpha_k \leq \lambda_1$ and $\alpha_k > \lambda_1$. When $\alpha_k \leq \lambda_1$, the ratio of exponentials is bounded by 1 because $e_1 e_2 e_3 = e_{k-1} e_k^2$ and we have $N^{-2\alpha_{k-1}(1-(1-\varepsilon)^2)} \leq N^{-\frac{1}{K}(1-(1-\varepsilon)^2)}$ since $\alpha_{k-1} \geq \alpha_r \geq 1/(2K)$. When $\alpha_k > \lambda_1$, the ratio of exponentials is bounded by $N^{\lambda_1(1-(1-\varepsilon)^2)}$ by choosing K large enough for a fixed ε and we have $N^{-2\alpha_{k-1}(1-(1-\varepsilon)^2)} \leq N^{-2\lambda_1(1-(1-\varepsilon)^2)}$ because $\alpha_{k-1} \geq \lambda_1$. Since $\lambda_1 \geq 1/K$, the right-hand side of (1.3.19) is always bounded by

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

With (1.3.18), this shows (1.3.14) and concludes the proof of the lemma. □

1.3.3. Log-number of high points

The proof of the upper bound for the log-number of high-points uses an argument based on the path at every scale λ^l similar to the one in Lemma 1.3.1. Recall the definition of the critical levels γ^l and the entropy $\mathcal{E}_\gamma$ in Theorem 1.1.3.

Lemma 1.3.4 (Upper bound on the log-number of high points). *Let $\{\psi_v\}_{v\in V_N}$ be the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF on V_N of Definition 1.1.1 and $\gamma^\star$ as defined in Theorem 1.1.2. Also, let $\gamma^{l-1} < \gamma \leq \gamma^l$ for some $l \in \{1, ..., m\}$. For all $0 < \varepsilon < (\gamma - \gamma^{l-1})/m$, there exists a constant $c = c(\gamma, \varepsilon, \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ such that*

$$\mathbb{P}\Big(|\mathcal{H}_N^\gamma| \geq N^{2\mathcal{E}_\gamma + \varepsilon}\Big) \leq N^{-c} \tag{1.3.20}$$

for N large enough.

Proof. Recall the definition of the optimal path L_N^γ from (1.2.6) and the notation R_{λ^j} for the set of representatives at scale λ^j. Consider

$$\mathcal{H}_{N,j}^{\gamma,\varepsilon} \mathrel{\hat=} \Big\{ v \in R_{\lambda^j} \,|\, \psi_v(\lambda^j) \geq L_N^{\gamma + j\varepsilon}(\lambda^j) \Big\}, \quad 1 \leq j \leq m.$$

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

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

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

$$\text{the paths } L_N^\star \text{ and } L_N^{\gamma - j\varepsilon} \text{ coincide on the interval } [0, \lambda^j]. \tag{1.3.21}$$

The idea is to split the probability that at least $N^{2\mathcal{E}_\gamma + \varepsilon}$ points in V_N reach the optimal height by looking at the first scale λ^j, $1 \le j \le l-1$, where the set $\mathcal{H}_{N,j}^{\gamma - j\varepsilon, \varepsilon}$ is not empty. As for the maximum, this yields the appropriate constraints along the scales to get the correct upper bound. A union bound in (1.3.20) gives

$$\begin{aligned} \mathbb{P}\left(|\mathcal{H}_N^\gamma| \ge N^{2\mathcal{E}_\gamma + \varepsilon}\right) &= \mathbb{P}\left(|\mathcal{H}_{N,m}^{\gamma - m\varepsilon, \varepsilon}| \ge N^{2\mathcal{E}_\gamma + \varepsilon}\right) \\ &\le \mathbb{P}\left(\begin{array}{l} |\mathcal{H}_{N,1}^{\gamma - 1\varepsilon, \varepsilon}| = ... = |\mathcal{H}_{N,l-1}^{\gamma - (l-1)\varepsilon, \varepsilon}| = 0 \\ \text{and } |\mathcal{H}_{N,m}^{\gamma - m\varepsilon, \varepsilon}| \ge N^{2\mathcal{E}_\gamma + \varepsilon} \end{array}\right) + \sum_{j=1}^{l-1} \mathbb{P}\left(|\mathcal{H}_{N,j}^{\gamma - j\varepsilon, \varepsilon}| \ge 1\right). \end{aligned} \tag{1.3.22}$$

Because of (1.3.21), the probabilities in the sum are bounded by $N^{-c(\varepsilon)}$ in exactly the same manner as $\mathbb{P}(|\mathcal{H}_{N,m}^{\star, \varepsilon}| \ge 1)$ in Lemma 1.3.1. The first probability in (1.3.22) is bounded by

$$\begin{aligned} &\mathbb{P}\left(\left|\left\{v \in V_N \mid \begin{array}{l} \psi_v \ge L_N^\gamma(1) \text{ and } \psi_{v_{\lambda^j}}(\lambda^j) < L_N^\gamma(\lambda^j) \\ \text{for all } 1 \le j \le l-1 \end{array}\right\}\right| \ge N^{2\mathcal{E}_\gamma + \varepsilon}\right) \\ &\quad \le C e^{-c(\eta_\varepsilon)(\log N)^2} + N^{-\varepsilon} N^{-2\mathcal{E}_\gamma} N^2 \max_{v \in V_N^o} \mathbb{P}\left(\begin{array}{l} \psi_v \ge L_N^{\gamma - \eta_\varepsilon}(1) \text{ and} \\ \psi_v(\lambda^j) < L_N^{\gamma + j\eta_\varepsilon}(\lambda^j) \\ \text{for all } 1 \le j \le l-1 \end{array}\right) \end{aligned} \tag{1.3.23}$$

using Markov's inequality and using the event A_ε as in (1.3.4), where we impose

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

this time around. See (1.3.26) for the definition of c_γ. See just below and also (1.3.27) for the justification of the constraints on η_ε. When $l = 1$, a Gaussian estimate and the bound $\max_{v \in V_N} \mathbb{V}(\psi_v) \le \mathcal{J}_{\sigma^2}(1) \log N + C$ from (1.3.2) yield

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

because $L_N^{\gamma-\eta_\varepsilon}(1) = 2(\gamma - \eta_\varepsilon)\log N$ and $\mathcal{E}_\gamma = 1 - \gamma^2/\mathcal{J}_{\sigma^2}(1)$ in this case. This proves that the second term in (1.3.23) decays like $N^{-c(\gamma,\varepsilon)}$, as needed.

It remains to show a similar estimate for a fixed $l \in \{2, ..., m\}$. To simplify the notation, denote $(X_v^1, ..., X_v^{l-1}, X_v^m) \circeq (\psi_v(\lambda^1), ..., \psi_v(\lambda^{l-1}), \psi_v)$. By conditioning on the value of the vector $\boldsymbol{X} \circeq (X_v^1, ..., X_v^{l-1})$, the probability in (1.3.23) is equal to

$$\int_{-\infty}^{L_N^{\gamma+1\eta_\varepsilon}(\lambda^1)} \cdots \int_{-\infty}^{L_N^{\gamma+(l-1)\eta_\varepsilon}(\lambda^{l-1})} \mathbb{P}\Big(X_v^m \geq L_N^{\gamma-\eta_\varepsilon}(1) \,|\, \boldsymbol{X} = \boldsymbol{x}\Big)\, f_v(\boldsymbol{x})\, d\boldsymbol{x}$$

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

$$\int_{-\infty}^{L_N^{\gamma+1\eta_\varepsilon}(\lambda^1)} \cdots \int_{-\infty}^{L_N^{\gamma+(l-1)\eta_\varepsilon}(\lambda^{l-1})} \mathbb{P}\Big(X_v^m - X_v^{l-1} \geq L_N^{\gamma-\eta_\varepsilon}(1) - x_{l-1}\Big)\, f_v(\boldsymbol{x})\, d\boldsymbol{x}. \tag{1.3.24}$$

The bound $\max_{v\in V_N} \mathbb{V}\Big(X_v^m - X_v^{l-1}\Big) \leq \mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)\log N + C$ from (1.3.2) and a Gaussian estimate show that

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

where we introduced $L_N^{\gamma}(\lambda^{l-1})$ and used (1.2.6). By definition of $\mathcal{E}_\gamma$ and the definition of γ^{l-1} in (1.1.5), this is equal to

$$\begin{aligned}
&\frac{C(\gamma) N^{-2+2\mathcal{E}_\gamma}}{\sqrt{\log N}} N^{2\lambda^{l-1}} \exp\left(-2\left[\frac{(\gamma-\gamma^{l-1})}{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)} + \frac{1}{\bar{\sigma}_{l-1}}\right](L_N^{\gamma-\eta_\varepsilon}(\lambda^{l-1}) - x_{l-1})\right) \\
&= \frac{C(\gamma) N^{-2+2\mathcal{E}_\gamma}}{\sqrt{\log N}} N^{2\lambda^{l-1}} N^{\frac{4lc_\gamma}{\bar{\sigma}_{l-1}}\eta_\varepsilon} \exp\left(-2\frac{c_\gamma}{\bar{\sigma}_{l-1}}(L_N^{\gamma+(l-1)\eta_\varepsilon}(\lambda^{l-1}) - x_{l-1})\right)
\end{aligned} \tag{1.3.25}$$

where

$$c_\gamma \circeq \frac{(\gamma - \gamma^{l-1})\bar{\sigma}_{l-1}}{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)} + 1 > 1. \tag{1.3.26}$$

Putting the bound (1.3.25) in (1.3.24) and in (1.3.23), we get that the first term in (1.3.22) decays like

$$N^{-\left(\varepsilon-\frac{4lc_\gamma}{\bar{\sigma}_{l-1}}\eta_\varepsilon\right)} \tag{1.3.27}$$

provided that

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

where $Y_{v,j} \doteq \eta_\varepsilon \log N^2 + \nabla L_N^\gamma(\lambda^j) - \nabla X_v^j$. Similarly to (1.3.7), the integral has the right decay as a consequence of Lemma 1.3.2, with $a_l \doteq c_\gamma/\bar{\sigma}_{l-1} > 1/\bar{\sigma}_{l-1}$, because $L_N^\star$ and L_N^γ coincide on the interval $[0, \lambda^{l-1}]$. □

Lemma 1.3.5 (Lower bound on the log-number of high points)**.** *Let $\{\psi_v\}_{v\in V_N}$ be the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF on V_N of Definition 1.1.1 and $\gamma^\star$ as in Theorem 1.1.2. Let $\gamma > 0$ be such that $\gamma^{l-1} \leq \gamma < \gamma^l$ for some $l \in \{1, ..., m\}$. For all $0 < \varepsilon < \min\{1/4, (\gamma^l - \gamma)/(4\gamma)\}$, there exists a constant $c = c(\gamma, \varepsilon, \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ such that*

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

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

We use the same notations as in the proof of Lemma 1.3.3. As before, we can assume, without loss of generality, that $\lambda_i K \in \mathbb{N}_0$ for all $\{0, ..., M\}$ so that the α_k's form a finer partition of $[0, 1]$ than the λ_i's. The parameter $K \in \mathbb{N}$ will be chosen large enough depending on γ and ε during the proof. Again, we restrict ourselves to V_N^δ to ensure that for all $k \in \{1, ..., K\}$ and for all $v \in V_N^\delta$:

$$|\mathbb{V}(\nabla\psi_v(\alpha_k)) - \sigma^2(\alpha_k)\nabla\alpha_k \log N| \leq C(\delta). \tag{1.3.28}$$

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

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

Since $\varepsilon < (\gamma^l - \gamma)/(4\gamma)$ by hypothesis, we have $\gamma^{l-1} \leq \gamma < \gamma(1 + 4\varepsilon) < \gamma^l$. This condition implies that the increments of the path $L_{N,\varepsilon}^\gamma$ are always bounded by the increments of the sub-optimal path $L^{\star}_{N,\varepsilon}$ (see Figure 1.2.3), namely

$$L^{\gamma}_{N,\varepsilon}(s_2) - L^{\gamma}_{N,\varepsilon}(s_1) \leq L^{\star}_{N,\varepsilon}(s_2) - L^{\star}_{N,\varepsilon}(s_1), \quad 0 \leq s_1 \leq s_2 \leq 1. \tag{1.3.29}$$

Indeed, the paths $L_N^{\gamma(1+4\varepsilon)}$ and $L_N^{\star}$ coincide on the interval $[0, \lambda^{l-1}]$. Moreover, when $s \in (\lambda^{l-1}, 1]$, we have, by the definition of the critic levels γ^l in (1.1.5) and the optimal path $L_N^{\gamma(1+4\varepsilon)}$ in (1.2.6),

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

This proves inequality (1.3.29). By hypothesis, we also have $\varepsilon < 1/4$, which yields

$$L^{\gamma}_{N,\varepsilon}(1) = (1-\varepsilon)(1+4\varepsilon)\gamma \log N^2 > (1+2\varepsilon)\gamma \log N^2. \tag{1.3.30}$$

The proof again relies on the Paley-Zygmund inequality applied to a *modified number of exceedances* where we consider only points in V_N^{δ} whose increments are almost optimal. We drop the first r increments to allow more independence which is needed for the second-moment method to work. We can choose $r \geq 3$ independently of any other variable as in the proof of Lemma 1.3.3. The case $l = 1$ is easier to deal with, so we omit the details. Assume $l \in \{2, ..., m\}$ and define

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

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

$$\mathbb{P}\left(|\mathcal{H}_N^{\gamma}| \geq N^{2\mathcal{E}_{\gamma} - \tilde{\varepsilon}}\right) \geq \mathbb{P}\left(\mathcal{N}_{\varepsilon}^{\gamma} \geq N^{2\mathcal{E}_{\gamma} - \tilde{\varepsilon}}\right) - O(N^{-c}). \tag{1.3.31}$$

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

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{N}_{\varepsilon}^{\gamma} \geq N^{2\mathcal{E}_{\gamma} - \tilde{\varepsilon}}, \min_{v \in V_N^{\delta}} \psi_v(\alpha_r) > -\varepsilon\gamma \log N^2\right) \\
&\quad + \mathbb{P}\left(\mathcal{N}_{\varepsilon}^{\gamma} \geq N^{2\mathcal{E}_{\gamma} - \tilde{\varepsilon}}, \min_{v \in V_N^{\delta}} \psi_v(\alpha_r) \leq -\varepsilon\gamma \log N^2\right)
\end{aligned} \tag{1.3.32}$$

To simplify the argument, assume from now on that K is large enough to ensure $\alpha_r \le \lambda^{l-1}$. The first probability in (1.3.32) is smaller than $\mathbb{P}\Big(|\mathcal{H}_N^\gamma| \ge N^{2\mathcal{E}_\gamma - \tilde{\varepsilon}}\Big)$ because the points $v \in V_N^\delta$ that are contributing to the sum $\mathcal{N}_\varepsilon^\gamma$, on the event $\{\min_{v\in V_N^\delta} \psi_v(\alpha_r) > -\varepsilon\gamma \log N^2\}$, are also in $\mathcal{H}_N^\gamma$. Indeed, when $\mathbf{1}_{A_v} = 1$,

$$\begin{aligned}\psi_v - \psi_v(\alpha_r) &\ge (1-\varepsilon) L_N^{\gamma(1+4\varepsilon)}(1) - (1-\varepsilon) L_N^\star(\alpha_r) \\ &= (1-\varepsilon)(1+4\varepsilon)\gamma \log N^2 - (1-\varepsilon)\mathcal{J}_{\sigma^2/\bar{\sigma}}(\alpha_r) \log N^2 \\ &\ge (1+\varepsilon)\gamma \log N^2 \end{aligned} \tag{1.3.33}$$

where we take K large enough that $(1-\varepsilon)\mathcal{J}_{\sigma^2/\bar{\sigma}}(\alpha_r) < \varepsilon\gamma$ and use (1.3.30) to obtain the last inequality in (1.3.33). The distribution of $\psi_v(\alpha_r)$ is symmetric, so the second probability in (1.3.32) is smaller than

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

where we used a union bound, a Gaussian estimate and (1.3.2) to get the inequality. This is $O(N^{-c})$ by choosing K large enough for a fixed ε and r. Therefore, we have (1.3.31) as claimed.

Proof of Lemma 1.3.5. In view of (1.3.31), it suffices to show that $\mathbb{P}\Big(\mathcal{N}_\varepsilon^\gamma \ge N^{2\mathcal{E}_\gamma - \tilde{\varepsilon}}\Big) = 1 - O(N^{-c})$. The Paley-Zygmund inequality (Lemma 1.4.8) implies

$$\mathbb{P}\Big(\mathcal{N}_\varepsilon^\gamma \ge N^{2\mathcal{E}_\gamma - \tilde{\varepsilon}}\Big) \ge \left(1 - \frac{N^{2\mathcal{E}_\gamma - \tilde{\varepsilon}}}{\mathbb{E}[\mathcal{N}_\varepsilon^\gamma]}\right)^2 \frac{(\mathbb{E}[\mathcal{N}_\varepsilon^\gamma])^2}{\mathbb{E}[(\mathcal{N}_\varepsilon^\gamma)^2]}. \tag{1.3.34}$$

First, we make sure that $N^{2\mathcal{E}_\gamma - \tilde{\varepsilon}}/\mathbb{E}[\mathcal{N}_\varepsilon^\gamma] \to 0$ as $N \to \infty$. By independence of the increments and the variance estimate (1.3.28), Gaussian estimates yield, for some constant $c = c(\gamma, \varepsilon, \delta) > 0$,

$$\begin{aligned}\mathbb{E}[\mathcal{N}_\varepsilon^\gamma] = \sum_{v\in V_N^\delta} \mathbb{P}(A_v) &= \sum_{v\in V_N^\delta} \prod_{j=r+1}^{K} \mathbb{P}\Big(\nabla\psi_v(\alpha_j) \ge \nabla L_{N,\varepsilon}^\gamma(\alpha_j)\Big) \\ &\ge c \cdot (\log N)^{-\frac{1}{2}(K-r)} N^{2(1-(1-\varepsilon)^2) + 2(1-\varepsilon)^2 \mathcal{E}_{\gamma(1+4\varepsilon)} + 2(1-\varepsilon)^2 \int_0^{\alpha_r} \frac{\sigma^2(s)}{\bar{\sigma}^2(s)} ds} \\ &\ge N^{2(1-(1-\varepsilon)^2) + 2(1-\varepsilon)^2 \mathcal{E}_{\gamma(1+4\varepsilon)}}. \end{aligned} \tag{1.3.35}$$

By the definition of $\mathcal{E}_\gamma$ in Theorem 1.1.3, and because $\gamma^{l-1} \leq \gamma < \gamma(1+4\varepsilon) < \gamma^l$,

$$\begin{aligned} \left|\mathcal{E}_{\gamma(1+4\varepsilon)} - \mathcal{E}_\gamma\right| &= \frac{(\gamma(1+4\varepsilon) - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))^2 - (\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))^2}{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)} \\ &= \frac{16\varepsilon^2\gamma^2 + 8\varepsilon\gamma(\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))}{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)} \leq \frac{12(\gamma^\star)^2}{\bar{\sigma}_m^2 \nabla\lambda^m}\varepsilon \mathrel{\hat{=}} \tilde{\varepsilon}/2 \end{aligned} \tag{1.3.36}$$

where we used $\varepsilon < 1/4$, $\gamma < \gamma^\star$ and $\mathcal{J}_{\sigma^2}(\lambda^{l-1},1) \geq \mathcal{J}_{\sigma^2}(\lambda^{m-1},1) = \bar{\sigma}_m^2\nabla\lambda^m$ to obtain the inequality. By inserting the bound (1.3.36) in (1.3.35), we get

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

Since $(1-\varepsilon)^2 < 1$ and $\mathcal{E}_\gamma \leq 1$, it proves the assertion that $N^{2\mathcal{E}_\gamma-\tilde{\varepsilon}}/\mathbb{E}[\mathcal{N}_\varepsilon^\gamma] \to 0$ and also justify the use of the Paley-Zygmund inequality. In view of (1.3.34), it suffices to show, like in Lemma 1.3.3, that

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

to prove the lemma. The proof is almost identical to the proof of Lemma 1.3.3. Indeed, by Gaussian estimates and the variance estimates in (1.3.28), the probabilities on the increments in A_v are for every j and $v \in V_N^\delta$:

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

where the $\tilde{e}_j$'s are the corresponding exponential factors. The proof is exactly the same up to (1.3.19) with $\tilde{e}_j$'s instead of e_j's. From there, the third factor in the decomposition is still bounded by 1 because of property (1.3.29), and the rest of the argument follows if we choose K large enough for a fixed ε and γ. This ends the proof of the lemma. □

1.4. Appendix

1.4.1. Technical lemmas

The Markov property of the GFF, which is a consequence of the strong Markov property of the simple random walk (in the covariance function in (1.1.1)), implies that the value of the field inside a neighborhood is independent of the field outside given the boundary, see e.g. Dynkin (1980). In particular, for the neighborhood $[v]_\lambda$, where $\lambda \in [0,1]$, this implies

$$\phi_v(\lambda) \doteq \mathbb{E}\Big[\phi_v \,|\, \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}\Big] = \mathbb{E}\Big[\phi_v \,|\, \mathcal{F}_{\partial[v]_\lambda}\Big]. \tag{1.4.1}$$

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

$$\phi_v(\lambda') - \phi_v(\lambda) \quad \text{is independent of} \quad \phi_{v'}(\mu') - \phi_{v'}(\mu). \tag{1.4.2}$$

This is because the shell $[v]_\lambda \cap [v]_{\lambda'}^c$ does not intersect the shell $[v']_\mu \cap [v']_{\mu'}^c$ in both cases, see Figure 1.2.2. We stress that, in general, the field ψ does not have the Markov property. However, by working with increments of the field ψ, the property analogous to (1.4.2) can be proved.

Lemma 1.4.1. *Let $v, v' \in V_N$, $\lambda < \lambda'$ and $\mu < \mu'$. If we have $\lambda, \mu > \rho(v,v')$ or $\lambda > \rho(v,v') > \mu'$, then*

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

Proof. Let $v \in V_N$ and $\lambda < \lambda'$. By Definition 1.1.1 of the field ψ and its conditional expectation, we have

$$\psi_v(\lambda) = \sum_{1 \le i \le M} \sigma_i \, \mathbb{E}\Big[\nabla \phi_v(\lambda_i) \,|\, \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}\Big] = \sum_{\substack{1 \le i \le M: \\ \lambda_{i-1} < \lambda}} \sigma_i \left(\phi_v(\lambda \wedge \lambda_i) - \phi_v(\lambda_{i-1})\right). \tag{1.4.3}$$

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

$$\psi_v(\lambda') - \psi_v(\lambda) = \sum_{\substack{1 \le i \le M: \\ \lambda \le \lambda_{i-1} < \lambda' \text{ or } \lambda < \lambda_i \le \lambda' \\ \text{or } \lambda_{i-1} \le \lambda < \lambda' \le \lambda_i}} \sigma_i \left(\phi_v(\lambda' \wedge \lambda_i) - \phi_v(\lambda \vee \lambda_{i-1})\right). \tag{1.4.4}$$

The conclusion of the lemma follows from (1.4.2). $\square$

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

Lemma 1.4.2. *Let $\delta \in (0, 1/2]$ and $\lambda_{i-1} \leq \lambda < \lambda' \leq \lambda_i$ for some $i \in \{1, ..., M\}$, then*

$$-C_1(\delta, \sigma_i) \leq \mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) - (\lambda' - \lambda)\sigma_i^2 \log N \leq C_2(\sigma_i) \tag{1.4.5}$$

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

Proof. The Markov property (1.4.1) yields $\mathbb{E}\Big[\phi_v - \phi_v(\lambda) \,|\, \mathcal{F}_{\partial[v]_{\lambda'}}\Big] = \phi_v(\lambda') - \phi_v(\lambda)$. Using the conditional variance formula and $\mathbb{V}(X \,|\, \mathcal{F}) \mathrel{\hat=} \mathbb{E}[(X - \mathbb{E}[X \,|\, \mathcal{F}])^2 \,|\, \mathcal{F}]$, we can compute the variance of (1.4.4) in the special case $\lambda_{i-1} \leq \lambda < \lambda' \leq \lambda_i$:

$$\begin{aligned}
\mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) &= \sigma_i^2 \; \mathbb{V}\Big(\mathbb{E}\Big[\phi_v - \phi_v(\lambda) \,|\, \mathcal{F}_{\partial[v]_{\lambda'}}\Big]\Big) \\
&= \sigma_i^2 \Big(\mathbb{V}(\phi_v - \phi_v(\lambda)) - \mathbb{E}\Big[\mathbb{V}\Big(\phi_v - \phi_v(\lambda) \,|\, \mathcal{F}_{\partial[v]_{\lambda'}}\Big)\Big]\Big) \\
&= \sigma_i^2 \left(\mathbb{V}(\phi_v - \phi_v(\lambda)) - \mathbb{V}(\phi_v - \phi_v(\lambda'))\right).
\end{aligned} \tag{1.4.6}$$

But, it is well known that $\{\phi_u - \mathbb{E}[\phi_u \,|\, \mathcal{F}_{\partial B}]\}_{u \in B}$ is a GFF when $B \subseteq \mathbb{Z}^2$ is a finite box, see e.g. Zeitouni (2017). Simply choose $B = [v]_s$, $s = \lambda, \lambda'$, in (1.4.6), then by the variance definition in (1.1.1),

$$\mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) = \sigma_i^2 \, (G_{[v]_\lambda}(v, v) - G_{[v]_{\lambda'}}(v, v)). \tag{1.4.7}$$

Using standard estimates for the discrete Green function, we can now evaluate the last expression. For every finite box $B \subseteq \mathbb{Z}^2$, Proposition 1.6.3 of Lawler (1991) shows that (keeping in mind our choice of normalization by $\pi/2$ in (1.1.1)) :

$$G_B(x, y) = \left[\sum_{z \in \partial B} \mathscr{P}_x(W_{\tau_{\partial B}} = z) \, a(z - y)\right] - a(y - x), \quad x, y \in B, \tag{1.4.8}$$

where

$$a(w) = \begin{cases} \log(\|w\|_2) + \text{const.} + O(\|w\|_2^{-2}), & \text{if } w \in \mathbb{Z}^2 \backslash \{\mathbf{0}\}, \\ 0, & \text{if } w = \mathbf{0}, \end{cases} \tag{1.4.9}$$

and $\mathscr{P}_x$ is the law of the simple random walk starting at $x \in \mathbb{Z}^2$. Using (1.4.8), we can rewrite the difference of Green functions in (1.4.7) as

$$\sum_{z \in \partial[v]_\lambda} \mathscr{P}_v\Big(W_{\tau_{\partial[v]_\lambda}} = z\Big) \, a(z - v) \; - \sum_{z \in \partial[v]_{\lambda'}} \mathscr{P}_v\Big(W_{\tau_{\partial[v]_{\lambda'}}} = z\Big) \, a(z - v). \tag{1.4.10}$$

When $\lambda' = 1$, we have $\|z - v\|_2 = 0$ for $z \in \partial[v]_{\lambda'}$. Otherwise, we assumed $v \in V_N^\delta$, so take N large enough (depending on δ) that $[v]_{\lambda'}$ is not cut off by ∂V_N. We have $\|z - v\|_2 \leq \sqrt{2}N^{1-\lambda}$ for $z \in \partial[v]_\lambda$ in general and $\|z - v\|_2 \geq \frac{1}{2}N^{1-\lambda'}$ for $z \in \partial[v]_{\lambda'}$ when $\lambda' \neq 1$. We deduce the following bound on the variance in (1.4.7) :

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

Similarly, we have $\|z-v\|_2 \geq \delta N$ for $z \in \partial[v]_\lambda$ when $\lambda = 0$. Otherwise, take N large enough (depending on δ) that $[v]_\lambda$ is not cut off by the boundary of V_N. We have $\|z-v\|_2 \geq \frac{1}{2}N^{1-\lambda}$ for $z \in \partial[v]_\lambda$ when $\lambda \neq 0$ and $\|z - v\|_2 \leq \frac{1}{\sqrt{2}}N^{1-\lambda'}$ for $z \in \partial[v]_{\lambda'}$ in general. We deduce the following bound on the variance in (1.4.7) :

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

This ends the proof of the lemma. □

Since the upper bound in Lemma 1.4.2 is only valid for N large enough depending on δ, we cannot immediately conclude that it holds for all $v \in V_N$. We show in the next lemma how to extend the bound.

Lemma 1.4.3. *Let $\lambda_{i-1} \leq \lambda < \lambda' \leq \lambda_i$ for a certain $i \in \{1, ..., M\}$, then*

$$\max_{v\in V_N} \mathbb{V}(\psi_v(\lambda') - \psi_v(\lambda)) \leq (\lambda' - \lambda)\sigma_i^2 \log N + C(\sigma_i) \tag{1.4.11}$$

for N large enough.

Proof. When $v \in \partial V_N$, the bound is trivial because $\psi_v = 0$. Therefore, let $v \in V_N^o$. To obtain the upper bound on the difference of Green functions in (1.4.7), we only used the fact that $[v]_{\lambda'}$ was not cut off by ∂V_N for N large enough depending on δ. Hence, we only need to show that when $[v]_{\lambda'}$ is cut off, there exists $u \in V_N^o$ such that $[u]_{\lambda'}$ is not cut off and for which

$$G_{[v]_\lambda}(v, v) - G_{[v]_{\lambda'}}(v, v) \leq G_{[u]_\lambda}(u, u) - G_{[u]_{\lambda'}}(u, u) + \tilde{C}(\sigma_i). \tag{1.4.12}$$

Assume that $[v]_{\lambda'}$ is cut off by ∂V_N and choose u to be the center of V_N. Clearly, the neighborhood $[u]_{\lambda'}$ is not cut off by the boundary of V_N. When $\lambda' = 1$, inequality (1.4.12) is trivial because $G_{[v]_{\lambda'}}(v,v) = G_{[u]_{\lambda'}}(u,u) = 0$ and $G_{[v]_\lambda}(v,v) \leq G_{[u]_\lambda}(u,u)$ since $[v]_\lambda$ is cut off and $[u]_\lambda$ is not. Now, assume $\lambda' < 1$. Denote $\theta(x) \doteqdot x + u - v$ the translation function that moves v to u, see Figure 1.4.4.

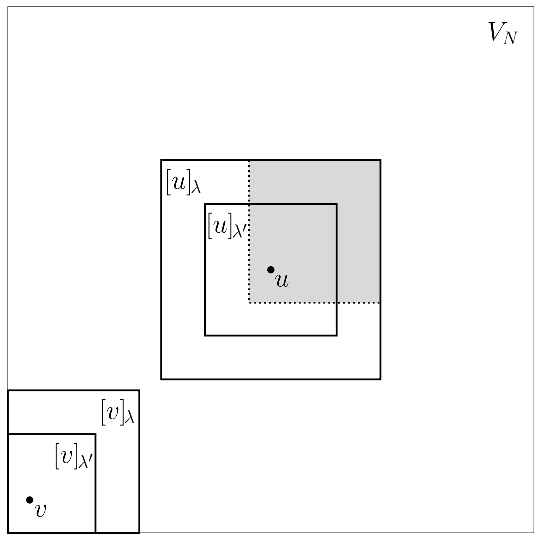


FIGURE 1.4.4. The grey area $\theta([v]_\lambda)$ is the translation of $[v]_\lambda$.

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

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

This proves (1.4.12) when $\lambda \neq 0$. Since $[v]_0 \doteqdot V_N$ throughout the article and we defined $[v]_0$ differently in this proof, it remains to show that

$$\max_{v \in V_N} G_{V_N}(v,v) - G_{[v]_0}(v,v) \leq \tilde{C}(\sigma_i) \tag{1.4.13}$$

for (1.4.12) to be true when $\lambda = 0$. By the strong Markov property and (1.4.8) :

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

But $\|w - v\|_2 \leq \sqrt{2}N$ for all $w \in \partial V_N$ and $\|v - z\|_2 \geq \frac{1}{2}N$ for all $z \in \partial [v]_0 \cap V_N^o$. We get the desired conclusion using (1.4.9). □

In order to approximate the branching structure of the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF in Lemma 1.3.1 and Lemma 1.3.4, we need to show that the variance of $\psi_v(\lambda) - \psi_{v_\lambda}(\lambda)$ is bounded by a constant, where v_λ denotes any representative in R_λ that is closest to v. Our final goal here is to show Lemma 1.4.6. We start by proving a more general version of Lemma 12 found in Bolthausen et al. (2001). We define

$$\phi_v(A) \mathrel{\hat{=}} \mathbb{E}\Big[\phi_v \,|\, \mathcal{F}_{\partial (A \cap V_N)}\Big]$$

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

Lemma 1.4.4. *Let $B \subseteq \mathbb{Z}^2$ be a square box of width smaller or equal to $N/2$ such that $B \cap V_N \neq \emptyset$. Moreover, let $0 \leq \eta < 1$ and $L \in \{1, 2, ..., N/4\}$, then there exists a constant $C = C(\eta) > 0$ such that*

$$\max_{\substack{u,v \in B \cap V_N \\ d(u, \partial B) = L \\ \|u - v\|_2 \leq \eta L}} \mathbb{V}(\phi_u(B) - \phi_v(B)) \leq C. \tag{1.4.14}$$

Proof. Let $u, v \in B \cap V_N$ be such that $d(u, \partial B) = L$ and $\|u - v\|_2 \leq \eta L$. Denote $\overline{B} \mathrel{\hat{=}} B \cap V_N$. Using the conditional variance formula as in (1.4.6), we have

$$
\begin{aligned}
\mathbb{V}(\phi_u(B) - \phi_v(B)) &= \mathbb{V}(\mathbb{E}[\phi_u - \phi_v \,|\, \mathcal{F}_{\partial \overline{B}}]) \\
&= \mathbb{V}(\phi_u - \phi_v) - \mathbb{E}[\mathbb{V}(\phi_u - \phi_v \,|\, \mathcal{F}_{\partial \overline{B}})] \\
&= (G_{V_N}(u,u) - G_{V_N}(u,v) + G_{V_N}(v,v) - G_{V_N}(v,u)) \\
&\quad - (G_{\overline{B}}(u,u) - G_{\overline{B}}(u,v) + G_{\overline{B}}(v,v) - G_{\overline{B}}(v,u)).
\end{aligned} \tag{1.4.15}
$$

For this proof, redefine $[u]_0$ as the square box of side length N centered at u that has been cut off by ∂V_N. From (1.4.13), we know $\max_{u \in V_N} G_{V_N}(u,u) - G_{[u]_0}(u,u) \leq C$. Using the exact same method, we can also easily show that

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

because we would have $\|v - z\|_2 \geq (1-\eta)N/2$ for all $z \in \partial[u]_0 \cap V_N^o$ in the reasoning below (1.4.13), where $\eta < 1$ by hypothesis. Finally, $-G_{V_N}(u,v) \leq -G_{[u]_0}(u,v)$, so proving (1.4.14) boils down to the proof of the following inequality :

$$(\clubsuit) \mathrel{\hat{=}} \left\{ \begin{array}{l} (G_{[u]_0}(u,u) - G_{[u]_0}(u,v)) - (G_{\overline{B}}(u,u) - G_{\overline{B}}(u,v)) \\ + (G_{[u]_0}(v,v) - G_{[u]_0}(v,u)) - (G_{\overline{B}}(v,v) - G_{\overline{B}}(v,u)) \end{array} \right\} \leq \tilde{C}(\eta). \tag{1.4.16}$$

To show (1.4.16), we consider two cases : $d(u, \partial V_N) \leq L$ and $d(u, \partial V_N) > L$.

$$\boxed{\text{Case 1 : } d(u, \partial V_N) \leq L}$$

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

$$(\clubsuit) \leq (G_{[u]_0}(u,u) - G_{\overline{B}}(u,u)) + (G_{[u]_0}(v,v) - G_{\overline{B}}(v,v)). \tag{1.4.17}$$

Note that the box B is cut off by ∂V_N in Case 1. By translating u, v, B together in such a way that u doesn't get closer to ∂V_N with respect to both axes, each difference of Green functions in (1.4.17) can only increase (see the argument below Figure 1.4.4). Therefore, it is sufficient to bound (1.4.17) when $d(u, \partial V_N) = L$. Assume $d(u, \partial V_N) = L$ for the rest of Case 1. Since $d(u, \partial B) = L$ by hypothesis, we have $d(u, \partial \overline{B}) = L$ and we get $d(v, \partial \overline{B}) \geq \lceil (1-\eta)L \rceil \geq 1$ by the triangle inequality. Consequently,

$$(\clubsuit) \leq G_{[u]_0}(u,u) + G_{[u]_0}(v,v) - 2\log L + C(\eta) \tag{1.4.18}$$

using (1.4.8) and (1.4.9).

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

$$(\clubsuit) \leq 2 \max_{\substack{z \in \mathcal{H} \\ \lceil (1-\eta)L \rceil \leq d(z, \partial\mathcal{H}) \leq (1+\eta)L}} G_{\mathcal{H}}(z, z) - 2 \log L + C(\eta). \tag{1.4.19}$$

From Proposition 8.1.1 of Lawler and Limic (2010),

$$G_{\mathcal{H}}(z, z) = a(z - \bar{z}) \overset{(1.4.9)}{=} \log(\|z - \bar{z}\|_2) + \text{const.} + O(\|z - \bar{z}\|_2^{-2}) \tag{1.4.20}$$

where $z = (z_1, z_2)$ and $\bar{z} \mathrel{\hat{=}} (z_1, -z_2)$. The conclusion for Case 1 follows from (1.4.19) because $2 \leq 2\lceil (1 - \eta)L \rceil \leq \|z - \bar{z}\|_2 \leq 2(1 + \eta)L$ in (1.4.20).

$$\boxed{\text{Case 2}: \ d(u, \partial V_N) > L}$$

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

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

The inequality comes from the fact that each pair of braces in the infinite sum is equal to $\mathbb{V}_{[u]_0^k}(\mathbb{E}[\phi_u - \phi_v \,|\, \mathcal{F}_{\partial [u]_0^{k-1}}]) \geq 0$ by steps analogous to (1.4.15). The equality follows because the infinite sum is telescopic.

By conditioning on the point $z \in \partial\overline{B}$ where the simple random walk starting at u or v

will be when hitting the boundary of $\overline{B}$, and using the strong Markov property, we deduce

$$
\begin{aligned}
(\clubsuit) &\leq \sum_{z\in\partial\overline{B}} \Big(\mathscr{P}_u(W_{\tau_{\partial\overline{B}}} = z) - \mathscr{P}_v(W_{\tau_{\partial\overline{B}}} = z)\Big) \cdot \frac{\pi}{2} \cdot \mathscr{E}_z\left[\sum_{k=0}^{\infty}(\mathbf{1}_{\{W_k=u\}} - \mathbf{1}_{\{W_k=v\}})\right] \\
&= \sum_{z\in\partial\overline{B}} \Big(\mathscr{P}_u(W_{\tau_{\partial\overline{B}}} = z) - \mathscr{P}_v(W_{\tau_{\partial\overline{B}}} = z)\Big) \cdot (a(v-z) - a(u-z))
\end{aligned} \tag{1.4.21}
$$

where " a ", the *potential kernel* (see p.37 in Lawler (1991)), is defined by

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

Theorem 1.6.2 in Lawler (1991) shows that this is the same function as in (1.4.9). Therefore, we can evaluate (1.4.21) :

$$
a(v-z) - a(u-z) = \log\left(\frac{\|v-z\|_2}{\|u-z\|_2}\right) + O(\|v-z\|_2^{-2}) - O(\|u-z\|_2^{-2}). \tag{1.4.22}
$$

By the triangle inequality, we have

$$
\log\left(1 - \frac{\|u-v\|_2}{\|u-z\|_2}\right) \leq \log\left(\frac{\|v-z\|_2}{\|u-z\|_2}\right) \leq \log\left(1 + \frac{\|u-v\|_2}{\|u-z\|_2}\right). \tag{1.4.23}
$$

Now, notice that

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

Using the three bullets in (1.4.22) and (1.4.23), we have

$$
\log(1-\eta) - \frac{C_1}{\lceil(1-\eta)L\rceil^2} \leq (1.4.22) \leq \log(1+\eta) + \frac{C_2}{\lceil(1-\eta)L\rceil^2} \tag{1.4.24}
$$

for appropriate constants $C_1, C_2 > 0$. Since $L \geq 1$ and $\lceil(1-\eta)L\rceil \geq 1$, inequality (1.4.16) follows by regrouping (1.4.21), (1.4.22) and (1.4.24). $\square$

Lemma 1.4.5. *Let $0 \leq \lambda' < 1$ and $d \geq 1/\sqrt{2}$. For all $v \in V_N$, define $S_{v,d}$ to be the set of finite boxes $B \subseteq \mathbb{Z}^2$ such that $[v]_{\lambda'} \subseteq B \cap V_N$ and $\max_{z\in\partial B}\|v-z\|_2 \leq dN^{1-\lambda'}$, then there exists a constant $C = C(d) > 0$ such that, for N large enough,*

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

Proof. This follows directly from the calculations in Lemma 1.4.2 and Lemma 1.4.3 where $B \cap V_N$ plays the same role as $[v]_\lambda$. $\square$

The next lemma is used in equation (1.3.4) of Lemma 1.3.1 and equation (1.3.23) of Lemma 1.3.4 to show that the error coming from the approximation of the branching structure of ψ is small enough that the problem of finding the upper bound for the maximum and the log-number of γ-high points is the same (modulo the additional hurdle caused by the decay of variance near the edges of V_N) as in the context of branching random walks.

Lemma 1.4.6. *Let $\lambda_{j-1} < \lambda \leq \lambda_j$ for a certain $j \in \{1, ..., M\}$, then there exists a constant $C = C(\sigma_1, ..., \sigma_j) > 0$ such that*

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

for N large enough.

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

$$\mathbb{V}(\phi_v(\mu) - \phi_{v_\lambda}(\mu)) \leq 3 \cdot \left\{ \begin{array}{c} \mathbb{V}(\phi_v(\mu) - \phi_v(B_\mu)) \\ + \, \mathbb{V}(\phi_v(B_\mu) - \phi_{v_\lambda}(B_\mu)) \\ + \, \mathbb{V}(\phi_{v_\lambda}(B_\mu) - \phi_{v_\lambda}(\mu)) \end{array} \right\} \leq \tilde{C}. \tag{1.4.25}$$

To see the last inequality, bound the first and third variance term inside the braces using Lemma 1.4.5 with $d = 3/\sqrt{2}$ and bound the second variance term inside the braces using Lemma 1.4.4 with $\eta = 1/\sqrt{2}$ (since $\|v - v_\lambda\|_2 \leq N^{1-\lambda}/\sqrt{2} \leq N^{1-\mu}/\sqrt{2}$), $u = v_\lambda$ and $L = \lceil N^{1-\mu} \rceil$. Now, from (1.4.3) and Jensen's inequality, we get

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

Simply use (1.4.25) to bound each variance term inside the braces by a constant. This ends the proof of the lemma. $\square$

Lemma 1.4.7 (Gaussian estimates, see e.g. Adler and Taylor (2007)). *Suppose that $Z \sim \mathcal{N}(0, \sigma^2)$ where $\sigma > 0$, then for all $z > 0$,*

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

Lemma 1.4.8 (Paley and Zygmund (1932) inequality). *Let $0 \leq X \in L^2(\mathbb{P})$ be such that $\mathbb{P}(X > 0) > 0$, then for all $0 \leq \theta \leq 1$,*

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

1.4.2. Karush-Kuhn-Tucker theorem and applications

In this section, we state the Karush-Kuhn-Tucker theorem and the solutions to the two optimization problems posed in Section 1.2. The optimal path for the maximum, $\lambda \mapsto L_N^\star(\lambda)$, comes from the solution to the problem stated in Lemma 1.4.10 while the optimal path for γ-high points, $\lambda \mapsto L_N^\gamma(\lambda)$, comes from the solution to the problem stated in Lemma 1.4.11. The Karush-Kuhn-Tucker theorem only gives, a priori, necessary conditions for local optimality. However, the conditions are also sufficient for global optimality here because the objective function f_γ below and the constraint functions g_k are continuously differentiable and concave ($f_{\gamma^\star}$ is linear), see Hanson (1981). The proof of the two lemmas can be found in Appendix A of Ouimet (2014) and are direct applications of the theorem.

Theorem 1.4.9 (Karush-Kuhn-Tucker, see e.g. Delfour (2012)). *Let $f : \mathbb{R}^{n_1} \to \mathbb{R}$ be an objective function and let*

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

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

(a) f attains a local maximum at $\boldsymbol{x}^\star \in \mathcal{U}^{\geq}$ with respect to $\mathcal{U}^{\geq}$;

(b) f is Fréchet differentiable at $\boldsymbol{x}^\star$;

(c) the g_k's are Fréchet differentiable at $\boldsymbol{x}^\star$.

When the constraints qualify (they do in Lemma 1.4.10 and Lemma 1.4.11 because the g_k's are concave and $\mathbf{0} \in \mathcal{U}^{>}$, see Slater's condition in Delfour (2012)), then there exists

$(\mu_1, ..., \mu_{n_2}) \in \mathbb{R}^{n_2}$ *such that the following points hold for all* $k \in \{1, ..., n_2\}$ *(*∇ *is the gradient here) :*

(1) $\nabla f(\boldsymbol{x}^\star) + \sum_{k=1}^{n_2} \mu_k \nabla g_k(\boldsymbol{x}^\star) = 0$*;*

(2) $g_k(\boldsymbol{x}^\star) \geq 0$*;*

(3) $\mu_k \geq 0$*;*

(4) $\mu_k g_k(\boldsymbol{x}^\star) = 0$*.*

Lemma 1.4.10. *(Optimal path for the maximum) Let*

$$f_{\gamma^\star}(x_1, ..., x_M) \mathrel{\hat{=}} \sum_{i=1}^{M} x_i$$

be the objective function to maximize under the constraints

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

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

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

and the maximum is given by

$$f_{\gamma^\star}(x_1^\star, ..., x_M^\star) = \mathcal{J}_{\sigma^2/\bar{\sigma}}(1) \mathrel{\hat{=}} \gamma^\star.$$

Lemma 1.4.11. *(Optimal path for* γ*-high points) Let* $\gamma^{l-1} \leq \gamma < \gamma^l$ *for a certain* $l \in \{1, ..., m\}$*, where the critical levels* γ^l *are defined in (1.1.5). Furthermore, let*

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

be the objective function to maximize under the constraints

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

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

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

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

$$f_\gamma(x_1^\star, ..., x_{M-1}^\star) = (1-\lambda^{l-1}) - \frac{\left(\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1})\right)^2}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} \mathrel{\hat{=}} \mathcal{E}_\gamma.$$

Acknowledgements

The authors would like to thank an anonymous referee for a careful review of the preliminary version. We also gratefully acknowledge insightful discussions with David Belius, Anton Bovier, Nicola Kistler and Olivier Zindy. This work is supported by a NSERC Doctoral Program Alexander Graham Bell scholarship (CGS D3).

Bibliography

R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Monographs in Mathematics. Springer, New York, 2007. MR2319516.

L.-P. Arguin and O. Zindy. Poisson-Dirichlet statistics for the extremes of the two-dimensional discrete Gaussian free field. *Electron. J. Probab.*, 20:no. 59, 19, 2015. MR3354619.

M. Biskup and O. Louidor. Extreme local extrema of two-dimensional discrete Gaussian free field. *Comm. Math. Phys.*, 345(1):271–304, 2016. MR3509015.

E. Bolthausen, J.-D. Deuschel, and G. Giacomin. Entropic repulsion and the maximum of the two-dimensional harmonic crystal. *Ann. Probab.*, 29(4):1670–1692, 2001. MR1880237.

A. Bovier and L. Hartung. The extremal process of two-speed branching Brownian motion. *Electron. J. Probab.*, 19:no. 18, 28, 2014. MR3164771.

A. Bovier and L. Hartung. Variable speed branching Brownian motion 1. Extremal processes in the weak correlation regime. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12(1): 261–291, 2015. MR3351476.

A. Bovier and I. Kurkova. Derrida's generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):481–495, 2004. MR2070335.

M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.*, 69(1):62–123, 2016. MR3433630.

D. Capocaccia, M. Cassandro, and P. Picco. On the existence of thermodynamics for the generalized random energy model. *J. Statist. Phys.*, 46(3-4):493–505, 1987. MR883541.

O. Daviaud. Extremes of the discrete two-dimensional Gaussian free field. *Ann. Probab.*, 34(3):962–986, 2006. MR2243875.

M. C. Delfour. *Introduction to optimization and semidifferential calculus*, volume 12 of *MOS-SIAM Series on Optimization*. SIAM, Philadelphia, 2012. MR2954022.

B. Derrida. A generalization of the Random Energy Model which includes correlations between energies. *J. Physique Lett.*, 46(9):401–407, 1985. doi:10.1051/jphyslet:01985004609040100.

B. Duplantier and S. Sheffield. Liouville quantum gravity and KPZ. *Invent. Math.*, 185 (2):333–393, 2011. MR2819163.

E. B. Dynkin. Markov processes and random fields. *Bull. Amer. Math. Soc. (N.S.)*, 3(3): 975–999, 1980. MR585179.

M. Fang. Tightness for maxima of generalized branching random walks. *J. Appl. Probab.*, 49(3):652–670, 2012. MR3012090.

M. Fang and O. Zeitouni. Branching random walks in time inhomogeneous environments. *Electron. J. Probab.*, 17:no. 67, 18, 2012a. MR2968674.

M. Fang and O. Zeitouni. Slowdown for time inhomogeneous branching Brownian motion. *J. Stat. Phys.*, 149(1):1–9, 2012b. MR2981635.

M. A. Hanson. On sufficiency of the Kuhn-Tucker conditions. *J. Math. Anal. Appl.*, 80 (2):545–550, 1981. MR614849.

X. Hu, J. Miller, and Y. Peres. Thick points of the Gaussian free field. *Ann. Probab.*, 38 (2):896–926, 2010. MR2642894.

N. Kistler. Derrida's random energy models. From spin glasses to the extremes of correlated random fields. In *Correlated Random Systems: Five Different Methods*, volume 2143 of *Lecture Notes in Mathematics*. Springer, 2015. MR3380419.

N. Kistler and M. A. Schmidt. From Derrida's random energy model to branching random walks: from 1 to 3. *Elec. Comm. Prob.*, 20:no. 47, 12, 2015. MR3358969.

G. F. Lawler. *Intersections of random walks*. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, 1991. MR1117680.

G. F. Lawler and V. Limic. *Random walk: a modern introduction*, volume 123 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. MR2677157.

P. Maillard and O. Zeitouni. Slowdown in branching Brownian motion with inhomogeneous variance. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(3):1144–1160, 2016. MR3531703.

B. Mallein. Maximal displacement of a branching random walk in time-inhomogeneous environment. *Stoch. Proc. Appl.*, 125(10):3958–4019, 2015. MR3373310.

F. Ouimet. *Étude du maximum et des hauts points de la marche aléatoire branchante inhomogène et du champ libre gaussien inhomogène*. Master's thesis, Université de Montréal, 2014. URL `https://papyrus.bib.umontreal.ca/xmlui/handle/1866/11510`.

F. Ouimet. Maxima of branching random walks with piecewise constant variance. *Braz. J. Probab. Stat.*, 32(4):679–706, 2018. MR3845025.

R.E.A.C. Paley and A. Zygmund. A note on analytic functions in the unit circle. *Math. Proc. Cambridge Philos. Soc.*, 28(3):266–272, 1932. doi:10.1017/S0305004100010112.

S. Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields*, 139 (3-4):521–541, 2007. MR2322706.

O. Zeitouni. Gaussian fields. *Notes for Lectures*, 2017. URL `http://www.cims.nyu.edu/~zeitouni/notesGauss.pdf`.

Article 2

Geometry of the Gibbs measure for the discrete 2D Gaussian free field with scale-dependent variance

by

Frédéric Ouimet[1]

([1]) Université de Montréal, Montréal, QC H3T 1J4, Canada.

Here is the complete citation :

Ouimet, F. (2017). Geometry of the Gibbs measure for the discrete 2D Gaussian free field with scale-dependent variance. ALEA, Lat. Am. J. Probab. Math. Stat. 14, no. 2, 851-902. http://alea.impa.br/articles/v14/14-38.pdf.

Abstract. We continue our study of the scale-inhomogeneous Gaussian free field introduced in Arguin and Ouimet (2016). Firstly, we compute the limiting free energy on V_N and adapt a technique of Bovier and Kurkova (2004b) to determine the limiting two-overlap distribution. The adaptation was already successfully applied in the simpler case of Arguin and Zindy (2015), where the limiting free energy was computed for the field with two levels (in the center of V_N) and the limiting two-overlap distribution was determined in the homogeneous case. Our results agree with the analogous quantities for the Generalized Random Energy Model (GREM); see Capocaccia et al. (1987) and Bovier and Kurkova (2004a), respectively. Secondly, we show that the extended Ghirlanda-Guerra identities hold exactly in the limit. As a corollary, the limiting array of overlaps is ultrametric and the limiting Gibbs measure has the same law as a Ruelle probability cascade.

Keywords: Gaussian free field, Gibbs measure, inhomogeneous environment, Ghirlanda-Guerra identities, ultrametricity, spin glasses, Ruelle probability cascades

2.1. The model

Let $(W_k)_{k\geq 0}$ be a simple random walk starting at $u \in \mathbb{Z}^2$ with law $\mathscr{P}_u$. For every finite box $B \subseteq \mathbb{Z}^2$, the Gaussian free field (GFF) on B is a centered Gaussian field $\phi \circeq \{\phi_v\}_{v\in B}$ with covariance matrix

$$G_B(v,v') \circeq \frac{\pi}{2} \cdot \mathscr{E}_v\left[\sum_{k=0}^{\tau_{\partial B}-1} \mathbf{1}_{\{W_k=v'\}}\right], \quad v,v' \in B, \tag{2.1.1}$$

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

$$\partial B \circeq \{v \in B : \exists z \in \mathbb{Z}^2 \backslash B \ \text{ such that } \|v-z\|_2 = 1\}, \tag{2.1.2}$$

and $\|\cdot\|_2$ denotes the Euclidean distance in $\mathbb{Z}^2$. With this definition, B contains its boundary. We let $B^o \circeq B\backslash\partial B$. By convention, summations are zero when there are no indices, so ϕ is identically zero on ∂B. This is the *Dirichlet boundary condition*. The constant $\pi/2$ in (2.1.1) is a convenient normalization for the variance.

We build a family of Gaussian fields constructed from the GFF $\{\phi_v^N\}_{v\in V_N}$ on the square box $V_N \circeq \{0,1,...,N\}^2$. For $\lambda \in (0,1)$ and $v = (v_1,v_2) \in V_N$, consider the closed neighborhood $[v]_\lambda$ in V_N consisting of the square box of width $N^{1-\lambda}$ centered at v that has been cut off by the boundary of V_N :

$$[v]_\lambda \circeq \left((v_1,v_2) + \left[-\frac{1}{2}N^{1-\lambda}, \frac{1}{2}N^{1-\lambda}\right]^2\right) \bigcap V_N. \tag{2.1.3}$$

By convention, define $[v]_0 \circeq V_N$ and $[v]_1 \circeq \{v\}$. Let $\mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c} \circeq \sigma(\{\phi_v^N, v \notin [v]_\lambda^o\})$ be the σ-algebra generated by the variables on the boundary of the box $[v]_\lambda$ and those outside of it. Since the neighborhoods are shrinking when λ increases, for any $v \in V_N$, the collection $\mathbb{F}_v \circeq \{\mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}\}_{\lambda \in [0,1]}$ is a filtration. In particular, if we let

$$\phi_v^N(\lambda) \circeq \mathbb{E}\Big[\phi_v^N \,|\, \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}\Big], \tag{2.1.4}$$

then

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

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

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

$$\boldsymbol{\sigma} \circeq (\sigma_1, \sigma_2, ..., \sigma_M) \in (0, \infty)^M, \qquad \text{(variance parameters)} \tag{2.1.5}$$

$$\boldsymbol{\lambda} \circeq (\lambda_1, \lambda_2, ..., \lambda_M) \in (0, 1]^M, \qquad \text{(scale parameters)} \tag{2.1.6}$$

where

$$0 \circeq \lambda_0 < \lambda_1 < ... < \lambda_M \circeq 1. \tag{2.1.7}$$

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

$$\nabla \phi_v^N(\lambda_i) \circeq \phi_v^N(\lambda_i) - \phi_v^N(\lambda_{i-1}). \tag{2.1.8}$$

Definition 2.1.1 (Scale-inhomogeneous Gaussian free field)**.** *Let $\{\phi_v^N\}_{v \in V_N}$ be the GFF on V_N. The $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF on V_N is a Gaussian field $\{\psi_v^N\}_{v \in V_N}$ defined by*

$$\psi_v^N \circeq \sum_{i=1}^M \sigma_i \nabla \phi_v^N(\lambda_i) = \sum_{i=1}^M \sigma_i \Big(\phi_v^N(\lambda_i) - \phi_v^N(\lambda_{i-1})\Big). \tag{2.1.9}$$

Similarly to the GFF, we define

$$\psi_v^N(\lambda) \circeq \mathbb{E}\Big[\psi_v^N \,|\, \mathcal{F}_{\partial[v]_\lambda \cup [v]_\lambda^c}\Big] \quad \text{and} \quad \psi_v^N(\lambda, \lambda') \circeq \psi_v^N(\lambda') - \psi_v^N(\lambda). \tag{2.1.10}$$

From hereon, we make the dependence on N implicit everywhere for ϕ and ψ.

The $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF is the analogue of various types of inhomogeneous branching processes :

- The GREM, see e.g. Bovier and Kurkova (2004a,b); Capocaccia et al. (1987); Derrida (1985); Ruelle (1987);
- The non-hierarchical GREM, see Bolthausen and Kistler (2006, 2009);
- The perceptron GREM, see Bolthausen and Kistler (2012);
- The multi-scale logarithmic potential (also called multi-scale log-REM), see e.g. Fyodorov and Bouchaud (2008a); Cao et al. (2016);
- The branching random walk in time-inhomogeneous environment, see e.g. Fang and Zeitouni (2012a); Mallein (2015b,a); Ouimet (2018);
- The variable speed branching Brownian motion, see e.g. Bovier and Hartung (2014, 2015); Fang and Zeitouni (2012b); Maillard and Zeitouni (2016).

Remark 2.1.1. *Our model is most closely related to the multi-scale log-REM of Fyodorov and Bouchaud (2008a). In both cases :*

(1) There is a boundary effect, meaning that the variance of the variables of the field decays to 0 *as we approach the boundary;*

(2) There is no exact hierarchical structure;

(3) The covariance between two random variables of the field is directly tied to the distance between their index in a finite-dimensional Euclidean space;

(4) The number of possible covariance values grows with the size of the system.

Amongst the other models, not one has property 1 or 3, only the non-hierarchical GREM has property 2, and only the time-inhomogeneous branching random walk and the variable speed branching Brownian motion have property 4.

The only significant difference between the single-scale log-REM and the critical GFF is the fact that the field is indexed in a N'*-dimensional space (instead of 2-dimensional) and the covariance is defined to be logarithmic directly instead of it being a consequence of estimates on the Green function in two dimensions. Fyodorov and Bouchaud (2008a) calculate the limiting free energy and the limiting two-overlap distribution in the limit* $N' \to \infty$ *(for a finite system size) by using the replica trick and Parisi's hierarchical ansatz. In the thermodynamic limit, they recover the same structure as in the GREM case and argue that the results should also hold if* N' *is fixed instead and the system size grows* *to infinity. In this article, we put their argument on rigorous ground via the* $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$*-GFF. We go even further by showing that the limiting Gibbs measure has the same law as the one found in Bovier and Kurkova (2004a) for the GREM. For an introduction to log-REM models and physical motivations, see e.g. Carpentier and Le Doussal (2001); Fyodorov and Bouchaud (2008b,a); Fyodorov et al. (2009); Cao et al. (2016).*

2.2. Motivation for the scale-inhomogeneous GFF

In contrast with branching random walks (BRWs) :

— The branching structure is *approximate* in the sense that $\phi_v(\lambda)$ and $\phi_{v'}(\lambda)$ are not perfectly correlated when λ is smaller than the *branching scale*, namely the largest scale at which $[v]_\lambda$ and $[v']_\lambda$ intersect. The branching scale itself is arbitrarily defined since it is conceptually more of a transition interval : between the scale where v' is "well-inside" $[v]_\lambda$ and the smallest scale for which $[v]_\lambda \cap [v']_\lambda = \emptyset$.

— At a given scale, the covariance of the increments of the field decays near the boundary of the domain. In the context of BRWs, this means that at a given time, the law of each point process would depend on the position of the associated ancestors in the tree.

Several covariance estimates can be found in Appendix 2.9.1.

We are interested in the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF to see how the results on the extremes (and the methods of proof) are robust to perturbations in the correlation structure. This interest is amplified by the fact that many models in recent applications have underlying approximate branching structures. For example :

- Cover times, see e.g. Abe (2014); Belius (2013); Belius and Kistler (2017); Comets et al. (2013); Dembo et al. (2003, 2004, 2006); Ding (2012, 2014); Ding et al. (2012); Ding and Zeitouni (2012);
- The randomized Riemann zeta function on the critical line, see e.g. Arguin et al. (2017b); Arguin and Tai (2018); Harper (2013); Saksman and Webb (2018);
- The Riemann zeta function on random intervals of the critical line, see e.g. Arguin et al. (2019); Najnudel (2018);
- The characteristic polynomials of random unitary matrices, see e.g. Arguin et al. (2017a); Chhaibi et al. (2018); Paquette and Zeitouni (2018).

In particular, note that all these models are heavily correlated in the *critical regime*, that is when the correlation starts to affect the extremal statistics.

Generally, there are two ways to study the distribution of the extremes : via the *extremal process* and via the *Gibbs measure*. Since one of our goal here is to show the tree structure of the extremes in the limit $N \to \infty$ (this interest comes partly from the physicists and the Parisi ultrametricity conjecture for mean field spin glass models (see e.g. Bovier (2006); Mézard et al. (1987); Panchenko (2013b); Talagrand (2011a,b))), the mathematics in the latter approach is much simpler (at least in our case). A very important advancement was made recently in Panchenko (2013a) where it is shown that a random measure supported on the unit ball of a separable Hilbert space that satisfies the extended Ghirlanda-Guerra identities must have an ultrametric support with probability one. The summary in Section 2.3 gives a detailed description of the steps we will make to prove the extended Ghirlanda-Guerra identities and the consequences we can deduce from the work of Panchenko.

Remarkably, despite the imperfect branching structure of the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF and the growing number of scales as $N \to \infty$, the results of this paper show that the *limiting Gibbs measure* has the same tree structure as in the context of the GREM. More precisely, we show that the limiting Gibbs measure has the same law as a *Ruelle probability cascade* (see Ruelle (1987)) with *functional order parameter* ζ defined by the *limiting two-overlap distribution*. This is the content of Corollary 2.7.2. In the limit, this means, in particular, that the extremes of the model are clustered in a hierarchical way and in fact satisfy the *ultrametric inequality*, see Corollary 2.7.1.

Another reason why the study of the extremes via the Gibbs measure might be more desirable to prove ultrametricity results is its robustness, i.e. the applicability of the methods to other models. For instance, Jagannath (2017) defines the notion of *approximate ultrametricity* for finite system sizes by imposing conditions on the sequence of Gibbs measures. It is proved that if the sequence of two-overlap distributions converge weakly and the *approximate extended Ghirlanda-Guerra identities* are satisfied (in his sense, see Definition 1.4 in Jagannath (2017)), then the sequence of Gibbs measures (assuming they are supported on the unit ball of a separable Hilbert space) is *regularly approximately ultrametric*. His paper tie in very nicely with our approach since we prove the weak convergence of the two-overlap distribution in Theorem 2.6.3, we prove a slightly different version of the approximate extended Ghirlanda-Guerra identities in Theorem 2.6.4, and then we show that the identities must hold exactly in the limit (Theorem 2.6.5). Of course, this doesn't prove that our model is *regularly approximately ultrametric*, but it seems at least plausible that the notion of approximate ultrametricity could hold for a large class of non-hierarchical models and could be (part of) the grand explanation behind the phenomenon of ultrametricity of the extremes in the system size limit.

2.3. Structure of the paper

In order to make the logical structure of this article as clear as possible, the new results and their proof are stated and written in a linear fashion. Some technical lemmas are relegated to appendices. However, we emphasize that these lemmas are not at all necessary to understand the main structure and are sparsely used. Below, we summarize the main results of the paper, give the main ingredients of the proofs and indicate exactly where the lemmas in the appendices are needed.

Section 2.5 : We recall the main results of Arguin and Ouimet (2016) :

- Theorem 2.5.1 : $\max_{v\in V_N} \psi_v / \log N^2 \xrightarrow{\mathbb{P}} \gamma^\star$.
- Theorem 2.5.2 : $\log(|\{v \in V_N : \psi_v \geq \gamma \log N^2\}|)/\log N^2 \xrightarrow{\mathbb{P}} \mathcal{E}(\gamma)$.

Section 2.6 : The new main results are stated :

- Theorem 2.6.1 : Limit of the free energy on V_N :

  $$\tfrac{1}{\log N^2} \log \textstyle\sum_{v\in V_N} e^{\beta\psi_v} \xrightarrow{\mathbb{P} \text{ and } L^p} \max_{\gamma\in[0,\gamma^\star]} \beta\gamma + \mathcal{E}(\gamma) = f^\psi(\beta).$$

  The proof uses the results of Section 2.5 for the $\mathbb{P}$-convergence and we show that the powers of the free energy are uniformly integrable. We find the explicit form of the maximum on the right-hand side using

  — Lemma 2.9.7 : Differentiability of $\mathcal{E}$,

  — Lemma 2.9.8 : Solution of the maximization problem.
- Theorem 2.6.2 : Same as Theorem 2.6.1, but on a set far enough from the boundary of V_N, denoted $A_{N,\rho}$.

- Theorem 2.6.3 : If $\mathcal{G}_{\beta,N}$ denotes the Gibbs measure of ψ and $q^N(v,v')$ denotes the normalized covariance (overlap) between ψ_v and $\psi_{v'}$, then we compute the limit of the *two-overlap distribution*, namely $r \mapsto \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\left[\mathbf{1}_{\{q^N(v,v')\leq r\}}\right]$. The main ingredients of the proof are :
  — The Gibbs measure doesn't hold any weight outside $A_{N,\rho}$ in the limit,
  — The overlap estimates of Corollary 2.9.6,
  — Gaussian integration by parts,
  — The mean convergence of the derivative at $u = 0$ of a perturbed version of the free energy to $f^{\psi^u}(\beta)$. This is proved using
    - Theorem 2.6.2,
    - Lemma 2.9.9 : Convexity of the free energy with respect to u,
    - Lemma 2.9.10 : Differentiability of $u \mapsto f^{\psi^u}(\beta)$ for all $|u| < \delta$.
- Theorem 2.6.4 : As $N \to \infty$, the extended Ghirlanda-Guerra identities hold approximately. The main ingredients are the same as in the proof of Theorem 2.6.3, but we also need Theorem 2.6.2 and the main ingredients to get a concentration result (Lemma 2.8.9).
- Theorem 2.6.5 : In the limit, the extended Ghirlanda-Guerra identities hold exactly. The main ingredients of the proof are :
  — Theorem 2.6.3 and Theorem 2.6.4,
  — The representation theorem of Dovbysh and Sudakov (1982).

Section 2.7 : The consequences of Theorem 2.6.3 and Theorem 2.6.5 :

- Corollary 2.7.1 : The limiting array of overlaps is almost-surely *ultrametric* under the mean of the limiting Gibbs measure.
- Corollary 2.7.2 : The limiting Gibbs measure has the same law as a Ruelle probability cascade Ruelle (1987) with functional order parameter ζ defined by the limiting two-overlap distribution, which means that it samples extreme values from an *exact* tree structure, where the number of branching scales is finite and controlled by the *inverse temperature* β.

Section 2.8 : Proof of the main results.

Appendix 2.9.1 : The sole purpose of this appendix is to prove Corollary 2.9.6.

Appendix 2.9.2 : We indicated above where each of its four lemmas are needed.

2.4. Some notations

Now, we introduce some notations. The parameters $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$ can be encoded simultaneously in the left-continuous step function

$$\sigma(s) \mathrel{\hat{=}} \sigma_1 \mathbf{1}_{\{0\}}(s) + \sum_{i=1}^{M} \sigma_i \mathbf{1}_{(\lambda_{i-1}, \lambda_i]}(s), \quad s \in [0,1]. \tag{2.4.1}$$

For any positive measurable function $f : [0,1] \to \mathbb{R}$, define the integral operators

$$\mathcal{J}_f(s) \mathrel{\hat{=}} \int_0^s f(r)dr \quad \text{and} \quad \mathcal{J}_f(s_1, s_2) \mathrel{\hat{=}} \int_{s_1}^{s_2} f(r)dr. \tag{2.4.2}$$

We refer to $\mathcal{J}_{\sigma^2}(\cdot)$ as the *speed function*. The *concavification* of $\mathcal{J}_{\sigma^2}$, denoted $\hat{\mathcal{J}}_{\sigma^2}$, is the function whose graph is the concave hull of $\mathcal{J}_{\sigma^2}$. From Arguin and Ouimet (2016), we know that the asymptotics of the maximum and the log-number of high points of ψ are controlled by this function. Its graph is an increasing and concave polygonal line, see Figure 2.4.1 for an example.

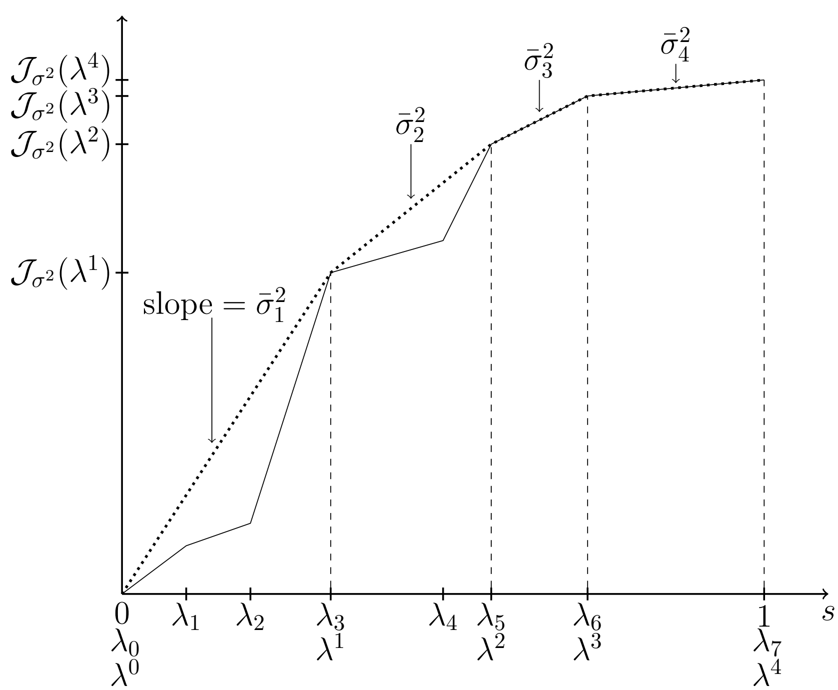


FIGURE 2.4.1. Example of $\mathcal{J}_{\sigma^2}$ (closed line) and $\hat{\mathcal{J}}_{\sigma^2}$ (dotted line) with 7 values for σ^2.

Clearly, there exists a unique non-increasing left-continuous step function $s \mapsto \bar{\sigma}(s)$ such that

$$\hat{\mathcal{J}}_{\sigma^2}(s) = \mathcal{J}_{\bar{\sigma}^2}(s) = \int_0^s \bar{\sigma}^2(r)\, dr \quad \text{for all } s \in (0,1]. \tag{2.4.3}$$

The scales in $[0,1]$ where $\bar{\sigma}$ jumps are denoted by

$$0 \mathrel{\hat{=}} \lambda^0 < \lambda^1 < \ldots < \lambda^m \mathrel{\hat{=}} 1, \tag{2.4.4}$$

where $m \le M$. To be consistent with previous notations, we set $\bar{\sigma}_l \mathrel{\hat{=}} \bar{\sigma}(\lambda^l)$. In particular, note that

$$\bar{\sigma}_1 > \bar{\sigma}_2 > \ldots > \bar{\sigma}_m. \tag{2.4.5}$$

We define $\bar{\sigma}_{m+1} \mathrel{\hat{=}} 0$ and interpret $\beta_c(\bar{\sigma}_{m+1}) \mathrel{\hat{=}} 2/\bar{\sigma}_{m+1}$ as $+\infty$ whenever it is encountered.

2.5. Previous results

In this section, we recall the main results from Arguin and Ouimet (2016) on the first order asymptotics of the maximum and the log-number of γ-high points of the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF. These results are needed to compute the limiting free energy.

Theorem 2.5.1. *Let $\{\psi_v\}_{v \in V_N}$ be the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF on V_N of Definition 2.1.1, then*

$$\lim_{N \to \infty} \frac{\max_{v \in V_N} \psi_v}{\log N^2} = \mathcal{J}_{\sigma^2/\bar{\sigma}}(1) \mathrel{\hat{=}} \gamma^\star \quad \text{in probability}. \tag{2.5.1}$$

In fact, from Lemma 3.1 and Lemma 3.3 in Arguin and Ouimet (2016), for any $\varepsilon > 0$, there exists a constant $c = c(\varepsilon, \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ such that for N large enough,

$$\mathbb{P}\left(\max_{v \in V_N} \psi_v \ge (1+\varepsilon)\gamma^\star \log N^2\right) \le N^{-c} \tag{2.5.2}$$

and

$$\mathbb{P}\left(\max_{v \in V_N} \psi_v \le (1-\varepsilon)\gamma^\star \log N^2\right) \le N^{-c}. \tag{2.5.3}$$

The set of γ-high points of the field ψ is defined as

$$\mathcal{H}_N(\gamma) \mathrel{\hat{=}} \{v \in V_N : \psi_v \ge \gamma \log N^2\}, \quad \text{for all } \gamma \in [0, \gamma^\star]. \tag{2.5.4}$$

The number of γ-high points depends on *critical levels* defined by

$$\gamma^l \mathrel{\hat{=}} \int_0^1 \frac{\sigma^2(s)}{\bar{\sigma}(s \wedge \lambda^l)} ds, \quad 1 \le l \le m, \qquad \gamma^0 \mathrel{\hat{=}} 0. \tag{2.5.5}$$

For $\gamma \in (\gamma^{l-1}, \gamma^l]$, define

$$\mathcal{E}(\gamma) \doteq (1 - \lambda^{l-1}) - \frac{(\gamma - \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))^2}{\mathcal{J}_{\sigma^2}(\lambda^{l-1}, 1)} \quad \text{and} \quad \mathcal{E}(0) \doteq 1. \tag{2.5.6}$$

Theorem 2.5.2. *Let $\{\psi_v\}_{v \in V_N}$ be the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF on V_N of Definition 2.1.1 and let $\gamma \in [0, \gamma^\star)$, then*

$$\lim_{N \to \infty} \frac{\log |\mathcal{H}_N(\gamma)|}{\log N^2} = \mathcal{E}(\gamma) \quad \textit{in probability}. \tag{2.5.7}$$

In fact, from Lemma 3.4 and Lemma 3.5 in Arguin and Ouimet (2016), for any $\gamma \in [0, \gamma^\star]$ and for any $\varepsilon > 0$, there exists a constant $c = c(\gamma, \varepsilon, \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ such that for N large enough,

$$\mathbb{P}\Big(|\mathcal{H}_N(\gamma)| \geq N^{2\mathcal{E}(\gamma) + \varepsilon}\Big) \leq N^{-c}, \tag{2.5.8}$$

and, when $\gamma \in [0, \gamma^\star)$,

$$\mathbb{P}\Big(|\mathcal{H}_N(\gamma)| < N^{2\mathcal{E}(\gamma) - \varepsilon}\Big) \leq N^{-c}. \tag{2.5.9}$$

Remark 2.5.1. *The case $\gamma = 0$ is not explicitly covered in Arguin and Ouimet (2016). In that case, (2.5.8) is trivial (the probability is 0) and (2.5.9) is a simple and direct application of the Paley-Zygmund inequality.*

2.6. New results

The first main result of this article concerns the *free energy* of the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF, which is defined by

$$f_N^\psi(\beta) \doteq \frac{1}{\log N^2} \log Z_N^\psi(\beta), \quad \beta > 0, \tag{2.6.1}$$

where $Z_N^\psi(\beta) \doteq \sum_{v \in V_N} e^{\beta \psi_v}$. The L^1-limit of the free energy will be central to obtain the limiting two-overlap distribution of the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF and the extended Ghirlanda-Guerra identities. This limit is better expressed in terms of the limiting free energy of the REM(σ) model consisting of N^2 i.i.d. Gaussian variables of variance $\sigma^2 \log N$. From Theorem 8.1 in Bolthausen and Sznitman (2002),

$$f^{\mathrm{REM}(\sigma)}(\beta) \doteq \lim_{N \to \infty} \frac{\log Z_N^{\mathrm{REM}(\sigma)}(\beta)}{\log N^2} \overset{\text{a.s.}}{=} \begin{cases} 2(\beta/\beta_c(\sigma)), & \text{if } \beta > \beta_c(\sigma), \\ 1 + (\beta/\beta_c(\sigma))^2, & \text{if } \beta \leq \beta_c(\sigma), \end{cases} \tag{2.6.2}$$

for all $\beta > 0$, where $\beta_c(\sigma) \doteq 2/\sigma$.

Theorem 2.6.1 (Limit of the free energy on V_N). *Let $\{\psi_v\}_{v\in V_N}$ be the $(\boldsymbol{\sigma},\boldsymbol{\lambda})$-GFF on V_N of Definition 2.1.1, then*

$$\lim_{N\to\infty} f_N^{\psi}(\beta) = \max_{\gamma\in[0,\gamma^\star]} (\beta\gamma + \mathcal{E}(\gamma)) = \sum_{j=1}^{m} f^{REM(\bar{\sigma}_j)}(\beta)\nabla\lambda^j \mathrel{\hat=} f^{\psi}(\beta), \tag{2.6.3}$$

where the limit holds in probability and in L^p, $1 \le p < \infty$.

Remark 2.6.1. *This result was first proved for the GREM by Capocaccia et al. (1987), although with vastly different notations. The expression for the GREM can also be recovered from Theorem 1.6 in Bovier and Kurkova (2004a).*

In Arguin and Zindy (2015), the convergence in probability (and in L^1) of the free energy of the scale-inhomogeneous GFF with two variance parameters was only proved on a subset of V_N that excludes the points that are too close to the boundary ∂V_N. This was done because the decay of variance near ∂V_N makes the asymptotics of the log-number of points in sets of the form

$$\mathcal{H}_{A_N}(\gamma) \mathrel{\hat=} \{v \in A_N : \psi_v \ge \gamma \log N^2\} \tag{2.6.4}$$

harder to determine when A_N includes points close to ∂V_N. In contrast, let

$$A_{N,\rho} \mathrel{\hat=} \left\{ v \in V_N : \min_{z\in\mathbb{Z}^2\setminus V_N} \|v - z\|_2 \ge N^{1-\rho} \right\}, \quad \rho \in (0,1]. \tag{2.6.5}$$

Then, it turns out that for $\rho > 0$ small enough, the variance of the increments of ψ on $A_{N,\rho}$ is within a uniform bound from the analogue quantity in the context of the GREM (except when the scale 0 is involved), see Lemma 2.9.2.

Theorem 2.6.1 not only generalizes Theorem 2.1 in Arguin and Zindy (2015), but is also stronger because it tells us that including points arbitrarily close to ∂V_N in the free energy has no impact on its limit, as long as we include the center of V_N. We are able to prove Theorem 2.6.1 here because the asymptotics of $|\mathcal{H}_N(\gamma)|$ were proved on V_N in Arguin and Ouimet (2016).

Even though Theorem 2.6.1 is interesting on its own, we will instead use the version on $A_{N,\rho}$ later in this article.

Theorem 2.6.2 (Limit of the free energy on $A_{N,\rho}$). *Let $\{\psi_v\}_{v\in V_N}$ be the $(\boldsymbol{\sigma},\boldsymbol{\lambda})$-GFF on V_N of Definition 2.1.1 and define*

$$f_{N,\rho}^{\psi}(\beta) \doteq \frac{1}{\log N^2} \log Z_{N,\rho}^{\psi}(\beta), \quad \beta > 0, \tag{2.6.6}$$

where $Z_{N,\rho}^{\psi}(\beta) \doteq \sum_{v\in A_{N,\rho}} e^{\beta\psi_v}$. Then, for all $\rho \in (0,1]$,

$$\lim_{N\to\infty} f_{N,\rho}^{\psi}(\beta) = f^{\psi}(\beta), \tag{2.6.7}$$

where the limit holds in probability and in L^p, $1 \le p < \infty$.

For the second main result of this article, consider the normalized covariances or *overlaps* of the $(\boldsymbol{\sigma},\boldsymbol{\lambda})$-GFF :

$$q^N(v,v') \doteq \frac{\mathbb{E}[\psi_v\psi_{v'}]}{\mathcal{J}_{\sigma^2}(1)\log N + C_0}, \quad v,v' \in V_N, \tag{2.6.8}$$

where C_0 is the constant introduced in Lemma 2.9.3. The overlap is the covariance divided by the uniform upper bound on the variance. From the Cauchy-Schwarz inequality, it is clear that $|q^N(v,v')| \le 1$, for any $v,v' \in V_N$.

We are concerned with the limiting distribution of the overlaps, when the variables are sampled from the *Gibbs measure*

$$\mathcal{G}_{\beta,N}(\{v\}) \doteq \frac{e^{\beta\psi_v}}{Z_N^{\psi}(\beta)}, \quad v \in V_N. \tag{2.6.9}$$

Since the Gibbs measure samples extreme values of the field ψ, the overlaps under the Gibbs measure can be interpreted as measures of relative distance between the extremes. In spin-glass theory, the relevant object to classify the extreme value statistics of strongly correlated variables is the *two-overlap distribution*

$$\mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\Big[\mathbf{1}_{\{q^N(v,v')\le r\}}\Big], \quad r \in [0,1]. \tag{2.6.10}$$

Since the overlaps are normalized, their asymptotics will also be normalized to lie in $[0,1]$. Define

$$\bar{\mathcal{J}}_{\sigma^2}(\cdot) \doteq \frac{\mathcal{J}_{\sigma^2}(\cdot)}{\mathcal{J}_{\sigma^2}(1)}. \tag{2.6.11}$$

This is motivated by the fact that if

$$b_N(v,v') \doteq \max\{\lambda \in [0,1] : [v]_\lambda \cap [v']_\lambda \neq \emptyset\} \tag{2.6.12}$$

denotes the *branching scale between* v *and* v' (the analogue of the normalized branching time for BRWs), then Corollary 2.9.6 in Appendix 2.9.1 shows (in particular) that for all $\rho \in (0,1]$ and N large enough,

$$\max_{v,v' \in A_{N,\rho}} \left| q^N(v,v') - \bar{\mathcal{J}}_{\sigma^2}(b_N(v,v')) \right| \leq \frac{C_7}{\sqrt{\log N}} + C_8\, \rho. \tag{2.6.13}$$

For any *inverse temperature* $\beta > 0$, denote

$$l_\beta \doteq \begin{cases} \min\{l \in \{1,...,m\} : \beta \leq \beta_c(\bar{\sigma}_l) \doteq 2/\bar{\sigma}_l\}, & \text{if } \beta \leq 2/\bar{\sigma}_m, \\ m+1, & \text{otherwise.} \end{cases} \tag{2.6.14}$$

This is the smallest index l for which a *critical inverse temperature* $\beta_c(\bar{\sigma}_l)$ is at least β.

Remark 2.6.2. *In Bovier and Kurkova (2004a),* $l(\beta)$ *is defined such that* $l(\beta) = l_\beta - 1$*. Our choice is more natural with the notation we used in* (2.5.5) *and* (2.5.6)*, see the proof of Lemma 2.9.8.*

We are now ready to state the second main result of this article.

Theorem 2.6.3 (Limiting two-overlap distribution)**.** *Let* $\{\psi_v\}_{v \in V_N}$ *be the* $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$*-GFF on* V_N *of Definition 2.1.1. Then, for* $\beta > 0$,

$$\lim_{N \to \infty} \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\left[\mathbf{1}_{\{q^N(v,v') \leq r\}}\right] = \begin{cases} 0, & \text{if } r < 0, \\ \beta_c(\bar{\sigma}_j)/\beta, & \text{if } r \in [x^{j-1}, x^j),\ j \leq l_\beta - 1, \\ 1, & \text{if } r \geq x^{l_\beta - 1}, \end{cases} \tag{2.6.15}$$

where $\beta_c(\bar{\sigma}_j) \doteq 2/\bar{\sigma}_j$ *and* $x^j \doteq \bar{\mathcal{J}}_{\sigma^2}(\lambda^j)$*.*

Remark 2.6.3. *This is the same expression as in the context of the GREM. Compare this to Proposition 1.11 in Bovier and Kurkova (2004a). In the homogeneous case, the theorem was proved by Arguin and Zindy (2014) for a certain class of non-hierarchical log-correlated Gaussian fields with no boundary effect, by Arguin and Zindy (2015) for the GFF (trivial adjustments of their proof show the same result for the BRW), and by Jagannath (2016) for the BRW (using an alternative method).*

Remark 2.6.4. *In the context of the GFF, we have to show that the Gibbs measure doesn't carry any weight outside* $A_{N,\rho}$ *in the limit. In Arguin and Zindy (2015), a crucial step was* *to use self-averaging and Slepian's lemma in order to compare the free energy outside $A_{N,\rho}$ with that of a REM. Here, we find an upper bound on the free energy outside $A_{N,\rho}$ through the optimization problems for the maximum and γ-high points, see the proof of Lemma 2.8.3. This approach is much more efficient when there are several effective scales λ^j.*

Remark 2.6.5. *Theorem 2.6.3 tells us that even though the overlap between the extremes can be (almost) anything between* 0 *and* 1 *for finite system sizes, this variability disappears in the limit. The extremes can only branch asymptotically at the effective distances N^{1-d}, $d \in \{0, \lambda^1, \lambda^2, ..., \lambda^{l_\beta - 1}\}$, where λ^j is defined in* (2.4.4)*. We see that the number of branching scales λ^j for the extremes is finite in the limit and increases as the inverse temperature β becomes larger than some of the critical thresholds $0 < \beta_c(\bar{\sigma}_1) < \beta_c(\bar{\sigma}_2) < ... < \beta_c(\bar{\sigma}_{l_\beta - 1}) < \infty$. In comparison, for homogeneous models (like the GFF and the BRW), there is only one critical inverse temperature*

- *above which the extremes only branch at scale* 0 *or* 1 *in the limit, and*
- *under which the extremes only branch at scale* 0 *in the limit (meaning that the extremes are all asymptotically uncorrelated).*

The third and final main result of this article concerns the *Ghirlanda-Guerra identities*. These identities were introduced in Ghirlanda and Guerra (1998) and an extended version of the identities was proved for a general class of models, called the mixed p-spin, in Panchenko (2010b). Before taking the limit, we have the following approximate version.

Theorem 2.6.4 (Approximate extended Ghirlanda-Guerra identities)**.** *Let $\beta > 0$, and let $\alpha < \alpha'$ be any pair of scales such that*

$$\lambda^{j-1} < \alpha < \alpha' < \lambda^{j + (m-j)\mathbf{1}_{\{j = l_\beta\}}} \tag{2.6.16}$$

for some $j \in \{1, 2, ..., l_\beta\}$. Denote $\boldsymbol{v} \mathrel{\hat{=}} (v^1, v^2, ..., v^s)$ and $S_{\alpha,\alpha'} \mathrel{\hat{=}} (\bar{\mathcal{J}}_{\sigma^2}(\alpha), \bar{\mathcal{J}}_{\sigma^2}(\alpha')]$. Then, for any $s \in \mathbb{N}$, any $k \in \{1, ..., s\}$, and any functions $h : V_N^s \to \mathbb{R}$ such that $\sup_N \|h\|_\infty < \infty$,

$$\lim_{N\to\infty} \left| \begin{array}{l} \mathbb{E}\mathcal{G}_{\beta,N}^{\times(s+1)}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^k, v^{s+1})\}} dr \ \ h(\boldsymbol{v})\Big] \\ - \left\{ \begin{array}{l} \frac{1}{s}\mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^1, v^2)\}} dr\Big] \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h(\boldsymbol{v})\Big] \\ + \frac{1}{s}\sum_{l \neq k}^s \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^k, v^l)\}} dr \ \ h(\boldsymbol{v})\Big] \end{array} \right\} \end{array} \right| = 0. \tag{2.6.17}$$

Remark 2.6.6. *The word "approximate" here is* **NOT** *to be understood in exactly the same sense as in Definition 1.4 of Jagannath (2017). It is approximate in the sense that the limit $N \to \infty$ is taken, but also because linear combinations of functions of the form $q \mapsto \int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r<q\}} dr$ do not describe all the bounded mesurable functions defined on* $[0,1]$.

We now show that the extended form of the Ghirlanda-Guerra identities hold exactly in the limit. Along with Theorem 2.6.3, these identities completely determine the law of the limiting array of overlaps when the variables are sampled by the Gibbs measure, see Section 2.7.

We follow closely the reasoning from page 101 in Panchenko (2013b) and page 1459 in Arguin and Zindy (2014). Let $(v_l)_{l\in\mathbb{N}}$ be an i.i.d. sequence sampled from the Gibbs measure $\mathcal{G}_{\beta,N}$ and let

$$R^N \mathrel{\hat{=}} (R^N_{l,l'})_{l,l'\in\mathbb{N}} \mathrel{\hat{=}} (q^N(v^l, v^{l'}))_{l,l'\in\mathbb{N}} \tag{2.6.18}$$

be the array of overlaps of this sample. Note that the array R^N is symmetric and non-negative definite because the entries are normalized covariances of the Gaussian field ψ. Since each point is sampled independently, it is also *weakly exchangeable*, namely, for any permutation π of a finite number of indices,

$$\left(R^N_{\pi(l),\pi(l')}\right) \overset{\text{law}}{=} \left(R^N_{l,l'}\right). \tag{2.6.19}$$

The push-forward of the probability measure $\mathbb{E}\mathcal{G}^{\times\infty}_{\beta,N}$ under the mapping

$$(v^l)_{l\in\mathbb{N}} \mapsto R^N \tag{2.6.20}$$

defines a probability measure on the space $\mathcal{C}$ of $\mathbb{N}\times\mathbb{N}$ arrays with entries in $[-1,1]$, endowed with the product topology. Since $\mathcal{C}$ is a compact metric space (by Tychonoff's theorem), the space $\mathcal{M}_1(\mathcal{C})$ of probability measures on $\mathcal{C}$ is compact under the weak topology. Therefore, there exists a subsequence $\{N_m\}_{m\in\mathbb{N}}$ under which the above push-forward measures converge weakly to the distribution of some infinite array $R \mathrel{\hat{=}} (R_{l,l'})_{l,l'\in\mathbb{N}}$ in the sense of convergence of all their finite dimensional marginals. In particular, the three properties of R^N mentioned above are preserved by the limit, meaning that R is also symmetric, non-negative definite and weakly exchangeable.

By the representation theorem of Dovbysh and Sudakov (1982) (see also the proof in Panchenko (2010c)) and the atoms in Theorem 2.6.3, we can assume that the limiting array R is a random element of some probability space with measure P (and corresponding expectation E), generated by

$$\big(R_{l,l'}\big)_{l,l'\in\mathbb{N}} = \Big((\rho^l,\rho^{l'})_{\mathcal{H}} + (1 - x^{l_\beta - 1})\mathbf{1}_{\{l=l'\}}\Big)_{l,l'\in\mathbb{N}}, \tag{2.6.21}$$

where $(\rho^l)_{l\in\mathbb{N}}$ is an i.i.d. sample of *replicas* from some random measure μ_β concentrated a.s. on the sphere of radius $\sqrt{x^{l_\beta-1}}$ of a separable Hilbert space $\mathcal{H}$.

By construction, there exists a subsequence $\{N_m\}_{m\in\mathbb{N}}$ such that for any $s\in\mathbb{N}$ and any continuous function $h : [-1,1]^{s(s-1)/2} \to \mathbb{R}$,

$$\lim_{m\to\infty} \mathbb{E}\mathcal{G}_{\beta,N_m}^{\times\infty}\Big[h((R_{l,l'}^{N_m})_{1\le l,l'\le s})\Big] = E\mu_\beta^{\times\infty}\Big[h((R_{l,l'})_{1\le l,l'\le s})\Big]. \tag{2.6.22}$$

In particular, from Theorem 2.6.3, we have

$$E\mu_\beta^{\times 2}\Big[\mathbf{1}_{\{R_{1,2}\le r\}}\Big] = \begin{cases} 0, & \text{if } r < 0,\\ \beta_c(\bar{\sigma}_j)/\beta, & \text{if } r \in [x^{j-1}, x^j),\ 1\le j\le l_\beta - 1,\\ 1, & \text{if } r \ge x^{l_\beta-1}, \end{cases} \tag{2.6.23}$$

where $\beta_c(\bar{\sigma}_j) \mathrel{\hat{=}} 2/\bar{\sigma}_j$ and $x^j \mathrel{\hat{=}} \bar{\mathcal{J}}_{\sigma^2}(\lambda^j)$.

Next, we show the consequence of taking the limit (2.6.22) in the statement of Theorem 2.6.4. Note that a bounded function $h : \{x^0, x^1, ..., x^{l_\beta-1}\}^{s(s-1)/2} \to \mathbb{R}$ can always be embedded in a continuous function defined on $[-1,1]^{s(s-1)/2}$.

Theorem 2.6.5 (Extended Ghirlanda-Guerra identities in the limit)**.** *Let $\beta > 0$ and let μ_β be a subsequential limit of $\{\mathcal{G}_{\beta,N}\}_{N\in\mathbb{N}}$ in the sense of (2.6.22). Then, for any $s\in\mathbb{N}$, any $k\in\{1,...,s\}$, and any functions $h : \{x^0,x^1,...,x^{l_\beta-1}\}^{s(s-1)/2} \to \mathbb{R}$ and $g : \{x^0,x^1,...,x^{l_\beta-1}\} \to \mathbb{R}$, we have*

$$\begin{aligned} E\mu_\beta^{\times(s+1)}\Big[g(R_{k,s+1})\, h((R_{i,i'})_{1\le i,i'\le s})\Big] &= \frac{1}{s} E\mu_\beta^{\times 2}\Big[g(R_{1,2})\Big] E\mu_\beta^{\times s}\Big[h((R_{i,i'})_{1\le i,i'\le s})\Big] \\ &\quad + \frac{1}{s}\sum_{l\neq k}^{s} E\mu_\beta^{\times s}\Big[g(R_{k,l})\, h((R_{i,i'})_{1\le i,i'\le s})\Big]. \end{aligned} \tag{2.6.24}$$

Proof of Theorem 2.6.5. Let $\alpha < \alpha'$ be a pair of scales such that

$$\lambda^{j-1} < \alpha < \alpha' < \lambda^{j+(m-j)\mathbf{1}_{\{j=l_\beta\}}} \tag{2.6.25}$$

for some $j \in \{1, 2, ..., l_\beta\}$. From (2.6.22) and from Theorem 2.6.4 (in the particular case where h is a function of the overlaps), we deduce

$$\begin{aligned} &E\mu_\beta^{\times(s+1)}\Big[\int_{(\bar{\mathcal{J}}_{\sigma^2}(\alpha),\bar{\mathcal{J}}_{\sigma^2}(\alpha')]}\mathbf{1}_{\{r<R_{k,s+1}\}}dr\ h((R_{i,i'})_{1\le i,i'\le s})\Big] \\ &= \tfrac{1}{s}E\mu_\beta^{\times 2}\Big[\int_{(\bar{\mathcal{J}}_{\sigma^2}(\alpha),\bar{\mathcal{J}}_{\sigma^2}(\alpha')]}\mathbf{1}_{\{r<R_{1,2}\}}dr\Big]E\mu_\beta^{\times s}\Big[h((R_{i,i'})_{1\le i,i'\le s})\Big] \\ &\quad+\tfrac{1}{s}\textstyle\sum_{l\neq k}^{s} E\mu_\beta^{\times s}\Big[\int_{(\bar{\mathcal{J}}_{\sigma^2}(\alpha),\bar{\mathcal{J}}_{\sigma^2}(\alpha')]}\mathbf{1}_{\{r<R_{k,l}\}}dr\ h((R_{i,i'})_{1\le i,i'\le s})\Big]. \end{aligned} \tag{2.6.26}$$

From (2.6.23), we know that $\mathbf{1}_{\{r<R_{i,i'}\}}$ is $E\mu_\beta^{\times 2}$-a.s. constant in r on the interval $[x^{j-1}, x^{j+(m-j)\mathbf{1}_{\{j=l_\beta\}}})$. Therefore, from (2.6.25) and (2.6.26), we get

$$\begin{aligned} &E\mu_\beta^{\times(s+1)}\Big[\mathbf{1}_{\{x^{j-1}<R_{k,s+1}\}}\ h((R_{i,i'})_{1\le i,i'\le s})\Big] \\ &= \tfrac{1}{s}E\mu_\beta^{\times 2}\Big[\mathbf{1}_{\{x^{j-1}<R_{1,2}\}}\Big]E\mu_\beta^{\times s}\Big[h((R_{i,i'})_{1\le i,i'\le s})\Big] \\ &\quad+\tfrac{1}{s}\textstyle\sum_{l\neq k}^{s} E\mu_\beta^{\times s}\Big[\mathbf{1}_{\{x^{j-1}<R_{k,l}\}}\ h((R_{i,i'})_{1\le i,i'\le s})\Big]. \end{aligned} \tag{2.6.27}$$

Since $R_{i,i'} \geq 0$ $E\mu_\beta^{\times 2}$-a.s. by (2.6.23), the last equation is also trivially satisfied with say $x^{-1} \circeq -1$. But, any function $g : \{x^0, x^1, ..., x^{l_\beta - 1}\} \to \mathbb{R}$ can be written as a linear combination of the indicator functions $\mathbf{1}_{\{x^{j-1}<\cdot\}}$, $j \in \{0\} \cup \{1, 2, ..., l_\beta\}$, so we get the conclusion (by the linearity of (2.6.27)). □

Remark 2.6.7. *The extended Ghirlanda-Guerra identities are still the subject of ongoing research in spin glass theory and the study of log-correlated random fields, so it is still not clear why these identities seem to be a property shared in the limit by such a vast collection of models. Perhaps even more universal could be the stochastic stability property of random overlap structures (ROSt's), which are defined and treated (for example) in Aizenman and Contucci (1998); Contucci and Giardinà (2005); Arguin and Aizenman (2009); Talagrand (2010); Arguin and Chatterjee (2013). For instance, it is conjectured in Arguin and Chatterjee (2013) that the laws of the ROSt's satisfying the Ghirlanda-Guerra identities correspond to the extremes of the convex set of laws of the stochastically stable ROSt's. It is also conjectured that the stochastic stability of a ROSt for a subsequence of p-th power cavity fields implies ultrametricity. In this sense, it is expected that the* *stochastic stability property is more universal then the Ghirlanda-Guerra identities but still implies ultrametricity under technical conditions. In Panchenko (2012), it is shown how the Aizenman-Contucci stochastic stability property can be combined with a specific form of the Ghirlanda-Guerra identities into a unified stability property analogous to the Bolthausen-Sznitman invariance property in the setting of Ruelle probability cascades (see Bolthausen and Sznitman (1998)).*

Remark 2.6.8. *It is expected that the results of Arguin and Ouimet (2016) on the first order asymptotics of the maximum and γ-high points can be extended to the more general case where the variance function σ in (2.4.1) is piecewise C^1. Therefore, it is also expected that the results in the present article could be generalised just like Bovier and Kurkova (2004b) did when they generalized the results of the GREM to the CREM (the GREM with a continuum of hierarchies).*

We could take this further by imposing σ to be piecewise C^1 and by working directly with the continuous version of the two-dimensional GFF instead of the discrete version. A formal definition of such a field is given in Section 1.3 of Arguin and Ouimet (2016) as well as a conjecture on the Hausdorff dimension of the γ-thick points (the analogue of the γ-high points). The field is a random distribution (i.e. generalized function), so it cannot be defined pointwise, but we can make sense of the collection of circle averages around a point $v \in [0,1]^2$ as a stochastic process. In fact, we would expect such a process, after a time-change, to be equal in law to $\int_0^{\cdot} \sigma(s) dB_v(s)$, where B_v is a Brownian motion adapted to a certain filtration $\mathbb{F}_v$. We could then ask if it is possible to characterize the limiting (with respect to the approximation procedure) law of the Liouville measure (the analogue of the Gibbs measure) of this new field. For an introduction to these concepts, see e.g. Berestycki (2016); Rhodes and Vargas (2014); Sheffield (2007).

Finally, another natural question is to ask if there is a way to introduce randomness in the function $s \mapsto \sigma(s)$ and still make sense of the questions above, although this is not clear since the process $(\sigma(s))_{s\in[0,1]}$ cannot be adapted (let alone predictable) simultaneously to all the filtrations $\mathbb{F}_v,\ v \in [0,1]^2$. Maybe there is a way around this problem if the filtrations share "information" in a very structured way.

2.7. Consequences of Theorem 2.6.3 and Theorem 2.6.5

The first consequence concerns the geometry of the overlaps in the limit. It was shown in Panchenko (2010a) (see also Panchenko (2011) for a simplified proof) that any limiting array of overlaps that takes finitely many values and satisfy (2.6.24) must be ultrametric under $E\mu_\beta^{\times\infty}$.

Corollary 2.7.1 (Ultrametricity in the limit)**.** *Let $\beta > 0$ and let μ_β be a subsequential limit of $\{\mathcal{G}_{\beta,N}\}_{N\in\mathbb{N}}$ in the sense of* (2.6.22)*. We must have*

$$E\mu_\beta^{\times 3}\Big(R_{1,2} \geq R_{1,3} \wedge R_{2,3}\Big) = 1. \tag{2.7.1}$$

Since the replicas ρ^l all have norm $\sqrt{x^{l_\beta - 1}}$ almost-surely in $\mathcal{H}$, then (2.7.1) *is equivalent to the ultrametric inequality*

$$E\mu_\beta^{\times 3}\Big(\|\rho^1 - \rho^2\| \leq \|\rho^1 - \rho^3\| \vee \|\rho^2 - \rho^3\|\Big) = 1. \tag{2.7.2}$$

Remark 2.7.1. *The random measure μ_β gives more weight to extreme values, so we can interpret this corollary as saying that the extremes are clustered in a hierarchical way. From* (2.6.23)*, the number of hierarchies increases as the inverse temperature β becomes larger than some of the critical thresholds $\beta_c(\bar{\sigma}_j)$.*

The second consequence makes the description of the structure of μ_β even more precise. Probability cascades were introduced in Ruelle (1987) to describe the limiting Gibbs measure of the GREM. Since the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF satisfies the extended Ghirlanda-Guerra identities and the limiting two-overlap distribution takes finitely many values, we can show that the limiting Gibbs measure μ_β is a Ruelle probability cascade. First, we define probability cascades by following Panchenko (2013b).

For a given $r \geq 1$, let

$$\mathcal{T} \doteq \{\emptyset\} \cup \mathbb{N} \cup \mathbb{N}^2 \cup ... \cup \mathbb{N}^r \tag{2.7.3}$$

be the vertex set of a tree rooted at $\emptyset$. Each vertex $v = (n_1, ..., n_p) \in \mathbb{N}^p$, for $p \leq r-1$, has children

$$vn \doteq (n_1, ..., n_p, n) \in \mathbb{N}^{p+1}, \quad n \in \mathbb{N}. \tag{2.7.4}$$

Each vertex $v \in \mathbb{N}^p$ is connected to the root by a path. Denote by $p(v)$ the set of vertices (excluding the root) on the shortest path from v to the root. Additionally, fix two sequences of parameters :

$$\begin{aligned} &0 \doteq \zeta_{-1} < \zeta_0 < \zeta_1 < ... < \zeta_{r-1} < \zeta_r \doteq 1, \\ &0 \doteq q_0 < q_1 < ... < q_{r-1} < q_r \leq 1. \end{aligned} \tag{2.7.5}$$

For all $v \in \mathcal{T}$, denote by $|v|$ its distance in the tree, namely $|v| \doteq \# p(v)$. Then, for all $v \in \mathcal{T} \backslash \mathbb{N}^r$, generate independent Poisson point processes, denoted by Π_v, with mean measure $\zeta_{|v|} x^{-1-\zeta_{|v|}} dx$ on $(0, \infty)$. We arrange the points in Π_v in decreasing order :

$$z_{v1} > z_{v2} > ... > z_{vn} > ... \tag{2.7.6}$$

For each vertex $v \in \mathcal{T} \backslash \mathbb{N}^r$, the relative weight of each point in Π_v is defined by

$$w_{vn} \doteq \frac{z_{vn}}{\sum_{i \in \mathbb{N}} z_{vi}}, \quad n \in \mathbb{N}. \tag{2.7.7}$$

Say we are on a separable Hilbert space $\mathscr{H}$ with orthonormal basis $\{e_v\}_{v \in \mathcal{T} \backslash \{\emptyset\}}$. Consider the vectors in $\mathscr{H}$

$$h_v \doteq \sum_{u \in p(v)} e_u \left(q_{|u|} - q_{|u|-1}\right)^{1/2}, \quad v \in \mathcal{T} \backslash \{\emptyset\}, \tag{2.7.8}$$

and define a random measure on them by

$$G(h_v) \doteq \prod_{u \in p(v)} w_u, \quad v \in \mathbb{N}^r. \tag{2.7.9}$$

The random measure G is called a Ruelle probability cascade (RPC) associated with the parameters in (2.7.5). It is defined up to an orthonormal change of basis.

We can think of $\mathbb{N}^r$ as the leaves in the tree structure. From (2.7.8), each element in $\{h_v\}_{v \in \mathbb{N}^r}$ has norm $\sqrt{q_r}$ and the scalar product between two such elements can only take values in the finite set

$$\{0, q_1, q_2, ..., q_r\}. \tag{2.7.10}$$

The weights (2.7.7), associated with each branch in the tree, are random. Hence, (2.7.9) defines a random probability measure and each instance of G samples elements in $\{h_v\}_{v \in \mathbb{N}^r}$ by choosing a branch independently at each step, between the root and a leaf, with probability given by the associated weight in (2.7.7). The tree structure is illustrated in Figure 2.7.2. Since the weights are ordered in decreasing order at each scale, the branches on the left are more likely to be selected at each step.

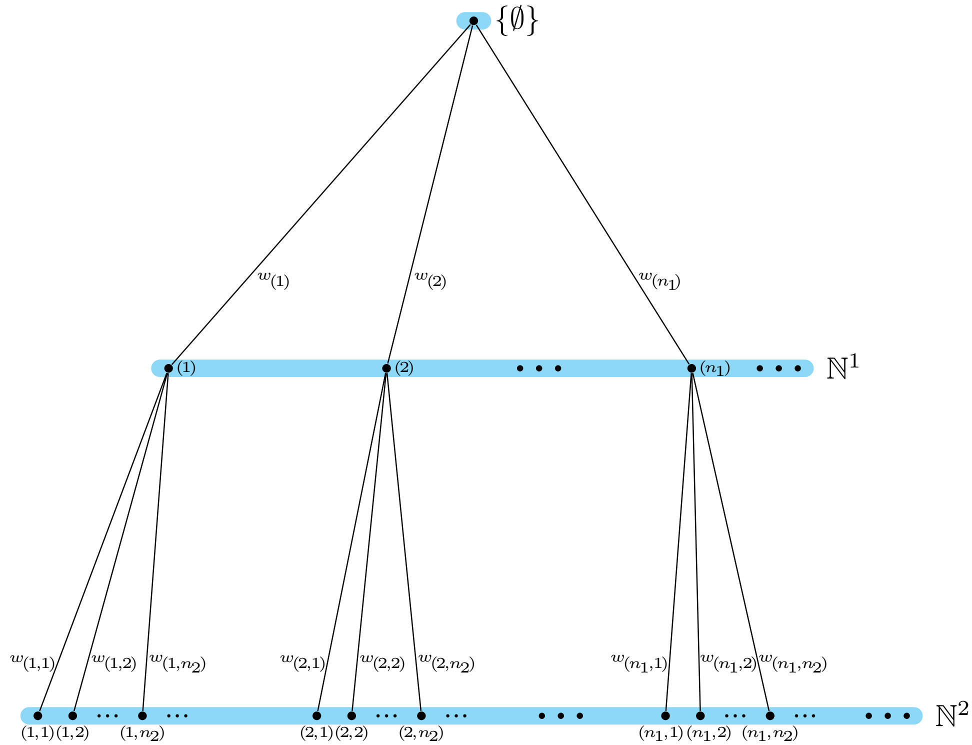


FIGURE 2.7.2. Exact tree structure of a Ruelle probability cascade with $r = 2$ levels. Given an instance $\omega \in \Omega$ of the random weights, the measure G samples elements in $\{h_v\}_{v\in\mathbb{N}^r}$ with probability equal to the product of the probabilities associated with the branches on the shortest path from the root to the leaf v. For example, we have $(G(\omega))(h_{(1,2)}) = w_{(1)}(\omega)w_{(1,2)}(\omega)$. For the scalar products, we have, for example, $(h_{(1,1)}, h_{(2,2)})_{\mathscr{H}} = 0$ and $(h_{(2,1)}, h_{(2,2)})_{\mathscr{H}} = q_1$.

The corollary below shows that the limiting Gibbs measure of the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF is a RPC, despite its underlying tree structure being only approximate for finite N.

Corollary 2.7.2. *Let $\beta > 0$ and let μ_β be a subsequential limit of $\{\mathcal{G}_{\beta,N}\}_{N\in\mathbb{N}}$ in the sense of (2.6.22). Then, μ_β has the same law as a RPC with parameters*

- $r = l_\beta - 1$,
- $\zeta_j = E\mu_\beta^{\times 2}\Big[\mathbf{1}_{\{R_{1,2}\leq x^j\}}\Big] = (2/\bar{\sigma}_{j+1})/\beta,$ *for all* $j \in \{0, 1, ..., r-1\}$,
- $q_j = x^j \doteq \bar{\mathcal{J}}_{\sigma^2}(\lambda^j),$ *for all* $j \in \{0, 1, ..., r\}$.

PROOF. The proof follows directly from Theorem 2.13 in Panchenko (2013b) or from the proof of Theorem 1.13 in Bovier and Kurkova (2004a) (once we have the ultrametricity). We simply need to match the parameters so that $\{q_j\}_{j=0}^r$ are the atoms of $E\mu_\beta^{\times 2}(R_{1,2} \in \cdot)$ and $\{\nabla\zeta_j\}_{j=0}^r$ are the corresponding probabilities. $\square$

2.8. Proofs of the main results

Throughout the proofs, $c, C, \widetilde{C}, etc.$, will denote positive constants whose value can change from line to line and can only depend on the parameters $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$, unless additional variables are specified. Equations are always implicitly stated to hold for N large enough when needed.

2.8.1. Computation of the limiting free energy

Theorem 2.6.1 is a direct consequence of Lemma 2.8.1, which shows $f_N^{\psi}(\beta) \to f^{\psi}(\beta)$ in probability, and Lemma 2.8.2, which shows the uniform integrability of the sequence $\{|f_N^{\psi}(\beta)|^p\}_{N\in\mathbb{N}}$ for all $p \in [1, \infty)$.

Lemma 2.8.1 (Convergence in probability of the free energy)**.** *Let $\eta > 0$ and $\beta > 0$. There exists $c = c(\eta, \beta, \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ such that for N large enough,*

$$\mathbb{P}\Big(|f_N^{\psi}(\beta) - f^{\psi}(\beta)| > \eta\Big) \le N^{-c}. \tag{2.8.1}$$

Proof. Fix $\eta > 0$ and $\beta > 0$. For all $i \in \{0, 1, ..., K+1\}$, define $\gamma_i \mathrel{\hat{=}} i\gamma^{\star}/K$. We will choose $K \in \mathbb{N}$ large enough later. We prove the upper bound first. Define

$$\mathcal{H}_N^{\mathrm{abs}}(\gamma) \mathrel{\hat{=}} \{v \in V_N : |\psi_v| \ge \gamma \log N^2\}. \tag{2.8.2}$$

From (2.5.2), (2.5.8), and the symmetry of Gaussian densities, the event

$$U_{N,K,\eta} \mathrel{\hat{=}} \bigcap_{i=0}^{K} \Big\{|\mathcal{H}_N^{\mathrm{abs}}(\gamma_i)| < N^{2\mathcal{E}(\gamma_i)+\eta}\Big\} \bigcap \Big\{\max_{v\in V_N} |\psi_v| < \frac{K+1}{K}\gamma^{\star} \log N^2\Big\} \tag{2.8.3}$$

satisfies $\mathbb{P}\Big(U_{N,K,\eta}^c\Big) \le N^{-c(K,\eta,\boldsymbol{\sigma},\boldsymbol{\lambda})}$ for any given K. On the event $U_{N,K,\eta}$,

$$\begin{aligned}
Z_N^{\psi}(\beta) \mathrel{\hat{=}} \sum_{v\in V_N} e^{\beta\psi_v} &\le \sum_{i=1}^{K+1} (|\mathcal{H}_N^{\mathrm{abs}}(\gamma_{i-1})| - |\mathcal{H}_N^{\mathrm{abs}}(\gamma_i)|) N^{2\beta\gamma_i} \\
&= \sum_{i=1}^{K} (N^{2\beta\gamma_{i+1}} - N^{2\beta\gamma_i}) |\mathcal{H}_N^{\mathrm{abs}}(\gamma_i)| + N^{2\beta\gamma_1} |\mathcal{H}_N^{\mathrm{abs}}(\gamma_0)| \\
&\le N^{2\beta\gamma^{\star}/K} \sum_{i=0}^{K} N^{2\beta\gamma_i} |\mathcal{H}_N^{\mathrm{abs}}(\gamma_i)|.
\end{aligned} \tag{2.8.4}$$

We used the fact that $|\mathcal{H}_N^{\mathrm{abs}}(\gamma_{K+1})| = 0$ to obtain the second equality. Therefore, on the event $U_{N,K,\eta}$,

$$\begin{aligned}
f_N^{\psi}(\beta) &\overset{(2.8.4)}{\leq} \frac{\beta\gamma^{\star}}{K} + \frac{\log(K+1)}{\log N^2} + \max_{0\leq i\leq K}(\beta\gamma_i + \mathcal{E}(\gamma_i)) + \frac{\eta}{2} \\
&\leq \max_{0\leq i\leq K}(\beta\gamma_i + \mathcal{E}(\gamma_i)) + \eta \qquad \begin{array}{l}\text{for } K \text{ large enough with respect} \\ \text{to } \eta \text{ and } \beta\text{, and } N \text{ large enough} \\ \text{with respect to } K \text{ and } \eta,\end{array} \\
&\leq \max_{\gamma\in[0,\gamma^{\star}]}(\beta\gamma + \mathcal{E}(\gamma)) + \eta \\
&= f^{\psi}(\beta) + \eta \qquad \text{by Lemma 2.9.8.}
\end{aligned} \tag{2.8.5}$$

Thus, for K large enough (fixed, depending on η and β), we have

$$\mathbb{P}\Big(f_N^{\psi}(\beta) > f^{\psi}(\beta) + \eta\Big) \leq \mathbb{P}\Big(U_{N,K,\eta}^c\Big) \leq N^{-c(K,\eta,\boldsymbol{\sigma},\boldsymbol{\lambda})}. \tag{2.8.6}$$

We now prove the lower bound. Recall that

$$\mathcal{H}_N(\gamma) \circeq \{v \in V_N : \psi_v \geq \gamma \log N^2\}. \tag{2.8.7}$$

From (2.5.2) and (2.5.9), the event

$$B_{N,K,\eta} \circeq \bigcap_{i=1}^{K-1} \Big\{|\mathcal{H}_N(\gamma_i)| \geq N^{2\mathcal{E}(\gamma_i)-\eta}\Big\} \bigcap \Big\{\max_{v\in V_N} \psi_v < \frac{K+1}{K}\gamma^{\star}\log N^2\Big\} \tag{2.8.8}$$

satisfies $\mathbb{P}\Big(B_{N,K,\eta}^c\Big) \leq N^{-c(K,\eta,\boldsymbol{\sigma},\boldsymbol{\lambda})}$ for any given K. On the event $B_{N,K,\eta}$,

$$\begin{aligned}
Z_N^{\psi}(\beta) \circeq \sum_{v\in V_N} e^{\beta\psi_v} &\geq \sum_{i=1}^{K+1}(|\mathcal{H}_N(\gamma_{i-1})| - |\mathcal{H}_N(\gamma_i)|)N^{2\beta\gamma_{i-1}} \\
&= \sum_{i=1}^{K}(N^{2\beta\gamma_i} - N^{2\beta\gamma_{i-1}})|\mathcal{H}_N(\gamma_i)| + N^{2\beta\gamma_0}|\mathcal{H}_N(\gamma_0)| \\
&\geq \frac{1}{2}\sum_{i=1}^{K-1} N^{2\beta\gamma_i}|\mathcal{H}_N(\gamma_i)|.
\end{aligned} \tag{2.8.9}$$

We used the fact that $|\mathcal{H}_N(\gamma_{K+1})| = 0$ to obtain the second equality. We dropped the 0-th and K-th summands to obtain the last inequality and took N large enough that $1 - N^{-2\beta\gamma^{\star}/K} \geq 1/2$.

Therefore, on $B_{N,K,\eta}$,

$$
\begin{aligned}
f_N^{\psi}(\beta) &\overset{(2.8.9)}{\geq} \max_{1\leq i\leq K-1}(\beta\gamma_i + \mathcal{E}(\gamma_i)) - \frac{\eta}{2} - \frac{\log 2}{\log N^2} && \\
&\geq \max_{1\leq i\leq K-1}(\beta\gamma_i + \mathcal{E}(\gamma_i)) - \frac{3\eta}{4} && \text{for } N \text{ large enough with respect to } \eta, \\
&\geq \max_{\gamma\in[0,\gamma^\star]}(\beta\gamma + \mathcal{E}(\gamma)) - \eta && \text{for } K \text{ large enough with respect to } \eta \text{ and } \beta \text{ since } \gamma \mapsto (\beta\gamma + \mathcal{E}(\gamma)) \text{ is continuous by Lemma 2.9.7}, \\
&= f^{\psi}(\beta) - \eta && \text{by Lemma 2.9.8.} \qquad (2.8.10)
\end{aligned}
$$

Thus, for K large enough (fixed, depending on η and β), we have

$$\mathbb{P}\Big(f_N^{\psi}(\beta) < f^{\psi}(\beta) - \eta\Big) \leq \mathbb{P}\Big(B_{N,K,\eta}^c\Big) \leq N^{-c(K,\eta,\boldsymbol{\sigma},\boldsymbol{\lambda})}. \tag{2.8.11}$$

Equations (2.8.6) and (2.8.11) together prove the lemma. □

For the uniform integrability, we follow the proof of Capocaccia et al. (1987), originally given in the context of the GREM.

Lemma 2.8.2 (Uniform integrability of $\{|f_N^{\psi}(\beta)|^p\}_{N\in\mathbb{N}}$)**.** *Let $\beta > 0$ and $1 \leq p < \infty$. Then,*

$$\lim_{\alpha\to\infty} \sup_{N\in\mathbb{N}} \mathbb{E}\Big[|f_N^{\psi}(\beta)|^p \, \mathbf{1}_{\{|f_N^{\psi}(\beta)|^p > \alpha\}}\Big] = 0. \tag{2.8.12}$$

Proof. By definition, $f_N^{\psi}(\beta) \mathrel{\hat{=}} \frac{1}{\log N^2} \log \sum_{v\in V_N} e^{\beta\psi_v}$. Bound from above every summand by the maximum summand and bound from below by keeping only the maximum summand. If $\xi_N \mathrel{\hat{=}} \max_{v\in V_N} \psi_v/(\log N^2)$, it is easily seen that for $N \geq 2$,

$$\beta\xi_N \leq f_N^{\psi}(\beta) \leq \beta\xi_N + \frac{\log(N+1)^2}{\log N^2} \leq \beta\xi_N + 2. \tag{2.8.13}$$

Assume that $\alpha^{1/p} - 2 > 0$. By splitting the event $\{|f_N^{\psi}(\beta)| > \alpha^{1/p}\}$ in two parts : $\{f_N^{\psi}(\beta) > \alpha^{1/p}\}$ and $\{-f_N^{\psi}(\beta) > \alpha^{1/p}\}$, and then using (2.8.13), we deduce

$$
\begin{aligned}
&\mathbb{E}\Big[|f_N^{\psi}(\beta)|^p \, \mathbf{1}_{\{|f_N^{\psi}(\beta)| > \alpha^{1/p}\}}\Big] \\
&\quad \leq \mathbb{E}\Big[(\beta\xi_N + 2)^p \, \mathbf{1}_{\{\beta\xi_N + 2 > \alpha^{1/p}\}}\Big] + \mathbb{E}\Big[(-\beta\xi_N)^p \, \mathbf{1}_{\{-\beta\xi_N > \alpha^{1/p}\}}\Big]
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{l=1}^{\infty} \mathbb{E}\Big[(\beta\xi_N + 2)^p \, \mathbf{1}_{\{(l+1)\alpha^{1/p} \geq \beta\xi_N + 2 > l\alpha^{1/p}\}}\Big] \\
&\qquad + \sum_{l=1}^{\infty} \mathbb{E}\Big[(-\beta\xi_N)^p \, \mathbf{1}_{\{(l+1)\alpha^{1/p} \geq -\beta\xi_N > l\alpha^{1/p}\}}\Big] \\
&\leq 2 \sum_{l=1}^{\infty} (l+1)^p \, \alpha \, \mathbb{P}\Big(|\xi_N| > \frac{1}{\beta}(l\alpha^{1/p} - 2)\Big).
\end{aligned} \tag{2.8.14}
$$

Note that $|\xi_N| \leq \max_{v \in V_N} |\psi_v| / (\log N^2)$, and $\max_{v \in V_N} \mathbb{V}(\psi_v) \leq \mathcal{J}_{\sigma^2}(1) \log N + C_0$ by Lemma 2.9.3. Therefore, for all $l \in \mathbb{N}$, a union bound and a standard Gaussian tail estimate yield (when N is large enough, say $N \geq N_0 \geq 2$)

$$
\begin{aligned}
\mathbb{P}\Big(|\xi_N| > \frac{1}{\beta}(l\alpha^{1/p} - 2)\Big) &\leq (N+1)^2 \max_{v \in V_N} 2\, \mathbb{P}\Big(\psi_v > \frac{1}{\beta}(l\alpha^{1/p} - 2) \log N^2\Big) \\
&\leq (N+1)^2 N^{-2\frac{(l\alpha^{1/p} - 2)^2}{\beta^2 \mathcal{J}_{\sigma^2}(1)}} \\
&\leq (N+1)^2 N^{-2\frac{(\alpha^{1/p} - 2)^2}{\beta^2 \mathcal{J}_{\sigma^2}(1)}} N^{-2\frac{(l-1)^2 \alpha^{2/p}}{\beta^2 \mathcal{J}_{\sigma^2}(1)}}.
\end{aligned} \tag{2.8.15}
$$

To obtain the last inequality, we wrote $(l\alpha^{1/p} - 2)^2 = (\alpha^{1/p} - 2 + (l-1)\alpha^{1/p})^2$ and used $(a+b)^2 \geq a^2 + b^2, \ a, b \geq 0$. If we further assume that $(\alpha^{1/p} - 2)^2 > \beta^2 \mathcal{J}_{\sigma^2}(1)$, the sum in (2.8.14) tends to 0 as $\alpha \to \infty$, uniformly for $N \geq N_0$. $\square$

Proof of Theorem 2.6.2. Since $x \mapsto \log x$ is an increasing function and $A_{N,\rho} \subseteq V_N$, we have the upper bound on the limit in probability :

$$
f_{N,\rho}^{\psi}(\beta) \leq \frac{1}{\log N^2} \log \sum_{v \in V_N} e^{\beta \psi_v} \mathrel{\hat{=}} f_N^{\psi}(\beta) \stackrel{N \to \infty}{\longrightarrow} f^{\psi}(\beta). \tag{2.8.16}
$$

On the other hand, from Lemma A.2 in Arguin and Ouimet (2016) and the independence of the increments, we know that for any $\delta \in (0, 1/2]$ and $j \in \{1, 2, ..., m\}$, then for N large enough and all $v \in V_N^{\delta} \mathrel{\hat{=}} \Big\{v \in V_N : \min_{z \in \partial V_N} \|v - z\|_2 \geq \delta N\Big\}$, we have

$$
-C_1(\delta, \boldsymbol{\sigma}) \leq \mathbb{V}\Big(\nabla \psi_v(\lambda^j)\Big) - \bar{\sigma}_j^2 \nabla \lambda^j \log N \leq C_2(\boldsymbol{\sigma}). \tag{2.8.17}
$$

Hence, from the remark at the end of Lemma 3.1 in Arguin and Ouimet (2016), we know that Theorem 2.5.1 and Theorem 2.5.2 (in this paper) hold on V_N^{δ}; the proof is in fact easier. Since $A_{N,\rho} \supseteq V_N^\delta$ for N large enough, we have

$$f_{N,\rho}^{\psi}(\beta) \doteq \frac{1}{\log N^2} \log \sum_{v \in A_{N,\rho}} e^{\beta \psi_v} \geq \frac{1}{\log N^2} \log \sum_{v \in V_N^\delta} e^{\beta \psi_v}. \tag{2.8.18}$$

A rerun of the proof of the lower bound in Lemma 2.8.1, with $\mathcal{H}_N(\gamma)$ restricted to V_N^δ, yields the conclusion. □

2.8.2. The Gibbs measure near the boundary

The first step in the proof of Theorem 2.6.3 is to show that the Gibbs measure does not carry any weight near the boundary of V_N in the limit $N \to \infty$. For this purpose, recall

$$A_{N,\rho} \doteq \left\{ v \in V_N : \min_{z \in \mathbb{Z}^2 \setminus V_N} \|v - z\|_2 \geq N^{1-\rho} \right\}, \quad \rho \in (0,1]. \tag{2.8.19}$$

This box contains the points in V_N that are at least at a distance of $N^{1-\rho}$ from the exterior. The Gibbs measure of the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-GFF restricted to $A_{N,\rho}$ is

$$\mathcal{G}_{\beta,N,\rho}(\{v\}) \doteq \frac{e^{\beta \psi_v}}{Z_{N,\rho}^{\psi}(\beta)}, \quad v \in A_{N,\rho}, \tag{2.8.20}$$

where $Z_{N,\rho}^{\psi}(\beta) \doteq \sum_{v \in A_{N,\rho}} e^{\beta \psi_v}$. We start by proving an upper bound on the following quantity :

$$\widetilde{f}_{N,\rho}^{\psi}(\beta) \doteq \frac{1}{\log N^2} \log \sum_{v \in A_{N,\rho}^c} e^{\beta \psi_v}. \tag{2.8.21}$$

Lemma 2.8.3. *Let $\eta > 0$, $\beta > 0$ and $\rho \in (0, \lambda_1)$. There exists $c = c(\eta, \beta, \rho, \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ such that*

$$\mathbb{P}\Big(\widetilde{f}_{N,\rho}^{\psi}(\beta) > (f^{\psi}(\beta) - \rho/2) + \eta \Big) \leq N^{-c} \tag{2.8.22}$$

for N large enough.

Proof. In Arguin and Ouimet (2016), Theorem 1.2 and Theorem 1.3 prove that

$$\lim_{N \to \infty} \frac{\max_{v \in V_N} \psi_v}{\log N^2} = \gamma^\star, \qquad \text{in probability,} \tag{2.8.23}$$

where

$\gamma^\star = \max_{\gamma_1, \gamma_2, \ldots, \gamma_M} \sum_{i=1}^M \nabla \gamma_i,$
under the constraints $\sum_{i=1}^k \left(\nabla \lambda_i - \frac{(\nabla \gamma_i)^2}{\sigma_i^2 \nabla \lambda_i} \right) \geq 0, \quad 1 \leq k \leq M,$

and, for all $0 \leq \gamma < \gamma^\star$,

$$\lim_{N\to\infty} \frac{\log(|\mathcal{H}_N(\gamma)|)}{\log N^2} = \mathcal{E}(\gamma), \qquad \text{in probability,} \tag{2.8.24}$$

where

$$\begin{aligned}
&\mathcal{E}(\gamma) = \max_{\gamma_1,\gamma_2,\ldots,\gamma_{M-1}} \textstyle\sum_{i=1}^{M-1} \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i}\right) + \left(\nabla\lambda_M - \frac{(\gamma-\gamma_{M-1})^2}{\sigma_M^2\nabla\lambda_M}\right), \\
&\text{under the constraints } \textstyle\sum_{i=1}^{k} \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2\nabla\lambda_i}\right) \geq 0, \quad 1 \leq k \leq M-1.
\end{aligned}$$

The unique solution of each optimization problem is rigorously found in Appendix A of Ouimet (2014) by using the Karush-Kuhn-Tucker theorem, and the solutions are shown to coincide with (2.8.23) and (2.8.24) in Arguin and Ouimet (2016).

Now, to bound $\widetilde{f}_{N,\rho}^{\psi}(\beta)$ from above, we need to find the analogues of (2.8.23) and (2.8.24) on $A_{N,\rho}^c$ instead of V_N. To this end, we recall the *set of representatives at scale* λ from Arguin and Ouimet (2016), denoted by R_λ. Loosely speaking, at a given scale λ, the points in $R_\lambda \subseteq V_N$ represent the $O(N^{2\lambda})$ nods of the underlying branching quaternary tree structure of the GFF, see Figure 2.8.3. This branching structure is motivated by the fact that if v_λ denotes the representative at scale $\lambda > 0$ that is closest to v, then, from Lemma A.6 in Arguin and Ouimet (2016), we know that $\max_{v\in V_N} \mathbb{V}(\psi_v(\lambda) - \psi_{v_\lambda}(\lambda)) \leq C$, for N large enough.

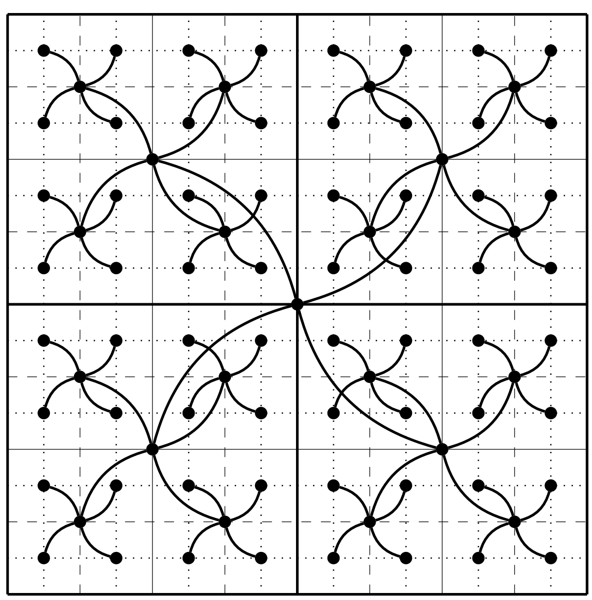


FIGURE 2.8.3. The representatives at scale $0, 1/4, 1/2$ and $3/4$.

More precisely, let $R_1 \triangleq V_N$, and for $\lambda \in [0,1)$, the set R_λ contains $\lfloor N^\lambda \rfloor^2$ v's with neighborhoods $[v]_\lambda$ that can only touch at their boundary (if they do touch) and are not cut off by ∂V_N. To remove any ambiguity, define R_λ in such a way that

$$\max_{v \in V_N} \min_{z \in R_\lambda} \|v - z\|_2 \quad \text{is minimized.}$$

For instance, if $N = 2^n$, $\lambda \in [0,1)$ and $\lambda n \in \mathbb{N}_0$, then divide V_N into a grid with $N^{2\lambda}$ squares of side length $N^{1-\lambda}$; the center point of each square is a representative at scale λ.

Since we assumed $\rho \in (0, \lambda_1)$, the only difference is that there are $O(N^{2(\lambda_i - \rho/2)})$ representatives at each scale λ_i on $A_{N,\rho}^c$ (ψ is still defined on V_N) instead of $O(N^{2\lambda_i})$. Therefore, a rerun of the proof of Lemma 3.1 and 3.4 in Arguin and Ouimet (2016) (note that only the upper bounds work) shows that for all $\varepsilon > 0$,

$$\mathbb{P}\left(\max_{v \in A_{N,\rho}^c} \psi_v \geq (1+\varepsilon)\gamma_\rho^\star \log N^2\right) \leq N^{-c(\varepsilon, \boldsymbol{\sigma}, \boldsymbol{\lambda})}, \tag{2.8.25}$$

for N large enough, where

$$\begin{aligned} &\gamma_\rho^\star \triangleq \max_{\gamma_1, \gamma_2, \dots, \gamma_M} \textstyle\sum_{i=1}^M \nabla\gamma_i, \\ &\text{under the constraints } \textstyle\sum_{i=1}^k \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i}\right) \geq \rho/2, \quad 1 \leq k \leq M, \end{aligned}$$

and, for all $0 \leq \gamma \leq \gamma_\rho^\star$,

$$\mathbb{P}\Big(|\{v \in A_{N,\rho}^c : \psi_v \geq \gamma \log N^2\}| \geq N^{2\mathcal{E}_\rho(\gamma) + \varepsilon}\Big) \leq N^{-c(\gamma, \varepsilon, \boldsymbol{\sigma}, \boldsymbol{\lambda})}, \tag{2.8.26}$$

for N large enough, where

$$\begin{aligned} &\mathcal{E}_\rho(\gamma) \triangleq \max_{\gamma_1, \gamma_2, \dots, \gamma_{M-1}} \textstyle\sum_{i=1}^{M-1} \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i}\right) + \left(\nabla\lambda_M - \frac{(\gamma - \gamma_{M-1})^2}{\sigma_M^2 \nabla\lambda_M}\right) - \rho/2, \\ &\text{under the constraints } \textstyle\sum_{i=1}^k \left(\nabla\lambda_i - \frac{(\nabla\gamma_i)^2}{\sigma_i^2 \nabla\lambda_i}\right) \geq \rho/2, \quad 1 \leq k \leq M-1. \end{aligned}$$

A rerun of the upper bound in the proof of Lemma 2.8.1 shows that for all $\eta > 0$, there exists a constant $c = c(\eta, \beta, \rho, \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ such that for N large enough,

$$\mathbb{P}\left(\widetilde{f}_{N,\rho}^{\psi}(\beta) > \max_{\gamma \in [0, \gamma_\rho^\star]} (\beta\gamma + \mathcal{E}_\rho(\gamma)) + \eta\right) \leq N^{-c}. \tag{2.8.27}$$

The constraints associated with $\gamma_\rho^\star$ and $\mathcal{E}_\rho(\gamma)$ are respectively more restrictive than the constraints associated with $\gamma^\star$ and $\mathcal{E}(\gamma)$, so we obviously have

$$\gamma_\rho^\star \leq \gamma^\star \quad \text{and} \quad \mathcal{E}_\rho(\gamma) \leq \mathcal{E}(\gamma) - \rho/2, \quad 0 \leq \gamma \leq \gamma_\rho^\star. \tag{2.8.28}$$

Therefore,

$$\max_{\gamma\in[0,\gamma_\rho^\star]}(\beta\gamma+\mathcal{E}_\rho(\gamma))\le\max_{\gamma\in[0,\gamma^\star]}(\beta\gamma+\mathcal{E}(\gamma)-\rho/2)\overset{(2.6.3)}{=}f^\psi(\beta)-\rho/2. \tag{2.8.29}$$

The conclusion of the lemma follows directly from (2.8.27) and (2.8.29). □

Lemma 2.8.4. *Let $\beta>0$ and $\rho\in(0,\lambda_1)$. Then,*

$$\lim_{N\to\infty}\mathcal{G}_{\beta,N}(A_{N,\rho}^c)=0, \tag{2.8.30}$$

where the limit holds in $\mathbb{P}$-probability and in L^p, $1\le p<\infty$.

Remark 2.8.1. *The result in Lemma 2.8.4 would not hold if we considered instead the complement of V_N^δ, which is much larger than the complement of $A_{N,\rho}$.*

Proof of Lemma 2.8.4. Fix $\rho\in(0,\lambda_1)$ and $\varepsilon\in(0,1)$, and let $\widetilde{\eta}>0$ depend on ρ. We have

$$\begin{aligned}\mathbb{P}\Big(\mathcal{G}_{\beta,N}(A_{N,\rho}^c)>\varepsilon\Big)&\le\mathbb{P}\bigg(\mathcal{G}_{\beta,N}(A_{N,\rho}^c)>\varepsilon,\frac{1}{\log N^2}\log Z_N^\psi(\beta)\ge f^\psi(\beta)-\widetilde{\eta}\bigg)\\&\quad+\mathbb{P}\bigg(\frac{1}{\log N^2}\log Z_N^\psi(\beta)<f^\psi(\beta)-\widetilde{\eta}\bigg)\\&\mathrel{\hat{=}}(1)+(2).\end{aligned} \tag{2.8.31}$$

For any $\widetilde{\eta}>0$, we have $(2)\to 0$ by (2.8.11). Furthermore, since

$$\Big\{\mathcal{G}_{\beta,N}(A_{N,\rho}^c)>\varepsilon\Big\}\subseteq\bigg\{\log\sum_{v\in A_{N,\rho}^c}e^{\beta\psi_v}>\log Z_N^\psi(\beta)+\log\varepsilon\bigg\}, \tag{2.8.32}$$

then

$$(1)\le\mathbb{P}\Bigg(\frac{1}{\log N^2}\log\sum_{v\in A_{N,\rho}^c}e^{\beta\psi_v}>f^\psi(\beta)-\rho/2+\eta\Bigg), \tag{2.8.33}$$

where

$$\eta\mathrel{\hat{=}}\rho/2-\widetilde{\eta}+\frac{\log\varepsilon}{\log N^2}. \tag{2.8.34}$$

Choose $\widetilde{\eta}>0$ small enough, with respect to ρ, and N large enough, with respect to ρ and ε, that $\eta>0$. The right-hand side of (2.8.33) converges to 0 by Lemma 2.8.3. This proves $\lim_{N\to\infty}\mathcal{G}_{\beta,N}(A_{N,\rho}^c)=0$ in $\mathbb{P}$-probability. Since

$$\sup_{N\in\mathbb{N}}|\mathcal{G}_{\beta,N}(A_{N,\rho}^c)|^p\le 1, \tag{2.8.35}$$

the L^p convergence follows trivially. □

The fact that the Gibbs measure does not carry any weight on $A_{N,\rho}^c$ in the limit generalizes to expectations of bounded functions of s vertices in V_N sampled from the product of Gibbs measures. In Section 2.8.5, this will be used to obtain the approximate extended Ghirlanda-Guerra identities on V_N from the ones on $A_{N,\rho}$.

Proposition 2.8.5. *Let $\beta > 0$ and $\rho \in (0, \lambda_1)$. Denote $\boldsymbol{v} \mathrel{\hat{=}} (v^1, v^2, ..., v^s)$. Then, for any $s \in \mathbb{N}$ and any functions $h : V_N^s \to \mathbb{R}$ such that $\sup_N \|h\|_\infty < \infty$,*

$$\lim_{N\to\infty} \left| \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h(\boldsymbol{v})\Big] - \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[h(\boldsymbol{v})\Big] \right| = 0. \tag{2.8.36}$$

Proof. Introducing an auxiliary term,

$$\begin{aligned} \left| \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h(\boldsymbol{v})\Big] - \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[h(\boldsymbol{v})\Big] \right| &\leq \left| \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h(\boldsymbol{v})\Big] - \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h(\boldsymbol{v})\, \mathbf{1}_{\{\boldsymbol{v}\in A_{N,\rho}^{\times s}\}}\Big] \right| \\ &\quad + \left| \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h(\boldsymbol{v})\, \mathbf{1}_{\{\boldsymbol{v}\in A_{N,\rho}^{\times s}\}}\Big] - \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[h(\boldsymbol{v})\Big] \right| \\ &\mathrel{\hat{=}} (1) + (2). \end{aligned} \tag{2.8.37}$$

Now, by monotonicity and sub-additivity,

$$(1) = \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h(\boldsymbol{v})\mathbf{1}_{\{\exists i\in\{1,\dots,s\} \text{ s.t. } v^i\in A_{N,\rho}^c\}}\Big] \leq s\ \mathbb{E}\mathcal{G}_{\beta,N}(A_{N,\rho}^c) \cdot \sup_N \|h\|_\infty. \tag{2.8.38}$$

Similarly, for the second term,

$$\begin{aligned} (2) &= \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[h(\boldsymbol{v})\Big] - \mathbb{E}\mathcal{G}_{\beta,N}^{\times s}\Big[h(\boldsymbol{v})\, \mathbf{1}_{\{\boldsymbol{v}\in A_{N,\rho}^{\times s}\}}\Big] \\ &= \mathbb{E}\left[\frac{\mathcal{G}_{\beta,N}^{\times s}\Big[h(\boldsymbol{v})\, \mathbf{1}_{\{\boldsymbol{v}\in A_{N,\rho}^{\times s}\}}\Big]}{\mathcal{G}_{\beta,N}^{\times s}(\boldsymbol{v}\in A_{N,\rho}^{\times s})} \Big(1 - \mathcal{G}_{\beta,N}^{\times s}(\boldsymbol{v}\in A_{N,\rho}^{\times s})\Big) \right] \\ &\leq \mathbb{E}\Big[1 - \mathcal{G}_{\beta,N}^{\times s}(\boldsymbol{v}\in A_{N,\rho}^{\times s})\Big] \cdot \sup_N \|h\|_\infty \leq s\ \mathbb{E}\mathcal{G}_{\beta,N}(A_{N,\rho}^c) \cdot \sup_N \|h\|_\infty. \end{aligned} \tag{2.8.39}$$

By Lemma 2.8.4, $(1) + (2) \to 0$ as $N \to \infty$. This ends the proof. $\square$

When $s = 2$ and $h(v, v') \mathrel{\hat{=}} \mathbf{1}_{\{q^N(v,v')\leq r\}}$, Proposition 2.8.5 tells us that we can compute the limiting two-overlap distribution of Theorem 2.6.3 by only considering a restricted version, where the points are sampled from $A_{N,\rho}^2$ instead of V_N^2. This property will be crucial to control the covariance of the increments in the next section, where we adapt the Bovier-Kurkova technique. The proof of Theorem 2.6.3 will be given right after, in Section 2.8.4.

Corollary 2.8.6. *Let $\beta > 0$ and $\rho \in (0, \lambda_1)$. Then, for any $r \in \mathbb{R}$,*

$$\lim_{N\to\infty} \left| \mathbb{E}\mathcal{G}_{\beta,N}^{\times 2}\left[\mathbf{1}_{\{q^N(v,v')\leq r\}}\right] - \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times 2}\left[\mathbf{1}_{\{q^N(v,v')\leq r\}}\right] \right| = 0. \tag{2.8.40}$$

Note that (2.8.40) is valid even if r depends on ρ.

2.8.3. An adaptation of the Bovier-Kurkova technique

The Bovier-Kurkova technique is a way to compute the two-overlap distribution of a model in terms of the free energy of a perturbed version of that model. In the context of this paper, this connection is established by Proposition 2.8.8 below in the case $(s = 1, k = 1, h \equiv 1)$. One difficulty in the present case is the fact that the covariance between the increments of the field depends on their position relative to the boundary. The restriction to the set $A_{N,\rho}$ is a way to control this, cf. Lemma 2.8.7.

To simplify the notation, recall

$$\bar{\mathcal{J}}_{\sigma^2}(\cdot) \mathrel{\hat{=}} \frac{\mathcal{J}_{\sigma^2}(\cdot)}{\mathcal{J}_{\sigma^2}(1)}, \tag{2.8.41}$$

and denote the increments of overlaps by

$$q_{\alpha,\alpha'}^N(v,v') \mathrel{\hat{=}} \frac{\mathbb{E}[\psi_v(\alpha,\alpha')\psi_{v'}]}{\mathcal{J}_{\sigma^2}(1)\log N + C_0}, \quad v,v' \in V_N, \tag{2.8.42}$$

where C_0 is the constant introduced in Lemma 2.9.3. Estimates on (2.8.42) are given in terms of (2.8.41) in Corollary 2.9.6 of Appendix 2.9.1. The following lemma uses these estimates in order to compare $q^N(v,v')$ and $q_{\alpha,\alpha'}^N(v,v')$.

Lemma 2.8.7. *Let $0 \leq \alpha < \alpha' \leq 1$ and $\rho \in (0,1]$. Then, for all $v,v' \in A_{N,\rho}$, for all*

$$\varepsilon \geq \frac{C_7}{\sqrt{\log N}} + C_8\,\rho, \qquad (C_7, C_8 \text{ are from (2.9.44)}) \tag{2.8.43}$$

and for N large enough (dependent on α and α', but independent from v,v', and independent from ρ (except when $\alpha = 0$)) :

(1) If $q^N(v,v') \leq \bar{\mathcal{J}}_{\sigma^2}(\alpha) - \varepsilon$, then

$$q_{\alpha,\alpha'}^N(v,v') = O\left((\log N)^{-1/2}\right) + O(\rho).$$

(2) If $\bar{\mathcal{J}}_{\sigma^2}(\alpha) + \varepsilon \leq q^N(v,v') \leq \bar{\mathcal{J}}_{\sigma^2}(\alpha') - \varepsilon$, then

$$q_{\alpha,\alpha'}^N(v,v') = q^N(v,v') - \bar{\mathcal{J}}_{\sigma^2}(\alpha) + O\left((\log N)^{-1/2}\right) + O(\rho).$$

(3) If $\bar{\mathcal{J}}_{\sigma^2}(\alpha') + \varepsilon \leq q^N(v,v')$, then

$$q^N_{\alpha,\alpha'}(v,v') = \bar{\mathcal{J}}_{\sigma^2}(\alpha,\alpha') + O\left((\log N)^{-1/2}\right) + O(\rho).$$

In all three cases, $O(\rho)$ is uniform in N.

Proof. From (2.9.44), we know that $|q^N(v,v') - \bar{\mathcal{J}}_{\sigma^2}(b_N(v,v'))| \leq \varepsilon$. Thus, in each case respectively, we deduce $(1): b_N \leq \alpha$, $(2): \alpha \leq b_N \leq \alpha'$, and $(3): \alpha' \leq b_N$. Use (2.9.44) again to get the appropriate bounds on $q^N_{\alpha,\alpha'}(v,v')$. □

Here is the main result of this section.

Proposition 2.8.8. *Let $0 \leq \alpha < \alpha' \leq 1$, $\rho \in (0,1]$ and $S_{\alpha,\alpha'} \mathrel{\hat{=}} (\bar{\mathcal{J}}_{\sigma^2}(\alpha), \bar{\mathcal{J}}_{\sigma^2}(\alpha')]$. Let $\beta > 0$, $s \in \mathbb{N}$, $k \in \{1,...,s\}$, and let $h : V_N^s \to \mathbb{R}$ be such that $\sup_N \|h\|_\infty < \infty$. Then, for all*

$$\varepsilon \geq \frac{C_7}{\sqrt{\log N}} + C_8\,\rho, \qquad (C_7, C_8 \text{ are from (2.9.44)}) \tag{2.8.44}$$

and for N large enough (dependent on α and α', but independent from v, v', and independent from ρ (except when $\alpha = 0$)), we have

$$\begin{aligned} &\left| \frac{\mathbb{E}\mathcal{G}^{\times s}_{\beta,N,\rho}[\psi_{v^k}(\alpha,\alpha')h(\boldsymbol{v})]}{\beta\,(\mathcal{J}_{\sigma^2}(1)\log N + C_0)} - \left\{ \begin{array}{l} \sum_{l=1}^s \mathbb{E}\mathcal{G}^{\times s}_{\beta,N,\rho}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r<q^N(v^k,v^l)\}} dr\, h(\boldsymbol{v})\Big] \\ -s\,\mathbb{E}\mathcal{G}^{\times(s+1)}_{\beta,N,\rho}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r<q^N(v^k,v^{s+1})\}} dr\, h(\boldsymbol{v})\Big] \end{array} \right\} \right| \\ &\leq C \cdot s \cdot \sup_N \|h\|_\infty \cdot \left\{ \begin{array}{l} \mathbb{E}\mathcal{G}^{\times 2}_{\beta,N,\rho}\Big[\mathbf{1}_{\{\bar{\mathcal{J}}_{\sigma^2}(\alpha)-\varepsilon \leq q^N(v,v') \leq \bar{\mathcal{J}}_{\sigma^2}(\alpha)+\varepsilon\}}\Big] \\ \mathbb{E}\mathcal{G}^{\times 2}_{\beta,N,\rho}\Big[\mathbf{1}_{\{\bar{\mathcal{J}}_{\sigma^2}(\alpha')-\varepsilon \leq q^N(v,v') \leq \bar{\mathcal{J}}_{\sigma^2}(\alpha')+\varepsilon\}}\Big] \\ +O\left((\log N)^{-1/2}\right) + O(\rho) \end{array} \right\}, \end{aligned} \tag{2.8.45}$$

where $O(\rho)$ is uniform in N and $C > 0$ is a universal constant.

Proof. For any $l \in \{1,...,s+1\}$,

$$\begin{aligned} &\mathbb{E}\mathcal{G}^{\times(s+1)}_{\beta,N,\rho}\left[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r<q^N(v^k,v^l)\}} dr\ h(\boldsymbol{v})\right] \\ &\quad = \mathbb{E}\mathcal{G}^{\times(s+1)}_{\beta,N,\rho}\Big[(q^N(v^k,v^l) - \bar{\mathcal{J}}_{\sigma^2}(\alpha))\mathbf{1}_{\{\bar{\mathcal{J}}_{\sigma^2}(\alpha) < q^N(v^k,v^l) \leq \bar{\mathcal{J}}_{\sigma^2}(\alpha')\}}\, h(\boldsymbol{v})\Big] \\ &\quad\quad + \mathbb{E}\mathcal{G}^{\times(s+1)}_{\beta,N,\rho}\Big[\bar{\mathcal{J}}_{\sigma^2}(\alpha,\alpha')\mathbf{1}_{\{\bar{\mathcal{J}}_{\sigma^2}(\alpha') < q^N(v^k,v^l)\}}\, h(\boldsymbol{v})\Big]. \end{aligned} \tag{2.8.46}$$

On the other hand,

$$\mathbb{E}\mathcal{G}^{\times s}_{\beta,N,\rho}\Big[\psi_{v^k}(\alpha,\alpha')h(\boldsymbol{v})\Big] = \sum_{\boldsymbol{v}\in A^{\times s}_{N,\rho}} \mathbb{E}\left[\frac{\psi_{v^k}(\alpha,\alpha')h(\boldsymbol{v})\prod_{l=1}^s \exp(\beta\psi_{v^l})}{\prod_{l'=1}^s \sum_{v^{l'}\in A_{N,\rho}} \exp(\beta\psi_{v^{l'}})}\right]. \tag{2.8.47}$$

For a centered Gaussian vector $\boldsymbol{X} \circeq (X_1, ..., X_n)$ and a twice-continuously differentiable function F on $\mathbb{R}^n$, of moderate growth at infinity, we have the formula $\mathbb{E}[X_i F(\boldsymbol{X})] = \sum_{j=1}^n \mathbb{E}[X_i X_j]\, \mathbb{E}\Big[\partial_{X_j} F(\boldsymbol{X})\Big]$. Here, for any $\boldsymbol{v} \in A_{N,\rho}^{\times s}$, the relevant Gaussian vector is

$$\Big(\psi_{v^k}(\alpha, \alpha') \;;\; \psi_{v^l}, l \in \{1, ..., s\} \;;\; \psi_{v^{l'}}, v^{l'} \in A_{N,\rho}, l' \in \{1, ..., s\}\Big), \tag{2.8.48}$$

where $X_i \circeq \psi_{v^k}(\alpha, \alpha')$ and $F \circeq h(\boldsymbol{v}) \prod_{l=1}^s \exp(\beta \psi_{v^l}) / \prod_{l'=1}^s \sum_{v^{l'} \in A_{N,\rho}} \exp(\beta \psi_{v^{l'}})$. Applying the formula to the right-hand side of (2.8.47) yields

$$\begin{aligned}(2.8.47) = \sum_{l=1}^s \beta\, \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}&\Big[\mathbb{E}\Big[\psi_{v^k}(\alpha, \alpha')\psi_{v^l}\Big] h(\boldsymbol{v})\Big] \\ &- s\, \beta\, \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times (s+1)}\Big[\mathbb{E}\Big[\psi_{v^k}(\alpha, \alpha')\psi_{v^{s+1}}\Big] h(\boldsymbol{v})\Big].\end{aligned} \tag{2.8.49}$$

If we divide (2.8.49) on both sides by $\beta\, (\mathcal{J}_{\sigma^2}(1) \log N + C_0)$, we deduce

$$\frac{\mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[\psi_{v^k}(\alpha, \alpha') h(\boldsymbol{v})\Big]}{\beta\, (\mathcal{J}_{\sigma^2}(1) \log N + C_0)} = \left\{\begin{array}{l} \sum_{l=1}^s \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[q_{\alpha,\alpha'}^N(v^k, v^l) h(\boldsymbol{v})\Big] \\ - s\, \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times (s+1)}\Big[q_{\alpha,\alpha'}^N(v^k, v^{s+1}) h(\boldsymbol{v})\Big] \end{array}\right\}. \tag{2.8.50}$$

Now, one by one, take the difference in absolute value between each of the $s+1$ expectations inside the braces in (2.8.50) and the corresponding expectation on the left-hand side of (2.8.46). We obtain the bound (2.8.45) by using Lemma 2.8.7. $\square$

2.8.4. Computation of the limiting two-overlap distribution

Let $\alpha, \alpha' \in [0, 1]$ be such that

$$\lambda^{j^\star - 1} \leq \lambda_{i^\star - 1} \leq \alpha < \alpha' \leq \lambda_{i^\star} \leq \lambda^{j^\star} \tag{2.8.51}$$

for some $i^\star$ and $j^\star$. Define ψ^u, the *perturbed scale-inhomogeneous GFF*, mentioned in the previous section, by

$$\psi_v^u \circeq u\, \phi_v(\alpha, \alpha') + \psi_v, \quad \text{where } u > -\sigma_{i^\star}. \tag{2.8.52}$$

The dependence on α and α' is made implicit to lighten the notation. In the proof of Theorem 2.6.3, Proposition 2.8.8 will be used to link the limiting two-overlap distribution of ψ to the derivative of the limiting free energy of ψ^u with respect to the perturbation parameter u.

Proof of Theorem 2.6.3. By Corollary 2.8.6, it suffices to prove that

$$\lim_{\rho\to 0}\lim_{N\to\infty}\mathbb{E}\mathcal{G}^{\times 2}_{\beta,N,\rho}\Big[\mathbf{1}_{\{q^N(v,v')\le r\}}\Big] \\ = \begin{cases} 0, & \text{if } r<0,\\ (2/\bar\sigma_j)/\beta, & \text{if } r\in[\bar{\mathcal{J}}_{\sigma^2}(\lambda^{j-1}),\bar{\mathcal{J}}_{\sigma^2}(\lambda^{j})),\ j\le l_\beta-1,\\ 1, & \text{if } r\ge \bar{\mathcal{J}}_{\sigma^2}(\lambda^{l_\beta-1}).\end{cases} \tag{2.8.53}$$

Since $[0,1]\subseteq\mathbb{R}$ is compact, the space $\mathcal{M}_1([0,1])$ of probability measures on $[0,1]$ is compact under the weak topology. Thus, any subsequence of the cumulative distribution functions on the left-hand side of (2.8.53) has a subsequence converging to a cumulative distribution function. Pick any converging sub-subsequence and denote its limit by $r\mapsto Q_\beta(r)$. Since $\mathcal{M}_1([0,1])$ is a metric space, the proof is reduced to showing that Q_β is given by the right-hand side of (2.8.53).

We already know that $Q_\beta(r)=0$ for all $r<0$ since Corollary 2.9.6 implies

$$\liminf_{\rho\to 0}\liminf_{N\to\infty}\min_{v,v'\in A_{N,\rho}} q^N(v,v')\ge 0. \tag{2.8.54}$$

We also have $Q_\beta(r)=1$ for all $r\ge 1$ since $\max_{v,v'\in V_N} q^N(v,v')\le 1$ by Lemma 2.9.3 and the Cauchy-Schwarz inequality.

To determine Q_β on $[0,1)$, let $\alpha,\alpha'\in[0,1]$ be such that $\bar{\mathcal{J}}_{\sigma^2}(\alpha),\bar{\mathcal{J}}_{\sigma^2}(\alpha')$ are continuity points of Q_β and (2.8.51) is satisfied. Direct differentiation gives

$$\frac{2\sigma_{i^\star}}{\beta^2\mathcal{J}_{\sigma^2}(1)}\,\frac{\partial}{\partial u}\mathbb{E}\Big[f^{\psi^u}_{N,\rho}(\beta)\Big]\bigg|_{u=0}=\frac{\mathbb{E}\mathcal{G}_{\beta,N,\rho}\Big[\psi_v(\alpha,\alpha')\Big]}{\beta\mathcal{J}_{\sigma^2}(1)\log N}. \tag{2.8.55}$$

Combine this result with Proposition 2.8.8 in the special case ($s=1$, $k=1$, $h\equiv 1$). After taking the limits $N\to\infty$ (use Corollary 2.8.6 on the right-hand side of (2.8.45)), $\rho\to 0$ and then $\varepsilon\to 0$, we find

$$\int_{(\bar{\mathcal{J}}_{\sigma^2}(\alpha),\bar{\mathcal{J}}_{\sigma^2}(\alpha')]} Q_\beta(r)dr=\lim_{\rho\to 0}\lim_{N\to\infty}\frac{2\sigma_{i^\star}}{\beta^2\mathcal{J}_{\sigma^2}(1)}\,\frac{\partial}{\partial u}\mathbb{E}\Big[f^{\psi^u}_{N,\rho}(\beta)\Big]\bigg|_{u=0}. \tag{2.8.56}$$

For all $\rho\in(0,1]$, the function $u\mapsto\mathbb{E}\Big[f^{\psi^u}_{N,\rho}(\beta)\Big]$ is convex by Lemma 2.9.9, and by Theorem 2.6.2,

$$\lim_{N\to\infty}\mathbb{E}\Big[f^{\psi^u}_{N,\rho}(\beta)\Big]=f^{\psi^u}(\beta). \tag{2.8.57}$$

Pointwise limits preserve convexity, so $u \mapsto f^{\psi^u}(\beta)$ is convex. From Lemma 2.9.10, , we also know that $u \mapsto f^{\psi^u}(\beta)$ is differentiable on an open interval $(-\delta, \delta)$, for $\delta = \delta(\beta, \alpha, \alpha', \boldsymbol{\sigma}, \boldsymbol{\lambda})$ small enough. In particular, by another standard result of convexity (see e.g. Theorem 25.7 in Rockafellar (1970)),

$$\lim_{N\to\infty} \frac{\partial}{\partial u} \mathbb{E}\left[f_{N,\rho}^{\psi^u}(\beta)\right] = \frac{\partial}{\partial u} f^{\psi^u}(\beta), \tag{2.8.58}$$

for all $u \in (-\delta, \delta)$ (and all $\rho \in (0,1]$). The derivative of $u \mapsto f^{\psi^u}(\beta)$ at $u = 0$ is given by (2.9.76). Thus, from (2.8.56), we get

$$\int_{(\bar{\mathcal{J}}_{\sigma^2}(\alpha), \bar{\mathcal{J}}_{\sigma^2}(\alpha')]} Q_\beta(r) dr = \begin{cases} \bar{\mathcal{J}}_{\sigma^2}(\alpha, \alpha') \frac{(2/\bar{\sigma}_{j^\star})}{\beta}, & \text{if } j^\star \leq l_\beta - 1, \\ \bar{\mathcal{J}}_{\sigma^2}(\alpha, \alpha'), & \text{if } j^\star \geq l_\beta. \end{cases} \tag{2.8.59}$$

But Q_β is right-continuous (it's a cumulative distribution function) and (2.8.59) holds for all pairs $\bar{\mathcal{J}}_{\sigma^2}(\alpha), \bar{\mathcal{J}}_{\sigma^2}(\alpha')$ of continuity points satisfying (2.8.51), so Q_β must be equal to the right-hand side of (2.8.53). This ends the proof. □

2.8.5. Proof of the approximate extended Ghirlanda-Guerra identities

We start by proving a concentration result. Denote $\boldsymbol{v} \mathrel{\hat=} (v^1, ..., v^s)$ in this section.

Lemma 2.8.9. *Let $\lambda_{i^\star - 1} \leq \alpha < \alpha' \leq \lambda_{i^\star}$ for some $i^\star$, and let $\beta > 0$ and $\rho \in (0,1]$. Then, for any $s \in \mathbb{N}$, any $k \in \{1, ..., s\}$ and any functions $h : V_N^s \to \mathbb{R}$ such that $\sup_N \|h\|_\infty < \infty$,*

$$\lim_{N\to\infty} \frac{\left|\mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\left[\psi_{v^k}(\alpha, \alpha') h(\boldsymbol{v})\right] - \mathbb{E}\mathcal{G}_{\beta,N,\rho}\left[\psi_{v^k}(\alpha, \alpha')\right] \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\left[h(\boldsymbol{v})\right]\right|}{\beta\left(\mathcal{J}_{\sigma^2}(1) \log N + C_0\right)} = 0. \tag{2.8.60}$$

Proof. If we apply Jensen's inequality to the expectation $\mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}[\cdot]$, followed by the triangle inequality, we have

$$\begin{aligned} &\left|\mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\left[\psi_{v^k}(\alpha, \alpha') h(\boldsymbol{v})\right] - \mathbb{E}\mathcal{G}_{\beta,N,\rho}\left[\psi_{v^k}(\alpha, \alpha')\right] \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\left[h(\boldsymbol{v})\right]\right| \\ &\quad \leq \mathbb{E}\mathcal{G}_{\beta,N,\rho}\left|\psi_{v^k}(\alpha, \alpha') - \mathbb{E}\mathcal{G}_{\beta,N,\rho}\left[\psi_{v^k}(\alpha, \alpha')\right]\right| \cdot \sup_N \|h\|_\infty \\ &\quad \leq ((a) + (b)) \cdot \sup_N \|h\|_\infty, \end{aligned} \tag{2.8.61}$$

where

$$\begin{aligned} (a) + (b) \mathrel{\hat=} \; &\mathbb{E}\mathcal{G}_{\beta,N,\rho}\left|\psi_{v^k}(\alpha, \alpha') - \mathcal{G}_{\beta,N,\rho}\left[\psi_{v^k}(\alpha, \alpha')\right]\right| \\ &+ \mathbb{E}\left|\mathcal{G}_{\beta,N,\rho}\left[\psi_{v^k}(\alpha, \alpha')\right] - \mathbb{E}\mathcal{G}_{\beta,N,\rho}\left[\psi_{v^k}(\alpha, \alpha')\right]\right|. \end{aligned} \tag{2.8.62}$$

In the remainder, we follow the strategy developed in the proof of Theorem 3.8 in Panchenko (2013b), where the same concentration result was proved for the mixed p-spin model. We show that, for all $\rho \in (0,1]$,

$$\lim_{N\to\infty} \frac{(a)}{\log N} = 0 \quad \text{and} \quad \lim_{N\to\infty} \frac{(b)}{\log N} = 0. \tag{2.8.63}$$

$$\boxed{\textbf{Step 1 :} \text{ For all } \rho \in (0,1],\ \lim_{N\to\infty} \tfrac{(a)}{\log N} = 0.}$$

Note that

$$\begin{aligned}(a) &= \mathbb{E}\mathcal{G}_{\beta,N,\rho}\left|\sum_{v^2\in A_{N,\rho}} (\psi_{v^1}(\alpha,\alpha') - \psi_{v^2}(\alpha,\alpha'))\frac{\exp(\beta\psi_{v^2})}{\sum_{z^2\in A_{N,\rho}}\exp(\beta\psi_{z^2})}\right| \\ &\leq \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times 2}\Big|\psi_{v^1}(\alpha,\alpha') - \psi_{v^2}(\alpha,\alpha')\Big|. \end{aligned} \tag{2.8.64}$$

For $u \geq 0$, we define a perturbed version of the last quantity (where the Gibbs measure $\mathcal{G}_{\beta,N,\rho,u}$ is now defined with respect to ψ^u) :

$$D(u) \mathrel{\hat=} \mathbb{E}\mathcal{G}_{\beta,N,\rho,u}^{\times 2}\Big|\psi_{v^1}(\alpha,\alpha') - \psi_{v^2}(\alpha,\alpha')\Big| \tag{2.8.65}$$

We can easily verify that

$$uD(0) = \int_0^u D(y)dy - \int_0^u \int_0^x \frac{\partial}{\partial y} D(y)dydx, \tag{2.8.66}$$

and also that

$$\frac{\partial}{\partial y} D(y) = \frac{\beta}{\sigma_{i^\star}} \mathbb{E}\mathcal{G}_{\beta,N,\rho,y}^{\times 3}\left[\begin{aligned}&\Big|\psi_{v^1}(\alpha,\alpha') - \psi_{v^2}(\alpha,\alpha')\Big| \\ &\cdot \Big(\psi_{v^1}(\alpha,\alpha') + \psi_{v^2}(\alpha,\alpha') - 2\psi_{v^3}(\alpha,\alpha')\Big)\end{aligned}\right]. \tag{2.8.67}$$

If we separate the last expectation in two parts and apply the Cauchy-Schwarz inequality to each one of them, we find (for $y \geq 0$) :

$$\begin{aligned}\left|\frac{\partial}{\partial y} D(y)\right| &\leq \frac{\beta}{\sigma_{i^\star}}\left\{\begin{aligned}&\mathbb{E}\mathcal{G}_{\beta,N,\rho,y}^{\times 3}\Big|\psi_{v^1}(\alpha,\alpha') - \psi_{v^2}(\alpha,\alpha')\Big|\Big|\psi_{v^1}(\alpha,\alpha') - \psi_{v^3}(\alpha,\alpha')\Big| \\ &+ \mathbb{E}\mathcal{G}_{\beta,N,\rho,y}^{\times 3}\Big|\psi_{v^1}(\alpha,\alpha') - \psi_{v^2}(\alpha,\alpha')\Big|\Big|\psi_{v^2}(\alpha,\alpha') - \psi_{v^3}(\alpha,\alpha')\Big|\end{aligned}\right\} \\ &\leq \frac{\beta}{\sigma_{i^\star}} \cdot 2\, \mathbb{E}\mathcal{G}_{\beta,N,\rho,y}^{\times 2}\Big[\Big(\psi_{v^1}(\alpha,\alpha') - \psi_{v^2}(\alpha,\alpha')\Big)^2\Big]. \end{aligned} \tag{2.8.68}$$

From the elementary inequality $(c+d)^2 \leq 2c^2 + 2d^2$, we also have

$$\begin{aligned} &2\,\mathbb{E}\mathcal{G}^{\times 2}_{\beta,N,\rho,y}\Big[\Big(\psi_{v^1}(\alpha,\alpha') - \psi_{v^2}(\alpha,\alpha')\Big)^2\Big] \\ &\quad \leq 8\,\mathbb{E}\mathcal{G}_{\beta,N,\rho,y}\Big[\Big(\psi_v(\alpha,\alpha') - \mathcal{G}_{\beta,N,\rho,y}\big[\psi_v(\alpha,\alpha')\big]\Big)^2\Big]. \end{aligned} \tag{2.8.69}$$

By putting (2.8.68) and (2.8.69) together in (2.8.66), we obtain (for $u > 0$) :

$$\begin{aligned} D(0) &\leq \frac{1}{u}\int_0^u D(y)dy + \int_0^u \left|\frac{\partial}{\partial y} D(y)\right| dy \\ &\leq 2\left(\frac{1}{u}\int_0^u \mathbb{E}\mathcal{G}_{\beta,N,\rho,y}\Big[\Big(\psi_v(\alpha,\alpha') - \mathcal{G}_{\beta,N,\rho,y}\big[\psi_v(\alpha,\alpha')\big]\Big)^2\Big] dy\right)^{1/2} \\ &\qquad + \frac{8\beta}{\sigma_{i^\star}}\int_0^u \mathbb{E}\mathcal{G}_{\beta,N,\rho,y}\Big[\Big(\psi_v(\alpha,\alpha') - \mathcal{G}_{\beta,N,\rho,y}\big[\psi_v(\alpha,\alpha')\big]\Big)^2\Big] dy. \end{aligned} \tag{2.8.70}$$

In order to bound $\frac{1}{u}\int_0^u D(y)dy$, we separated $D(y)$ in two parts (with the triangle inequality) and we applied the Cauchy-Schwarz inequality to the two resulting expectations $\frac{1}{u}\int_0^u \mathbb{E}\mathcal{G}_{\beta,N,\rho,y}[\,\cdot\,]\,dy$. Denote

$$\varepsilon_{N,\rho}(u) \mathrel{\hat{=}} \frac{1}{\log N}\int_0^u \mathbb{E}\mathcal{G}_{\beta,N,\rho,y}\Big[\Big(\psi_v(\alpha,\alpha') - \mathcal{G}_{\beta,N,\rho,y}\big[\psi_v(\alpha,\alpha')\big]\Big)^2\Big] dy. \tag{2.8.71}$$

So far, we have shown that

$$\frac{(a)}{\log N} \leq \frac{D(0)}{\log N} \leq 2\sqrt{\frac{\varepsilon_{N,\rho}(u)}{u\log N}} + \frac{8\beta}{\sigma_{i^\star}}\,\varepsilon_{N,\rho}(u). \tag{2.8.72}$$

Let

$$F(u) \mathrel{\hat{=}} f^{\psi^u}_{N,\rho}(\beta) = \frac{1}{\log N^2}\log\sum_{v\in A_{N,\rho}} e^{\beta(u\phi_v(\alpha,\alpha') + \psi_v)}, \tag{2.8.73}$$

and note that

$$\begin{aligned} \mathbb{E}[F''(y)] &= \frac{\beta^2}{\sigma^2_{i^\star}\log N^2}\,\mathbb{E}\Big[\mathcal{G}_{\beta,N,\rho,y}\big[\big(\psi_v(\alpha,\alpha')\big)^2\big] - \Big(\mathcal{G}_{\beta,N,\rho,y}\big[\psi_v(\alpha,\alpha')\big]\Big)^2\Big] \\ &= \frac{\beta^2}{2\sigma^2_{i^\star}}\cdot\frac{1}{\log N}\,\mathbb{E}\mathcal{G}_{\beta,N,\rho,y}\Big[\Big(\psi_v(\alpha,\alpha') - \mathcal{G}_{\beta,N,\rho,y}\big[\psi_v(\alpha,\alpha')\big]\Big)^2\Big]. \end{aligned} \tag{2.8.74}$$

From (2.8.71) and the convexity of F (see Lemma 2.9.9), we have, for all $y \in (0,\sigma_{i^\star})$,

$$\begin{aligned} \varepsilon_{N,\rho}(u) &= \frac{2\sigma^2_{i^\star}}{\beta^2}\int_0^u \mathbb{E}[F''(y)]\,dy = \frac{2\sigma^2_{i^\star}}{\beta^2}\,\mathbb{E}[F'(u) - F'(0)] \\ &\leq \frac{2\sigma^2_{i^\star}}{\beta^2}\,\mathbb{E}\left[\frac{F(u+y) - F(u)}{y} - \frac{F(0) - F(-y)}{y}\right]. \end{aligned} \tag{2.8.75}$$

By putting (2.8.75) in (2.8.72) and by using the mean convergence in Theorem 2.6.2, we get, for all $\rho \in (0,1]$ and all $u > 0$ and $y \in (0, \sigma_{i^\star})$,

$$\limsup_{N\to\infty} \frac{(a)}{\log N} \leq \frac{8\beta}{\sigma_{i^\star}} \cdot \frac{2\sigma_{i^\star}^2}{\beta^2} \left(\frac{f(u+y) - f(u)}{y} - \frac{f(0) - f(-y)}{y} \right), \tag{2.8.76}$$

where $f(u) \mathrel{\hat=} f^{\psi^u}(\beta)$. From Lemma 2.9.10, there exists $\delta = \delta(\beta, \alpha, \alpha', \boldsymbol{\sigma}, \boldsymbol{\lambda})$ such that f is differentiable on $(-\delta, \delta)$. Therefore, take $u \to 0^+$ and then $y \to 0^+$ in the above equation, the right-hand side goes to 0. The left-hand side does not depend on u or y, so we conclude that for all $\rho \in (0,1]$, $\lim_{N\to\infty} (a)/\log N = 0$.

$$\boxed{\textbf{Step 2 :} \text{ For all } \rho \in (0,1],\ \lim_{N\to\infty} \tfrac{(b)}{\log N} = 0.}$$

Let $F(u) \mathrel{\hat=} f_{N,\rho}^{\psi^u}(\beta)$ as in (2.8.73) and, for $u \in (0, \sigma_{i^\star})$, let

$$\eta(u) \mathrel{\hat=} \Big|F(-u) - \mathbb{E}[F(-u)]\Big| + \Big|F(0) - \mathbb{E}[F(0)]\Big| + \Big|F(u) - \mathbb{E}[F(u)]\Big|. \tag{2.8.77}$$

Differentiation of the free energy gives

$$(b) = \frac{\sigma_{i^\star} \log N^2}{\beta} \mathbb{E}\Big|F'(0) - \mathbb{E}[F'(0)]\Big|. \tag{2.8.78}$$

From the convexity of F (see Lemma 2.9.9),

$$\begin{aligned} F'(0) - \mathbb{E}[F'(0)] &\leq \frac{F(u) - F(0)}{u} - \mathbb{E}[F'(0)] \\ &\leq \left| \frac{\mathbb{E}[F(u)] - \mathbb{E}[F(0)]}{u} - \mathbb{E}[F'(0)] \right| + \frac{\eta(u)}{u}, \end{aligned} \tag{2.8.79}$$

$$\begin{aligned} F'(0) - \mathbb{E}[F'(0)] &\geq \frac{F(0) - F(-u)}{u} - \mathbb{E}[F'(0)] \\ &\geq - \left| \frac{\mathbb{E}[F(0)] - \mathbb{E}[F(-u)]}{u} - \mathbb{E}[F'(0)] \right| - \frac{\eta(u)}{u}. \end{aligned} \tag{2.8.80}$$

By taking the absolute value and the expectation, we get

$$\begin{aligned} \frac{\beta}{2\sigma_{i^\star}} \cdot \frac{(b)}{\log N} &\leq \left| \frac{\mathbb{E}[F(u)] - \mathbb{E}[F(0)]}{u} - \mathbb{E}[F'(0)] \right| \\ &\quad + \left| \frac{\mathbb{E}[F(0)] - \mathbb{E}[F(-u)]}{u} - \mathbb{E}[F'(0)] \right| + \frac{\mathbb{E}[\eta(u)]}{u}. \end{aligned} \tag{2.8.81}$$

Recall that F and η are functions of N and ρ by definition. From Theorem 2.6.2, we know that for all $\rho \in (0,1]$ and all $u \in (0, \sigma_{i^\star})$,

$$\lim_{N\to\infty} \mathbb{E}[\eta(u)] = 0. \tag{2.8.82}$$

Using (2.8.57) and (2.8.58) in (2.8.81), we get, for all $\rho \in (0,1]$ and all $u \in (0, \sigma_{i^\star})$,

$$\limsup_{N\to\infty} \left\{ \frac{\beta}{2\sigma_{i^\star}} \cdot \frac{(b)}{\log N} \right\} \leq \left| \frac{f(u) - f(0)}{u} - f'(0) \right| + \left| \frac{f(0) - f(-u)}{u} - f'(0) \right|, \tag{2.8.83}$$

where $f(u) \circeq f^{\psi^u}(\beta)$. Finally, take $u \to 0^+$ in the last equation, the differentiability of f at 0 (from Lemma 2.9.10) implies that for all $\rho \in (0,1]$, $\lim_{N\to\infty} (b)/\log N = 0$. This ends the proof of Lemma 2.8.9. □

Finally, we can prove the approximate extended Ghirlanda-Guerra identities.

Proof of Theorem 2.6.4. In addition to (2.6.16), assume that $\lambda_{i^\star-1} \leq \alpha < \alpha' \leq \lambda_{i^\star}$ for some $i^\star$. Also, let $\rho \in (0, \lambda_1)$. If we combine Lemma 2.8.9 and Proposition 2.8.8 with the triangle inequality, we get

$$\left| \begin{array}{l} \frac{\mathbb{E}\mathcal{G}_{\beta,N,\rho}[\psi_{\boldsymbol{v}^k}(\alpha,\alpha')]}{\beta\,(\mathcal{J}_{\sigma^2}(1)\log N + C_0)}\, \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[h(\boldsymbol{v})\Big] \\ - \left\{ \begin{array}{l} \sum_{l=1}^{s} \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^k, v^l)\}} dr\, h(\boldsymbol{v})\Big] \\ - s\, \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times (s+1)}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^k, v^{s+1})\}} dr\, h(\boldsymbol{v})\Big] \end{array} \right\} \end{array} \right| \leq \mathrm{RHS}(2.8.45) + o_N(1), \tag{2.8.84}$$

where RHS means "right-hand side of". Furthermore, from Proposition 2.8.8 in the special case ($s = 1$, $k = 1$, $h \equiv 1$),

$$\left| \begin{array}{l} \frac{\mathbb{E}\mathcal{G}_{\beta,N,\rho}[\psi_{\boldsymbol{v}^k}(\alpha,\alpha')]}{\beta\,(\mathcal{J}_{\sigma^2}(1)\log N + C_0)} \\ - \left\{ \begin{array}{l} \mathbb{E}\mathcal{G}_{\beta,N,\rho}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^k, v^k)\}} dr\Big] \\ - \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times 2}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^1, v^2)\}} dr\Big] \end{array} \right\} \end{array} \right| \leq \mathrm{RHS}_{(s=1,h\equiv 1)}(2.8.45). \tag{2.8.85}$$

By combining the last two bounds with the triangle inequality, we find

$$\limsup_{N\to\infty} \left| \begin{array}{l} \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times (s+1)}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^k, v^{s+1})\}} dr\ h(\boldsymbol{v})\Big] \\ - \left\{ \begin{array}{l} \frac{1}{s}\mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times 2}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^1, v^2)\}} dr\Big] \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[h(\boldsymbol{v})\Big] \\ + \frac{1}{s} \sum_{l\neq k}^{s} \mathbb{E}\mathcal{G}_{\beta,N,\rho}^{\times s}\Big[\int_{S_{\alpha,\alpha'}} \mathbf{1}_{\{r < q^N(v^k, v^l)\}} dr\ h(\boldsymbol{v})\Big] \end{array} \right\} \end{array} \right| \tag{2.8.86}$$

$$\le \widetilde{C} \cdot \sup_N \|h\|_\infty \cdot \left\{ \begin{array}{l} \limsup_{N\to\infty} \mathbb{E}\mathcal{G}^{\times 2}_{\beta,N,\rho}\Big[\mathbf{1}_{\{\bar{\mathcal{J}}_{\sigma^2}(\alpha)-\varepsilon\le q^N(v,v')\le \bar{\mathcal{J}}_{\sigma^2}(\alpha)+\varepsilon\}}\Big] \\ \limsup_{N\to\infty} \mathbb{E}\mathcal{G}^{\times 2}_{\beta,N,\rho}\Big[\mathbf{1}_{\{\bar{\mathcal{J}}_{\sigma^2}(\alpha')-\varepsilon\le q^N(v,v')\le \bar{\mathcal{J}}_{\sigma^2}(\alpha')+\varepsilon\}}\Big] \\ +O(\rho) \end{array} \right\}.$$

Using the triangle inequality, Proposition 2.8.5 and Corollary 2.8.6 in (2.8.86), it is easy to show that inequality (2.8.86) is also true if $\mathcal{G}_{\beta,N,\rho}$ is replaced everywhere by $\mathcal{G}_{\beta,N}$. From Theorem 2.6.3, condition (2.6.16) guarantees that $\bar{\mathcal{J}}_{\sigma^2}(\alpha)$ and $\bar{\mathcal{J}}_{\sigma^2}(\alpha')$ are continuity points of the limiting two-overlap distribution. Thus, after the replacement of the Gibbs measures in (2.8.86), take $\rho \to 0$ and then $\varepsilon \to 0$ to deduce (2.6.17).

If we only assume (2.6.16), note that $\lambda_{i-1} \le \alpha < \lambda_i$ and $\lambda_{i'-1} < \alpha' \le \lambda_{i'}$ for some i, i'. By the above argument, we have (2.6.17) for each pair of scales

$$\alpha < \lambda_i \;\; ; \;\; \lambda_i < \lambda_{i+1} \;\; ; \ldots ; \;\; \lambda_{i'-2} < \lambda_{i'-1} \;\; ; \;\; \lambda_{i'-1} < \alpha'. \tag{2.8.87}$$

Add all the limits together and use the triangle inequality to conclude. □

2.9. Appendix

2.9.1. Covariance estimates

The Markov property of the GFF, which is a consequence of the strong Markov property of the simple random walk (in the covariance function in (2.1.1)), implies that the value of the field inside a neighborhood is independent of the field outside given the boundary, see e.g. Dynkin (1980). In particular, for the neighborhood $[v]_\lambda$, this implies

$$\phi_v(\lambda) \doteq \mathbb{E}\Big[\phi_v \,|\, \mathcal{F}_{\partial[v]_\lambda \cup [v]^c_\lambda}\Big] = \mathbb{E}\Big[\phi_v \,|\, \mathcal{F}_{\partial[v]_\lambda}\Big]. \tag{2.9.1}$$

Define the *branching scale between v and v'* in V_N :

$$b_N(v,v') \doteq \max\Big\{\lambda \in [0,1] : [v]_\lambda \cap [v']_\lambda \neq \emptyset\Big\}. \tag{2.9.2}$$

This is the largest λ for which the two neighborhoods $[v]_\lambda$ and $[v']_\lambda$ intersect. We always have by definition that $\|v - v'\|_2$ is of order $N^{1-b_N(v,v')}$. The branching scale plays the same role as the branching time (normalized to lie in $[0,1]$) in branching random walk.

Define

$$\varepsilon_N \mathrel{\hat{=}} \frac{\log 4}{\log N}.$$

For all $v, v' \in V_N$ $(v \neq v')$, this definition guarantees that for all $N \in \mathbb{N}$,

$$\begin{aligned} &[v]_{1\wedge(b_N+\varepsilon_N)} \cap [v']_{1\wedge(b_N+\varepsilon_N)} = \emptyset \\ \text{and} \quad & \\ &[v]_{b_N} \cup [v']_{b_N} \subseteq [v]_{0\vee(b_N-\varepsilon_N)} \cap [v']_{0\vee(b_N-\varepsilon_N)}. \end{aligned} \tag{2.9.3}$$

To convince the reader, see Figure 2.9.4 below and note that $N^{\varepsilon_N} = 4$.

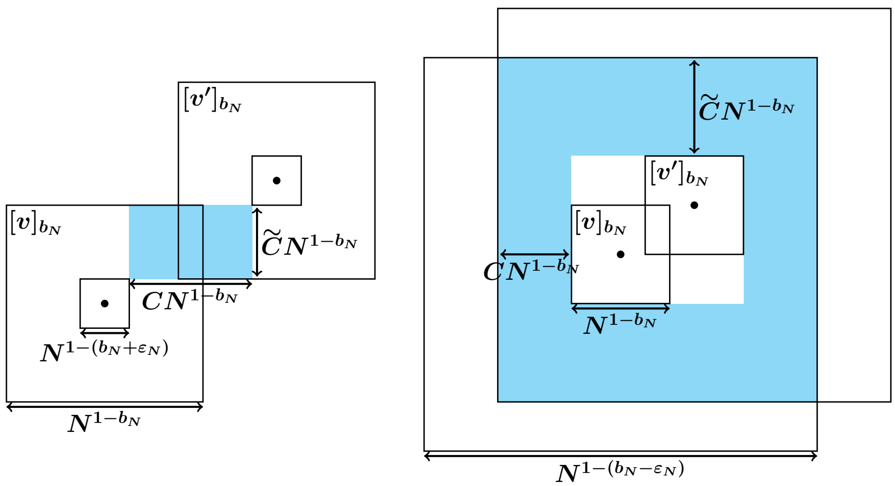


FIGURE 2.9.4. Illustration of Equation (2.9.3).

If $\lambda < \lambda'$ and $\mu < \mu'$, a direct consequence of (2.9.3) and the Markov property of the GFF is the fact that when

$$v \neq v' \quad \text{and} \quad \begin{cases} \quad (1): \lambda, \mu \geq b_N(v,v') + \varepsilon_N, \\ \text{or } (2): \lambda \geq b_N(v,v') + \varepsilon_N > b_N(v,v') - \varepsilon_N \geq \mu', \\ \text{or } (3): b_N(v,v') - \varepsilon_N \geq \lambda \geq \mu' + \varepsilon_N, \end{cases}$$

or

$$v = v' \quad \text{and} \quad \lambda \geq \mu',$$

then

$$\phi_v(\lambda, \lambda') \quad \text{is independent of} \quad \phi_{v'}(\mu, \mu'). \tag{2.9.4}$$

This is because the shell $[v]_\lambda \cap [v]_{\lambda'}^c$ does not intersect the shell $[v']_\mu \cap [v']_{\mu'}^c$ in all cases, see Figure 2.2 in Arguin and Ouimet (2016). The "spacing" ε_N is not optimal but sufficient for our purpose. We stress that, in general, the field ψ does not have the Markov property.

However, by working with increments of the field ψ, the property analogous to (2.9.4) can be proved. The following lemma is a refinement of Lemma A.1 in Arguin and Ouimet (2016), where the error term ε_N is introduced to make the statement hold for all N, not only N large enough.

Lemma 2.9.1. *Let $v, v' \in V_N$, $\lambda < \lambda'$, $\mu < \mu'$ and $\varepsilon_N \circeq (\log 4)/(\log N)$. If*

$$v \neq v' \quad \textit{and} \quad \begin{cases} \quad (1) : \lambda, \mu \geq b_N(v,v') + \varepsilon_N, \\ \textit{or } (2) : \lambda \geq b_N(v,v') + \varepsilon_N > b_N(v,v') - \varepsilon_N \geq \mu', \\ \textit{or } (3) : b_N(v,v') - \varepsilon_N \geq \lambda \geq \mu' + \varepsilon_N, \end{cases}$$

or

$$v = v' \quad \textit{and} \quad \lambda \geq \mu',$$

then

$$\psi_v(\lambda, \lambda') \quad \textit{is independent of} \quad \psi_{v'}(\mu, \mu'). \tag{2.9.5}$$

Proof. Using the tower property of conditional expectations, we have the following decomposition (see (A.4) in Arguin and Ouimet (2016)) :

$$\psi_v(\lambda, \lambda') = \sum_{\substack{1 \leq i \leq M: \\ \lambda \leq \lambda_{i-1} < \lambda' \text{ or } \lambda < \lambda_i \leq \lambda' \\ \text{or } \lambda_{i-1} \leq \lambda < \lambda' \leq \lambda_i}} \sigma_i \, \phi_v(\lambda \vee \lambda_{i-1}, \lambda' \wedge \lambda_i). \tag{2.9.6}$$

The conclusion follows directly from (2.9.4) above. $\square$

The next lemma gives upper and lower bounds on the variance of the increments of the field ψ in $A_{N,\rho}$. Recall from (2.8.19) that

$$A_{N,\rho} \circeq \Big\{ v \in V_N : \min_{z \in \mathbb{Z}^2 \setminus V_N} \|v - z\|_2 \geq N^{1-\rho} \Big\}, \quad \rho \in (0,1]. \tag{2.9.7}$$

Lemma 2.9.2. *Let $\lambda_{i-1} \leq \alpha < \alpha' \leq \lambda_i$ for some $i \in \{1, ..., M\}$, $\alpha \neq 0$ and $\rho \in (0, \alpha]$. Then, for N large enough (dependent on α, but independent from ρ),*

$$\max_{v \in A_{N,\rho}} \Big| \mathbb{E}\Big[\psi_v(\alpha, \alpha')^2\Big] - (\alpha' - \alpha)\sigma_i^2 \log N \Big| \leq C \sigma_i^2. \tag{2.9.8}$$

Proof. This is Lemma A.2 in Arguin and Ouimet (2016) with $v \in A_{N,\rho}$ instead of $v \in V_N^\delta$. The proof is exactly the same and the constant C is independent of α because $\rho \in (0, \alpha]$ implies that the boxes $[v]_\alpha$ are not cut off by ∂V_N. $\square$

The next lemma shows that the upper bound on the variance of the increments in (2.9.8) is in fact uniform on V_N. We extend the statement to include all combinations of scales $\alpha < \alpha'$.

Lemma 2.9.3. *There exists a constant $C_0 = C_0(\boldsymbol{\sigma}) > 0$ such that for all scales $0 \leq \alpha < \alpha' \leq 1$ and N large enough (independent from α and α'),*

$$\max_{v \in V_N} \mathbb{E}\Big[\psi_v(\alpha, \alpha')^2\Big] \leq \mathcal{J}_{\sigma^2}(\alpha, \alpha') \log N + C_0. \tag{2.9.9}$$

Proof. This follows immediately from Lemma A.3 in Arguin and Ouimet (2016) and the independence of the increments. □

In Section 2.8.3, estimates on the covariance of the increments are needed to bound certain overlaps and adapt the Bovier-Kurkova technique. The next two lemmas take care of this problem. To simplify the notation, define

$$\phi_v(A) \mathrel{\hat=} \mathbb{E}\Big[\phi_v \,\big|\, \mathcal{F}_{\partial(A \cap V_N)}\Big], \tag{2.9.10}$$

$$\phi_v(A_1, A_2) \mathrel{\hat=} \phi_v(A_2) - \phi_v(A_1), \tag{2.9.11}$$

for any sets $A, A_1, A_2 \subseteq \mathbb{Z}^2$. With this notation, we can also mix sets and scales with the obvious meaning. For example,

$$\phi_v(A, \lambda) \mathrel{\hat=} \phi_v(\lambda) - \phi_v(A). \tag{2.9.12}$$

For simplicity, we write b_N instead of $b_N(v, v')$ in the remaining of this section.

Lemma 2.9.4. *Let $\lambda_{i-1} \leq \alpha < \alpha' \leq \lambda_i$ for some $i \in \{1, ..., M\}$, $\alpha \neq 0$, $\rho \in (0, \alpha/2]$, and $\varepsilon_N \mathrel{\hat=} (\log 4)/(\log N)$. All four equations below hold for N large enough (dependent on α and α', but independent from ρ and v, v'). All the constants C_i, $1 \leq i \leq 4$, depend only on $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$. For all $v, v' \in A_{N,\rho}$ such that $1 \wedge (\alpha' + 2\varepsilon_N) \leq b_N \leq 1$,*

$$\Big|\mathbb{E}\Big[\psi_v(\alpha, \alpha')\psi_{v'}\Big] - (\alpha' - \alpha)\sigma_i^2 \log N\Big| \leq C_1 \sqrt{\log N}. \tag{2.9.13}$$

For all $v, v' \in A_{N,\rho}$ such that $\alpha' - 2\varepsilon_N \leq b_N \leq 1 \wedge (\alpha' + 2\varepsilon_N)$,

$$\Big|\mathbb{E}\Big[\psi_v(\alpha, \alpha')\psi_{v'}\Big] - (\alpha' - \alpha)\sigma_i^2 \log N\Big| \leq C_2 \sqrt{\log N}. \tag{2.9.14}$$

For all $v, v' \in A_{N,\rho}$ such that $\alpha + 2\varepsilon_N \leq b_N \leq \alpha' - 2\varepsilon_N$,

$$\left|\mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}\Big] - (b_N - \alpha)\,\sigma_i^2 \log N\right| \leq C_3\sqrt{\log N}. \tag{2.9.15}$$

For all $v, v' \in V_N$ such that $b_N \leq \alpha + 2\varepsilon_N$,

$$\left|\mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}\Big]\right| \leq C_4\sqrt{\log N}. \tag{2.9.16}$$

Proof of Equation (2.9.13). Let $v, v' \in A_{N,\rho}$ be such that $1 \wedge (\alpha' + 2\varepsilon_N) \leq b_N \leq 1$. The case $b_N = 1$ (i.e. $v = v'$) is covered by Lemma 2.9.2. Therefore, assume

$$\alpha' + 2\varepsilon_N \leq b_N < 1.$$

From $(1) - (3)$ in Lemma 2.9.1 :

$$\begin{aligned}
&(2): \quad \mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}(1\wedge(b_N+\varepsilon_N),1)\Big] = 0,\\
&(3): \quad \mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}(\alpha'+\varepsilon_N, b_N-\varepsilon_N)\Big] = 0,\\
&(3): \quad \mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}(\alpha-\varepsilon_N)\Big] = 0.
\end{aligned} \tag{2.9.17}$$

Moreover, by the Cauchy-Schwarz inequality and Lemma 2.9.3,

$$\left.\begin{aligned}
&\left|\mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}(b_N-\varepsilon_N,1\wedge(b_N+\varepsilon_N))\Big]\right|\\
&\left|\mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}(\alpha',\alpha'+\varepsilon_N)\Big]\right|\\
&\left|\mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}(\alpha-\varepsilon_N,\alpha)\Big]\right|
\end{aligned}\right\} \leq C\sqrt{\varepsilon_N}\log N. \tag{2.9.18}$$

From the last six equations, it thus suffices to prove

$$\left|\mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}(\alpha,\alpha')\Big] - (\alpha'-\alpha)\sigma_i^2\log N\right| \leq C\sqrt{\log N}. \tag{2.9.19}$$

But, from Definition 2.1.1 and the tower property of conditional expectations, it is easily shown (see (2.9.6)) that when $\lambda_{i-1} \leq \alpha < \alpha' \leq \lambda_i$,

$$\psi_u(\alpha,\alpha') = \sigma_i\phi_u(\alpha,\alpha'), \quad u \in V_N. \tag{2.9.20}$$

Therefore, to show (2.9.19), it suffices to prove

$$\left|\mathbb{E}\Big[\phi_v(\alpha,\alpha')\phi_{v'}(\alpha,\alpha')\Big] - (\alpha'-\alpha)\log N\right| \leq C\sqrt{\log N}. \tag{2.9.21}$$

Since $b_N \geq \alpha' + 2\varepsilon_N$ by hypothesis, we have

$$[v]_\alpha \cup [v']_\alpha \subseteq [v]_{\alpha-\varepsilon_N} \quad \text{and} \quad [v]_{\alpha'} \cup [v']_{\alpha'} \subseteq [v]_{\alpha'-\varepsilon_N}. \tag{2.9.22}$$

From (2.9.22) and Lemma A.5 in Arguin and Ouimet (2016), we deduce

$$\mathbb{E}\Big[\phi_u(\lambda, [v]_{\lambda-\varepsilon_N})^2\Big] \leq C, \quad \text{for all } u \in \{v, v'\}, \ \lambda \in \{\alpha, \alpha'\}. \tag{2.9.23}$$

By combining these four inequalities in (2.9.21) with the Cauchy-Schwarz inequality and Lemma 2.9.3, it suffices to prove

$$\Big|\mathbb{E}\Big[\phi_v([v]_{\alpha-\varepsilon_N}, [v]_{\alpha'-\varepsilon_N})\phi_{v'}([v]_{\alpha-\varepsilon_N}, [v]_{\alpha'-\varepsilon_N})\Big] - (\alpha' - \alpha)\log N\Big| \leq C. \tag{2.9.24}$$

For $u \in \{v, v'\}$, the Markov property (2.9.1) yields

$$\mathbb{E}\Big[\phi_u([v]_{\alpha-\varepsilon_N}, 1) \,|\, \mathcal{F}_{\partial[v]_{\alpha'-\varepsilon_N}}\Big] = \phi_u([v]_{\alpha-\varepsilon_N}, [v]_{\alpha'-\varepsilon_N}). \tag{2.9.25}$$

Using $(\clubsuit)$: $\mathbb{E}[\mathbb{E}[X \,|\, \mathcal{F}]\, \mathbb{E}[Y \,|\, \mathcal{F}]] = \mathbb{E}[XY] - \mathbb{E}[(X - \mathbb{E}[X \,|\, \mathcal{F}])(Y - \mathbb{E}[Y \,|\, \mathcal{F}])]$ together with (2.9.25), we can compute the covariance in (2.9.24) :

$$\begin{aligned}
&\mathbb{E}\Big[\phi_v([v]_{\alpha-\varepsilon_N}, [v]_{\alpha'-\varepsilon_N})\phi_{v'}([v]_{\alpha-\varepsilon_N}, [v]_{\alpha'-\varepsilon_N})\Big] \\
&\overset{(2.9.25)}{=} \mathbb{E}\Big[\mathbb{E}\Big[\phi_v([v]_{\alpha-\varepsilon_N}, 1) \,|\, \mathcal{F}_{\partial[v]_{\alpha'-\varepsilon_N}}\Big]\mathbb{E}\Big[\phi_{v'}([v]_{\alpha-\varepsilon_N}, 1) \,|\, \mathcal{F}_{\partial[v]_{\alpha'-\varepsilon_N}}\Big]\Big] \\
&\overset{(\clubsuit)}{=} \mathbb{E}\Big[\phi_v([v]_{\alpha-\varepsilon_N}, 1)\phi_{v'}([v]_{\alpha-\varepsilon_N}, 1)\Big] - \mathbb{E}\Big[\phi_v([v]_{\alpha'-\varepsilon_N}, 1)\phi_{v'}([v]_{\alpha'-\varepsilon_N}, 1)\Big].
\end{aligned} \tag{2.9.26}$$

But, it is well known that $\{\phi_u(B, 1)\}_{u \in B}$ is a GFF on B when $B \subseteq \mathbb{Z}^2$ is a finite box, see e.g. Zeitouni (2017). Simply choose $B = [v]_{\lambda-\varepsilon_N}$, $\lambda = \alpha, \alpha'$, in (2.9.26), then by the covariance definition in (2.1.1),

$$(2.9.26) = G_{[v]_{\alpha-\varepsilon_N}}(v, v') - G_{[v]_{\alpha'-\varepsilon_N}}(v, v'). \tag{2.9.27}$$

Using standard estimates for the discrete Green function, we can now evaluate the last expression. For every finite box $B \subseteq \mathbb{Z}^2$, Proposition 1.6.3 of Lawler (1991) shows that (keeping in mind our normalization by $\pi/2$ in (2.1.1)) :

$$G_B(x, y) = \left[\sum_{z \in \partial B} \mathscr{P}_x(W_{\tau_{\partial B}} = z)\, a(z - y)\right] - a(y - x), \quad x, y \in B, \tag{2.9.28}$$

where

$$a(w) = \begin{cases} \log(\|w\|_2) + \text{const.} + O(\|w\|_2^{-2}), & \text{if } w \in \mathbb{Z}^2 \backslash \{\mathbf{0}\}, \\ 0, & \text{if } w = \mathbf{0}, \end{cases} \tag{2.9.29}$$

and $\mathscr{P}_x$ is the law of the simple random walk starting at $x \in \mathbb{Z}^2$. Using (2.9.28), we can rewrite the difference of Green functions in (2.9.27) as

$$\sum_{z \in \partial[v]_{\alpha - \varepsilon_N}} \mathscr{P}_v\Big(W_{\tau_{\partial[v]_{\alpha-\varepsilon_N}}} = z\Big) a(z - v') - \sum_{z \in \partial[v]_{\alpha' - \varepsilon_N}} \mathscr{P}_v\Big(W_{\tau_{\partial[v]_{\alpha'-\varepsilon_N}}} = z\Big) a(z - v'). \tag{2.9.30}$$

Since $\rho \leq \alpha/2 < \alpha - \varepsilon_N < \alpha' - \varepsilon_N$ by hypothesis, the boxes $[v]_{\alpha-\varepsilon_N}$ and $[v]_{\alpha'-\varepsilon_N}$ are not cut off by ∂V_N for N large enough. Furthermore, $\alpha' \leq b_N$ implies that $\|v - v'\|_\infty \leq N^{1-\alpha'}$, so it is easily seen that $N^{1-\lambda} \leq \|z - v'\|_2 \leq 4\sqrt{2} N^{1-\lambda}$ for all $z \in \partial[v]_{\lambda - \varepsilon_N}$ and $\lambda \in \{\alpha, \alpha'\}$. Then, (2.9.24) follows immediately by using (2.9.29) in (2.9.30). This proves (2.9.13). $\square$

Proof of Equation (2.9.14). Let $v, v' \in A_{N,\rho}$ be such that

$$\alpha' - 2\varepsilon_N \leq b_N \leq 1 \wedge (\alpha' + 2\varepsilon_N). \tag{2.9.31}$$

Define $\widetilde{\alpha}' \triangleq \alpha' - 4\varepsilon_N$. For N large enough (independent from v, v' and ρ), we have $\lambda_{i-1} \leq \alpha < \widetilde{\alpha}' < \alpha' \leq \lambda_i$ and $1 \wedge (\widetilde{\alpha}' + 2\varepsilon_N) \leq b_N \leq 1$. From Equation (2.9.13),

$$\Big|\mathbb{E}\Big[\psi_v(\alpha, \widetilde{\alpha}')\psi_{v'}\Big] - (\alpha' - \alpha - 4\varepsilon_N)\sigma_i^2 \log N\Big| \leq C_1 \sqrt{\log N}, \tag{2.9.32}$$

and from the Cauchy-Schwarz inequality and Lemma 2.9.3,

$$\Big|\mathbb{E}\Big[\psi_v(\widetilde{\alpha}', \alpha')\psi_{v'}\Big]\Big| \leq C\sqrt{\varepsilon_N} \log N. \tag{2.9.33}$$

This proves Equation (2.9.14). $\square$

Proof of Equation (2.9.15). Let $v, v' \in A_{N,\rho}$ be such that $\alpha + 2\varepsilon_N \leq b_N \leq \alpha' - 2\varepsilon_N$. From $(1) - (3)$ in Lemma 2.9.1 :

$$\begin{aligned}
(1): &\quad \mathbb{E}\Big[\psi_v(b_N + \varepsilon_N, \alpha')\psi_{v'}(b_N + \varepsilon_N, 1)\Big] = 0, \\
(2): &\quad \mathbb{E}\Big[\psi_v(\alpha, b_N - \varepsilon_N)\psi_{v'}(b_N + \varepsilon_N, 1)\Big] = 0, \\
(2): &\quad \mathbb{E}\Big[\psi_v(b_N + \varepsilon_N, \alpha')\psi_{v'}(\alpha, b_N - \varepsilon_N)\Big] = 0, \\
(2): &\quad \mathbb{E}\Big[\psi_v(b_N + \varepsilon_N, \alpha')\psi_{v'}(\alpha - \varepsilon_N)\Big] = 0, \\
(3): &\quad \mathbb{E}\Big[\psi_v(\alpha, b_N - \varepsilon_N)\psi_{v'}(\alpha - \varepsilon_N)\Big] = 0.
\end{aligned} \tag{2.9.34}$$

Moreover, by the Cauchy-Schwarz inequality and Lemma 2.9.3,

$$\left.\begin{array}{l}\left|\mathbb{E}\Big[\psi_v(b_N-\varepsilon_N,b_N+\varepsilon_N)\psi_{v'}(b_N+\varepsilon_N,1)\Big]\right|\\ \left|\mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}(b_N,b_N+\varepsilon_N)\Big]\right|\\ \left|\mathbb{E}\Big[\psi_v(b_N,\alpha')\psi_{v'}(b_N-\varepsilon_N,b_N)\Big]\right|\\ \left|\mathbb{E}\Big[\psi_v(b_N,b_N+\varepsilon_N)\psi_{v'}(\alpha,b_N-\varepsilon_N)\Big]\right|\\ \left|\mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}(\alpha-\varepsilon_N,\alpha)\Big]\right|\\ \left|\mathbb{E}\Big[\psi_v(b_N-\varepsilon_N,b_N+\varepsilon_N)\psi_{v'}(\alpha-\varepsilon_N)\Big]\right|\end{array}\right\}\leq C\sqrt{\varepsilon_N}\log N. \tag{2.9.35}$$

From the last eleven equations, it thus suffices to prove

$$\left|\mathbb{E}\Big[\psi_v(\alpha,b_N)\psi_{v'}(\alpha,b_N)\Big]-(b_N-\alpha)\,\sigma_i^2\log N\right|\leq C\sqrt{\log N}. \tag{2.9.36}$$

The conclusion follows from the exact same argument used after (2.9.19) in the proof of Equation (2.9.13), with b_N replacing α' everywhere. □

Proof of Equation (2.9.16). Let $v,v'\in V_N$ be such that $b_N\leq\alpha+2\varepsilon_N\leq\alpha'-2\varepsilon_N$. From $(1)-(3)$ in Lemma 2.9.1 :

$$\begin{aligned}(1):&\quad \mathbb{E}\Big[\psi_v(\alpha+3\varepsilon_N,\alpha')\psi_{v'}(\alpha+3\varepsilon_N,1)\Big]=0,\\ (1):&\quad \mathbb{E}\Big[\psi_v(\alpha+3\varepsilon_N,\alpha')\psi_{v'}(\alpha\wedge(b_N+\varepsilon_N),\alpha)\Big]=0,\\ (2):&\quad \mathbb{E}\Big[\psi_v(\alpha+3\varepsilon_N,\alpha')\psi_{v'}(\alpha\wedge(0\vee(b_N-\varepsilon_N)))\Big]=0.\end{aligned} \tag{2.9.37}$$

Moreover, by the Cauchy-Schwarz inequality and Lemma 2.9.3,

$$\left.\begin{array}{l}\left|\mathbb{E}\Big[\psi_v(\alpha+3\varepsilon_N,\alpha')\psi_{v'}(\alpha,\alpha+3\varepsilon_N)\Big]\right|\\ \left|\mathbb{E}\Big[\psi_v(\alpha+3\varepsilon_N,\alpha')\psi_{v'}(\alpha\wedge(0\vee(b_N-\varepsilon_N)),\alpha\wedge(b_N+\varepsilon_N))\Big]\right|\\ \left|\mathbb{E}\Big[\psi_v(\alpha,\alpha+3\varepsilon_N)\psi_{v'}\Big]\right|\end{array}\right\}\leq C\sqrt{\varepsilon_N}\log N.$$

The last six equations together yield Equation (2.9.16). □

We summarize the results of the previous lemma and extend the statement to include all combinations of scales $\alpha<\alpha'$ and all $\rho\in(0,1]$.

Lemma 2.9.5. *Let $0\leq\alpha<\alpha'\leq 1$ and let $\rho\in(0,1]$. Then, for N large enough (dependent on α and α', but independent from ρ (except when $\alpha=0$)),*

$$\begin{aligned}\max_{v,v'\in A_{N,\rho}}&\left|\mathbb{E}\Big[\psi_v(\alpha,\alpha')\psi_{v'}\Big]-\mathcal{J}_{\sigma^2}(\alpha\wedge b_N,\alpha'\wedge b_N)\log N\right|\\ &\leq C_5(\boldsymbol{\sigma},\boldsymbol{\lambda})\sqrt{\log N}+C_6(\alpha,\alpha',\boldsymbol{\sigma},\boldsymbol{\lambda})\,\rho\log N.\end{aligned} \tag{2.9.38}$$

PROOF. If $\alpha \neq 0$ and $\rho \leq \alpha/2$, then write the decomposition from (2.9.6),

$$\psi_v(\alpha, \alpha') = \sum_{\substack{1 \leq i \leq M: \\ \alpha \leq \lambda_{i-1} < \alpha' \text{ or } \alpha < \lambda_i \leq \alpha' \\ \text{or } \lambda_{i-1} \leq \alpha < \alpha' \leq \lambda_i}} \psi_v(\alpha \vee \lambda_{i-1}, \alpha' \wedge \lambda_i), \tag{2.9.39}$$

and apply Lemma 2.9.4 to each increment ($C_6 = 0$). If $\alpha \neq 0$ and $\rho > \alpha/2$, or if $\alpha = 0$ and $\rho \geq \alpha'/2$, then simply choose C_6 big enough (depending on α or α') that (2.9.38) is satisfied. This is always possible since $\mathcal{J}_{\sigma^2}(\cdot, \cdot)$ is bounded and since

$$\max_{v,v' \in V_N} \frac{\left|\mathbb{E}\left[\psi_v(\alpha, \alpha')\psi_{v'}\right]\right|}{\log N} \leq C, \tag{2.9.40}$$

by Lemma 2.9.3. Finally, if $\alpha = 0$ and $\rho < \alpha'/2$, then define $\widetilde{\alpha} \mathrel{\hat{=}} 2\rho$ and apply (2.9.38) in the first case ($0 \neq \widetilde{\alpha} < \alpha'$ and $\rho \leq \widetilde{\alpha}/2$), we have

$$\max_{v,v' \in A_{N,\rho}} \left|\mathbb{E}\left[\psi_v(\widetilde{\alpha}, \alpha')\psi_{v'}\right] - \mathcal{J}_{\sigma^2}(\widetilde{\alpha} \wedge b_N, \alpha' \wedge b_N) \log N\right| \leq C_5(\boldsymbol{\sigma}, \boldsymbol{\lambda})\sqrt{\log N}. \tag{2.9.41}$$

On the other hand, if we "cut" the increments with small covariance contributions like we did multiple times in the proof of the previous lemma (using Lemma 2.9.1, Lemma 2.9.3 and the Cauchy-Schwarz inequality), then

$$\begin{aligned}
\max_{v,v' \in V_N} \left|\mathbb{E}\left[\psi_v(\widetilde{\alpha})\psi_{v'}\right]\right| &\leq \max_{v,v' \in V_N} \left|\mathbb{E}\left[\psi_v(\widetilde{\alpha} \wedge b_N)\psi_{v'}(\widetilde{\alpha} \wedge b_N)\right]\right| + C\sqrt{\varepsilon_N} \log N \\
&\leq \widetilde{C}\,(\widetilde{\alpha} \wedge b_N) \log N + C_0 + C\sqrt{\varepsilon_N} \log N \\
&\leq \widetilde{C}\,\rho \log N + C\sqrt{\log N}.
\end{aligned} \tag{2.9.42}$$

Combining (2.9.41) and (2.9.42) proves (2.9.38) in the last case. $\square$

The following corollary gives estimates on the increments of overlaps. For convenience, we recall their definition from (2.8.42) :

$$q^N_{\alpha,\alpha'}(v, v') \mathrel{\hat{=}} \frac{\mathbb{E}\left[\psi_v(\alpha, \alpha')\psi_{v'}\right]}{\mathcal{J}_{\sigma^2}(1) \log N + C_0}, \quad v, v' \in V_N, \tag{2.9.43}$$

where C_0 is the constant introduced in Lemma 2.9.3. The estimates are crucial in Section 2.8.3 to adapt the Bovier-Kurkova technique and prove Proposition 2.8.8.

Corollary 2.9.6. *Let $0 \leq \alpha < \alpha' \leq 1$ and let $\rho \in (0,1]$. Then, for N large enough (dependent on α and α', but independent from ρ (except when $\alpha = 0$)),*

$$\max_{v,v' \in A_{N,\rho}} \left| q^N_{\alpha,\alpha'}(v,v') - \bar{\mathcal{J}}_{\sigma^2}(\alpha \wedge b_N, \alpha' \wedge b_N) \right| \leq \frac{C_7(\boldsymbol{\sigma},\boldsymbol{\lambda})}{\sqrt{\log N}} + C_8(\alpha,\alpha',\boldsymbol{\sigma},\boldsymbol{\lambda})\, \rho. \tag{2.9.44}$$

2.9.2. Technical lemmas

Lemma 2.9.7. *The function $\mathcal{E} : [0,\gamma^\star] \to \mathbb{R}$ defined in (2.5.6) is in $C^1([0,\gamma^\star])$.*

PROOF. The function $\mathcal{E}$ is clearly continuously differentiable at $\gamma \in [0,\gamma^\star] \backslash \{\gamma^l\}_{l=0}^m$. Furthermore, for $0 < h < \gamma^1$,

$$\lim_{h\to 0^+} \frac{\mathcal{E}(h) - \mathcal{E}(0)}{h} = \lim_{h\to 0^+} \frac{-h}{\mathcal{J}_{\sigma^2}(1)} = 0 = \lim_{h\to 0^+} \frac{-2h}{\mathcal{J}_{\sigma^2}(1)} = \lim_{h\to 0^+} \mathcal{E}'(h). \tag{2.9.45}$$

Therefore, $\mathcal{E}$ is continuously differentiable at $\gamma = 0 \mathrel{\hat{=}} \gamma^0$ (from the right).

For $\gamma = \gamma^\star$, we can write

$$\gamma^\star = \gamma^m = \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{m-1}) + \frac{\mathcal{J}_{\sigma^2}(\lambda^{m-1},1)}{\bar{\sigma}_m}, \tag{2.9.46}$$

where $\mathcal{J}_{\sigma^2}(\lambda^{m-1},1) = \bar{\sigma}_m^2 \nabla\lambda^m$. Thus, for $0 < h < \nabla\gamma^m$,

$$\lim_{h\to 0^+} \frac{\mathcal{E}(\gamma^\star - h) - \mathcal{E}(\gamma^\star)}{-h} = \lim_{h\to 0^+} \frac{-1}{h}\left[\nabla\lambda^m - \frac{(\bar{\sigma}_m\nabla\lambda^m - h)^2}{\bar{\sigma}_m^2\nabla\lambda^m}\right] = \frac{-2}{\bar{\sigma}_m}, \tag{2.9.47}$$

and

$$\lim_{h\to 0^+} \mathcal{E}'(\gamma^\star - h) = \lim_{h\to 0^+} \frac{-2(\bar{\sigma}_m\nabla\lambda^m - h)}{\bar{\sigma}_m^2\nabla\lambda^m} = \frac{-2}{\bar{\sigma}_m}. \tag{2.9.48}$$

Therefore, $\mathcal{E}$ is continuously differentiable at $\gamma = \gamma^\star$ (from the left).

For the remaining points $\gamma = \gamma^l$, fix $l \in \{1,...,m-1\}$. The critical level γ^l from (2.5.5) can be expressed in two ways :

$$\gamma^l = \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}) + \frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}{\bar{\sigma}_l} \tag{2.9.49}$$

$$= \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l}) + \frac{\mathcal{J}_{\sigma^2}(\lambda^{l},1)}{\bar{\sigma}_l}. \tag{2.9.50}$$

Also, note that

$$\mathcal{E}(\gamma^l) \overset{(2.9.49)}{=} (1-\lambda^{l-1}) - \frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}{\bar{\sigma}_l^2} \tag{2.9.51}$$

$$= \quad (1-\lambda^{l}) - \frac{\mathcal{J}_{\sigma^2}(\lambda^{l},1)}{\bar{\sigma}_l^2}, \tag{2.9.52}$$

where the last equality follows from $\mathcal{J}_{\sigma^2}(\lambda^{l-1},\lambda^l)=\bar{\sigma}_l^2\nabla\lambda^l$. For $0<h<\min_j\nabla\gamma^j$,

$$\lim_{h\to 0^+}\frac{\mathcal{E}((2.9.49)-h)-\mathcal{E}((2.9.49))}{-h}\overset{(2.9.51)}{=}\lim_{h\to 0^+}\frac{+h}{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}+\frac{-2}{\bar{\sigma}_l}=\frac{-2}{\bar{\sigma}_l},\tag{2.9.53}$$

$$\lim_{h\to 0^+}\frac{\mathcal{E}((2.9.50)+h)-\mathcal{E}((2.9.50))}{h}\overset{(2.9.52)}{=}\lim_{h\to 0^+}\frac{-h}{\mathcal{J}_{\sigma^2}(\lambda^{l},1)}+\frac{-2}{\bar{\sigma}_l}=\frac{-2}{\bar{\sigma}_l}.\tag{2.9.54}$$

and

$$\lim_{h\to 0^+}\mathcal{E}'(\gamma^l-h)\overset{(2.9.49)}{=}\lim_{h\to 0^+}\frac{-2(\frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}{\bar{\sigma}_l}-h)}{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}=\frac{-2}{\bar{\sigma}_l}\tag{2.9.55}$$

$$\lim_{h\to 0^+}\mathcal{E}'(\gamma^l+h)\overset{(2.9.50)}{=}\lim_{h\to 0^+}\frac{-2(\frac{\mathcal{J}_{\sigma^2}(\lambda^{l},1)}{\bar{\sigma}_l}+h)}{\mathcal{J}_{\sigma^2}(\lambda^{l},1)}=\frac{-2}{\bar{\sigma}_l}.\tag{2.9.56}$$

Hence, $\mathcal{E}$ is continuously differentiable at $\gamma=\gamma^l$, for all $l\in\{1,...,m-1\}$. $\square$

Lemma 2.9.8. *Let $\beta>0$. Define $P_\beta(\gamma)\doteq\beta\gamma+\mathcal{E}(\gamma)$, and recall*

$$l_\beta\doteq\begin{cases}\min\{l\in\{1,...,m\}:\beta\le\beta_c(\bar{\sigma}_l)\doteq 2/\bar{\sigma}_l\}, & \text{if } \beta\le 2/\bar{\sigma}_m,\\ m+1, & \text{otherwise},\end{cases}\tag{2.9.57}$$

from (2.6.14). *Then,*

$$\max_{\gamma\in[0,\gamma^\star]}P_\beta(\gamma)=\sum_{j=1}^{l_\beta-1}\left\{2\frac{\beta}{(2/\bar{\sigma}_j)}\right\}\nabla\lambda^j+\sum_{j=l_\beta}^{m}\left\{1+\frac{\beta^2}{(2/\bar{\sigma}_j)^2}\right\}\nabla\lambda^j\doteq f^\psi(\beta).\tag{2.9.58}$$

Proof. We consider three cases :

$$(1)\ \ l_\beta=m+1;\qquad (2)\ \ l_\beta=1;\qquad (3)\ \ l_\beta\in\{2,...,m\}.$$

Since $\bar{\sigma}_1>\bar{\sigma}_2>...>\bar{\sigma}_m$, these three cases imply (respectively) :

(i) $\beta>2/\bar{\sigma}_j$ for all $j\in\{1,...,m\}$;
(ii) $\beta\le 2/\bar{\sigma}_j$ for all $j\in\{1,...,m\}$;
(iii) $\beta\in(2/\bar{\sigma}_{l_\beta-1},2/\bar{\sigma}_{l_\beta}]$.

Case (1) : For any $\gamma\in(\gamma^{l-1},\gamma^l]\backslash\{\gamma^\star\}$, we have

$$P_\beta'(\gamma)=\beta-2\frac{(\gamma-\mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}))}{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}.\tag{2.9.59}$$

Any solution to $P_\beta'(\gamma)=0$ must satisfy

$$\gamma=\mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1})+\frac{\beta}{2}\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)\overset{(i)}{>}\mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1})+\frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}{\bar{\sigma}_l}\overset{(2.9.49)}{=}\gamma^l,\tag{2.9.60}$$

which is impossible. Therefore, the maximum $\max_{\gamma\in[0,\gamma^\star]} P_\beta(\gamma)$ must be achieved at the boundary of $[0,\gamma^\star]$. We have

$$P_\beta(\gamma^\star) \doteq \beta\gamma^\star + 0 = \sum_{j=1}^{m} \left\{ 2\frac{\beta}{(2/\bar{\sigma}_j)} \right\} \nabla\lambda^j \overset{(i)}{>} 2 > \beta\cdot 0 + 1 \doteq P_\beta(0), \tag{2.9.61}$$

which proves (2.9.58) when $l_\beta = m+1$.

Case (2) : From (2.9.59), any solution $\gamma \in (\gamma^{l-1},\gamma^l]\backslash\{\gamma^\star\}$ to $P'_\beta(\gamma) = 0$ must satisfy

$$\gamma = \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}) + \frac{\beta}{2}\mathcal{J}_{\sigma^2}(\lambda^{l-1},1) \quad \text{and} \quad l = 1, \tag{2.9.62}$$

because $l \geq 2$ and the restriction (ii) would otherwise imply $\gamma \leq \gamma^{l-1}$, from (2.9.50). In other words, the maximum $\max_{\gamma\in[0,\gamma^\star]} P_\beta(\gamma)$ must be achieved at the boundary of $[0,\gamma^\star]$ or at $\bar{\gamma} \doteq \frac{\beta}{2}\mathcal{J}_{\sigma^2}(1) \in (0,\gamma^1]$. Since $\beta > 0$, we have

$$P_\beta(\bar{\gamma}) = \frac{\beta^2}{2}\mathcal{J}_{\sigma^2}(1) + 1 - \frac{\beta^2}{4}\mathcal{J}_{\sigma^2}(1) = 1 + \frac{\beta^2}{4}\mathcal{J}_{\sigma^2}(1) > 1 = P_\beta(0), \tag{2.9.63}$$

and the identity $1 + x^2 \geq 2x$ yields

$$P_\beta(\bar{\gamma}) = \sum_{j=1}^{m} \left\{ 1 + \frac{\beta^2}{(2/\bar{\sigma}_j)^2} \right\} \nabla\lambda^j \geq \sum_{j=1}^{m} \left\{ 2\frac{\beta}{(2/\bar{\sigma}_j)} \right\} \nabla\lambda^j = P_\beta(\gamma^\star). \tag{2.9.64}$$

This proves (2.9.58) when $l_\beta = 1$.

Case (3) : From (2.9.59), any solution $\gamma \in (\gamma^{l-1},\gamma^l]\backslash\{\gamma^\star\}$ to $P'_\beta(\gamma) = 0$ must satisfy

$$\gamma = \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}) + \frac{\beta}{2}\mathcal{J}_{\sigma^2}(\lambda^{l-1},1) \quad \text{and} \quad l = l_\beta. \tag{2.9.65}$$

We must have the restriction $l = l_\beta$ since $\gamma \in (\gamma^{l-1},\gamma^l]$ and $\beta \in (2/\bar{\sigma}_{l_\beta-1}, 2/\bar{\sigma}_{l_\beta}]$ from (iii) imply

$$\begin{aligned} &\left\{ \begin{array}{l} \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}) + \frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}{\bar{\sigma}_{l-1}} \overset{(2.9.50)}{=} \gamma^{l-1} < \gamma \overset{(iii)}{\leq} \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}) + \frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}{\bar{\sigma}_{l_\beta}} \\ \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}) + \frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}{\bar{\sigma}_{l_\beta-1}} \overset{(iii)}{<} \gamma \leq \gamma^l \overset{(2.9.49)}{=} \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l-1}) + \frac{\mathcal{J}_{\sigma^2}(\lambda^{l-1},1)}{\bar{\sigma}_l} \end{array} \right\} \\ \Longrightarrow\ & \left\{ \begin{array}{l} \bar{\sigma}_{l_\beta} < \bar{\sigma}_{l-1} \\ \bar{\sigma}_l < \bar{\sigma}_{l_\beta-1} \end{array} \right\} \\ \Longrightarrow\ & \{\ l = l_\beta\ \}, \end{aligned} \tag{2.9.66}$$

where the last implication holds because $\bar{\sigma}_1 > \bar{\sigma}_2 > \ldots > \bar{\sigma}_m$.

When we evaluate P_β at $\bar{\gamma} \mathrel{\mathring{=}} \mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l_\beta-1}) + \frac{\beta}{2}\mathcal{J}_{\sigma^2}(\lambda^{l_\beta-1},1)$, we get

$$
\begin{aligned}
P_\beta(\bar{\gamma}) &= \beta\mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l_\beta-1}) + \frac{\beta^2}{2}\mathcal{J}_{\sigma^2}(\lambda^{l_\beta-1},1) + (1-\lambda^{l_\beta-1}) - \frac{\beta^2}{4}\mathcal{J}_{\sigma^2}(\lambda^{l_\beta-1},1) \\
&= \beta\mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^{l_\beta-1}) + \left\{(1-\lambda^{l_\beta-1}) + \frac{\beta^2}{4}\mathcal{J}_{\sigma^2}(\lambda^{l_\beta-1},1)\right\} \\
&= \sum_{j=1}^{l_\beta-1}\left\{2\frac{\beta}{(2/\bar{\sigma}_j)}\right\}\left[\frac{\nabla\mathcal{J}_{\sigma^2/\bar{\sigma}}(\lambda^j)}{\bar{\sigma}_j\nabla\lambda^j}\right]\nabla\lambda^j + \sum_{j=l_\beta}^{m}\left\{1 + \frac{\beta^2}{(2/\bar{\sigma}_j)^2}\left[\frac{\nabla\mathcal{J}_{\sigma^2}(\lambda^j)}{\bar{\sigma}_j^2\nabla\lambda^j}\right]\right\}\nabla\lambda^j \\
&= \sum_{j=1}^{l_\beta-1}\left\{2\frac{\beta}{(2/\bar{\sigma}_j)}\right\}\nabla\lambda^j + \sum_{j=l_\beta}^{m}\left\{1 + \frac{\beta^2}{(2/\bar{\sigma}_j)^2}\right\}\nabla\lambda^j. \qquad (2.9.67)
\end{aligned}
$$

The last equality holds because the pairs of brackets $[\,\cdot\,]$ on the second and third to last line are equal to 1. Since $\beta > 2/\bar{\sigma}_{l_\beta-1} > 0$ by (iii), we have

$$
P_\beta(\bar{\gamma}) \overset{(2.9.67)}{>} \sum_{j=1}^{l_\beta-1}\{2\}\nabla\lambda^j + \sum_{j=l_\beta}^{m}\{1\}\nabla\lambda^j > 1 = P_\beta(0), \qquad (2.9.68)
$$

and the identity $1 + x^2 \geq 2x$ yields

$$
P_\beta(\bar{\gamma}) \overset{(2.9.67)}{\geq} \sum_{j=1}^{m}\left\{2\frac{\beta}{(2/\bar{\sigma}_j)}\right\}\nabla\lambda^j = P_\beta(\gamma^\star). \qquad (2.9.69)
$$

This proves (2.9.58) when $l_\beta \in \{2,...,m\}$, and end the proof of Lemma 2.9.8. $\square$

We recall the definition of the perturbed field ψ^u. Let $\lambda_{i^\star-1} \leq \alpha < \alpha' \leq \lambda_{i^\star}$ for a given $i^\star \in \{1,...,M\}$, and let $u > -\sigma_{i^\star}$. Then,

$$
\psi_v^u \mathrel{\mathring{=}} u\,\phi_v(\alpha,\alpha') + \psi_v, \quad v \in V_N. \qquad (2.9.70)
$$

Lemma 2.9.9. *Let $\beta > 0$, $\rho \in (0,1]$ and let $\lambda_{i^\star-1} \leq \alpha < \alpha' \leq \lambda_{i^\star}$ for some $i^\star$. Then, $u \mapsto f_{N,\rho}^{\psi^u}(\beta)$ is almost-surely convex and $u \mapsto \mathbb{E}\big[f_{N,\rho}^{\psi^u}(\beta)\big]$ is convex.*

Proof. By definition, we have

$$
F(u) \mathrel{\mathring{=}} f_{N,\rho}^{\psi^u}(\beta) = \frac{1}{\log N^2}\log\left(\int_{A_{N,\rho}}(g(v))^u d\mu(v)\right), \qquad (2.9.71)
$$

where $g(v) \mathrel{\mathring{=}} \exp(\beta\phi_v(\alpha,\alpha'))$ and $\mu(A) \mathrel{\mathring{=}} \sum_{v\in A}\exp(\beta\psi_v)$ for any $A \subseteq V_N$. By standard properties of logarithms, we see that $u \mapsto F(u)$ is convex almost-surely since, for all $\lambda \in [0,1]$ and all $u, u' > -\sigma_{i^\star}$, we have

$$
\begin{aligned}
&F(\lambda u+(1-\lambda)u') \le \lambda F(u)+(1-\lambda)F(u') \\
&\qquad \Longleftrightarrow \quad \int_{A_{N,\rho}} (g(v))^{\lambda u}(g(v))^{(1-\lambda)u'} d\mu(v) \\
&\qquad\qquad\qquad \le \left(\int_{A_{N,\rho}} (g(v))^{u} d\mu(v)\right)^{\lambda} \left(\int_{A_{N,\rho}} (g(v))^{u'} d\mu(v)\right)^{1-\lambda}, \qquad (2.9.72)
\end{aligned}
$$

and the last inequality is true by Holder's inequality ($p \mathrel{\hat{=}} 1/\lambda$, $q \mathrel{\hat{=}} 1/(1-\lambda)$ and $1/p + 1/q = 1$). The fact that $u \mapsto \mathbb{E}[F(u)]$ is also convex follows immediately from the linearity and monotonicity of expectations. $\square$

The parameters of ψ^u can be encoded simultaneously in the left-continuous step function

$$
\sigma_u(r) \mathrel{\hat{=}} \begin{cases} \sigma(r), & \text{for all } r \in [0,1]\backslash(\alpha,\alpha'], \\ \sigma_{i^\star} + u, & \text{for all } r \in (\alpha,\alpha']. \end{cases} \qquad (2.9.73)
$$

Since $\mathcal{J}_{\sigma_u^2}(\cdot)$ is an increasing polygonal line, there exists a unique non-increasing left-continuous step function $r \mapsto \bar{\sigma}_u(r)$ such that the *concavification* of $\mathcal{J}_{\sigma_u^2}$ can be expressed as the integral of $r \mapsto \bar{\sigma}_u^2(r)$:

$$
\hat{\mathcal{J}}_{\sigma_u^2}(s) = \mathcal{J}_{\bar{\sigma}_u^2}(s) = \int_0^s \bar{\sigma}_u^2(r)\, dr \quad \text{for all } s \in (0,1]. \qquad (2.9.74)
$$

As for the field ψ,

- $\bar{\sigma}_{u,j},\ 1 \le j \le m_u$, denote the heights of the steps of $r \mapsto \bar{\sigma}_u(r)$,
- m_u denotes the number of steps,
- λ_u^j denote the scales at which $r \mapsto \bar{\sigma}_u(r)$ jumps.

Recall l_β from (2.9.57) and define the analogue for ψ^u :

$$
l_{\beta,u} \mathrel{\hat{=}} \begin{cases} \min\{l \in \{1,\dots,m_u\} : \beta \le \beta_c(\bar{\sigma}_{u,l}) \mathrel{\hat{=}} 2/\bar{\sigma}_{u,l}\}, & \text{if } \beta \le 2/\bar{\sigma}_{u,m_u}, \\ m_u + 1, & \text{otherwise.} \end{cases} \qquad (2.9.75)
$$

The following lemma studies the differentiability of the limiting free energy of ψ^u with respect to the perturbation parameter u.

Lemma 2.9.10. *Let $\beta > 0$ and let $\lambda^{j^\star-1} \le \lambda_{i^\star-1} \le \alpha < \alpha' \le \lambda_{i^\star} \le \lambda^{j^\star}$ for some $i^\star, j^\star$. There exists $\delta = \delta(\beta, \alpha, \alpha', \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ such that $u \mapsto f^{\psi^u}(\beta)$ is differentiable on $(-\delta, \delta)$. The derivative at $u = 0$ is given by*

$$\frac{\partial}{\partial u} f^{\psi^0}(\beta) = \begin{cases} \frac{\beta \sigma_{i^\star}(\alpha'-\alpha)}{\bar{\sigma}_{j^\star}}, & \text{if } j^\star \le l_\beta - 1, \\ \frac{\beta^2 \sigma_{i^\star}(\alpha'-\alpha)}{2}, & \text{if } j^\star \ge l_\beta. \end{cases} \tag{2.9.76}$$

Proof. We separate the proof in two cases :

Case (i) : $\mathcal{J}_{\sigma^2}(r) < \mathcal{J}_{\bar{\sigma}^2}(r)$ for all $r \in (\lambda^{j^\star-1}, \lambda^{j^\star})$;

Case (ii) : $\exists r \in (\lambda^{j^\star-1}, \lambda^{j^\star})$ such that $\mathcal{J}_{\sigma^2}(r) = \mathcal{J}_{\bar{\sigma}^2}(r)$.

Case (i) : The function $u \mapsto \bar{\sigma}_u(r)$ is continuous, uniformly in $r \in [0,1]$. Hence, we can choose $\delta = \delta(\alpha, \alpha', \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ small enough that for all $u \in (-\delta, \delta)$:

- $\bar{\sigma}_{u,j} = \bar{\sigma}_j$ for all $j \ne j^\star$;
- $\lambda_u^j = \lambda^j$ for all $j \in \{1, ..., m_u\}$;
- $m_u = m$.

Figure 2.9.5 below illustrates this point more clearly.

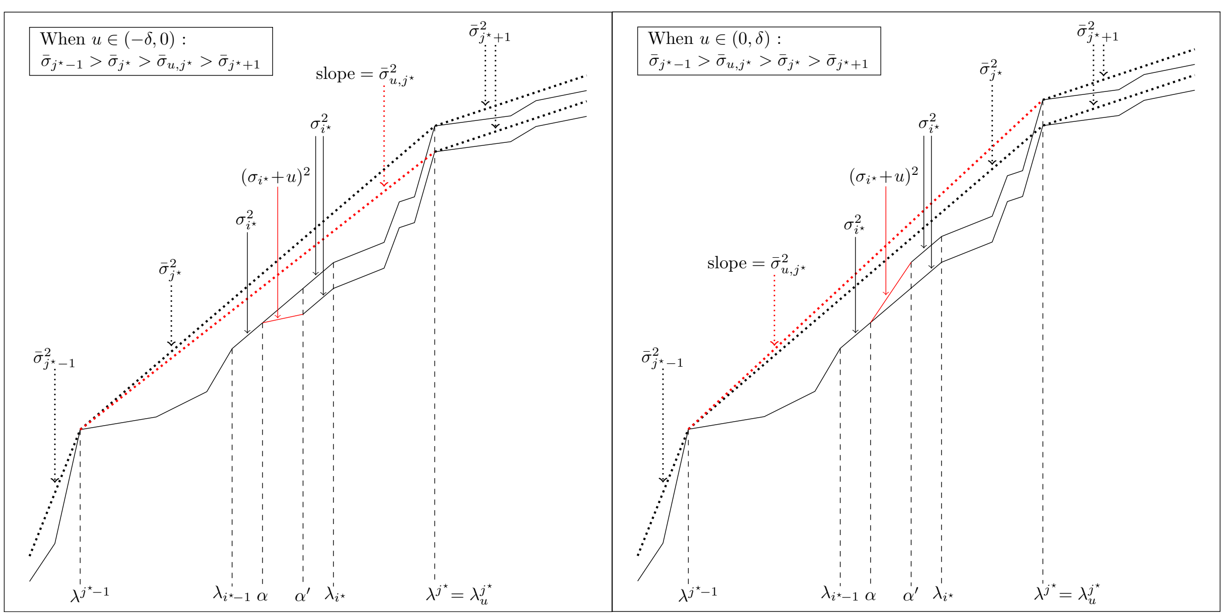


Figure 2.9.5. The dotted paths represent $\mathcal{J}_{\bar{\sigma}^2}$ and $\mathcal{J}_{\bar{\sigma}_u^2}$. The closed paths represent $\mathcal{J}_{\sigma^2}$ and $\mathcal{J}_{\sigma_u^2}$. The paths containing a red part are the ones for the perturbed field ψ^u.

Note that $l_{\beta,0} = l_\beta$ and also $2/\bar{\sigma}_{l_\beta - 1} < \beta \le 2/\bar{\sigma}_{l_\beta}$ ($\beta > 2/\bar{\sigma}_m$ when $l_\beta = m+1$). We can choose $\delta = \delta(\beta, \alpha, \alpha', \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ small enough that

$$\text{when } \beta \ne 2/\bar{\sigma}_{j^\star} (\Longrightarrow\ 1 \le j^\star \le m), \quad \text{then } l_{\beta,u} = l_\beta \text{ for all } u \in (-\delta, \delta), \tag{2.9.77}$$

$$\text{when } \beta = 2/\bar{\sigma}_{j^\star} (\Longrightarrow\ j^\star = l_\beta), \quad \text{then } l_{\beta,u} = \begin{cases} l_\beta, & \text{if } u \in (-\delta, 0], \\ l_\beta + 1, & \text{if } u \in (0, \delta). \end{cases} \tag{2.9.78}$$

From (2.9.58),

$$f^{\psi^u}(\beta) - f^{\psi}(\beta) = \begin{cases} \beta(\bar{\sigma}_{u,j^\star} - \bar{\sigma}_{j^\star})\nabla\lambda^{j^\star}, & \text{if } j^\star \le l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta, \\ \left[\beta\bar{\sigma}_{u,j^\star} - \left(1 + \frac{\beta^2\bar{\sigma}_{j^\star}^2}{4}\right)\right]\nabla\lambda^{j^\star}, & \text{if } j^\star = l_\beta \text{ and } l_{\beta,u} = l_\beta + 1, \\ \frac{\beta^2}{4}(\bar{\sigma}_{u,j^\star}^2 - \bar{\sigma}_{j^\star}^2)\nabla\lambda^{j^\star}, & \text{if } j^\star \ge l_\beta \text{ and } l_{\beta,u} = l_\beta, \end{cases}$$

$$= \begin{cases} \beta\left(\sqrt{\bar{\sigma}_{u,j^\star}^2\nabla\lambda^{j^\star}} - \sqrt{\bar{\sigma}_{j^\star}^2\nabla\lambda^{j^\star}}\right)\sqrt{\nabla\lambda^{j^\star}}, & \text{if } j^\star \le l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta, \\ \beta(\bar{\sigma}_{u,j^\star} - \bar{\sigma}_{j^\star})\nabla\lambda^{j^\star} - (1 - \frac{\beta}{2}\bar{\sigma}_{j^\star})^2\,\nabla\lambda^{j^\star}, & \text{if } j^\star = l_\beta \text{ and } l_{\beta,u} = l_\beta + 1, \\ \frac{\beta^2}{4}(\bar{\sigma}_{u,j^\star}^2 - \bar{\sigma}_{j^\star}^2)\nabla\lambda^{j^\star}, & \text{if } j^\star \ge l_\beta \text{ and } l_{\beta,u} = l_\beta, \end{cases}$$

$$= \begin{cases} (1^*):\ \beta\left\{\frac{(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha)}{2\bar{\sigma}_{j^\star}} + O(u^2)\right\}, & \text{if } j^\star \le l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta, \\ (2^*):\ \beta\left\{\frac{(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha)}{2\bar{\sigma}_{j^\star}} + O(u^2)\right\} & \\ \qquad\qquad -(1 - \frac{\beta}{2}\bar{\sigma}_{j^\star})^2\,\nabla\lambda^{j^\star}, & \text{if } j^\star = l_\beta \text{ and } l_{\beta,u} = l_\beta + 1, \\ (3^*):\ \frac{\beta^2}{4}(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha), & \text{if } j^\star \ge l_\beta \text{ and } l_{\beta,u} = l_\beta. \end{cases} \tag{2.9.79}$$

To get the last equality, we used

$$(\bar{\sigma}_{u,j^\star}^2 - \bar{\sigma}_{j^\star}^2)\nabla\lambda^{j^\star} = ((\sigma_{i^\star} + u)^2 - \sigma_{i^\star}^2)(\alpha' - \alpha). \tag{2.9.80}$$

The function $u \mapsto f^{\psi^u}(\beta)$ is always differentiable on $(-\delta, \delta)\backslash\{0\}$. Furthermore,

$$\text{when } \beta > 2/\bar{\sigma}_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(1^*)}{=} \frac{\beta\sigma_{i^\star}(\alpha' - \alpha)}{\bar{\sigma}_{j^\star}} \overset{(1^*)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta), \tag{2.9.81}$$

$$\text{when } \beta = 2/\bar{\sigma}_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(3^*)}{=} \frac{\beta^2\sigma_{i^\star}(\alpha' - \alpha)}{2} \overset{(2^*)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta), \tag{2.9.82}$$

$$\text{when } \beta < 2/\bar{\sigma}_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(3^*)}{=} \frac{\beta^2\sigma_{i^\star}(\alpha' - \alpha)}{2} \overset{(3^*)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta). \tag{2.9.83}$$

Thus, $u \mapsto f^{\psi^u}(\beta)$ is also differentiable at $u = 0$.

Case (ii) : Here are all the possible subcases of Case (ii) :

(ii.1) • $\mathcal{J}_{\sigma^2}(r) < \mathcal{J}_{\bar{\sigma}^2}(r)$ for all $r \in [\alpha', \lambda^{j^\star})$;

• $\exists s \in (\lambda^{j^\star-1}, \alpha]$ such that $\mathcal{J}_{\sigma^2}(s) = \mathcal{J}_{\bar{\sigma}^2}(s)$;

(ii.2) • $\mathcal{J}_{\sigma^2}(r) < \mathcal{J}_{\bar{\sigma}^2}(r)$ for all $r \in (\lambda^{j^\star-1}, \alpha]$;

• $\exists t \in [\alpha', \lambda^{j^\star})$ such that $\mathcal{J}_{\sigma^2}(t) = \mathcal{J}_{\bar{\sigma}^2}(t)$;

(ii.3) • $\exists s \in (\lambda^{j^\star-1}, \alpha]$ such that $\mathcal{J}_{\sigma^2}(s) = \mathcal{J}_{\bar{\sigma}^2}(s)$;

• $\exists t \in [\alpha', \lambda^{j^\star})$ such that $\mathcal{J}_{\sigma^2}(t) = \mathcal{J}_{\bar{\sigma}^2}(t)$.

Denote

$$s^\star \mathrel{\hat{=}} \max\Big\{r \in [\lambda^{j^\star-1}, \alpha] : \mathcal{J}_{\sigma^2}(r) = \mathcal{J}_{\bar{\sigma}^2}(r)\Big\}, \tag{2.9.84}$$

$$t^\star \mathrel{\hat{=}} \min\Big\{r \in [\alpha', \lambda^{j^\star}] : \mathcal{J}_{\sigma^2}(r) = \mathcal{J}_{\bar{\sigma}^2}(r)\Big\}. \tag{2.9.85}$$

Again, the function $u \mapsto \bar{\sigma}_u(r)$ is continuous, uniformly in $r \in [0,1]$. Hence, we can choose $\delta = \delta(\alpha, \alpha', \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ small enough that for all $u \in (-\delta, \delta)$:

Case	$u < 0$	$u > 0$
(ii.1)	• $\bar{\sigma}_{u,j} = \begin{cases} \bar{\sigma}_j & \text{for } j \le j^\star \\ \bar{\sigma}_{j-1} & \text{for } j \ge j^\star + 2 \end{cases}$ • $\lambda_u^j = \begin{cases} \lambda^j & \text{for } j \le j^\star - 1 \\ s^\star & \text{for } j = j^\star \\ \lambda^{j-1} & \text{for } j \ge j^\star + 1 \end{cases}$ • $m_u = m + 1$	• $\bar{\sigma}_{u,j} = \bar{\sigma}_j$ for $j \neq j^\star$ • $\lambda_u^j = \lambda^j$ for all j • $m_u = m$
(ii.2)	• $\bar{\sigma}_{u,j} = \bar{\sigma}_j$ for $j \neq j^\star$ • $\lambda_u^j = \lambda^j$ for all j • $m_u = m$	• $\bar{\sigma}_{u,j} = \begin{cases} \bar{\sigma}_j & \text{for } j \le j^\star - 1 \\ \bar{\sigma}_{j-1} & \text{for } j \ge j^\star + 1 \end{cases}$ • $\lambda_u^j = \begin{cases} \lambda^j & \text{for } j \le j^\star - 1 \\ t^\star & \text{for } j = j^\star \\ \lambda^{j-1} & \text{for } j \ge j^\star + 1 \end{cases}$ • $m_u = m + 1$
(ii.3)	• $\bar{\sigma}_{u,j} = \begin{cases} \bar{\sigma}_j & \text{for } j \le j^\star \\ \bar{\sigma}_{j-1} & \text{for } j \ge j^\star + 2 \end{cases}$ • $\lambda_u^j = \begin{cases} \lambda^j & \text{for } j \le j^\star - 1 \\ s^\star & \text{for } j = j^\star \\ \lambda^{j-1} & \text{for } j \ge j^\star + 1 \end{cases}$ • $m_u = m + 1$	• $\bar{\sigma}_{u,j} = \begin{cases} \bar{\sigma}_j & \text{for } j \le j^\star - 1 \\ \bar{\sigma}_{j-1} & \text{for } j \ge j^\star + 1 \end{cases}$ • $\lambda_u^j = \begin{cases} \lambda^j & \text{for } j \le j^\star - 1 \\ t^\star & \text{for } j = j^\star \\ \lambda^{j-1} & \text{for } j \ge j^\star + 1 \end{cases}$ • $m_u = m + 1$

In other words, the parameter δ is chosen small enough that, on $(\lambda^{j^\star-1}, \lambda^{j^\star}]$, the field ψ^u has either one or two effective variance parameters (depending on the subcase) and they remain strictly between $\bar{\sigma}_{j^\star-1}$ and $\bar{\sigma}_{j^\star+1}$. If there is only one effective slope, then $\lambda_u^{j^\star} = \lambda^{j^\star}$. If there are two effective slopes, the segments meet at $\lambda_u^{j^\star} \in \{s^\star, t^\star\}$. Figure 2.9.6 on the next page (analogous to Figure 2.9.5) illustrates this point more clearly.

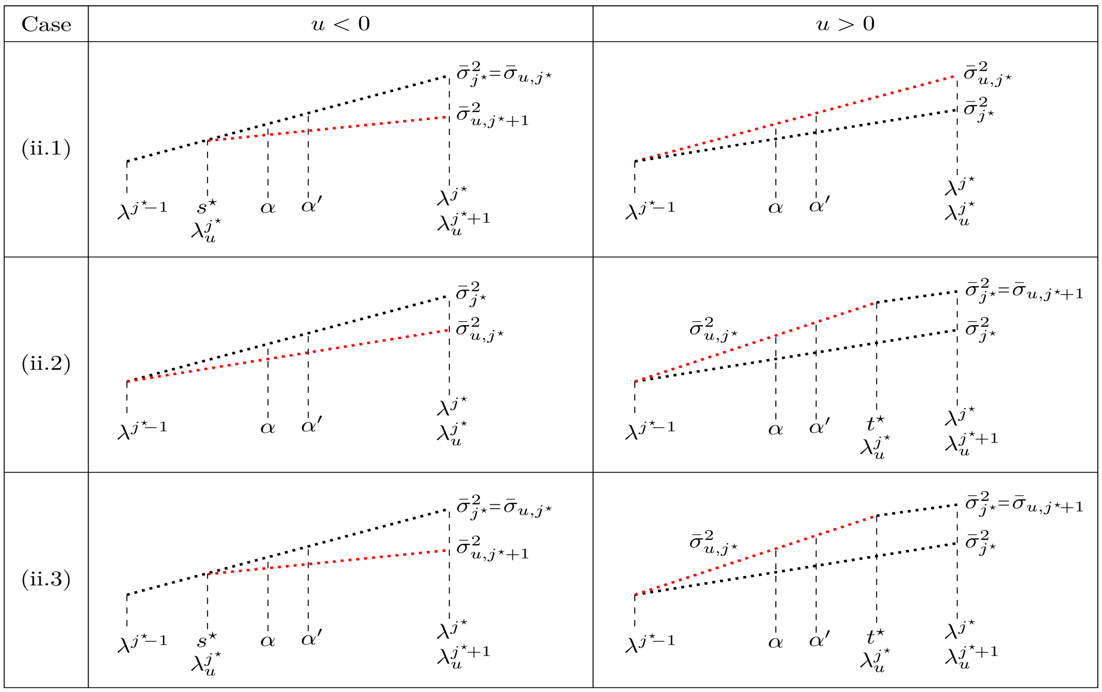


FIGURE 2.9.6. General form of $\mathcal{J}_{\bar{\sigma}^2}$ and $\mathcal{J}_{\bar{\sigma}_u^2}$ on $(\lambda^{j^\star-1}, \lambda^{j^\star}]$. The effective slopes of ψ and ψ^u are the quantities $\bar{\sigma}_j^2$ and $\bar{\sigma}_{u,j}^2$, respectively. The dotted paths containing a red part are the ones for the perturbed field ψ^u.

Case (ii.1) : In this case, $s^\star \in (\lambda^{j^\star-1}, \alpha]$. We can choose $\delta = \delta(\beta, \alpha, \alpha', \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ small enough that

$$\text{when } \beta > 2/\bar{\sigma}_{j^\star} (\Longrightarrow j^\star \leq l_\beta - 1), \quad \text{then } l_{\beta,u} = \begin{cases} l_\beta + 1, & \text{if } u \in (-\delta, 0), \\ l_\beta, & \text{if } u \in [0, \delta), \end{cases} \tag{2.9.86}$$

$$\text{when } \beta = 2/\bar{\sigma}_{j^\star} (\Longrightarrow j^\star = l_\beta), \quad \text{then } l_{\beta,u} = \begin{cases} l_\beta, & \text{if } u \in (-\delta, 0], \\ l_\beta + 1, & \text{if } u \in (0, \delta), \end{cases} \tag{2.9.87}$$

$$\text{when } \beta < 2/\bar{\sigma}_{j^\star} (\Longrightarrow j^\star \geq l_\beta), \quad \text{then } l_{\beta,u} = l_\beta \text{ for all } u \in (-\delta, \delta). \tag{2.9.88}$$

When $u < 0$,

$$
\begin{aligned}
& f^{\psi^u}(\beta) - f^{\psi}(\beta) \\
&= \begin{cases} \beta(\bar\sigma_{u,j^\star+1} - \bar\sigma_{j^\star})(\lambda^{j^\star} - s^\star), & \text{if } j^\star \le l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta + 1, \\ \frac{\beta^2}{4}(\bar\sigma^2_{u,j^\star+1} - \bar\sigma^2_{j^\star})(\lambda^{j^\star} - s^\star), & \text{if } j^\star \ge l_\beta \text{ and } l_{\beta,u} = l_\beta, \end{cases} \\
&= \begin{cases} \beta\left(\sqrt{\bar\sigma^2_{u,j^\star+1}(\lambda^{j^\star} - s^\star)} - \sqrt{\bar\sigma^2_{j^\star}(\lambda^{j^\star} - s^\star)}\right)\sqrt{\lambda^{j^\star} - s^\star}, & \\ & \text{if } j^\star \le l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta + 1, \\ \frac{\beta^2}{4}(\bar\sigma^2_{u,j^\star+1} - \bar\sigma^2_{j^\star})(\lambda^{j^\star} - s^\star), & \text{if } j^\star \ge l_\beta \text{ and } l_{\beta,u} = l_\beta, \end{cases} \\
&= \begin{cases} (1^-):\ \beta\left\{\frac{(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha)}{2\bar\sigma_{j^\star}} + O(u^2)\right\}, & \text{if } j^\star \le l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta + 1, \\ (2^-):\ \frac{\beta^2}{4}(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha), & \text{if } j^\star \ge l_\beta \text{ and } l_{\beta,u} = l_\beta. \end{cases}
\end{aligned} \tag{2.9.89}
$$

To get the last equality, we used

$$
(\bar\sigma^2_{u,j^\star+1} - \bar\sigma^2_{j^\star})(\lambda^{j^\star} - s^\star) = ((\sigma_{i^\star} + u)^2 - \sigma^2_{i^\star})(\alpha' - \alpha). \tag{2.9.90}
$$

When $u > 0$, it is the same as in (2.9.79) :

$$
\begin{aligned}
& f^{\psi^u}(\beta) - f^{\psi}(\beta) \\
&= \begin{cases} (1^+):\ \beta\left\{\frac{(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha)}{2\bar\sigma_{j^\star}} + O(u^2)\right\}, & \text{if } j^\star \le l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta, \\ (2^+):\ \beta\left\{\frac{(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha)}{2\bar\sigma_{j^\star}} + O(u^2)\right\} & \\ \qquad\qquad -(1 - \frac{\beta}{2}\bar\sigma_{j^\star})^2\,\nabla\lambda^{j^\star}, & \text{if } j^\star = l_\beta \text{ and } l_{\beta,u} = l_\beta + 1, \\ (3^+):\ \frac{\beta^2}{4}(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha), & \text{if } j^\star \ge l_\beta \text{ and } l_{\beta,u} = l_\beta. \end{cases}
\end{aligned} \tag{2.9.91}
$$

The function $u \mapsto f^{\psi^u}(\beta)$ is always differentiable on $(-\delta, \delta)\backslash\{0\}$. Furthermore,

$$
\text{when } \beta > 2/\bar\sigma_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(1^-)}{=} \frac{\beta\sigma_{i^\star}(\alpha' - \alpha)}{\bar\sigma_{j^\star}} \overset{(1^+)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta), \tag{2.9.92}
$$

$$
\text{when } \beta = 2/\bar\sigma_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(2^-)}{=} \frac{\beta^2\sigma_{i^\star}(\alpha' - \alpha)}{2} \overset{(2^+)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta), \tag{2.9.93}
$$

$$
\text{when } \beta < 2/\bar\sigma_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(2^-)}{=} \frac{\beta^2\sigma_{i^\star}(\alpha' - \alpha)}{2} \overset{(3^+)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta). \tag{2.9.94}
$$

Thus, $u \mapsto f^{\psi^u}(\beta)$ is also differentiable at $u = 0$.

Case (ii.2) : In this case, $t^\star \in [\alpha', \lambda^{j^\star})$. We can choose $\delta = \delta(\beta, \alpha, \alpha', \boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$ small enough that

$$\text{when } \beta > 2/\bar{\sigma}_{j^\star} (\Longrightarrow j^\star \leq l_\beta - 1), \quad \text{then } l_{\beta,u} = \begin{cases} l_\beta, & \text{if } u \in (-\delta, 0], \\ l_\beta + 1, & \text{if } u \in (0, \delta), \end{cases} \tag{2.9.95}$$

$$\text{when } \beta = 2/\bar{\sigma}_{j^\star} (\Longrightarrow j^\star = l_\beta), \quad \text{then } l_{\beta,u} = \begin{cases} l_\beta, & \text{if } u \in (-\delta, 0], \\ l_\beta + 1, & \text{if } u \in (0, \delta), \end{cases} \tag{2.9.96}$$

$$\text{when } \beta < 2/\bar{\sigma}_{j^\star} (\Longrightarrow j^\star \geq l_\beta), \quad \text{then } l_{\beta,u} = l_\beta \text{ for all } u \in (-\delta, \delta). \tag{2.9.97}$$

When $u < 0$, it is the same as in (2.9.79) (without (2^*)) :

$$\begin{aligned} & f^{\psi^u}(\beta) - f^\psi(\beta) \\ &= \begin{cases} (1^-): \ \beta \left\{ \frac{(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha)}{2\bar{\sigma}_{j^\star}} + O(u^2) \right\}, & \text{if } j^\star \leq l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta, \\ (2^-): \ \frac{\beta^2}{4}(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha), & \text{if } j^\star \geq l_\beta \text{ and } l_{\beta,u} = l_\beta. \end{cases} \end{aligned} \tag{2.9.98}$$

When $u > 0$,

$$\begin{aligned} & f^{\psi^u}(\beta) - f^\psi(\beta) \\ &= \begin{cases} \beta(\bar{\sigma}_{u,j^\star} - \bar{\sigma}_{j^\star})(t^\star - \lambda^{j^\star - 1}), & \text{if } j^\star \leq l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta + 1, \\ \left[\beta\bar{\sigma}_{u,j^\star} - \left(1 + \frac{\beta^2\bar{\sigma}_{j^\star}^2}{4}\right)\right](t^\star - \lambda^{j^\star - 1}), & \text{if } j^\star = l_\beta \text{ and } l_{\beta,u} = l_\beta + 1, \\ \frac{\beta^2}{4}(\bar{\sigma}_{u,j^\star+1}^2 - \bar{\sigma}_{j^\star}^2)(t^\star - \lambda^{j^\star - 1}), & \text{if } j^\star \geq l_\beta \text{ and } l_{\beta,u} = l_\beta, \end{cases} \\ &= \begin{cases} \beta\left(\sqrt{\bar{\sigma}_{u,j^\star}^2(t^\star - \lambda^{j^\star-1})} - \sqrt{\bar{\sigma}_{j^\star}^2(t^\star - \lambda^{j^\star-1})}\right)\sqrt{t^\star - \lambda^{j^\star-1}}, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } j^\star \leq l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta + 1, \\ \beta(\bar{\sigma}_{u,j^\star} - \bar{\sigma}_{j^\star})(t^\star - \lambda^{j^\star-1}) \\ \qquad -(1 - \frac{\beta}{2}\bar{\sigma}_{j^\star})^2\,(t^\star - \lambda^{j^\star-1}), \qquad \text{if } j^\star = l_\beta \text{ and } l_{\beta,u} = l_\beta + 1, \\ \frac{\beta^2}{4}(\bar{\sigma}_{u,j^\star+1}^2 - \bar{\sigma}_{j^\star}^2)(t^\star - \lambda^{j^\star-1}), \qquad \text{if } j^\star \geq l_\beta \text{ and } l_{\beta,u} = l_\beta, \end{cases} \end{aligned}$$

$$= \begin{cases} (1^+): \ \beta\left\{\frac{(2u\sigma_{i^\star}+u^2)(\alpha'-\alpha)}{2\bar{\sigma}_{j^\star}} + O(u^2)\right\}, & \text{if } j^\star \le l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta + 1, \\ (2^+): \ \beta\left\{\frac{(2u\sigma_{i^\star}+u^2)(\alpha'-\alpha)}{2\bar{\sigma}_{j^\star}} + O(u^2)\right\} & \\ \qquad\qquad -(1-\frac{\beta}{2}\bar{\sigma}_{j^\star})^2\,(t^\star - \lambda^{j^\star-1}), & \text{if } j^\star = l_\beta \text{ and } l_{\beta,u} = l_\beta + 1, \\ (3^+): \ \frac{\beta^2}{4}(2u\sigma_{i^\star}+u^2)(\alpha'-\alpha), & \text{if } j^\star \ge l_\beta \text{ and } l_{\beta,u} = l_\beta. \end{cases} \tag{2.9.99}$$

To get the last equality, we used

$$(\bar{\sigma}_{u,j^\star}^2 - \bar{\sigma}_{j^\star}^2)(t^\star - \lambda^{j^\star-1}) = ((\sigma_{i^\star}+u)^2 - \sigma_{i^\star}^2)(\alpha'-\alpha). \tag{2.9.100}$$

The function $u \mapsto f^{\psi^u}(\beta)$ is always differentiable on $(-\delta,\delta)\backslash\{0\}$. Furthermore,

$$\text{when } \beta > 2/\bar{\sigma}_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(1^-)}{=} \frac{\beta\sigma_{i^\star}(\alpha'-\alpha)}{\bar{\sigma}_{j^\star}} \overset{(1^+)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta), \tag{2.9.101}$$

$$\text{when } \beta = 2/\bar{\sigma}_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(2^-)}{=} \frac{\beta^2\sigma_{i^\star}(\alpha'-\alpha)}{2} \overset{(2^+)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta), \tag{2.9.102}$$

$$\text{when } \beta < 2/\bar{\sigma}_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(2^-)}{=} \frac{\beta^2\sigma_{i^\star}(\alpha'-\alpha)}{2} \overset{(3^+)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta). \tag{2.9.103}$$

Thus, $u \mapsto f^{\psi^u}(\beta)$ is also differentiable at $u = 0$.

Case (ii.3) : In this case, $s^\star \in (\lambda^{j^\star-1}, \alpha]$ and $t^\star \in [\alpha', \lambda^{j^\star})$. We can choose $\delta = \delta(\beta,\alpha,\alpha',\boldsymbol{\sigma},\boldsymbol{\lambda}) > 0$ small enough that

$$\text{when } \beta > 2/\bar{\sigma}_{j^\star} (\Longrightarrow j^\star \le l_\beta - 1), \quad \text{then } l_{\beta,u} = \begin{cases} l_\beta + 1, & \text{if } u \in (-\delta, 0], \\ l_\beta + 1, & \text{if } u \in (0,\delta), \end{cases} \tag{2.9.104}$$

$$\text{when } \beta = 2/\bar{\sigma}_{j^\star} (\Longrightarrow j^\star = l_\beta), \quad \text{then } l_{\beta,u} = \begin{cases} l_\beta, & \text{if } u \in (-\delta, 0], \\ l_\beta + 1, & \text{if } u \in (0,\delta), \end{cases} \tag{2.9.105}$$

$$\text{when } \beta < 2/\bar{\sigma}_{j^\star} (\Longrightarrow j^\star \ge l_\beta), \quad \text{then } l_{\beta,u} = l_\beta \text{ for all } u \in (-\delta,\delta). \tag{2.9.106}$$

When $u < 0$, it is the same as in (2.9.89) :

$$f^{\psi^u}(\beta) - f^{\psi}(\beta)$$

$$= \begin{cases} (1^-): \ \beta\left\{\frac{(2u\sigma_{i^\star}+u^2)(\alpha'-\alpha)}{2\bar{\sigma}_{j^\star}} + O(u^2)\right\}, & \text{if } j^\star \le l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta + 1, \\ (2^-): \ \frac{\beta^2}{4}(2u\sigma_{i^\star}+u^2)(\alpha'-\alpha), & \text{if } j^\star \ge l_\beta \text{ and } l_{\beta,u} = l_\beta. \end{cases} \tag{2.9.107}$$

When $u > 0$, it is the same as in (2.9.99) :

$$
\begin{aligned}
&f^{\psi^u}(\beta) - f^{\psi}(\beta) \\
&= \begin{cases}
(1^+): \ \beta \left\{ \frac{(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha)}{2\bar{\sigma}_{j^\star}} + O(u^2) \right\}, & \text{if } j^\star \leq l_\beta - 1 \text{ and } l_{\beta,u} = l_\beta + 1, \\
(2^+): \ \beta \left\{ \frac{(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha)}{2\bar{\sigma}_{j^\star}} + O(u^2) \right\} & \\
\qquad\qquad -(1 - \frac{\beta}{2}\bar{\sigma}_{j^\star})^2 \, (t^\star - \lambda^{j^\star - 1}), & \text{if } j^\star = l_\beta \text{ and } l_{\beta,u} = l_\beta + 1, \\
(3^+): \ \frac{\beta^2}{4}(2u\sigma_{i^\star} + u^2)(\alpha' - \alpha), & \text{if } j^\star \geq l_\beta \text{ and } l_{\beta,u} = l_\beta.
\end{cases}
\end{aligned} \tag{2.9.108}
$$

The function $u \mapsto f^{\psi^u}(\beta)$ is always differentiable on $(-\delta, \delta)\backslash\{0\}$. Furthermore,

$$
\text{when } \beta > 2/\bar{\sigma}_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(1^-)}{=} \frac{\beta \sigma_{i^\star}(\alpha' - \alpha)}{\bar{\sigma}_{j^\star}} \overset{(1^+)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta), \tag{2.9.109}
$$

$$
\text{when } \beta = 2/\bar{\sigma}_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(2^-)}{=} \frac{\beta^2 \sigma_{i^\star}(\alpha' - \alpha)}{2} \overset{(2^+)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta), \tag{2.9.110}
$$

$$
\text{when } \beta < 2/\bar{\sigma}_{j^\star}, \quad \frac{\partial}{\partial u^-} f^{\psi^0}(\beta) \overset{(2^-)}{=} \frac{\beta^2 \sigma_{i^\star}(\alpha' - \alpha)}{2} \overset{(3^+)}{=} \frac{\partial}{\partial u^+} f^{\psi^0}(\beta). \tag{2.9.111}
$$

Thus, $u \mapsto f^{\psi^u}(\beta)$ is also differentiable at $u = 0$. This ends the proof of Lemma 2.9.10. □

Acknowledgements

First, I would like to thank an anonymous referee and my advisor, Louis-Pierre Arguin, for their valuable comments that led to improvements in the presentation of this paper. I also thank Yan Fyodorov for helpful correspondences and for pointing out the literature on the multi-scale log-REM model. This work is supported by a NSERC Doctoral Program Alexander Graham Bell scholarship (CGS D3).

Bibliography

Y. Abe. Cover times for sequences of reversible Markov chains on random graphs. *Kyoto J. Math.*, 54(3):555–576, 2014. MR3263552.

M. Aizenman and P. Contucci. On the stability of the quenched state in mean-field spin-glass models. *J. Statist. Phys.*, 92(5-6):765–783, 1998. MR1657840.

L.-P. Arguin and M. Aizenman. On the structure of quasi-stationary competing particle systems. *Ann. Probab.*, 37(3):1080–1113, 2009. MR2537550.

L.-P. Arguin and S. Chatterjee. Random overlap structures: properties and applications to spin glasses. *Probab. Theory Related Fields*, 156(1-2):375–413, 2013. MR3055263.

L.-P. Arguin and F. Ouimet. Extremes of the two-dimensional Gaussian free field with scale-dependent variance. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(2):779–808, 2016. MR3541850.

L.-P. Arguin and W. Tai. Is the Riemann zeta function in a short interval a 1-RSB spin glass ? *Preprint. To appear in Sojourns in probability and statistical physics, in honor of the 70th birthday of C. Newman*, pages 1–20, 2018. arXiv:1706.08462.

L.-P. Arguin and O. Zindy. Poisson-Dirichlet statistics for the extremes of a log-correlated Gaussian field. *Ann. Appl. Probab.*, 24(4):1446–1481, 2014. MR3211001.

L.-P. Arguin and O. Zindy. Poisson-Dirichlet statistics for the extremes of the two-dimensional discrete Gaussian free field. *Electron. J. Probab.*, 20:no. 59, 19, 2015. MR3354619.

L.-P. Arguin, D. Belius, and P. Bourgade. Maximum of the characteristic polynomial of random unitary matrices. *Comm. Math. Phys.*, 349(2):703–751, 2017a. MR3594368.

L.-P. Arguin, D. Belius, and A. J. Harper. Maxima of a randomized Riemann zeta function, and branching random walks. *Ann. Appl. Probab.*, 27(1):178–215, 2017b. MR3619786.

L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł, and K. Soundararajan. Maximum of the Riemann zeta function on a short interval of the critical line. *Comm. Pure Appl. Math.*, 72(3):500–535, 2019. MR3911893.

D. Belius. Gumbel fluctuations for cover times in the discrete torus. *Probab. Theory Related Fields*, 157(3-4):635–689, 2013. MR3129800.

D. Belius and N. Kistler. The subleading order of two dimensional cover times. *Probab. Theory Related Fields*, 167(1-2):461–552, 2017. MR3602852.

N. Berestycki. Introduction to the Gaussian Free Field and Liouville Quantum Gravity. *Lecture notes*, 2016. URL `http://www.statslab.cam.ac.uk/~beresty/Articles/oxford4.pdf`.

E. Bolthausen and N. Kistler. On a nonhierarchical version of the generalized random energy model. *Ann. Appl. Probab.*, 16(1):1–14, 2006. MR2209333.

E. Bolthausen and N. Kistler. On a nonhierarchical version of the generalized random energy model. II. Ultrametricity. *Stochastic Process. Appl.*, 119(7):2357–2386, 2009. MR2531095.

E. Bolthausen and N. Kistler. A quenched large deviation principle and a Parisi formula for a perceptron version of the GREM. In *Probability in complex physical systems*, volume 11 of *Springer Proc. Math.*, pages 425–442. Springer, Heidelberg, 2012. MR3372858.

E. Bolthausen and A.-S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197(2):247–276, 1998. MR1652734.

E. Bolthausen and A.-S. Sznitman. *Ten lectures on random media*, volume 32 of *DMV Seminar*. Birkhäuser Verlag, Basel, 2002. MR1890289.

A. Bovier. *Statistical mechanics of disordered systems*, volume 18 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2006. A mathematical perspective. MR2252929.

A. Bovier and L. Hartung. The extremal process of two-speed branching Brownian motion. *Electron. J. Probab.*, 19:no. 18, 28, 2014. MR3164771.

A. Bovier and L. Hartung. Variable speed branching Brownian motion 1. Extremal processes in the weak correlation regime. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12(1): 261–291, 2015. MR3351476.

A. Bovier and I. Kurkova. Derrida's generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):439–480, 2004a. MR2070334.

A. Bovier and I. Kurkova. Derrida's generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):481–495, 2004b. MR2070335.

X. Cao, Y. V. Fyodorov, and P. Le Doussal. One step replica symmetry breaking and extreme order statistics of logarithmic REMs. *SciPost Phys.*, 1(2):011, 1–58, 2016. doi:10.21468/SciPostPhys.1.2.011.

D. Capocaccia, M. Cassandro, and P. Picco. On the existence of thermodynamics for the generalized random energy model. *J. Statist. Phys.*, 46(3-4):493–505, 1987. MR883541.

D. Carpentier and P. Le Doussal. Glass transition of a particle in a random potential, front selection in nonlinear renormalization group, and entropic phenomena in Liouville and sinh-Gordon models. *Phys. Rev. E*, 63(026110):1–33, 2001. doi:10.1103/PhysRevE.63.026110.

R. Chhaibi, T. Madaule, and J. Najnudel. On the maximum of the CβE field. *Duke Math. J.*, 167(12):2243–2345, 2018. MR3848391.

F. Comets, C. Gallesco, S. Popov, and M. Vachkovskaia. On large deviations for the cover time of two-dimensional torus. *Electron. J. Probab.*, 18:no. 96, 18, 2013. MR3126579.

P. Contucci and C. Giardinà. Spin-glass stochastic stability: a rigorous proof. *Ann. Henri Poincaré*, 6(5):915–923, 2005. MR2219862.

A. Dembo, Y. Peres, and J. Rosen. Brownian motion on compact manifolds: cover time and late points. *Electron. J. Probab.*, 8:no. 15, 14, 2003. MR1998762.

A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Cover times for Brownian motion and random walks in two dimensions. *Ann. of Math. (2)*, 160(2):433–464, 2004. MR2123929.

A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Late points for random walks in two dimensions. *Ann. Probab.*, 34(1):219–263, 2006. MR2206347.

B. Derrida. A generalization of the Random Energy Model which includes correlations between energies. *J. Physique Lett.*, 46(9):401–407, 1985. doi:10.1051/jphyslet:01985004609040100.

J. Ding. On cover times for 2D lattices. *Electron. J. Probab.*, 17:no. 45, 18, 2012. MR2946152.

J. Ding. Asymptotics of cover times via Gaussian free fields: bounded-degree graphs and general trees. *Ann. Probab.*, 42(2):464–496, 2014. MR3178464.

J. Ding and O. Zeitouni. A sharp estimate for cover times on binary trees. *Stochastic Process. Appl.*, 122(5):2117–2133, 2012. MR2921974.

J. Ding, J. R. Lee, and Y. Peres. Cover times, blanket times, and majorizing measures. *Ann. of Math. (2)*, 175(3):1409–1471, 2012. MR2912708.

L. N. Dovbysh and V. N. Sudakov. Gram-de Finetti matrices. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 119:77–86, 1982. MR666087.

E. B. Dynkin. Markov processes and random fields. *Bull. Amer. Math. Soc. (N.S.)*, 3(3):

975–999, 1980. MR585179.

M. Fang and O. Zeitouni. Branching random walks in time inhomogeneous environments. *Electron. J. Probab.*, 17:no. 67, 18, 2012a. MR2968674.

M. Fang and O. Zeitouni. Slowdown for time inhomogeneous branching Brownian motion. *J. Stat. Phys.*, 149(1):1–9, 2012b. MR2981635.

Y. V. Fyodorov and J.-P. Bouchaud. Statistical mechanics of a single particle in a multiscale random potential: Parisi landscapes in finite-dimensional Euclidean spaces. *J. Phys. A*, 41(32):324009, 25, 2008a. MR2425780.

Y. V. Fyodorov and J.-P. Bouchaud. Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential. *J. Phys. A*, 41(37):372001, 12, 2008b. MR2430565.

Y. V. Fyodorov, P. Le Doussal, and A. Rosso. Statistical mechanics of logarithmic REM: duality, freezing and extreme value statistics of $1/f$ noises generated by Gaussian free fields. *J. Stat. Mech. Theory Exp.*, 2009(10):P10005, 32, 2009. MR2882779.

S. Ghirlanda and F. Guerra. General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A*, 31(46):9149–9155, 1998. MR1662161.

A. J. Harper. A note on the maximum of the Riemann zeta function, and log-correlated random variables. *Preprint*, pages 1–26, 2013. arXiv:1304.0677.

A. Jagannath. On the overlap distribution of branching random walks. *Electron. J. Probab.*, 21:Paper No. 50 16, 2016. MR3539644.

A. Jagannath. Approximate ultrametricity for random measures and applications to spin glasses. *Comm. Pure Appl. Math.*, 70(4):611–664, 2017. MR3628881.

G. F. Lawler. *Intersections of random walks*. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, 1991. MR1117680.

P. Maillard and O. Zeitouni. Slowdown in branching Brownian motion with inhomogeneous variance. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(3):1144–1160, 2016. MR3531703.

B. Mallein. Maximal displacement in a branching random walk through interfaces. *Electron. J. Probab.*, 20:no. 68, 40, 2015a. MR3361256.

B. Mallein. Maximal displacement of a branching random walk in time-inhomogeneous environment. *Stoch. Proc. Appl.*, 125(10):3958–4019, 2015b. MR3373310.

M. Mézard, G. Parisi, and M. A. Virasoro. *Spin glass theory and beyond*, volume 9 of *World Scientific Lecture Notes in Physics*. World Scientific Publishing Co., Inc., Teaneck, NJ, 1987. MR1026102.

J. Najnudel. On the extreme values of the Riemann zeta function on random intervals of the critical line. *Probab. Theory Related Fields*, 172(1-2):387–452, 2018. MR3851835.

F. Ouimet. *Étude du maximum et des hauts points de la marche aléatoire branchante inhomogène et du champ libre gaussien inhomogène*. Master's thesis, Université de Montréal, 2014. URL `https://papyrus.bib.umontreal.ca/xmlui/handle/1866/11510`.

F. Ouimet. Maxima of branching random walks with piecewise constant variance. *Braz. J. Probab. Stat.*, 32(4):679–706, 2018. MR3845025.

D. Panchenko. A connection between the Ghirlanda-Guerra identities and ultrametricity. *Ann. Probab.*, 38(1):327–347, 2010a. MR2599202.

D. Panchenko. The Ghirlanda-Guerra identities for mixed p-spin model. *C. R. Math. Acad. Sci. Paris*, 348(3-4):189–192, 2010b. MR2600075.

D. Panchenko. On the Dovbysh-Sudakov representation result. *Electron. Commun. Probab.*, 15:330–338, 2010c. MR2679002.

D. Panchenko. Ghirlanda-Guerra identities and ultrametricity: an elementary proof in the discrete case. *C. R. Math. Acad. Sci. Paris*, 349(13-14):813–816, 2011. MR2825947.

D. Panchenko. A unified stability property in spin glasses. *Comm. Math. Phys.*, 313(3): 781–790, 2012. MR2945621.

D. Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013a. MR2999044.

D. Panchenko. *The Sherrington-Kirkpatrick Model*. Springer Monographs in Mathematics. Springer, New York, 2013b. MR3052333.

E. Paquette and O. Zeitouni. The maximum of the CUE field. *Int. Math. Res. Not. IMRN*, 2018(16):5028–5119, 2018. MR3848227.

R. Rhodes and V. Vargas. Gaussian multiplicative chaos and applications: a review. *Probab. Surv.*, 11:315–392, 2014. MR3274356.

R. T. Rockafellar. *Convex Analysis*. Princeton Mathematical Series, No. 28. Princeton

University Press, Princeton, N.J., 1970. MR0274683.

D. Ruelle. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.*, 108(2):225–239, 1987. MR875300.

E. Saksman and C. Webb. The Riemann zeta function and Gaussian multiplicative chaos: statistics on the critical line. *Preprint*, pages 1–85, 2018. arXiv:1609.00027.

S. Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields*, 139 (3-4):521–541, 2007. MR2322706.

M. Talagrand. Construction of pure states in mean field models for spin glasses. *Probab. Theory Related Fields*, 148(3-4):601–643, 2010. MR2678900.

M. Talagrand. *Mean Field Models for Spin Glasses. Volume I: Basic Examples*, volume 54. Springer-Verlag, Berlin, 2011a. MR2731561.

M. Talagrand. *Mean Field Models for Spin Glasses. Volume II: Advanced Replica-Symmetry and Low Temperature*, volume 55. Springer, Heidelberg, 2011b. MR3024566.

O. Zeitouni. Gaussian fields. *Notes for Lectures*, 2017. URL `http://www.cims.nyu.edu/~zeitouni/notesGauss.pdf`.

Article 3

Maxima of branching random walks with piecewise constant variance

by

Frédéric Ouimet[1]

([1]) Université de Montréal, Montréal, QC H3T 1J4, Canada.

Here is the complete citation :

Ouimet, F. (2018). Maxima of branching random walks with piecewise constant variance. Braz. J. Probab. Stat. 32, no. 4, 679-706. https://doi.org/10.1214/17-BJPS358.

Innovations with respect to my master's thesis

This article corrects my master's thesis by specifying that the main result and its proof (Theorem 3.1.3) are only valid under Restriction 3.1.2. The general strategy is explained in Section 3.2.2, following an email correspondence with Bastien Mallein. I added the proofs of the discrete Brownian bridge estimates (Lemma 3.2.4 and Lemma 3.2.5) by adapting Bramson's argument for the continuous version in Bramson (1978). The presentation of the proof of Lemma 3.2.6 was reworked following a referee's recommendation.

Abstract. This article extends the results of Fang and Zeitouni (2012a) on branching random walks (BRWs) with Gaussian increments in time inhomogeneous environments. We treat the case where the variance of the increments changes a finite number of times at different scales in $[0,1]$ under a slight restriction. We find the asymptotics of the maximum up to an $O_{\mathbb{P}}(1)$ error and show how the profile of the variance influences the leading order and the logarithmic correction term. A more general result was independently obtained by Mallein (2015a) when the law of the increments is not necessarily Gaussian. However, the proof we present here generalizes the approach of Fang and Zeitouni (2012a) instead of using the spinal decomposition of the BRW. As such, the proof is easier to understand and more robust in the presence of an approximate branching structure.
Keywords: extreme value theory, branching random walks, time inhomogeneous environments

3.1. Introduction

3.1.1. The model

The tree underlying the branching process we are interested in can be described as follows. At time $k = 0$, there exists only one particle o, called the *origin*, and we set $\mathbb{D}_0 \mathrel{\hat{=}} \{o\}$. At time $k = 1$, there are $b = 2$ particles and each of them is linked to o by an edge. Denote by $\mathbb{D}_1$ the set of particles at time 1. At time $k = 2$, there are four particles, two of which are linked to the first particle in $\mathbb{D}_1$ and the other two are linked to the second particle in $\mathbb{D}_1$. The set of particles at time 2 is denoted by $\mathbb{D}_2$. The tree is defined iteratively in this manner up to time $k = n$, where $\mathbb{D}_k$ denotes the set of all particles at time k and $|\mathbb{D}_k| = 2^k$. Figure 3.1.1 illustrates the tree structure.

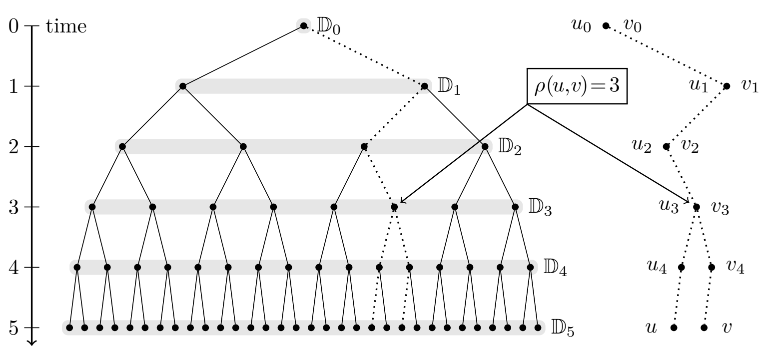


Figure 3.1.1. The tree structure with a branching factor $b = 2$.

For all $v \in \mathbb{D}_n$, we denote by v_k the *ancestor of v at time k*, namely the unique particle in $\mathbb{D}_k$ that intersects the shortest path from o to v. The *branching time* $\rho(u,v)$ is the latest time at which $u, v \in \mathbb{D}_n$ have the same ancestor. Formally,

$$\rho(u,v) \mathrel{\hat{=}} \max\{k \in \{0,1,\ldots,n\} : u_k = v_k\}.$$

In the standard branching random walk (BRW) setting, i.i.d. Gaussian random variables $\mathcal{N}(0,\sigma^2)$ are assigned to each branch of the tree structure and the field of interest is $\{S_v\}_{v\in\mathbb{D}_n}$, where S_v is the sum of the Gaussian variables along the shortest path from o to v. In the time-inhomogeneous context, the variance of the Gaussian variables depends on time. Fix $M \in \mathbb{N}$ and consider the parameters

$$\begin{aligned}
\boldsymbol{\sigma} &\mathrel{\hat{=}} (\sigma_1, \sigma_2, \ldots, \sigma_M) \in (0,\infty)^M && \text{(variance parameters)}\\
\boldsymbol{\lambda} &\mathrel{\hat{=}} (\lambda_1, \lambda_2, \ldots, \lambda_M) \in (0,1]^M && \text{(scale parameters)}
\end{aligned}$$

where $0 \mathrel{\hat{=}} \lambda_0 < \lambda_1 < \ldots < \lambda_M \mathrel{\hat{=}} 1$. The parameters $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$ can be encoded simultaneously in the left-continuous step function

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

The following definition and the results of this paper are easily extended to BRWs with other *branching factors* $b \in \mathbb{N}$.

Definition 3.1.1. *The $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-BRW of length n is a collection of positively correlated random walks $\{\{S_v(t)\}_{t=0}^n\}_{v\in\mathbb{D}_n}$ defined by*

$$S_v(t) \mathrel{\hat{=}} \sum_{i=1}^{M} \sum_{k=\lfloor \lambda_{i-1} n\rfloor+1}^{\lfloor \lambda_i n\rfloor \wedge t} \sigma_i Z_{v_k}, \quad t \in \{0,1,\ldots,n\},\ v \in \mathbb{D}_n, \tag{3.1.1}$$

where $\{Z_{v_k}\}_{k\in\{1,\ldots,n\};v\in\mathbb{D}_n}$ are i.i.d. $\mathcal{N}(0,1)$ random variables and $b = 2$.

Remark 3.1.1. *By convention, summations are zero when there are no indices. To avoid writing trivial corrections in the proofs, always assume, without loss of generality, that $t_i \mathrel{\hat{=}} \lambda_i n \in \mathbb{N}_0$ for all $i \in \{0,1,\ldots,M\}$. Therefore, the floor functions can be dropped in (3.1.1). For simplicity, we set $S_v \mathrel{\hat{=}} S_v(n)$.*

3.1.2. Main result

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

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

The first order of the maximum for the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-BRW is merely the solution to an optimization problem involving the *concave hull* of $\mathcal{J}_{\sigma^2}(\cdot)$, which we denote by $\hat{\mathcal{J}}_{\sigma^2}$. We refer the reader to Ouimet (2014) for a detailed heuristic and a rigorous proof, and to Arguin and Ouimet (2016) for the same results in the context of the scale-inhomogeneous Gaussian free field. By definition, the graph of $\hat{\mathcal{J}}_{\sigma^2}$ is an increasing and concave polygonal line, see Figure 3.1.2 below for some examples.

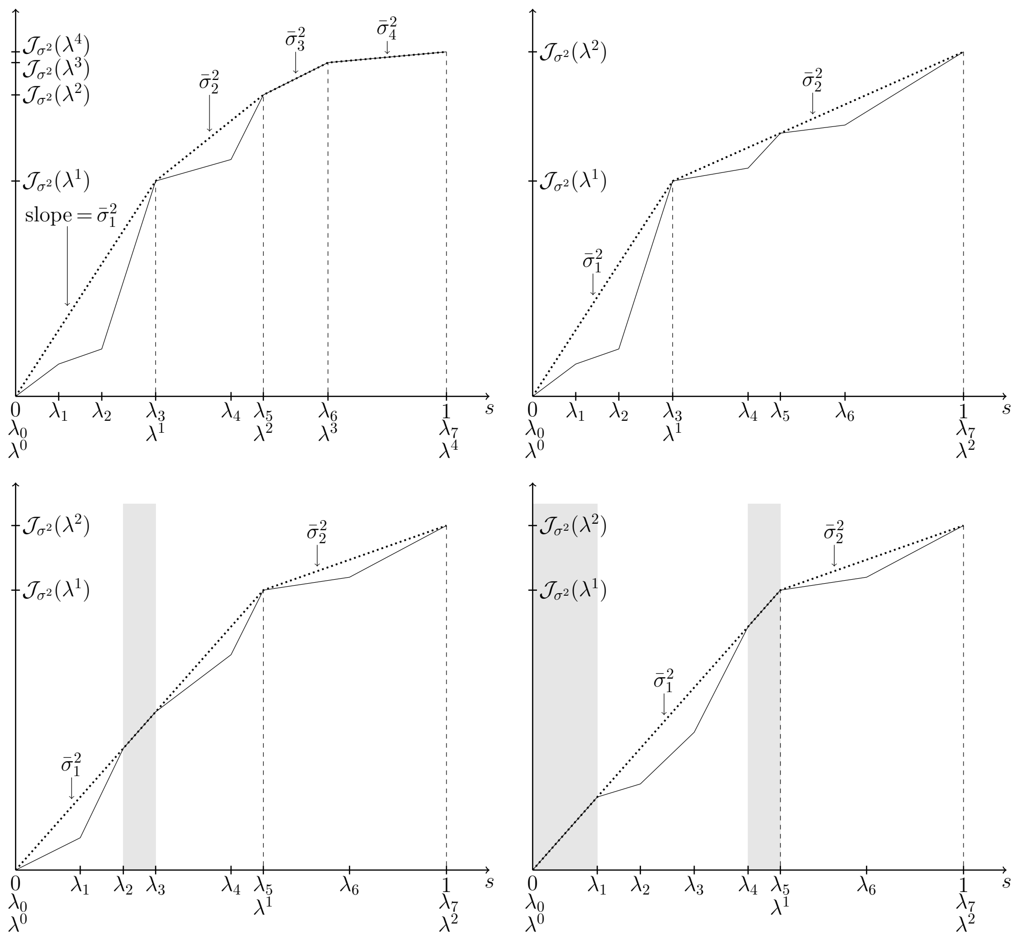


FIGURE 3.1.2. Examples of $\mathcal{J}_{\sigma^2}$ (closed lines) and $\hat{\mathcal{J}}_{\sigma^2}$ (dotted lines).

It is easy to see that there exists a unique non-increasing left-continuous step function $s \mapsto \bar{\sigma}(s)$ such that

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

The scales in $[0,1]$ where $\bar{\sigma}$ jumps are denoted by

$$0 \mathrel{\hat{=}} \lambda^0 < \lambda^1 < \ldots < \lambda^m \mathrel{\hat{=}} 1, \tag{3.1.2}$$

where $m \leq M$. As we will see in Theorem 3.1.3, the *effective scale parameters* λ^j and the *effective variance parameters* $\bar{\sigma}(\lambda^j)$ are the only parameters needed to fully determine the first and second order of the maximum for inhomogeneous branching random walks.

To be consistent with previous notations, we set $\bar{\sigma}_j \mathrel{\hat{=}} \bar{\sigma}(\lambda^j)$ and $t^j \mathrel{\hat{=}} \lambda^j n$. We write ∇_j for the difference operator with respect to the index j. When the index variable is obvious, we omit the subscript. For example, $\nabla t^j = t^j - t^{j-1}$.

To simplify the presentation of the proof of the main theorem, we impose a restriction on the variance parameters.

Restriction 3.1.2. *If $\mathcal{J}_{\sigma^2}$ and $\mathcal{J}_{\bar{\sigma}^2}$ coincide on a subinterval of $[\lambda^{j-1}, \lambda^j]$ for some j, then they must coincide everywhere on $[\lambda^{j-1}, \lambda^j]$.*

Remark 3.1.2. *Note that $\mathcal{J}_{\sigma^2}$ and $\mathcal{J}_{\bar{\sigma}^2}$ can still coincide at isolated points in $(\lambda^{j-1}, \lambda^j)$ when they do not coincide everywhere in that interval. The union of all the scales λ^j and all the isolated points where $\mathcal{J}_{\sigma^2}$ and $\mathcal{J}_{\bar{\sigma}^2}$ coincide form a subset of the scale parameters, say $\{\lambda_{i_d}\}_{0 \leq d \leq p}$, where $m \leq p \leq M$.*

For example, in Figure 3.1.2, the two models at the top satisfy Restriction 3.1.2, but the two models at the bottom do not. For the top models, the sets of scales described in Remark 3.1.2 are respectively $\{\lambda_0, \lambda_3, \lambda_5, \lambda_6, \lambda_7\}$ and $\{\lambda_0, \lambda_3, \lambda_5, \lambda_7\}$.

The main result of this paper is the derivation of the second order of the maximum (up to an $O_{\mathbb{P}}(1)$ error) for the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-BRW of Definition 3.1.1, under Restriction 3.1.2. This was an open problem in Fang and Zeitouni (2012a).

Theorem 3.1.3. *Let $\{S_v\}_{v\in\mathbb{D}_n}$ be as in Definition 3.1.1, under Restriction 3.1.2. Let $g \mathrel{\hat{=}} \sqrt{2\log 2}$. For all $\varepsilon > 0$, there exists $K_\varepsilon > 0$ such that for all $n \in \mathbb{N}$,*

$$\mathbb{P}\left(\left|\max_{v\in\mathbb{D}_n} S_v - \sum_{j=1}^{m}\left[g\bar{\sigma}_j \nabla t^j - \frac{(1+2\cdot\delta_j)\bar{\sigma}_j}{2g}\log(\nabla t^j)\right]\right| \geq K_\varepsilon\right) < \varepsilon,$$

where $\delta_j \mathrel{\hat{=}} 1$ when $\mathcal{J}_{\sigma^2}$ and $\mathcal{J}_{\bar{\sigma}^2}$ coincide on $[\lambda^{j-1}, \lambda^j]$, and $\delta_j \mathrel{\hat{=}} 0$ otherwise.

This theorem was proved in Fang and Zeitouni (2012a) for the case $M = 2$ and $\lambda_1 = 1/2$. Note that Restriction 3.1.2 is always satisfied when $M = 2$.

3.1.3. Related works

The first order of the maximum (without restriction),

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\max_{v\in\mathbb{D}_n} S_v - \sum_{j=1}^{m} g\bar{\sigma}_j \nabla t^j\right| > \varepsilon n\right) = 0, \quad \forall \varepsilon > 0,$$

was proved in Section 2 of Ouimet (2014) for the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-BRW and in Arguin and Ouimet (2016) for the analogous model of scale-inhomogeneous Gaussian free field (GFF). The proofs rely on an analysis of so-called "optimal paths" showing where the maximal particle must be at all times with high probability. These paths were found by a first moment heuristic and the resolution of a related optimisation problem (using the Karush-Kuhn-Tucker theorem).

The more involved question of finding the second order of the maximum was first solved by Fang and Zeitouni (2012a) for the case $M = 2$ and $\lambda_1 = 1/2$, and later by Mallein (2015a), when the law of the increments changes a finite number of times but is not necessarily Gaussian. In his proof, Mallein develops a time-inhomogeneous version of the spinal decomposition for the BRW. The argument presented in this paper was first developed, without the knowledge of Mallein's results, in Section 2.4 of Ouimet (2014) and instead generalizes the approach of Fang and Zeitouni (2012a). The proof rely on the control of the increments of high points at every *effective scale* λ^j.

One shortfall of the spinal decomposition is that it completely relies on the presence of an *exact* branching structure. Specifically, a crucial step in Mallein (2015a) is the proof of a time-inhomogeneous version of the classical many-to-one lemma, which is a direct consequence of his comparison between the size-biased law of the BRW (the usual change of measure) and a certain projection of a law on the set of planar rooted marked trees with spine.

In contrast, our method can be adapted to a number of cases where the branching structure is only *approximate*. For instance, although no explicit proof is written down, it can be applied to prove the second order of the maximum for the scale-inhomogeneous GFF of Arguin and Ouimet (2016). The model differs from the time-inhomogeneous BRW in two ways :

(1) The branching structure is *approximate* in the sense that increments of the field that are below the *branching scale* are not perfectly correlated and they decorrelate smoothly near the branching scale.
(2) At a given scale, the covariance of the increments of the field decays near the boundary of the domain. In the context of BRWs, this means that at a given time, the law of each point process would depend on the position of the associated ancestors in the tree.

The recent developments in the study of

- cover times (see e.g. Abe (2014, 2018); Belius (2013); Belius and Kistler (2017); Comets et al. (2013); Dembo et al. (2003, 2004, 2006); Ding (2012, 2014); Ding et al. (2012); Ding and Zeitouni (2012));
- the extremes of the randomized Riemann zeta function on the critical line (see e.g. Arguin et al. (2017b); Arguin and Ouimet (2019); Arguin and Tai (2018); Harper (2013); Ouimet (2018); Saksman and Webb (2018));
- the maxima of the Riemann zeta function on random intervals of the critical line (see e.g. Arguin et al. (2019); Najnudel (2018));
- the maxima of the characteristic polynomials of random unitary matrices (see e.g. Arguin et al. (2017a); Chhaibi et al. (2018); Paquette and Zeitouni (2018));
- etc.

show that approximate branching structures are present in a huge variety of models. Hence, the approach of this paper might become relevant in applications beyond the study of “pure” BRW.

For other recent and relevant results on branching processes in time-inhomogeneous environments, the reader is referred to Bovier and Hartung (2014, 2015); Bovier and

Kurkova (2004a,b); Chen (2018); Fang and Zeitouni (2012b); Maillard and Zeitouni (2016); Mallein (2015b); Mallein and Piotr (2018); Ouimet (2017).

3.2. Proof of the main result

3.2.1. Preliminaries

For all $v \in \mathbb{D}_n$ and $k, l \in \{1, \dots, M\}$, we can compute from Definition 3.1.1 :

$$\mathbb{V}(S_v(t_l) - S_v(t_{k-1})) = \sum_{i=k}^{l} \sigma_i^2 \nabla t_i = \mathcal{J}_{\sigma^2}(\lambda_{k-1}, \lambda_l)\, n. \tag{3.2.1}$$

The variance of the increments in (3.2.1) will be used repeatedly during the proofs in conjunction with the following lemma.

Lemma 3.2.1 (Gaussian estimates, see e.g. Adler and Taylor (2007))**.** *Suppose that* $Z \sim \mathcal{N}(0, \sigma^2)$ *where* $\sigma > 0$*, then for all* $z > 0$*,*

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

The particle achieving the maximum of the BRW at time n act like a Brownian bridge around the maximum level on all the intervals $[t^{j-1}, t^j]$ where $\mathcal{J}_{\sigma^2}(\cdot/n)$ and $\mathcal{J}_{\bar{\sigma}^2}(\cdot/n)$ coincide. The extra log terms in Theorem 3.1.3 (when $\delta_j = 1$) compensate for the "cost" of the Brownian bridge to stay below a certain logarithmic barrier. The sets $\mathcal{A}_l$ below identify the indices j of these intervals up to scale λ^l. The sets $\mathcal{T}_l$ consist of the *effective times* t^j, $1 \leq j \leq l$, and the integer times in $[t^{j-1}, t^j]$, $j \in \mathcal{A}_l$, where a Brownian bridge estimate will be needed. More precisely, for all $l \in \{1, \dots, m\}$,

$$\begin{aligned}
\mathcal{A}_l &\mathrel{\hat{=}} \{j \in \{1, \dots, l\} : \delta_j = 1\} = \left\{ j \in \{1, \dots, l\} : \mathcal{J}_{\sigma^2}|_{[\lambda^{j-1}, \lambda^j]} \equiv \mathcal{J}_{\bar{\sigma}^2}|_{[\lambda^{j-1}, \lambda^j]} \right\}, \\
\mathcal{T}_l &\mathrel{\hat{=}} \{t^1, t^2, \dots, t^l\} \cup \bigcup_{j \in \mathcal{A}_l} \{t^{j-1}, t^{j-1} + 1, \dots, t^j\}.
\end{aligned}$$

Let $\vartheta_k \in \{1, \dots, m\}$ be the index such that $t^{\vartheta_k - 1} < k \leq t^{\vartheta_k}$. For all $k \in \{0, \dots, n\}$, the **concave hull** of the *optimal path for the maximum* is

$$M_n^\star(k) \mathrel{\hat{=}} \sum_{j=1}^{\vartheta_k} \frac{(k \wedge t^j - t^{j-1})}{\nabla t^j} \left[g \bar{\sigma}_j \nabla t^j - \frac{(1 + 2 \cdot \delta_j) \bar{\sigma}_j}{2g} \log(\nabla t^j) \right] \tag{3.2.2}$$

where $g \circeq \sqrt{2\log 2}$, as in Theorem 3.1.3. We refer the reader to Ouimet (2014) or Arguin and Ouimet (2016) for a first moment heuristic. Note that $M_n^\star$ and the optimal path coincide on $\mathcal{T}_m$, see Figure 3.2.3 for an example of $M_n^\star$ under Restriction 3.1.2.

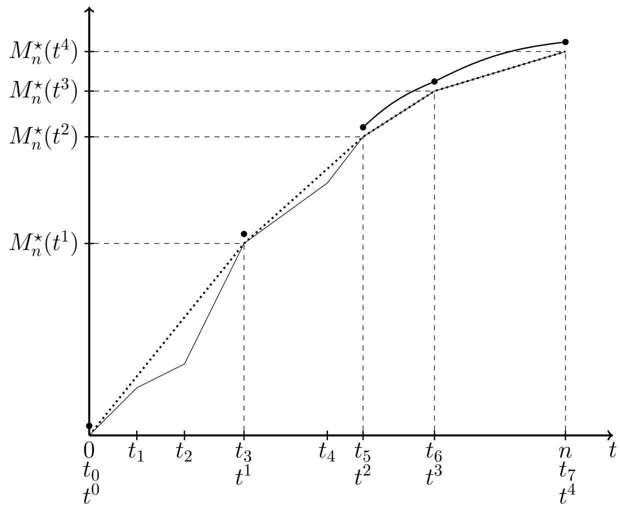


FIGURE 3.2.3. Example of the path $M_{n,x}^\star$ on the set $\mathcal{T}_m$ (in bold), the optimal path (thin line) and its concave hull $M_n^\star$ (dotted line).

For all $k \in \mathcal{T}_m$, define the logarithmic barrier as

$$b_n(k) \circeq \begin{cases} 0, & \text{if } k \in \{t^0, t^1, \dots, t^m\}, \\ \frac{5}{2}\frac{\bar{\sigma}_{\vartheta_k}}{g}\log(k - t^{\vartheta_k - 1}), & \text{if } \vartheta_k \in \mathcal{A}_m,\ t^{\vartheta_k-1} < k \leq \frac{t^{\vartheta_k-1}+t^{\vartheta_k}}{2}, \\ \frac{5}{2}\frac{\bar{\sigma}_{\vartheta_k}}{g}\log(t^{\vartheta_k} - k), & \text{if } \vartheta_k \in \mathcal{A}_m,\ \frac{t^{\vartheta_k-1}+t^{\vartheta_k}}{2} < k < t^{\vartheta_k}. \end{cases} \tag{3.2.3}$$

For all $x > 0$, denote

$$b_{n,x}(k) \circeq b_n(k) + x \qquad \text{and} \qquad M_{n,x}^\star(k) \circeq M_n^\star(k) + b_{n,x}(k).$$

Let us now define precisely what is meant by a Brownian bridge.

Definition 3.2.2 (Discrete Brownian bridge). *Let $0 \leq \lambda < \lambda' \leq 1$ be such that $\lambda n, \lambda' n \in \mathbb{N}_0$ and $\sigma > 0$. The discrete σ-Brownian bridge on the interval $[\lambda n, \lambda' n]$ is a centered Gaussian vector $(B_k)_{k=\lambda n}^{\lambda' n}$ such that*

(a) $B_{\lambda n} = B_{\lambda' n} = 0$,

(b) $Cov(B_k, B_{k'}) = \frac{(k\wedge k' - \lambda n)(\lambda' n - k \vee k')}{(\lambda' - \lambda)n}\sigma^2, \quad k, k' \in \{\lambda n, \lambda n + 1, \dots, \lambda' n\}$.

Here are relevant examples of discrete Brownian bridges constructed from a discrete random walk.

Lemma 3.2.3. *Let $v \in \mathbb{D}_n$ and $j \in \mathcal{A}_m$. Then, the centered Gaussian vector*

$$B^j_{v,i} \mathrel{\hat{=}} S_v(i) - S_v(t^{j-1}) - \frac{i - t^{j-1}}{\nabla t^j} \nabla S_v(t^j), \quad t^{j-1} \le i \le t^j, \tag{3.2.4}$$

is independent of $\{S_v(i')\}_{i' \notin (t^{j-1}, t^j)}$ and defines a discrete $\bar{\sigma}_j$-Brownian bridge under Definition 3.2.2. Similarly, when $l \in \mathcal{A}_m$ and $t^{l-1} < k \le t^l$, the centered Gaussian vector

$$B_{v,i} \mathrel{\hat{=}} S_v(i) - S_v(t^{l-1}) - \frac{i - t^{l-1}}{k - t^{l-1}} (S_v(k) - S_v(t^{l-1})), \quad t^{l-1} \le i \le k, \tag{3.2.5}$$

is independent of $\{S_v(i')\}_{i' \notin (t^{l-1}, k)}$ and defines a discrete $\bar{\sigma}_l$-Brownian bridge.

Proof. We only prove (3.2.4) since the proof of (3.2.5) is totally analogous. Assume $j \in \mathcal{A}_m$, meaning that $\sigma(s) = \bar{\sigma}_j$ for all $s \in (\lambda^{j-1}, \lambda^j]$. Then, for all $i' \in \{0, 1, \dots, t^{j-1}\} \cup \{t^j, t^j + 1, \dots, n\}$, $\mathrm{Cov}\Big(B^j_{v,i}, S_v(i')\Big)$ is equal to

$$\begin{aligned}
&\mathbb{V}(S_v(i \wedge i')) - \mathbb{V}\Big(S_v(t^{j-1} \wedge i')\Big) - \frac{i - t^{j-1}}{\nabla t^j} \nabla_j \mathbb{V}\Big(S_v(t^j \wedge i')\Big) \\
&= \left\{ \begin{array}{ll} \mathbb{V}(S_v(i)) - \mathbb{V}(S_v(t^{j-1})) - \frac{i - t^{j-1}}{\nabla t^j} \nabla_j \mathbb{V}(S_v(t^j)) & \text{if } t^j \le i' \le n \\ 0 - \frac{i - t^{j-1}}{\nabla t^j} 0 & \text{if } 0 \le i' \le t^{j-1} \end{array} \right\} \\
&\overset{(3.2.1)}{=} \left\{ \begin{array}{ll} \bar{\sigma}_j^2 (i - t^{j-1}) - \frac{i - t^{j-1}}{\nabla t^j} \bar{\sigma}_j^2 \nabla t^j & \text{if } t^j \le i' \le n \\ 0 & \text{if } 0 \le i' \le t^{j-1} \end{array} \right\} = 0.
\end{aligned}$$

The first claim follows since $\{B^j_{v,i}\}_{i \in \{t^{j-1}, \dots, t^j\}}$ and $\{S_v(i')\}_{i' \notin (t^{j-1}, t^j)}$ form a Gaussian vector together. For the second claim, we need to verify (a) and (b) in Definition 3.2.2 :

(a) We obviously have $B^j_{v,t^{j-1}} = B^j_{v,t^j} = 0$;

(b) For all $i, i' \in \{t^{j-1}, t^{j-1} + 1, \dots, t^j\}$,

$$\begin{aligned}
\mathrm{Cov}\Big(B^j_{v,i}, B^j_{v,i'}\Big) \;=\; &\mathrm{Cov}\Big(S_v(i) - S_v(t^{j-1}), S_v(i') - S_v(t^{j-1})\Big) \\
&- \frac{i - t^{j-1}}{\nabla t^j} \mathrm{Cov}\Big(\nabla S_v(t^j), S_v(i') - S_v(t^{j-1})\Big) \\
&- \frac{i' - t^{j-1}}{\nabla t^j} \mathrm{Cov}\Big(S_v(i) - S_v(t^{j-1}), \nabla S_v(t^j)\Big) \\
&+ \frac{(i - t^{j-1})(i' - t^{j-1})}{(\nabla t^j)^2} \mathbb{V}\Big(\nabla S_v(t^j)\Big)
\end{aligned}$$

$$
\begin{aligned}
&\overset{(3.2.1)}{=} (i \wedge i' - t^{j-1})\bar{\sigma}_j^2 - 2\frac{(i - t^{j-1})(i' - t^{j-1})}{\nabla t^j}\bar{\sigma}_j^2 \\
&\quad + \frac{(i - t^{j-1})(i' - t^{j-1})}{(\nabla t^j)^2}\bar{\sigma}_j^2 \nabla t^j \\
&= \frac{(i \wedge i' - t^{j-1})(t^j - i \vee i')}{\nabla t^j}\bar{\sigma}_j^2.
\end{aligned}
$$

This ends the proof of the lemma. □

Finally, to estimate the probability that a discrete Brownian bridge stays below a logarithmic barrier such as the one defined in (3.2.3), we adapt Proposition 1' of Bramson (1978).

Lemma 3.2.4 (Discrete Brownian bridge estimates)**.** *Let $0 \leq \lambda < \lambda' \leq 1$ be such that $\lambda n, \lambda' n \in \mathbb{N}_0$ and $\sigma > 0$. Let $(B_k)_{k=\lambda n}^{\lambda' n}$ be a discrete σ-Brownian bridge on the interval $[\lambda n, \lambda' n]$. For any constant $D = D(\lambda, \lambda', \sigma) > 0$ and the logarithmic barrier*

$$
b(k) = \begin{cases} 0, & \text{if } k \in \{\lambda n, \lambda' n\}, \\ D \log(k - \lambda n), & \text{if } \lambda n < k \leq \frac{\lambda n + \lambda' n}{2}, \\ D \log(\lambda' n - k), & \text{if } \frac{\lambda n + \lambda' n}{2} < k < \lambda' n, \end{cases}
$$

there exists a constant $C = C(D, \sigma) > 0$ such that for all $z > 0$ and all $n \in \mathbb{N}$,

$$
\mathbb{P}(B_k < b(k) + z, \ \lambda n \leq k \leq \lambda' n) \leq C\frac{(1+z)^2}{(\lambda' - \lambda)n}.
$$

In order to prove Lemma 3.2.4, we first need to prove that a random walk with Gaussian increments stays below the first part of the logarithmic barrier $b(\cdot) + z$ with probability $O(n^{-1/2})$. This is achieved through the following lemma, which is the analogue of Proposition 1 in Bramson (1978).

Lemma 3.2.5. *Let $\sigma > 0$ and let $(S_k)_{k=0}^t$ be a discrete random walk with $\mathcal{N}(0, \sigma^2)$ increments and $S_0 \stackrel{\circ}{=} 0$. For any constant $D = D(\lambda, \lambda', \sigma) > 0$ and the logarithmic barrier*

$$
\widetilde{b}(k) = \begin{cases} 0, & \text{if } k = 0, \\ D \log k, & \text{if } 0 < k \leq t, \end{cases}
$$

there exists a constant $C = C(D,\sigma) > 0$ such that for all $z > 0$ and all $t \in \mathbb{N}$,

$$\mathbb{P}\Big(S_k < \widetilde{b}(k) + z,\ 0 \leq k \leq t\Big) \leq C\frac{(1+z)}{t^{1/2}}.$$

Remark 3.2.1. *Throughout the proofs of this article, $c, C, \widetilde{C}$, etc., will denote positive constants whose value can change from line to line and can depend on the parameters $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$. For simplicity, equations are always implicitly stated to hold for n large enough when needed.*

Proof. Let $z > 0$ and $t \in \mathbb{N}$. When $t = 1$, the statement is trivially satisfied with $C \geq 1$. Therefore, assume $C \geq 1$ and $t \geq 2$ for the rest of the proof. Let $q_t = \lfloor D \log t \rfloor$ and for all $x > 0$, let $\tau_x \triangleq \inf\{k \geq 1 : S_k \geq x\}$. Then,

$$\begin{aligned}
&\mathbb{P}\Big(S_k < \widetilde{b}(k) + z,\ 0 \leq k \leq t\Big) \\
&\leq \mathbb{P}\Big(\max_{0 \leq k \leq t} S_k < z\Big) + \sum_{i=0}^{q_t} \mathbb{P}\begin{pmatrix} \lfloor e^{i/D} \rfloor \leq \tau_{z+i} \leq t \ \text{ and} \\ S_{\tau_{z+i}} < z + i + 1 \ \text{ and} \\ \max_{\tau_{z+i} \leq k \leq t}(S_k - S_{\tau_{z+i}}) < 1 \end{pmatrix}.
\end{aligned} \tag{3.2.6}$$

We bound the first probability in (3.2.6) using a standard gambler's ruin estimate. Indeed, from Theorem 5.1.7 in Lawler and Limic (2010), there exists a constant $C' = C'(\sigma) > 0$ such that for all $z > 0$ and all $t \in \mathbb{N}$,

$$\mathbb{P}\Big(\max_{0 \leq k \leq t} S_k < z\Big) \leq C' \frac{z+1}{t^{1/2}}. \tag{3.2.7}$$

We proceed to the individual summands in (3.2.6). The strong Markov property for the random walk implies

$$\begin{aligned}
&\mathbb{P}\begin{pmatrix} \lfloor e^{i/D} \rfloor \leq \tau_{z+i} \leq t \ \text{ and} \\ S_{\tau_{z+i}} < z + i + 1 \ \text{ and} \\ \max_{\tau_{z+i} \leq k \leq t}(S_k - S_{\tau_{z+i}}) < 1 \end{pmatrix} \\
&= \sum_{j=\lfloor e^{i/D} \rfloor}^{t} \mathbb{P}\Big(\tau_{z+i} = j,\ S_{\tau_{z+i}} < z + i + 1\Big) \cdot \mathbb{P}\Big(\max_{0 \leq k \leq t-j} S_k < 1\Big) \\
&= \sum_{j=\lfloor e^{i/D} \rfloor \wedge (1+\lfloor t/2 \rfloor)}^{\lfloor t/2 \rfloor} + \sum_{j=\lfloor e^{i/D} \rfloor \vee (1+\lfloor t/2 \rfloor)}^{t}.
\end{aligned} \tag{3.2.8}$$

Now, for the first summation in (3.2.8), we have

$$\sum_{j=\lfloor e^{i/D}\rfloor\wedge(1+\lfloor t/2\rfloor)}^{\lfloor t/2\rfloor} \mathbb{P}\Big(\tau_{z+i}=j,\ S_{\tau_{z+i}}<z+i+1\Big)\cdot\mathbb{P}\Big(\max_{0\leq k\leq t-j} S_k<1\Big)$$
$$\mathbb{P}\left(\max_{0\leq k\leq \lfloor e^{i/D}\rfloor} S_k<z+i+1\right)\cdot\mathbb{P}\left(\max_{0\leq k\leq t-\lfloor t/2\rfloor} S_k<1\right)\leq C\frac{z+i+1}{t^{1/2}}e^{-i/(2D)}. \tag{3.2.9}$$

We applied the estimate (3.2.7) to both terms on the second line and we used the fact that $(z+i+2)/(z+i+1)\leq 2$ for all $(z,i)\in(0,\infty)\times\mathbb{N}$ to obtain the last inequality.

For the second summation in (3.2.8), we can use an estimate closely related to the first hitting time distribution in the gambler's ruin problem. Indeed, from Lemma 3 in Mogul'skiĭ (2009), there exists a constant $C''=C''(\sigma)>0$ such that for all $x>0$ and all $j\in\mathbb{N}$,

$$\mathbb{P}(\tau_x=j,\ S_{\tau_x}<x+1)\leq\mathbb{P}\begin{pmatrix} S_j\in[x,x+1] \text{ and} \\ S_j=\max_{0\leq k\leq j}S_k\end{pmatrix}\leq C''\frac{x+1}{j^{3/2}}. \tag{3.2.10}$$

Using successively (3.2.10), the gambler's ruin estimate (3.2.7), the change of variable $j'=t-j$ and the fact that $s\mapsto s^{-1/2}$ is decreasing, we have

$$\begin{aligned}
&\sum_{j=\lfloor e^{i/D}\rfloor\vee(1+\lfloor t/2\rfloor)}^{t} \mathbb{P}\Big(\tau_{z+i}=j,\ S_{\tau_{z+i}}<z+i+1\Big)\cdot\mathbb{P}\Big(\max_{0\leq k\leq t-j} S_k<1\Big)\\
&\leq \sum_{j=\lfloor e^{i/D}\rfloor\vee(1+\lfloor t/2\rfloor)}^{t} C''\frac{z+i+1}{j^{3/2}}\cdot\mathbb{P}\Big(\max_{0\leq k\leq t-j} S_k<1\Big)\\
&\leq 2^{3/2}C''\frac{z+i+1}{t^{3/2}}\cdot\left\{1+\sum_{j'=1}^{\lfloor t/2\rfloor}C'\frac{2}{(j')^{1/2}}\right\}\\
&\leq 2^{3/2}C''\frac{z+i+1}{t^{3/2}}\cdot\left\{4(1+C')\int_0^t\frac{1}{2s^{1/2}}ds\right\}=C\frac{z+i+1}{t}.
\end{aligned} \tag{3.2.11}$$

From (3.2.8), (3.2.9) and (3.2.11), we deduce

$$\mathbb{P}\begin{pmatrix} \lfloor e^{i/D}\rfloor\leq\tau_{z+i}\leq t \ \text{ and} \\ S_{\tau_{z+i}}<z+i+1 \ \text{ and} \\ \max_{\tau_{z+i}\leq k\leq t}(S_k-S_{\tau_{z+i}})<1\end{pmatrix}\leq C^\star\frac{z+i+1}{t^{1/2}}e^{-i/(2D)} \tag{3.2.12}$$

for a certain constant $C^\star=C^\star(\sigma)>0$, since $t^{-1/2}\leq e^{-i/(2D)}$ for all $i\leq q_t$.

Note that $(z+i+1)\leq(z+1)(i+1)$ for all $i\geq 0$. Therefore, by applying (3.2.12) and

(3.2.7) in (3.2.6), we get

$$\mathbb{P}\Big(S_k < \widetilde{b}(k) + z,\ 0 \le k \le t\Big) \le C' \frac{z+1}{t^{1/2}} + C^\star \frac{z+1}{t^{1/2}} \sum_{i=0}^{q_t} (i+1) e^{-i/(2D)}.$$

The conclusion holds since $\sum_{i=0}^{\infty}(i+1)e^{-i/(2D)} < \infty$. $\square$

Now, the proof of Lemma 3.2.4 is exactly the same (except in discrete time) as the proof of Proposition 1' in Bramson (1978) for the case $s_0 = t$. We give the details for completeness.

Proof of Lemma 3.2.4. Without loss of generality, assume that $\lambda = 0$, $\lambda' = 1$ and $n/3 \in \mathbb{N}$. Let $(B_k)_{k=0}^n$ be a discrete σ-Brownian bridge and let $(S_k)_{k=0}^n$ be a discrete random walk with $\mathcal{N}(0,\sigma^2)$ increments and $S_0 \mathrel{\hat=} 0$. Denote by P_{b_1}, P_{b_2} and P_{b_3}, the sets of discrete paths in $\{0,1,\dots,n\}$ lying below the barrier $b(\cdot)+z$ on the sets $\{0,\dots,n/3\}$, $\{n/3,\dots,2n/3\}$ and $\{2n/3,\dots,n\}$ respectively. Using the Markov property of B and S,

$$\begin{aligned}
(\star) &\mathrel{\hat=} \mathbb{P}(B_k < b(k) + z,\ 0 \le k \le n) \\
&= \int_{-\infty}^{b(n/3)+z} \int_{-\infty}^{b(2n/3)+z} \mathbb{P}\Big(B \in P_{b_1} | B_{n/3} = x_1\Big) f_{B_{n/3}}(x_1) \\
&\qquad\qquad \times \mathbb{P}\Big(B \in P_{b_2} | B_{2n/3} = x_2, B_{n/3} = x_1\Big) f_{B_{2n/3}|B_{n/3}}(x_2|x_1) \\
&\qquad\qquad \times \mathbb{P}\Big(B \in P_{b_3} | B_{2n/3} = x_2\Big)\, dx_1 dx_2 \\
&= \frac{1}{f_{S_n}(0)} \left\{ \begin{array}{l}
\int_{-\infty}^{b(n/3)+z} \int_{-\infty}^{b(2n/3)+z} \mathbb{P}\Big(S \in P_{b_1} | S_{n/3} = x_1\Big) f_{S_{n/3}}(x_1) \\
\times \mathbb{P}\Big(S \in P_{b_2} | S_{2n/3} = x_2, S_{n/3} = x_1\Big) f_{S_{2n/3}|S_{n/3}}(x_2|x_1) \\
\times \mathbb{P}\Big(S \in P_{b_3} | S_n = 0, S_{2n/3} = x_2\Big) f_{S_n|S_{2n/3}}(0|x_2)\ dx_1 dx_2
\end{array} \right\}.
\end{aligned}$$

But, we have $\mathbb{P}\Big(S \in P_{b_2} | S_{2n/3} = x_2, S_{n/3} = x_1\Big) \le 1$, and $f_{S_{2n/3}|S_{n/3}}(x_2|x_1) \le f_{S_{n/3}}(0)$, and $f_{S_{n/3}}(0)/f_{S_n}(0) = \sqrt{3}$, so

$$(\star) \le \sqrt{3} \left\{ \begin{array}{l}
\int_{-\infty}^{b(n/3)+z} \mathbb{P}\Big(S \in P_{b_1} | S_{n/3} = x_1\Big) f_{S_{n/3}}(x_1) dx_1 \\
\times \int_{-\infty}^{b(2n/3)+z} \mathbb{P}\left(S \in P_{b_3} \,\middle|\, \begin{array}{l} S_n = 0, \\ S_{2n/3} = x_2 \end{array}\right) f_{S_n|S_{2n/3}}(0|x_2) dx_2
\end{array} \right\}.$$

By the symmetry of $b(\cdot)$ around $n/2$, both integrals are exactly the same. Thus, the right-hand side is equal to

$$\sqrt{3}\Big(\mathbb{P}\Big(S_k < \widetilde{b}(k) + z,\ 0 \le k \le n/3\Big)\Big)^2.$$

The conclusion follows directly from Lemma 3.2.5. □

3.2.2. Why Restriction 3.1.2 ?

Let $\pi_j \in \{0, 1, \dots, M\}$ denote the indice such that $\lambda_{\pi_j} = \lambda^j$. When the continuous and piecewise linear functions $\mathcal{J}_{\sigma^2}$ and $\mathcal{J}_{\bar{\sigma}^2}$ coincide on a subinterval of $[\lambda^{j-1}, \lambda^j]$, they either coincide

(1) everywhere on $[\lambda^{j-1}, \lambda^j]$;

(2) everywhere on the left and right end, meaning on $[\lambda^{j-1}, \lambda_{\pi_{j-1}+1}]$ and $[\lambda_{\pi_j - 1}, \lambda^j]$ respectively, but not somewhere in $(\lambda_{\pi_{j-1}+1}, \lambda_{\pi_j - 1})$;

(3) everywhere on the left end, but not on the right end;

(4) everywhere on the right end, but not on the left end.

Imposing Restriction 3.1.2 means that we only deal with the first case. The only reason we do this is to avoid overburdening the notation in the proof of Theorem 3.1.3 by dividing each interval $[t^{j-1}, t^j]$, $j \in \mathcal{A}_m$, in three parts like we did in the proof of Lemma 3.2.4.

From Lemma 3.2.5, the probability that the left (resp. right) end of a Brownian bridge stays below the left (resp. right) end of the logarithmic barrier $b(\cdot) + z$ is $O(n^{-1/2})$. The probability that the middle part of the Brownian bridge stays below the middle part of the logarithmic barrier is $O(1)$. Thus, it should now be obvious how to modify the statement of Theorem 3.1.3 when there is no restriction. Simply replace $2 \cdot \delta_j$ by $\delta_j^{\text{left}} + \delta_j^{\text{right}}$, where

$$\delta_j^{\text{left}} \mathrel{\hat{=}} \begin{cases} 1, & \text{when } \mathcal{J}_{\sigma^2} \text{ and } \mathcal{J}_{\bar{\sigma}^2} \text{ coincide on } [\lambda^{j-1}, \lambda_{\pi_{j-1}+1}], \\ 0, & \text{otherwise,} \end{cases}$$

$$\delta_j^{\text{right}} \mathrel{\hat{=}} \begin{cases} 1, & \text{when } \mathcal{J}_{\sigma^2} \text{ and } \mathcal{J}_{\bar{\sigma}^2} \text{ coincide on } [\lambda_{\pi_j - 1}, \lambda^j], \\ 0, & \text{otherwise.} \end{cases}$$

For confirmation, the reader is referred to Theorem 1.4 in Mallein (2015a), where a more general statement is given.

3.2.3. Second order of the maximum and tension

Theorem 3.1.3 is a direct consequence of Lemma 3.2.6, which proves the exponential decay of the probability that the recentered maximum is above a certain level, and Lemma 3.2.8, which shows the corresponding lower bound.

Lemma 3.2.6 (Upper bound)**.** *Let $\{S_v\}_{v\in\mathbb{D}_n}$ be the $(\boldsymbol{\sigma},\boldsymbol{\lambda})$-BRW at time n of Definition 3.1.1, under Restriction 3.1.2. Recall the definition of $M_n^\star$ from (3.2.2). There exists a constant $C = C(\boldsymbol{\sigma},\boldsymbol{\lambda}) > 0$ such that for all $x > 0$,*

$$\mathbb{P}\left(\max_{v\in\mathbb{D}_n} S_v \geq M_n^\star(n) + x\right) \leq C\,(1+x)^{2\sum_{j=1}^m \delta_j} e^{-x\frac{g}{\bar{\sigma}_1}}$$

for n large enough, where $\delta_j \mathrel{\hat{=}} \mathbf{1}_{\{j\in\mathcal{A}_m\}}$.

The proof of Lemma 3.2.6 is separated in two parts with a technical lemma in between them (Lemma 3.2.7).

Proof of Lemma 3.2.6 (first part). Define the set of particles that are above the path $M_{n,x}^\star$ at time k :

$$\mathcal{H}_{k,n,x} \mathrel{\hat{=}} \{v \in \mathbb{D}_k : S_v(k) \geq M_{n,x}^\star(k)\}, \quad k \in \mathcal{T}_m.$$

The idea of the proof is to split the probability that at least one particle at time n exceeds $M_{n,x}^\star(n)$ by looking at the first time $k \in \mathcal{T}_m$ when the set $\mathcal{H}_{k,n,x}$ is non-empty. Using sub-additivity, we have the following upper bound on the probability of the lemma :

$$\begin{aligned}\mathbb{P}(|\mathcal{H}_{n,n,x}| \geq 1) &\leq \sum_{k\in\mathcal{T}_m} \mathbb{P}\begin{pmatrix} |\mathcal{H}_{k,n,x}| \geq 1 \text{ and } |\mathcal{H}_{i,n,x}| = 0 \\ \forall i \in \mathcal{T}_m \text{ such that } i < k \end{pmatrix} \\ &\leq \sum_{k\in\mathcal{T}_m} 2^k \max_{v\in\mathbb{D}_k} \mathbb{P}\begin{pmatrix} S_v(k) \geq M_{n,x}^\star(k) \\ \text{and } S_v(i) < M_{n,x}^\star(i) \\ \forall i \in \mathcal{T}_m \text{ such that } i < k \end{pmatrix}. \end{aligned} \tag{3.2.13}$$

We only discuss the case $k > t^1$ from hereon. The case $k \leq t^1$ is easier (there is no conditioning in (3.2.14)), so we omit the details. Fix $l \in \{2,\dots,m\}$ and $t^{l-1} < k \leq t^l$ for the remaining of the proof. By conditioning on the event

$$E_v \mathrel{\hat{=}} \{(S_v(t^1),\dots,S_v(t^{l-1})) = (x_1,\dots,x_{l-1}) \mathrel{\hat{=}} \boldsymbol{x}\},$$

the probability in the maximum in (3.2.13) is equal to

$$\int_{-\infty}^{M^\star_{n,x}(t^1)} \cdots \int_{-\infty}^{M^\star_{n,x}(t^{l-1})} \underbrace{\mathbb{P}\left(\begin{array}{l} S_v(k) \geq M^\star_{n,x}(k) \\ \text{and } S_v(i) < M^\star_{n,x}(i) \\ \forall i \in \mathcal{T}_m \text{ such that } i < k \end{array} \middle| E_v \right)}_{\mathrel{\hat=} (\clubsuit)} f_v(\boldsymbol{x})\, d\boldsymbol{x}, \tag{3.2.14}$$

where f_v is the density function of $(S_v(t^1), \ldots, S_v(t^{l-1}))$.

Now, make the convenient change of variables

$$Y_{v,j} \mathrel{\hat=} \nabla S_v(t^j) - \nabla M^\star_n(t^j), \quad j \in \{1, \ldots, l-1\}.$$

By the independence of the increments, the density of the vector $(S_v(t^j))_{j=1}^{l-1}$ is the product of the densities of the $Y_{v,j}$'s, namely

$$f_v(\boldsymbol{x}) \mathrel{\hat=} f_v(x_1, \ldots, x_{l-1}) = f_{Y_{v,1}}(y_1) \cdot \ldots \cdot f_{Y_{v,l-1}}(y_{l-1}).$$

Since $\mathbb{V}(Y_{v,j}) = \mathbb{V}(\nabla S_v(t^j)) = \bar{\sigma}_j^2 \nabla t^j$, we can bound each density :

$$f_{Y_{v,j}}(y_j) = \frac{e^{-\frac{\left(y_j + \nabla M^\star_n(t^j)\right)^2}{2\bar{\sigma}_j^2 \nabla t^j}}}{\sqrt{2\pi}\sqrt{\bar{\sigma}_j^2 \nabla t^j}} \leq C 2^{-\nabla t^j} \frac{e^{\frac{(1+2\cdot\delta_j)}{2}\log(\nabla t^j)}}{\sqrt{\nabla t^j}} e^{-y_j \frac{g}{\bar{\sigma}_j}} = C 2^{-\nabla t^j} (\nabla t^j)^{\delta_j} e^{-y_j \frac{g}{\bar{\sigma}_j}}.$$

We deduce that the integral in (3.2.14) is smaller than

$$C 2^{-t^{l-1}} \int_{-\infty}^{x} \int_{-\infty}^{x-y_1} \cdots \int_{-\infty}^{x - \sum_{j=1}^{l-2} y_j} (\clubsuit) \cdot \prod_{j=1}^{l-1} (\nabla t^j)^{\delta_j} e^{-y_j \frac{g}{\bar{\sigma}_j}}\, d\boldsymbol{y}. \tag{3.2.15}$$

From Lemma 3.2.3, we know that for all $j \in \mathcal{A}_{l-1}$, the process

$$B^j_{v,i} \mathrel{\hat=} S_v(i) - S_v(t^{j-1}) - \frac{i - t^{j-1}}{\nabla t^j} \nabla S_v(t^j), \quad t^{j-1} \leq i \leq t^j, \tag{3.2.16}$$

is independent of $\{S_v(i')\}_{i' \notin (t^{j-1}, t^j)}$ and defines a discrete $\bar{\sigma}_j$-Brownian bridge. Similarly, when $l \in \mathcal{A}_m$, the process

$$B_{v,i} \mathrel{\hat=} S_v(i) - S_v(t^{l-1}) - \frac{i - t^{l-1}}{k - t^{l-1}} (S_v(k) - S_v(t^{l-1})), \quad t^{l-1} \leq i \leq k, \tag{3.2.17}$$

is independent of $\{S_v(i')\}_{i' \notin (t^{l-1}, k)}$ and defines a discrete $\bar{\sigma}_l$-Brownian bridge.

Using the independence of $S_v(k) - S_v(t^{l-1})$ with respect to $(S_v(t^j))_{j=1}^{l-1}$ and the processes in (3.2.16) and (3.2.17), we get

$$\begin{aligned}
(\clubsuit) &\leq \mathbb{P}\Big(S_v(k) - S_v(t^{l-1}) \geq M_{n,x}^\star(k) - x_{l-1}\Big) \\
&\quad \times \prod_{j \in \mathcal{A}_{l-1}} \mathbb{P}\left(\begin{array}{l} B_{v,i}^j < M_{n,x}^\star(i) - x_{j-1} - \frac{i - t^{j-1}}{\nabla t^j} \nabla x_j \\ \text{for all } i \text{ such that } t^{j-1} < i < t^j \end{array}\right) \\
&\quad \times \mathbb{P}\left(\begin{array}{l} B_{v,i} < (M_{n,x}^\star(i) - x_{l-1}) - \frac{i - t^{l-1}}{k - t^{l-1}}(M_{n,x}^\star(k) - x_{l-1}) \\ \text{for all } i \text{ such that } t^{l-1} < i < k \end{array}\right)^{\mathbf{1}_{\{l \in \mathcal{A}_m\}}} \\
&\mathrel{\hat{=}} (1) \times \prod_{j \in \mathcal{A}_{l-1}} (2)_j \times (3).
\end{aligned} \tag{3.2.18}$$

We bound (1) using a Gaussian estimate, and $(2)_j$ and (3) using the Brownian bridge estimates of Lemma 3.2.4. We pause the proof of Lemma 3.2.6 to state and prove these bounds in Lemma 3.2.7. $\square$

Lemma 3.2.7. *Let $l \in \{2, \dots, m\}$ and $t^{l-1} < k \leq t^l$. As in (3.2.14), we make the change of variables*

$$Y_{v,j} \mathrel{\hat{=}} \nabla S_v(t^j) - \nabla M_n^\star(t^j), \quad j \in \{1, \dots, l-1\}. \tag{3.2.19}$$

In (3.2.18), there exist constants $C, D > 0$, only depending on $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$, such that for n large enough,

$$(1) \leq C 2^{-(k - t^{l-1})} h_l(k) (k - t^{l-1})^{\mathbf{1}_{\{l \in \mathcal{A}_m \text{ and } (t^{l-1} + t^l)/2 < k \leq t^l\}}} e^{-\frac{x - \sum_{j=1}^{l-1} y_j}{g^{-1} \bar{\sigma}_l}} \tag{3.2.20}$$

where

$$h_l(k) \mathrel{\hat{=}} \begin{cases} (k - t^{l-1})^{-3/2}, & \text{when } l \in \mathcal{A}_m \text{ and } t^{l-1} < k \leq \frac{t^{l-1} + t^l}{2}, \\ (t^l - k)^{-5/2}, & \text{when } l \in \mathcal{A}_m \text{ and } \frac{t^{l-1} + t^l}{2} < k < t^l, \\ 1, & \text{when } k = t^l, \end{cases}$$

and

$$(2)_j \leq C \frac{(1 + D + 2x - 2\sum_{j'=1}^{j-1} y_{j'} - y_j)^2}{\nabla t^j}, \quad j \in \mathcal{A}_{l-1}, \tag{3.2.21}$$

and

$$(3) \leq \begin{cases} C \frac{(1 + D + 2x - 2\sum_{j'=1}^{l-2} y_{j'} - y_{l-1})^2}{k - t^{l-1}}, & \text{if } l \in \mathcal{A}_m \text{ and } \frac{t^{l-1} + t^l}{2} < k \leq t^l, \\ 1, & \text{otherwise.} \end{cases} \tag{3.2.22}$$

Proof of inequality (3.2.20). Since $\mathbb{V}\Big(S_v(k) - S_v(t^{l-1})\Big) = (k - t^{l-1})\bar{\sigma}_l^2$ when $k \in \mathcal{T}_m$, a Gaussian estimate yields

$$
\begin{aligned}
(1) &\mathrel{\hat{=}} \mathbb{P}\Big(S_v(k) - S_v(t^{l-1}) \geq M^\star_{n,x}(k) - x_{l-1}\Big) \\
&\leq \frac{\sqrt{(k-t^{l-1})\bar{\sigma}_l^2}}{\sqrt{2\pi}(M^\star_{n,x}(k) - x_{l-1})} e^{-\frac{(M^\star_{n,x}(k) - M^\star_{n,x}(t^{l-1}) + M^\star_{n,x}(t^{l-1}) - x_{l-1})^2}{2(k-t^{l-1})\bar{\sigma}_l^2}}. \qquad (3.2.23)
\end{aligned}
$$

Use successively $x_{l-1} \leq M^\star_{n,x}(t^{l-1})$ from (3.2.14), the definition of $M^\star_n$ in (3.2.2), the fact that $b_{n,x}(k) \geq x$ and $x \mapsto (\log x)/x$ is decreasing for $x \geq e$, to show

$$
\begin{aligned}
M^\star_{n,x}(k) - x_{l-1} &\geq M^\star_{n,x}(k) - M^\star_{n,x}(t^{l-1}) \\
&= g(k - t^{l-1})\bar{\sigma}_l - \frac{(1 + 2\cdot\delta_l)\bar{\sigma}_l}{2g} \frac{(k - t^{l-1})}{\nabla t^l} \log(\nabla t^l) + b_{n,x}(k) - x \\
&\geq g(k - t^{l-1})\bar{\sigma}_l - \frac{(1 + 2\cdot\delta_l)\bar{\sigma}_l}{2g} \log(e \vee (k - t^{l-1})). \qquad (3.2.24)
\end{aligned}
$$

Plugging inequality (3.2.24) in (3.2.23) and using the definition of $b_{n,x}$ from (3.2.3) and the fact that $M^\star_{n,x}(t^{l-1}) - x_{l-1} = x - \sum_{j=1}^{l-1} y_j$, we have

$$
\begin{aligned}
(1) &\leq C 2^{-(k-t^{l-1})} \frac{e^{\frac{(1+2\cdot\delta_l)}{2}\log(e\vee(k-t^{l-1})) - \frac{b_{n,x}(k)-x}{g^{-1}\bar{\sigma}_l}}}{\sqrt{k-t^{l-1}}} e^{-\frac{M^\star_{n,x}(t^{l-1}) - x_{l-1}}{g^{-1}\bar{\sigma}_l}} \\
&\leq \widetilde{C} 2^{-(k-t^{l-1})} h_l(k) (k - t^{l-1})^{\mathbf{1}_{\{l \in \mathcal{A}_m \text{ and } (t^{l-1}+t^l)/2 < k \leq t^l\}}} e^{-\frac{x - \sum_{j=1}^{l-1} y_j}{g^{-1}\bar{\sigma}_l}}
\end{aligned}
$$

where

$$
h_l(k) \mathrel{\hat{=}} \begin{cases} (k - t^{l-1})^{-3/2}, & \text{when } l \in \mathcal{A}_m \text{ and } t^{l-1} < k \leq \frac{t^{l-1}+t^l}{2}, \\ (t^l - k)^{-5/2}, & \text{when } l \in \mathcal{A}_m \text{ and } \frac{t^{l-1}+t^l}{2} < k < t^l, \\ 1, & \text{when } k = t^l. \end{cases}
$$

Note that the last inequality is an equality with $\widetilde{C} = C$ whenever $k - t^{l-1} \geq e$. When $k - t^{l-1} \in \{1, 2\}$, taking $\widetilde{C} = e^{3/2} \cdot C$ is sufficient to "absorb" the terms that do not cancel out exactly. □

Proof of inequality (3.2.21). Let $j \in \mathcal{A}_{l-1}$ and define

$$
z_{i,j} \mathrel{\hat{=}} M^\star_{n,x}(i) - x_{j-1} - \frac{i - t^{j-1}}{\nabla t^j} \nabla x_j, \quad t^{j-1} < i < t^j.
$$

We have

$$\begin{aligned} z_{i,j} &= b_{n,x}(i) + M_n^\star(i) + \left\{ \frac{i - t^{j-1}}{\nabla t^j} x_{j-1} + \frac{t^j - i}{\nabla t^j} x_j \right\} - x_{j-1} - x_j \\ &= b_{n,x}(i) + \left[M_n^\star(i) - \frac{t^j - i}{\nabla t^j} M_n^\star(t^{j-1}) - \frac{i - t^{j-1}}{\nabla t^j} M_n^\star(t^j) \right] - (x_{j-1} - M_n^\star(t^{j-1})) \\ &\quad + \left\{ \frac{i - t^{j-1}}{\nabla t^j} (x_{j-1} - M_n^\star(t^{j-1})) + \frac{t^j - i}{\nabla t^j} (x_j - M_n^\star(t^j)) \right\} - (x_j - M_n^\star(t^j)). \end{aligned}$$

Now, bound the braces using $(x_{j-1} - M_n^\star(t^{j-1})) \vee (x_j - M_n^\star(t^j)) \leq x$ from the integration limits of x_{j-1} and x_j in (3.2.14). The quantity between the brackets is zero because $M_n^\star$ is affine on $[t^{j-1}, t^j]$. Consequently,

$$\begin{aligned} z_{i,j} &\leq b_{n,x}(i) + x - (x_{j-1} - M_n^\star(t^{j-1})) - (x_j - M_n^\star(t^j)) \\ &\overset{(3.2.19)}{=} b_n(i) + 2x - \sum_{j'=1}^{j-1} y_{j'} - \sum_{j'=1}^{j} y_{j'}. \end{aligned} \tag{3.2.25}$$

Since $(2)_j \mathrel{\hat{=}} \mathbb{P}(B_{v,i}^j < z_{i,j},\ t^{j-1} < i < t^j)$, where B_v^j is a discrete $\bar{\sigma}_j$-Brownian bridge on $[t^{j-1}, t^j]$, the conclusion follows from Lemma 3.2.4 and (3.2.25). $\square$

Proof of inequality (3.2.22). Assume $l \in \mathcal{A}_m$ and $(t^{l-1} + t^l)/2 < k \leq t^l$. The other cases are trivial because (3) is a probability. Now, define

$$z_i \mathrel{\hat{=}} (M_{n,x}^\star(i) - x_{l-1}) - \frac{i - t^{l-1}}{k - t^{l-1}} (M_{n,x}^\star(k) - x_{l-1}), \quad t^{l-1} < i < k.$$

Similarly to the proof of (3.2.21), the path $M_n^\star$ is affine on $[t^{l-1}, t^l] \supseteq [t^{l-1}, k]$ and $x_{l-1} - M_n^\star(t^{l-1}) \leq x$ from the integration limits of x_{l-1} in (3.2.14), so

$$\begin{aligned} z_i &= b_{n,x}(i) - \frac{i - t^{l-1}}{k - t^{l-1}} b_{n,x}(k) - \frac{k - i}{k - t^{l-1}} (x_{l-1} - M_n^\star(t^{l-1})) \\ &\quad + \left[M_n^\star(i) - \frac{k - i}{k - t^{l-1}} M_n^\star(t^{l-1}) - \frac{i - t^{l-1}}{k - t^{l-1}} M_n^\star(k) \right] \\ &= b_n(i) - \frac{i - t^{l-1}}{k - t^{l-1}} b_n(k) + \left(1 - \frac{i - t^{l-1}}{k - t^{l-1}} \right) x \\ &\quad + \frac{i - t^{l-1}}{k - t^{l-1}} (x_{l-1} - M_n^\star(t^{l-1})) - (x_{l-1} - M_n^\star(t^{l-1})) \\ &\leq b_n(i) - \frac{i - t^{l-1}}{k - t^{l-1}} b_n(k) + x - \sum_{j'=1}^{l-1} y_{j'}. \end{aligned} \tag{3.2.26}$$

In order to use Lemma 3.2.4, it remains to show that the first two terms in (3.2.26) are bounded by an appropriate logarithmic barrier. Assume for now that $k \neq t^l$. There are three cases to consider.

$\boxed{\textbf{Case 1}: \text{All } i \text{ such that } t^{l-1} < i \leq (t^{l-1}+k)/2 < (t^{l-1}+t^l)/2 < k < t^l}$

Clearly,

$$b_n(i) - \frac{i - t^{l-1}}{k - t^{l-1}} b_n(k) \;\leq\; b_n(i) \stackrel{(3.2.3)}{=} \frac{5}{2}\frac{\bar{\sigma}_l}{g} \log(i - t^{l-1}). \tag{3.2.27}$$

$\boxed{\textbf{Case 2}: \text{All } i \text{ such that } t^{l-1} < (t^{l-1}+k)/2 < i \leq (t^{l-1}+t^l)/2 < k < t^l}$

Observe that $i - t^{l-1} \leq t^l - i$ and $t^l - k \leq k - t^{l-1}$ and $x \mapsto (\log x)/x$ is decreasing for $x \geq e$. Also, we have $(t^l - i) = (t^l - k) + (k - i) \leq 2(t^l - k)(k - i)$ because $a + b \leq 2ab$ for $a, b \geq 1$. Using all this (in that order), we get

$$\begin{aligned} b_n(i) - \frac{i - t^{l-1}}{k - t^{l-1}} b_n(k) &\stackrel{(3.2.3)}{=} \frac{5}{2}\frac{\bar{\sigma}_l}{g} \left\{ \log\left(\frac{i - t^{l-1}}{t^l - k}\right) + \frac{k - i}{k - t^{l-1}} \log(t^l - k) \right\} \\ &\leq \frac{5}{2}\frac{\bar{\sigma}_l}{g} \left\{ \log\left(\frac{t^l - i}{t^l - k}\right) + \log(e \vee (k - i)) \right\} \\ &\leq \frac{5}{2}\frac{\bar{\sigma}_l}{g} \left\{ \log 2 + 2\log(e \vee (k - i)) \right\}. \end{aligned} \tag{3.2.28}$$

$\boxed{\textbf{Case 3}: \text{All } i \text{ such that } t^{l-1} < (t^{l-1}+t^l)/2 < i < k < t^l}$

By the same reasoning as in Case 2 (without $i - t^{l-1} \leq t^l - i$), we get

$$\begin{aligned} b_n(i) - \frac{i - t^{l-1}}{k - t^{l-1}} b_n(k) &\stackrel{(3.2.3)}{=} \frac{5}{2}\frac{\bar{\sigma}_l}{g} \left\{ \log\left(\frac{t^l - i}{t^l - k}\right) + \frac{k - i}{k - t^{l-1}} \log(t^l - k) \right\} \\ &\leq \frac{5}{2}\frac{\bar{\sigma}_l}{g} \left\{ \log 2 + 2\log(e \vee (k - i)) \right\}. \end{aligned} \tag{3.2.29}$$

Finally, when $k = t^l$, the inequalities (3.2.27), (3.2.28) and (3.2.29) are trivial because $b_n(k) = 0$. Therefore, applying all three inequalities in (3.2.26), there exist appropriate constants $D, \widetilde{D} > 0$, depending only on $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$, for which

$$z_i \leq \begin{cases} \widetilde{D}\log(i - t^{l-1}) + D + x - \sum_{j'=1}^{l-1} y_{j'}, & \text{if } t^{l-1} < i \leq \frac{t^{l-1}+k}{2}, \\ \widetilde{D}\log(k - i) + D + x - \sum_{j'=1}^{l-1} y_{j'}, & \text{if } \frac{t^{l-1}+k}{2} < i < k, \end{cases}$$

$$\leq \begin{cases} \widetilde{D}\log(i - t^{l-1}) + D + 2x - 2\sum_{j'=1}^{l-2} y_{j'} - y_{l-1}, & \text{if } t^{l-1} < i \leq \frac{t^{l-1}+k}{2}, \\ \widetilde{D}\log(k - i) + D + 2x - 2\sum_{j'=1}^{l-2} y_{j'} - y_{l-1}, & \text{if } \frac{t^{l-1}+k}{2} < i < k. \end{cases}$$

We used $\sum_{j'=1}^{l-2} y_{j'} \leq x$ from the integration limits of y_{l-2} in (3.2.15) to get the last inequality. When $l \in \mathcal{A}_m$, recall that $(3) \doteq \mathbb{P}(B_{v,i} < z_i,\ t^{l-1} < i < k)$, where B_v is a discrete $\bar{\sigma}_l$-Brownian bridge on $[t^{l-1}, k]$. Applying Lemma 3.2.4 yields the conclusion. □

Proof of Lemma 3.2.6 (last part). By applying Lemma 3.2.7 in (3.2.18), the integral in (3.2.15) is smaller than

$$\begin{aligned} &C\,2^{-k}\,h_l(k)\,e^{-x\frac{g}{\bar{\sigma}_l}} \int_{-\infty}^{x}\int_{-\infty}^{x-y_1}\cdots\int_{-\infty}^{x-\sum_{j=1}^{l-2} y_j} (1 + D + 2x - 2\sum_{j'=1}^{l-2} y_{j'} - y_{l-1})^{2\cdot\delta_l} \\ &\times \left[\prod_{j\in\mathcal{A}_{l-1}} (1 + D + 2x - 2\sum_{j'=1}^{j-1} y_{j'} - y_j)^2\right] \cdot \prod_{j=1}^{l-1} e^{y_j\left[\frac{g}{\bar{\sigma}_l} - \frac{g}{\bar{\sigma}_j}\right]}\, d\boldsymbol{y} \end{aligned} \tag{3.2.30}$$

for an appropriate constant $D = D(\boldsymbol{\sigma}, \boldsymbol{\lambda}) > 0$. To obtain (3.2.30), the terms (∇t^j) in (3.2.15) canceled with the factors $1/(\nabla t^j)$ in (3.2.21), for all $j \in \mathcal{A}_{l-1}$. Similarly, the term $(k - t^{l-1})$ in (3.2.20) canceled with the factor $1/(k - t^{l-1})$ in (3.2.22), when $l \in \mathcal{A}_m$ and $(t^{l-1} + t^l)/2 < k \leq t^l$.

To bound the integral in (3.2.30), it is crucial to observe that the brackets in the exponentials are always strictly positive because $\bar{\sigma}_1 > \bar{\sigma}_2 > \ldots > \bar{\sigma}_m$ by definition. Denote these brackets by $\beta_{j,l},\ 1 \leq j \leq l-1$. We evaluate the integral iteratively. Note that $\sum_{j=1}^{l-2} y_j \leq x$ and $\sum_{j=1}^{l-3} y_j \leq x$ from the integration limits of y_{l-2} and y_{l-3} in (3.2.30). By integrating by parts, it is easy to show that the first integral (from the interior) have the property

$$\begin{aligned} &\int_{-\infty}^{x-\sum_{j=1}^{l-2} y_j} (1 + D + 2x - 2\sum_{j'=1}^{l-2} y_{j'} - y_{l-1})^a\ e^{y_{l-1}\beta_{l-1,l}}\ dy_{l-1} \\ &\quad\leq \frac{(a+1)!}{(1 \wedge \beta_{l-1,l})^{a+1}}\ (1 + D + 2x - 2\sum_{j'=1}^{l-3} y_{j'} - y_{l-2})^a\ e^{(x-\sum_{j=1}^{l-2} y_j)\beta_{l-1,l}} \end{aligned}$$

for any exponent $a \in \mathbb{N}_0$. Therefore, iterating this reasoning in (3.2.30) gives

$$\begin{aligned}(3.2.30) &\leq \widetilde{C}\, 2^{-k}\, h_l(k)\, e^{-x\frac{g}{\bar{\sigma}_l}} \cdot (1 + D + x)^{2\sum_{j=1}^{l} \delta_j} e^{x\sum_{j=1}^{l-1} \beta_{j,j+1}} \\ &= \widetilde{C}\, 2^{-k}\, h_l(k)\, e^{-x\frac{g}{\bar{\sigma}_1}} \cdot (1 + D + x)^{2\sum_{j=1}^{l} \delta_j}.\end{aligned}$$

Applying this bound in (3.2.13) yields the conclusion since

$$\sum_{l=1}^{m} \sum_{\substack{k\in\mathcal{T}_m \\ t^{l-1}<k\leq t^l}} h_l(k) < \infty.$$

This ends the proof of Lemma 3.2.6. □

Lemma 3.2.8 (Lower bound)**.** *Let $\{S_v\}_{v\in\mathbb{D}_n}$ be the $(\boldsymbol{\sigma}, \boldsymbol{\lambda})$-BRW at time n of Definition 3.1.1, under Restriction 3.1.2. Recall the definition of $M_n^\star$ from (3.2.2). For all $\varepsilon > 0$, there exists $K_\varepsilon > 0$ such that for all $n \in \mathbb{N}$,*

$$\mathbb{P}\left(\max_{v\in\mathbb{D}_n} S_v \leq M_n^\star(n) - K_\varepsilon\right) < \varepsilon.$$

Proof. Let $S_n^\star \circeq \max_{v\in\mathbb{D}_n} S_v$. From Theorem 1 of Fang (2012), we know that the family $\{S_n^\star - \mathrm{Med}(S_n^\star)\}_{n\in\mathbb{N}}$ is tight, that is for all $\varepsilon > 0$, there exists $\widetilde{K}_\varepsilon > 0$ such that for all $n \in \mathbb{N}$,

$$\mathbb{P}\left(|S_n^\star - \mathrm{Med}(S_n^\star)| \geq \widetilde{K}_\varepsilon\right) < \varepsilon. \tag{3.2.31}$$

We claim that there exist $c, C > 0$ and $n_0, \tilde{n}_0 \in \mathbb{N}$ such that

$$\left\{\begin{array}{l}\mathbb{P}(S_n^\star \geq M_n^\star(n) - C) \geq c \\ \text{for all } n \geq n_0\end{array}\right\} \Longrightarrow \left\{\begin{array}{l}\mathrm{Med}(S_n^\star) \geq M_n^\star(n) - C - \widetilde{K}_c \\ \text{for all } n \geq \tilde{n}_0\end{array}\right\}. \tag{3.2.32}$$

Otherwise, by (3.2.31), for each choice of $c, C > 0$, there would exist a subsequence $\{n_i\}_{i\in\mathbb{N}}$ such that

$$c \leq \mathbb{P}\left(S_{n_i}^\star \geq M_{n_i}^\star(n_i) - C\right) \leq \mathbb{P}\left(S_{n_i}^\star \geq \mathrm{Med}(S_{n_i}^\star) + \widetilde{K}_c\right) < c,$$

which is impossible. If the left side of (3.2.32) was satisfied for some constants $c, C > 0$, we could define $K_\varepsilon \circeq \widetilde{K}_\varepsilon + C + \widetilde{K}_c$, and (3.2.31) would give

$$\mathbb{P}(S_n^\star \leq M_n^\star(n) - K_\varepsilon) \leq \mathbb{P}\left(S_n^\star \leq \mathrm{Med}(S_n^\star) - \widetilde{K}_\varepsilon\right) < \varepsilon, \quad n \geq \tilde{n}_0,$$

and the proof of the lemma would be over.

To conclude, it remains to show the left side of (3.2.32). We now use Restriction 3.1.2. Recall from Remark 3.1.2 that $\{\lambda_{i_d}\}_{0\leq d\leq p}$ is the union of all the scales λ^j and all the isolated points where $\mathcal{J}_{\sigma^2}$ and $\mathcal{J}_{\bar{\sigma}^2}$ coincide. By independence of the increments, the left side of (3.2.32) is satisfied if there exist constants $c, C > 0$ such that

$$\mathbb{P}\left(\max_{v\in\mathbb{D}_{\nabla_d t_{i_d}}} S_v^{i_d} \geq \nabla_d M_n^\star(t_{i_d}) - (C/p)\right) \geq c^{1/p}, \quad 1 \leq d \leq p, \tag{3.2.33}$$

where each field $\{S_v^{i_d}\}_v$ consists of the end points of an inhomogeneous BRW on the time interval $[0, \nabla_d t_{i_d}]$ with variance parameters given by the step function $s \mapsto \sigma(s)$ on $(\lambda_{i_{d-1}}, \lambda_{i_d}]$.

It suffices to show (3.2.33) for the subinterval(s) $[t_{i_{d-1}}, t_{i_d}] \subseteq [0, t^1]$ since we did not assume anything on the other intervals $[t^{j-1}, t^j]$. When $1 \in \mathcal{A}_m$, that is when there is only one variance parameter $\sigma_1 = \bar{\sigma}_1$ on $(0, \lambda^1]$, then (3.2.33) follows from Theorem 3 of Addario-Berry and Reed (2009) by choosing $C > 0$ large enough and $c > 0$ small enough. Since $M_n^\star(\cdot)$ is linear on $[0, t^1]$ and the argument presented below could be applied for each subinterval of the partition (independently of d), we can assume, without loss of generality, that $t_{i_1} = t^1$, namely that

$$\begin{gathered}\mathcal{J}_{\sigma^2} \text{ lies strictly below its concave hull } \mathcal{J}_{\bar{\sigma}^2} \\ \text{everywhere on } (0, t^1).\end{gathered} \tag{3.2.34}$$

The usual trick to prove a lower bound in the BRW setting is the Paley-Zygmund inequality. If we naively try to apply the Paley-Zygmund inequality to the number of particles that stay above the optimal path, the method will not work because the correlations of the BRW inflate the second moment too much, see (3.2.36). Instead, we need to add a barrier condition that eliminates the overly large number of particles that are too far off the optimal path during their lifetime. For simplicity, we omit the superscript i_1 for $S_v^{i_1}$ in the remaining of the proof. Define $S_v \mathrel{\hat{=}} S_v(t^1)$ and let

$$\begin{aligned}
I_n &\mathrel{\hat{=}} [M_n^\star(t^1), M_n^\star(t^1) + 1], \\
I_{k,n}(x) &\mathrel{\hat{=}} [s_{k,n}(x) - f_{k,n}, s_{k,n}(x) + f_{k,n}], \\
\mathcal{N}_n &\mathrel{\hat{=}} \#\{v \in \mathbb{D}_{t^1} : S_v \in I_n, S_v(k) \in I_{k,n}(S_v) \ \ \forall 0 < k < t^1\},
\end{aligned}$$

where $s_{k,n}(x)$ is a path leading to $x \in \mathbb{R}$ and $f_{k,n}$ is a concave barrier. The definition we give to $s_{k,n}$ could seem strange at first, but is actually quite natural. It is argued in Arguin and Ouimet (2016) and proved in Appendix A of Ouimet (2014) that the log-number of particles that are above the path

$$s_{k,n}(x) \mathrel{\hat{=}} \frac{\mathcal{J}_{\sigma^2}(k/n)}{\mathcal{J}_{\sigma^2}(\lambda^1)} x, \quad 0 \leq k \leq t^1,$$

during their lifetime is asymptotically the same as the log-number of particles above x at time t^1. In particular, for particles reaching $x = M_n^\star(t^1)$ at time t^1, this path is optimal (for the first order). The barrier is

$$f_{k,n} \mathrel{\hat{=}} \begin{cases} C_f \, (\mathcal{J}_{\sigma^2}(k/n)n)^{2/3}, & \text{if } 0 \leq k \leq t_1, \\ C_f \, (\mathcal{J}_{\sigma^2}(k/n, \lambda^1)n)^{2/3}, & \text{if } t_1 < k \leq t^1, \end{cases} \tag{3.2.35}$$

where the constant $C_f > 0$ will be chosen large enough later in the proof. The exponent $2/3$ is not essential here (any exponent in $(1/2, 1)$ works), but this definition is useful for the Gaussian estimates.

Under assumption (3.2.34), the Paley-Zygmund inequality yields that the probability in (3.2.33) (when $d = 1$) is bounded from below by

$$\mathbb{P}\left(\max_{v \in \mathbb{D}_{t^1}} S_v \geq M_n^\star(t^1)\right) \geq \mathbb{P}(\mathcal{N}_n \geq 1) \overset{\textbf{P-Z}}{\geq} \frac{(\mathbb{E}[\mathcal{N}_n])^2}{\mathbb{E}[(\mathcal{N}_n)^2]}. \tag{3.2.36}$$

To conclude, we show $\mathbb{E}[\mathcal{N}_n] \geq c_\star$ and $\mathbb{E}[(\mathcal{N}_n)^2] \leq (\mathbb{E}[\mathcal{N}_n])^2 + (1 + C_\star)\mathbb{E}[\mathcal{N}_n]$ for some constants $c_\star, C_\star > 0$.

Lower bound on the first moment

By the linearity of expectation, we have the lower bound

$$\begin{aligned} \mathbb{E}[\mathcal{N}_n] &= 2^{t^1} \mathbb{P}\Big(S_v \in I_n, S_v(k) \in I_{k,n}(S_v) \ \ \forall 0 < k < t^1\Big) \\ &= 2^{t^1} \mathbb{P}(S_v \in I_n) \, \mathbb{P}\Big(S_v(k) \in I_{k,n}(S_v) \ \ \forall 0 < k < t^1\Big) \geq c_\star, \end{aligned} \tag{3.2.37}$$

provided that there exist constants $c_1, c_2 > 0$ such that

(1) S_v is independent of $\{S_v(k) - s_{k,n}(S_v)\}_{k=0}^{t^1}$,
(2) $2^{t^1} \mathbb{P}(S_v \in I_n) \geq c_1$,
(3) $\mathbb{P}(S_v(k) \in I_{k,n}(S_v) \ \ \forall 0 < k < t^1) \geq c_2$.

To show (1), observe that $\mathbb{V}(S_v(k)) = \mathcal{J}_{\sigma^2}(k/n)n$ and $\mathbb{V}(S_v) = \mathcal{J}_{\sigma^2}(\lambda^1)n$ from (3.2.1), so the independence between $S_v(k)$ and $S_v - S_v(k)$ gives

$$\mathrm{Cov}(S_v, S_v(k) - s_{k,n}(S_v)) = \mathbb{V}(S_v(k)) - \frac{\mathcal{J}_{\sigma^2}(k/n)}{\mathcal{J}_{\sigma^2}(\lambda^1)}\mathbb{V}(S_v) = 0.$$

To show (2), note that $M_n^\star(t^1) = g\bar{\sigma}_1 t^1 - \frac{1}{2}\frac{\bar{\sigma}_1}{g}\log(t^1)$, under assumption (3.2.34), and $\mathbb{V}(S_v) = \bar{\sigma}_1^2 t^1$. Therefore,

$$\mathbb{P}(S_v \in I_n) = \int_{M_n^\star(t^1)}^{M_n^\star(t^1)+1} \frac{e^{-\frac{z^2}{2\bar{\sigma}_1^2 t^1}}}{\sqrt{2\pi\bar{\sigma}_1^2 t^1}}\, dz \geq 1 \cdot \frac{c}{\sqrt{t^1}} e^{-\frac{(M_n^\star(t^1)+1)^2}{2\bar{\sigma}_1^2 t^1}} \geq c_1\, 2^{-t^1}.$$

To show (3), note that $\mathrm{Cov}(s_{k,n}(S_v), S_v(k) - s_{k,n}(S_v)) = 0$, by the independence in (1), and thus

$$\begin{aligned}\mathbb{V}(S_v(k) - s_{k,n}(S_v)) &= \mathrm{Cov}(S_v(k), S_v(k) - s_{k,n}(S_v)) \\ &= \mathcal{J}_{\sigma^2}(k/n)n\left[1 - \frac{\mathcal{J}_{\sigma^2}(k/n)}{\mathcal{J}_{\sigma^2}(\lambda^1)}\right].\end{aligned}$$

Then, sub-additivity followed by Gaussian estimates yield

$$\begin{aligned}\mathbb{P}\begin{pmatrix} S_v(k) \in I_{k,n}(S_v) \\ \forall 0 < k < t^1 \end{pmatrix} &\geq 1 - 2\sum_{k=1}^{t^1-1} \mathbb{P}(S_v(k) - s_{k,n}(S_v) > f_{k,n}) \\ &\geq 1 - 2\sum_{k=1}^{t^1-1} C\exp\left(-\frac{1}{2}\frac{(f_{k,n})^2}{\mathcal{J}_{\sigma^2}(k/n)n\left[1 - \frac{\mathcal{J}_{\sigma^2}(k/n)}{\mathcal{J}_{\sigma^2}(\lambda^1)}\right]}\right).\end{aligned}$$

By considering the cases $0 < k \leq t_1$ and $t_1 < k < t^1$ separately, the last sum is bounded from above by

$$\sum_{k=1}^{t_1} Ce^{-\frac{1}{2}C_f^2\, \sigma_1^{2/3} k^{1/3}} + \sum_{k=t_1+1}^{t^1-1} Ce^{-\frac{1}{2}C_f^2 \min_{i\in\{2,3,\dots,\pi_1\}} \sigma_i^{2/3}(t^1-k)^{1/3}}.$$

For C_f large enough, this is strictly smaller than $1/2$, independently of n, which proves (3).

Upper bound on the second moment

To estimate the second moment, we split $\mathbb{E}[(\mathcal{N}_n)^2]$ according to the branching time $\rho(u,v) \mathrel{\hat=} \max\{r \in \{0,1,\dots,t^1\} : u_r = v_r\}$ of each pair of particles :

$$\mathbb{E}\left[(\mathcal{N}_n)^2\right] = \sum_{r=0}^{t^1} \sum_{\substack{u,v\in\mathbb{D}_{t^1} \\ \rho(u,v)=r}} \mathbb{P}\begin{pmatrix} S_u, S_v \in I_n \text{ and } S_u(k) \in I_{k,n}(S_u), \\ S_v(k) \in I_{k,n}(S_v) \text{ for all } 0 < k < t^1 \end{pmatrix}.$$

When $\rho(u,v)=0$, the processes $\{S_u(k)\}_k$ and $\{S_v(k)\}_k$ are independent. Therefore, in the case $r=0$, the second sum above is bounded by $(\mathbb{E}[\mathcal{N}_n])^2$ by adding the missing terms. In the case $r=t^1$, the second sum is equal to $\mathbb{E}[\mathcal{N}_n]$ because u and v coincide. In the remaining cases $0<r<t^1$, the increment $S_v - S_v(r)$ is independent of $\{S_u(k)\}_k$, and $S_u(k)=S_v(k)$ for all $k\le r$. Therefore, $\mathbb{E}[(\mathcal{N}_n)^2]$ is bounded from above by

$$(\mathbb{E}[\mathcal{N}_n])^2 + \mathbb{E}[\mathcal{N}_n] + \sum_{r=1}^{t^1-1} \sum_{\substack{u,v\in\mathbb{D}_{t^1}\\ \rho(u,v)=r}} \mathbb{P}\begin{pmatrix} S_u \in I_n \text{ and } S_u(k)\in I_{k,n}(S_u) \\ \text{for all } 0<k<t^1 \end{pmatrix} \cdot \max_{x\in I_n} \mathbb{P}\Big(S_v - S_v(r) \in x - I_{r,n}(x) \Big). \tag{3.2.38}$$

There are at most $2^{t^1}\cdot 2^{t^1-r}$ pairs $(u,v)\in\mathbb{D}_{t^1}^2$ with branching time equal to r, so the double sum in (3.2.38) is bounded from above by

$$\mathbb{E}[\mathcal{N}_n] \times \sum_{r=1}^{t^1-1} 2^{t^1-r} \max_{\substack{x\in I_n\\ v\in\mathbb{D}_{t^1}}} \underbrace{\mathbb{P}\Big(S_v - S_v(r) \in x - I_{r,n}(x) \Big)}_{(\spadesuit)_r}. \tag{3.2.39}$$

It remains to estimate the probabilities $(\spadesuit)_r$ in (3.2.39). From (3.2.1), we know that $\mathbb{V}(S_v - S_v(r)) = \mathcal{J}_{\sigma^2}(r/n,\lambda^1)n$ for all $v\in\mathbb{D}_{t^1}$.

In the case $0<r\le t_1$, we have $f_{r,n} = C_f\,(\sigma_1^2 r)^{2/3}$. Thus, for $x\in I_n$,

$$\begin{aligned}
(\spadesuit)_r &= \int_{x-I_{r,n}(x)} \frac{e^{-\frac{1}{2}\frac{z^2}{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n}}}{\sqrt{2\pi\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n}}\,dz \le 2f_{r,n}\frac{e^{-\frac{1}{2}\frac{(M_n^\star(t^1)-s_{r,n}(M_n^\star(t^1))-f_{r,n})^2}{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n}}}{\sqrt{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n}} \\
&\le C\,r^{2/3}\,2^{-\frac{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)t^1}{\mathcal{J}_{\sigma^2}(\lambda^1)}}\,\frac{e^{\frac{1}{2}\frac{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)}{\mathcal{J}_{\sigma^2}(\lambda^1)}\log(t^1)}}{\sqrt{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n}}\,e^{\frac{C_f\,(\sigma_1^2 r)^{2/3}}{g^{-1}\bar{\sigma}_1}} &\quad (3.2.40)\\
&\le C\,r^{2/3}\,2^{-(t^1-\eta_1 r)}\,e^{\widetilde{C}r^{2/3}}. &\quad (3.2.41)
\end{aligned}$$

To obtain the last bound, we use two crucial observations. Since the function $x\mapsto(\log x)/x$ is decreasing for $x\ge e$, the ratio of the exponential over the square root in (3.2.40) is bounded by a constant independent of r and n. Also, under assumption (3.2.34) and for $0<r\le t_1$,

$$\frac{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)t^1}{\mathcal{J}_{\sigma^2}(\lambda^1)} = t^1 - \frac{\frac{1}{r/n}\mathcal{J}_{\sigma^2}(r/n)}{\frac{1}{\lambda^1}\mathcal{J}_{\sigma^2}(\lambda^1)}r = t^1 - \frac{\frac{1}{\lambda_1}\mathcal{J}_{\sigma^2}(\lambda_1)}{\frac{1}{\lambda^1}\mathcal{J}_{\sigma^2}(\lambda^1)}r \mathrel{\hat{=}} t^1 - \eta_1 r$$

where $\eta_1<1$ independently of r and n. See Figure 3.2.4 below for an example.

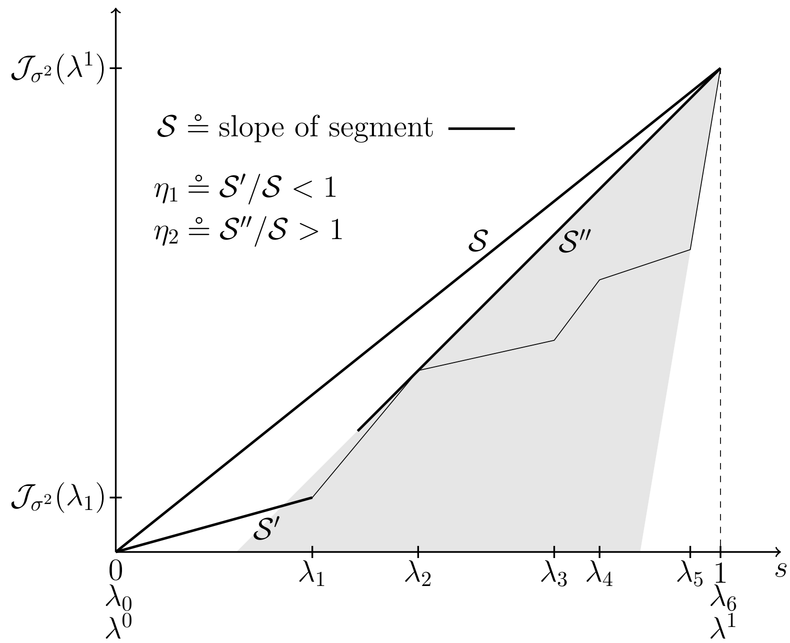


FIGURE 3.2.4. Example of η_1 and η_2 under assumption (3.2.34). The thin line represents $\mathcal{J}_{\sigma^2}$.

Similarly, in the case $t_1 < r < t^1$, we have $f_{r,n} = C_f\,(\mathcal{J}_{\sigma^2}(r/n, \lambda^1)n)^{2/3}$. Thus, for $x \in I_n$,

$$(\spadesuit)_r = \int_{x-I_{r,n}(x)} \frac{e^{-\frac{1}{2}\frac{z^2}{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n}}}{\sqrt{2\pi\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n}}\,dz \leq 2f_{r,n}\frac{e^{-\frac{1}{2}\frac{(M_n^\star(t^1)-s_{r,n}(M_n^\star(t^1))-f_{r,n})^2}{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n}}}{\sqrt{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n}}$$

$$\leq C\;2^{-\frac{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)t^1}{\mathcal{J}_{\sigma^2}(\lambda^1)}}\;\frac{e^{\frac{1}{2}\frac{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)}{\mathcal{J}_{\sigma^2}(\lambda^1)}\log(t^1)}}{(\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n)^{-1/6}}\;e^{\frac{C_f\,(\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n)^{2/3}}{g^{-1}\bar{\sigma}_1}} \tag{3.2.42}$$

$$\leq C\;2^{-\eta_2(t^1-r)}\;(\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n)^{2/3}\;e^{\widetilde{C}\,(t^1-r)^{2/3}}. \tag{3.2.43}$$

Again, to obtain the last bound, we use two crucial observations. The first exponential in (3.2.42) is bounded by $C(\mathcal{J}_{\sigma^2}(r/n,\lambda^1)n)^{1/2}$, where C is independent of r and n, using the fact that $x \mapsto (\log x)/x$ is decreasing for $x \geq e$. Also, under assumption (3.2.34) and for $t_1 < r < t^1$,

$$\frac{\mathcal{J}_{\sigma^2}(r/n,\lambda^1)t^1}{\mathcal{J}_{\sigma^2}(\lambda^1)} = \frac{\frac{1}{\lambda^1-r/n}\mathcal{J}_{\sigma^2}(r/n,\lambda^1)}{\frac{1}{\lambda^1}\mathcal{J}_{\sigma^2}(\lambda^1)}(t^1-r) \geq \eta_2(t^1-r)$$

where η_2 is the minimum of the last ratio with respect to $r \in \{t_1,\ldots,t^1-1\}$. Note that $\eta_2 > 1$ independently of r and n, see Figure 3.2.4 above.

By combining the bounds on $(\spadesuit)_r$ in (3.2.41) and (3.2.43), the sum in (3.2.39) is bounded from above by

$$C\left[\sum_{r=1}^{t_1} 2^{-(1-\eta_1)r+o(r)} + \sum_{r=t_1+1}^{t^1-1} 2^{(1-\eta_2)(t^1-r)+o(t^1-r)}\right] \leq C_\star$$

where $\eta_1 < 1$ and $\eta_2 > 1$ independently of r and n. By applying this bound in (3.2.39) and back in (3.2.38), we have

$$\frac{\mathbb{E}[(\mathcal{N}_n)^2]}{(\mathbb{E}[\mathcal{N}_n])^2} \leq 1 + \frac{1+C_\star}{\mathbb{E}[\mathcal{N}_n]} \stackrel{(3.2.37)}{\leq} 1 + \frac{1+C_\star}{c_\star}. \tag{3.2.44}$$

Using (3.2.44) in (3.2.36) yields (3.2.33) when $d = 1$, under assumption (3.2.34). This ends the proof of Lemma 3.2.8. $\square$

Acknowledgements

First, I would like to thank an anonymous referee for his valuable comments that led to improvements in the presentation of this paper. I also gratefully acknowledge insightful discussions with Bastien Mallein and my advisor, Louis-Pierre Arguin. This work is supported by a NSERC Doctoral Program Alexander Graham Bell scholarship (CGS D3).

Bibliography

Y. Abe. Cover times for sequences of reversible Markov chains on random graphs. *Kyoto J. Math.*, 54(3):555–576, 2014. MR3263552.

Y. Abe. Second order term of cover time for planar simple random walk. *Preprint*, pages 1–51, 2018. arXiv:1709.08151.

L. Addario-Berry and B. Reed. Minima in branching random walks. *Ann. Probab.*, 37(3): 1044–1079, 2009. MR2537549.

R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Monographs in Mathematics. Springer, New York, 2007. MR2319516.

L.-P. Arguin and F. Ouimet. Extremes of the two-dimensional Gaussian free field with scale-dependent variance. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(2):779–808, 2016. MR3541850.

L.-P. Arguin and F. Ouimet. Large deviations and continuity estimates for the derivative of a random model of $\log |\zeta|$ on the critical line. *J. Math. Anal. Appl.*, 472(1):687–695, 2019. MR3906393.

L.-P. Arguin and W. Tai. Is the Riemann zeta function in a short interval a 1-RSB spin glass ? *Preprint. To appear in Sojourns in probability and statistical physics, in honor of the 70th birthday of C. Newman*, pages 1–20, 2018. arXiv:1706.08462.

L.-P. Arguin, D. Belius, and P. Bourgade. Maximum of the characteristic polynomial of random unitary matrices. *Comm. Math. Phys.*, 349(2):703–751, 2017a. MR3594368.

L.-P. Arguin, D. Belius, and A. J. Harper. Maxima of a randomized Riemann zeta function, and branching random walks. *Ann. Appl. Probab.*, 27(1):178–215, 2017b. MR3619786.

L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł, and K. Soundararajan. Maximum of the Riemann zeta function on a short interval of the critical line. *Comm. Pure Appl. Math.*, 72(3):500–535, 2019. MR3911893.

D. Belius. Gumbel fluctuations for cover times in the discrete torus. *Probab. Theory Related Fields*, 157(3-4):635–689, 2013. MR3129800.

D. Belius and N. Kistler. The subleading order of two dimensional cover times. *Probab. Theory Related Fields*, 167(1-2):461–552, 2017. MR3602852.

A. Bovier and L. Hartung. The extremal process of two-speed branching Brownian motion. *Electron. J. Probab.*, 19:no. 18, 28, 2014. MR3164771.

A. Bovier and L. Hartung. Variable speed branching Brownian motion 1. Extremal processes in the weak correlation regime. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12(1):261–291, 2015. MR3351476.

A. Bovier and I. Kurkova. Derrida's generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):439–480, 2004a. MR2070334.

A. Bovier and I. Kurkova. Derrida's generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):481–495, 2004b. MR2070335.

M. Bramson. Maximal displacement of branching Brownian motion. *Comm. Pure Appl. Math.*, 31(5):531–581, 1978. MR0494541.

L. Chen. Steep points of Gaussian free fields in any dimension. *Preprint*, pages 1–43, 2018. arXiv:1802.03034.

R. Chhaibi, T. Madaule, and J. Najnudel. On the maximum of the CβE field. *Duke Math. J.*, 167(12):2243–2345, 2018. MR3848391.

F. Comets, C. Gallesco, S. Popov, and M. Vachkovskaia. On large deviations for the cover time of two-dimensional torus. *Electron. J. Probab.*, 18:no. 96, 18, 2013. MR3126579.

A. Dembo, Y. Peres, and J. Rosen. Brownian motion on compact manifolds: cover time and late points. *Electron. J. Probab.*, 8:no. 15, 14, 2003. MR1998762.

A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Cover times for Brownian motion and random walks in two dimensions. *Ann. of Math. (2)*, 160(2):433–464, 2004. MR2123929.

A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Late points for random walks in two dimensions. *Ann. Probab.*, 34(1):219–263, 2006. MR2206347.

J. Ding. On cover times for 2D lattices. *Electron. J. Probab.*, 17:no. 45, 18, 2012. MR2946152.

J. Ding. Asymptotics of cover times via Gaussian free fields: bounded-degree graphs and general trees. *Ann. Probab.*, 42(2):464–496, 2014. MR3178464.

J. Ding and O. Zeitouni. A sharp estimate for cover times on binary trees. *Stochastic Process. Appl.*, 122(5):2117–2133, 2012. MR2921974.

J. Ding, J. R. Lee, and Y. Peres. Cover times, blanket times, and majorizing measures. *Ann. of Math. (2)*, 175(3):1409–1471, 2012. MR2912708.

M. Fang. Tightness for maxima of generalized branching random walks. *J. Appl. Probab.*, 49(3):652–670, 2012. MR3012090.

M. Fang and O. Zeitouni. Branching random walks in time inhomogeneous environments. *Electron. J. Probab.*, 17:no. 67, 18, 2012a. MR2968674.

M. Fang and O. Zeitouni. Slowdown for time inhomogeneous branching Brownian motion. *J. Stat. Phys.*, 149(1):1–9, 2012b. MR2981635.

A. J. Harper. A note on the maximum of the Riemann zeta function, and log-correlated random variables. *Preprint*, pages 1–26, 2013. arXiv:1304.0677.

G. F. Lawler and V. Limic. *Random walk: a modern introduction*, volume 123 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. MR2677157.

P. Maillard and O. Zeitouni. Slowdown in branching Brownian motion with inhomogeneous variance. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(3):1144–1160, 2016. MR3531703.

B. Mallein. Maximal displacement in a branching random walk through interfaces. *Electron. J. Probab.*, 20:no. 68, 40, 2015a. MR3361256.

B. Mallein. Maximal displacement of a branching random walk in time-inhomogeneous environment. *Stoch. Proc. Appl.*, 125(10):3958–4019, 2015b. MR3373310.

B. Mallein and M. Piotr. Maximal displacement of a supercritical branching random walk in a time-inhomogeneous random environment. *Preprint, To appear in Stochastic Processes and their Applications*, pages 1–22, 2018. arXiv:1507.08835.

A. A. Mogul′skiĭ. A local limit theorem for the first hitting time of a fixed level by a random walk. *Mat. Tr.*, 12(2):126–138, 2009. MR2599428.

J. Najnudel. On the extreme values of the Riemann zeta function on random intervals of the critical line. *Probab. Theory Related Fields*, 172(1-2):387–452, 2018. MR3851835.

F. Ouimet. *Étude du maximum et des hauts points de la marche aléatoire branchante inhomogène et du champ libre gaussien inhomogène*. Master's thesis, Université de Montréal, 2014. URL `https://papyrus.bib.umontreal.ca/xmlui/handle/1866/11510`.

F. Ouimet. Geometry of the Gibbs measure for the discrete 2D Gaussian free field with scale-dependent variance. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(2):851–902, 2017. MR3731796.

F. Ouimet. Poisson-Dirichlet statistics for the extremes of a randomized Riemann zeta function. *Electron. Commun. Probab.*, 23:Paper No. 46, 15, 2018. MR3841407.

E. Paquette and O. Zeitouni. The maximum of the CUE field. *Int. Math. Res. Not. IMRN*, 2018(16):5028–5119, 2018. MR3848227.

E. Saksman and C. Webb. The Riemann zeta function and Gaussian multiplicative chaos: statistics on the critical line. *Preprint*, pages 1–85, 2018. arXiv:1609.00027.

Part 2 : The Riemann zeta function

Article 4

Poisson-Dirichlet statistics for the extremes of a randomized Riemann zeta function

by

Frédéric Ouimet[1]

([1]) Université de Montréal, Montréal, QC H3T 1J4, Canada.

Here is the complete citation :

Ouimet, F. (2018). Poisson-Dirichlet statistics for the extremes of a randomized Riemann zeta function. Electron. Commun. Probab. 23, no. 46, 1-15.
https://doi.org/10.1214/18-ecp154.

Abstract. In Arguin and Tai (2018), the authors prove the convergence of the two-overlap distribution at low temperature for a randomized Riemann zeta function on the critical line. We extend their results to prove the Ghirlanda-Guerra identities. As a consequence, we find the joint law of the overlaps under the limiting mean Gibbs measure in terms of Poisson-Dirichlet variables. It is expected that we can adapt the approach to prove the same result for the Riemann zeta function itself.
Keywords: extreme value theory, Riemann zeta function, Ghirlanda-Guerra identities, Gibbs measure, Poisson-Dirichlet variable, ultrametricity, spin glasses

4.1. Introduction

Following recent conjectures of Fyodorov et al. (2012) and Fyodorov and Keating (2014) about the limiting law of the *Gibbs measure* and the limiting law of the maximum for the Riemann zeta function on bounded random intervals of the critical line, progress have been made in the mathematics literature. If τ is sampled uniformly in $[T, 2T]$ for some large T, then it is expected that the limiting law of the Gibbs measure (see (4.2.6)) at low temperature for the field $(\log|\zeta(\frac{1}{2}+i(\tau+h))|, h \in [0,1])$ is a one-level Ruelle probability cascade (see e.g. Ruelle (1987)) and the law of the maximum is asymptotic to $\log\log T - \frac{3}{4}\log\log\log T + \mathcal{M}_T$ where $(\mathcal{M}_T, T \geq 2)$ is a sequence of random variables converging in distribution. For a randomized version of the Riemann zeta function (see (4.2.1)), the first order of the maximum was proved in Harper (2013), the second order of the maximum was proved in Arguin et al. (2017), and the limiting two-overlap distribution was found in Arguin and Tai (2018) (see Theorem 4.3.1 below). The tightness of the recentered maximum is still open (see Arguin and Ouimet (2019)). In this short paper, we complete the analysis of Arguin and Tai (2018) by proving the Ghirlanda-Guerra (GG) identities in the limit $T \to \infty$ (see Theorem 4.5.8). As is well known in the spin glass literature (see e.g. Chapter 2 in Panchenko (2013b)), the limiting law of the two-overlap distribution, with a finite support, together with the GG identities allow a complete description of the limiting law of the Gibbs measure as a *Ruelle probability cascade* with finitely many levels (a random measure with a tree structure and Poisson-Dirichlet weights at each level). Our main result (Theorem 4.3.2) describes the joint law of the overlaps under the limiting mean Gibbs measure in terms of Poisson-Dirichlet weights. It is expected that the approach presented here, which mostly stems from the work of Arguin and Zindy (2014), Bovier and Kurkova (2004a) and Panchenko (2013b) on other models, can be adapted to prove the same result for the (true) Riemann zeta function on bounded random intervals of the critical line. At present, for the (true) Riemann zeta function, the first order of the maximum is proved conditionally on the Riemann hypothesis in Najnudel (2018) and unconditionally in Arguin et al. (2019).

The paper is organised as follows. In Section 4.2, we give a few definitions. In Section 4.3, the main result is stated and shown to be a consequence of the GG identities and the main result from Arguin and Tai (2018) about the limiting two-overlap distribution. In Section 4.4, we state known results from Arguin and Tai (2018) that we will use to prove the GG identities. The GG identities are proven in Section 4.5 along with other preliminary results, see the structure of the proof in Figure 4.5.1. For an explanation of the consequences of the GG identities and their conjectured universality for mean field spin glass models, we refer the reader to Jagannath (2017), Panchenko (2013b) and Talagrand (2011).

4.2. Some definitions

Let $(U_p, p \text{ primes})$ be an i.i.d. sequence of uniform random variables on the unit circle in $\mathbb{C}$. The random field of interest is

$$X_h \mathrel{\hat{=}} \sum_{p\leq T} W_p(h) \mathrel{\hat{=}} \sum_{p\leq T} \frac{\mathrm{Re}(U_p\, p^{-ih})}{p^{1/2}}, \quad h\in[0,1]. \tag{4.2.1}$$

This is a good model for the large values of $(\log|\zeta(\frac{1}{2}+i(\tau+h))|, h\in[0,1])$ for the following reason. Proposition 1 in Harper (2013) proves that, assuming the Riemann hypothesis, and for T large enough, there exists a set $B\subseteq[T,T+1]$, of Lebesgue measure at least 0.99, such that

$$\log|\zeta(\frac{1}{2}+it)| = \mathrm{Re}\left(\sum_{p\leq T}\frac{1}{p^{1/2+it}}\frac{\log(T/p)}{\log T}\right) + O(1), \quad t\in B. \tag{4.2.2}$$

If we ignore the smoothing term $\log(T/p)/\log T$ and note that the process $(p^{-i\tau}, p \text{ primes})$, where τ is sampled uniformly in $[T,2T]$, converges (in the sense of convergence of its finite-dimensional distributions), as $T\to\infty$, to a sequence of independent random variables distributed uniformly on the unit circle (by computing the moments), then the model (4.2.1) follows. For more information, see Section 1.1 in Arguin et al. (2017).

For simplicity, the dependence in T will be implicit everywhere for X. Summations over p's and q's always mean that we sum over primes. For $\alpha \in [0,1]$, we denote truncated sums of X as follows :

$$X_h(\alpha) \doteq \sum_{p \leq \exp((\log T)^\alpha)} W_p(h), \quad h \in [0,1], \tag{4.2.3}$$

where $\sum_\emptyset \doteq 0$. Define the *overlap* between two points of the field by

$$\rho(h,h') \doteq \frac{\mathbb{E}[X_h X_{h'}]}{\sqrt{\mathbb{E}[X_h^2]\mathbb{E}[X_{h'}^2]}}, \quad h,h' \in [0,1]. \tag{4.2.4}$$

For any $\alpha \in [0,1]$ and any $\beta > 0$, define the *(normalized) free energy of the perturbed model* by

$$f_{\alpha,\beta,T}(u) \doteq \frac{1}{\log\log T} \log \int_0^1 e^{\beta(uX_h(\alpha)+X_h)} dh, \quad u > -1. \tag{4.2.5}$$

The parameter u is there to allow perturbations in the correlation structure of the model. When $u = 0$, we recover the *free energy*. Finally, for any Borel set $A \in \mathcal{B}([0,1])$, define the *Gibbs measure* by

$$G_{\beta,T}(A) = \int_A \frac{e^{\beta X_h}}{\int_{[0,1]} e^{\beta X_{h'}} dh'} dh. \tag{4.2.6}$$

The parameter β is called the *inverse temperature* in statistical mechanics.

4.3. Main result

The main result of this article is to present a complete description of the joint law of the overlaps for the model (4.2.1), under the *limiting mean Gibbs measure*

$$\lim_{T\to\infty} \mathbb{E} G_{\beta,T}^{\times\infty}. \tag{4.3.1}$$

We will show that, when $\beta > \beta_c \doteq 2$, this measure is the expectation E of a random measure μ_β sampling orthonormal vectors in an infinite-dimensional separable Hilbert space, where the probability weights follow a

Poisson-Dirichlet distribution of parameter β_c/β.

This is done through what is called the *Ghirlanda-Guerra identities.* These identities first appeared in Ghirlanda and Guerra (1998) and, 15 years later, it was proved in a celebrated work of Panchenko Panchenko (2013a) (a simple proof is given in Panchenko

(2011) when $E\mu_\beta$ has a finite support) that if a random measure on the unit ball of a separable Hilbert space satisfies an extended version of the Ghirlanda-Guerra identities, then we must have ultrametricity (a tree-like structure) of the overlaps under the mean of this random measure. This was an important step because it was well-known following the publication of Ghirlanda and Guerra (1998) that the Ghirlanda-Guerra identities and ultrametricity together completely determine the joint law of the overlaps, up to the distribution of one overlap. See e.g., Theorem 6.1 in Baffioni and Rosati (2000), Section 1.2 in Talagrand (2003) (in the context of the REM model from Derrida (1980)) and Theorem 1.13 in Bovier and Kurkova (2004a) (in the context of the GREM model from Derrida (1985)).

Thus, from the work of Panchenko, proving the (extended) Ghirlanda-Guerra identities under (4.3.1) implies ultrametricity and, consequently, determines the joint law of the overlaps, up to the *limiting two-overlap distribution*

$$\lim_{T\to\infty} \mathbb{E} G_{\beta,T}^{\times 2}[\mathbf{1}_{\{\rho(h,h')\in\,\cdot\,\}}], \tag{4.3.2}$$

which Arguin and Tai (2018) already determined for the model (4.2.1).

Theorem 4.3.1 (Theorem 1 in Arguin and Tai (2018))**.** *For any $\beta > \beta_c \triangleq 2$ and any Borel set $A \in \mathcal{B}([0,1])$,*

$$\lim_{T\to\infty} \mathbb{E} G_{\beta,T}^{\times 2}\left[\mathbf{1}_{\{\rho(h,h')\in A\}}\right] = \frac{2}{\beta}\mathbf{1}_A(0) + \left(1 - \frac{2}{\beta}\right)\mathbf{1}_A(1). \tag{4.3.3}$$

Remark 4.3.1. *The limiting two-overlap distribution in (4.3.3) can be interpreted as a measure of relative distance between the extremes of the model.*

To state our main result, recall the definition of a Poisson-Dirichlet variable. For $0 < \theta < 1$, let $\eta = (\eta_i)_{i\in\mathbb{N}^*}$ be the atoms of a Poisson random measure on $(0,\infty)$ with intensity measure $\theta x^{-\theta-1}dx$. A *Poisson-Dirichlet variable* ξ of parameter θ is a random variable on the space of decreasing weights

$$\left\{(x_1,x_2,\dots)\in[0,1]^{\mathbb{N}^*} : \begin{array}{l} 1\geq x_1\geq x_2\geq\dots\geq 0 \\ \text{and } \sum_{i=1}^\infty x_i = 1\end{array}\right\} \tag{4.3.4}$$

which has the same law as

$$\xi \stackrel{\text{law}}{=} \left(\frac{\eta_i}{\sum_{j=1}^\infty \eta_j},\ i\in\mathbb{N}^*\right)_{\downarrow}, \tag{4.3.5}$$

where $\downarrow$ stands for the decreasing rearrangement.

Here is the main result.

Theorem 4.3.2 (Main result). *Let $\beta > \beta_c \mathrel{\hat{=}} 2$ and let $\xi = (\xi_k)_{k\in\mathbb{N}^*}$ be a Poisson-Dirichlet variable of parameter β_c/β. Denote by E the expectation with respect to ξ. For any continuous function $\phi : [0,1]^{s(s-1)/2} \to \mathbb{R}$ of the overlaps of s points,*

$$\begin{aligned}\lim_{T\to\infty} \mathbb{E}G_{\beta,T}^{\times s}\Big[\phi\Big(\big(\rho(h_l, h_{l'})\big)_{1\le l,l'\le s}\Big)\Big]\\ = E\left[\sum_{k_1,\dots,k_s\in\mathbb{N}} \xi_{k_1}\cdots\xi_{k_s}\phi\Big(\big(\mathbf{1}_{\{k_l=k_{l'}\}}\big)_{1\le l,l'\le s}\Big)\right].\end{aligned} \tag{4.3.6}$$

Remark 4.3.2. *The domain of ϕ is $[0,1]^{s(s-1)/2}$ here because the matrix $(\rho(h_l, h_{l'}))_{1\le l,l'\le s}$ is symmetric and has 1's on the diagonal.*

Remark 4.3.3. *The proof of Theorem 4.3.2 is given in Section 4.6. As mentioned earlier, it is a consequence of Theorem 4.3.1, Theorem 4.5.8 and the ultrametric structure of the overlaps under the limiting mean Gibbs measure. To prove the extended Ghirlanda-Guerra identities in Section 4.5, we will use the strategy developed in Bovier and Kurkova (2004a,b) and used in Arguin and Zindy (2014) and Arguin and Tai (2018) (see Remark 4.3.4). For an alternative strategy (which requires a stronger control on the path of the maximal particle in the tree structure), see Jagannath (2016).*

Remark 4.3.4. *In this paper, we state most of our results above the critical inverse temperature (i.e. at low temperature), namely when $\beta > \beta_c \mathrel{\hat{=}} 2$, because that's the only interesting case. The description of the joint law of the overlaps under the limiting mean Gibbs measure turns out to be trivial when $\beta < \beta_c$. Here's why.*

When $\beta > \beta_c$, the Gibbs measure gives a lot of weight to the "particles" h that are near the maximum's height in the tree structure underlying the model (4.2.1). The result of Theorem 4.3.1 simply says that if you sample two particles under the Gibbs measure, then, in the limit and on average, either the particles branched off "at the last moment" in the tree structure (there are clusters of points reaching near the level of the maximum) or they branched off in the beginning. They cannot branch at intermediate scales.

When $\beta < \beta_c$, the weights in the Gibbs measure are more spread out so that most *contributions to the free energy actually come from particles reaching heights that are well below the level of the maximum in the tree structure. Hence, when two particles are selected from this larger pool of contributors that are not clustering, it can be shown that, in the limit and on average, the particles necessarily branched off in the beginning of the tree. The proof would follow the exact same strategy used in Arguin and Tai (2018) :*

- *find the free energy of the perturbed model as a function of the perturbation parameter u,*
- *link the expectation of the derivative of the perturbed free energy at $u = 0$ with the two-overlap distribution by using an approximate integration by parts argument and the convexity of the free energy.*

(We refer to this strategy as the Bovier-Kurkova technique since it is adapted from the strategy introduced in Bovier and Kurkova (2004a,b) for the GREM model.) The computations would actually be easier in this case. One would find that

$$\lim_{T\to\infty} \mathbb{E}G_{\beta,T}^{\times 2}\left[\mathbf{1}_{\{\rho(h,h')\in A\}}\right] = \mathbf{1}_A(0). \tag{4.3.7}$$

In other words, when $\beta < \beta_c$, the limiting mean Gibbs measure only samples points that are uncorrelated (and thus far from each other) in the limiting tree structure. More generally, our main result (Theorem 4.3.2), which describes the joint law of the overlaps under the limiting mean Gibbs measure, would say that for any continuous function $\phi : [0,1]^{s^2} \to \mathbb{R}$ of the overlaps of s points,

$$\lim_{T\to\infty} \mathbb{E}G_{\beta,T}^{\times s}\left[\phi\Big(\big(\rho(h_l, h_{l'})\big)_{1\le l,l'\le s}\Big)\right] = \phi(I_s), \tag{4.3.8}$$

where I_s denotes the identity matrix of order s. In the critical case $\beta = \beta_c$, we obtain (4.3.7) and (4.3.8) with the same techniques.

4.4. Known results

In this section, we gather the results from Arguin and Tai (2018) that we will use in Section 4.5 to prove the extended Ghirlanda-Guerra identities. The two propositions below are known convergence results for $f_{\alpha,\beta,T}$ and its derivative (with respect to u). We slightly reformulate them for later use.

Proposition 4.4.1 (Proposition 3 in Arguin and Tai (2018)). *Let $\beta > \beta_c \doteq 2$ and $0 < \alpha < 1$. Then,*

$$\frac{2}{\beta^2} \cdot \mathbb{E}\Big[f'_{\alpha,\beta,T}(0)\Big] = \int_0^\alpha \mathbb{E}G_{\beta,T}^{\times 2}[\mathbf{1}_{\{\rho(h,h') \leq y\}}]dy + o_T(1). \tag{4.4.1}$$

Since $f'_{\alpha,\beta,T}(0) = \beta(\log\log T)^{-1} G_{\beta,T}[X_h(\alpha)]$, we can also write (4.4.1) as

$$\frac{1}{\beta} \cdot \frac{\mathbb{E}G_{\beta,T}[X_h(\alpha)]}{\frac{1}{2}\log\log T} = \alpha - \mathbb{E}G_{\beta,T}^{\times 2}[\int_0^\alpha \mathbf{1}_{\{y < \rho(h,h')\}}dy] + o_T(1). \tag{4.4.2}$$

Proposition 4.4.2 (Equation 13, Proposition 4 and Lemma 14 in Arguin and Tai (2018)). *Let $\beta > \beta_c \doteq 2$, $0 \leq \alpha \leq 1$ and $u > -1$. Then,*

$$\lim_{T\to\infty} f_{\alpha,\beta,T}(u) = f_{\alpha,\beta}(u) \doteq \begin{cases} \frac{\beta^2}{4} V_{\alpha,u}, & \text{if } u < 0,\ 2 < \beta \leq 2/\sqrt{V_{\alpha,u}}, \\ \beta\sqrt{V_{\alpha,u}} - 1, & \text{if } u < 0,\ \beta > 2/\sqrt{V_{\alpha,u}}, \\ \beta(\alpha u + 1) - 1, & \text{if } u \geq 0,\ \beta > 2, \end{cases} \tag{4.4.3}$$

where the limit holds in L^1, and where $V_{\alpha,u} \doteq (1+u)^2\alpha + (1-\alpha)$.

4.5. Proof of the extended Ghirlanda-Guerra identities

This section is dedicated to the proof of the extended Ghirlanda-Guerra identities (Theorem 4.5.8). We adopt a "bottom-up" style of presentation, where Theorem 4.5.8 is the end goal. Here is the structure of the proof :

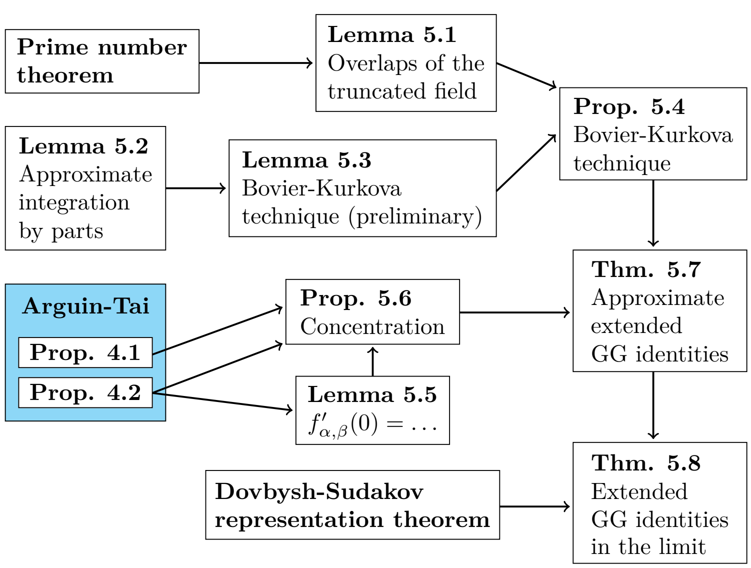


FIGURE 4.5.1. Structure of the proof

We start by relating the overlaps of the field X to the overlaps of the truncated field $X(\alpha)$.

Lemma 4.5.1 (Overlaps of the truncated field)**.** *Let $0 \leq \alpha \leq 1$. Then, for all $h, h' \in [0,1]$,*

$$\frac{\mathbb{E}[X_h(\alpha)X_{h'}(\alpha)]}{\frac{1}{2}\log\log T} = \begin{cases} \rho(h,h') + O((\log\log T)^{-1})\,, & \text{if } \rho(h,h') \leq \alpha, \\ \alpha + O((\log\log T)^{-1})\,, & \text{if } \rho(h,h') > \alpha. \end{cases} \tag{4.5.1}$$

In both cases, the $O((\log\log T)^{-1})$ term is uniform in α.

Proof. Since $\mathrm{Re}(z) = (z+\overline{z})/2$, $\mathbb{E}[U_p^2] = \mathbb{E}[(\overline{U_p})^2] = 0$ and $\mathbb{E}[U_p\overline{U_p}] = 1$, it is easily shown from (4.2.1) that, for any prime p,

$$\mathbb{E}[W_p(h)W_p(h')] = \frac{1}{2p}\cos(|h-h'|\log p), \quad h,h' \in [0,1]. \tag{4.5.2}$$

Thus, from the independence of the U_p's and (4.2.3),

$$\mathbb{E}[X_h(\alpha)X_{h'}(\alpha)] = \sum_{p\leq \exp((\log T)^\alpha)} \frac{1}{2p}\cos(|h-h'|\log p), \quad h,h' \in [0,1]. \tag{4.5.3}$$

Sums like the one on the right-hand side of (4.5.3) were estimated on page 20 of Appendix A in Harper (2013) by using the prime number theorem. In particular,

$$\rho(h,h') = \frac{\frac{1}{2}\log\Big((\log T)\wedge |h-h'|^{-1}\Big)}{\frac{1}{2}\log\log T} + O\Big((\log\log T)^{-1}\Big)\,, \tag{4.5.4}$$

and

$$\frac{\mathbb{E}[X_h(\alpha)X_{h'}(\alpha)]}{\frac{1}{2}\log\log T} = \begin{cases} \frac{\log|h-h'|^{-1}}{\log\log T} + O((\log\log T)^{-1})\,, & \text{if } 1 \leq |h-h'|^{-1} < (\log T)^\alpha, \\ \alpha + O((\log\log T)^{-1})\,, & \text{if } |h-h'|^{-1} \geq (\log T)^\alpha, \end{cases} \tag{4.5.5}$$

where the $O((\log\log T)^{-1})$ terms are all uniform in α. By comparing (4.5.4) and (4.5.5), we get

$$\begin{aligned} \frac{\mathbb{E}[X_h(\alpha)X_{h'}(\alpha)]}{\frac{1}{2}\log\log T} &= \begin{cases} \rho(h,h') + O((\log\log T)^{-1})\,, & \text{if } \rho(h,h') - O((\log\log T)^{-1}) < \alpha, \\ \alpha + O((\log\log T)^{-1})\,, & \text{if } \rho(h,h') - O((\log\log T)^{-1}) \geq \alpha, \end{cases} \\ &= \begin{cases} \rho(h,h') + O((\log\log T)^{-1})\,, & \text{if } \rho(h,h') \leq \alpha, \\ \alpha + O((\log\log T)^{-1})\,, & \text{if } \rho(h,h') > \alpha. \end{cases} \end{aligned} \tag{4.5.6}$$

This ends the proof. □

The next lemma is an approximate integration by parts result. It is a straightforward generalization of Lemma 9 in Arguin and Tai (2018).

Lemma 4.5.2 (Approximate integration by parts)**.** *Let $s \in \mathbb{N}^*$ and let $\boldsymbol{\xi} \mathrel{\hat{=}} (\xi_1, \xi_2, \dots, \xi_s)$ be a random vector taking values in $\mathbb{C}^s$, such that $\mathbb{E}[|\xi_j|^3] < \infty$ and $\mathbb{E}[\xi_j] = 0$ for all $j \in \{1, \dots, s\}$, and such that $\mathbb{E}[\xi_l \xi_j] = 0$ for all $l, j \in \{1, \dots, s\}$. Let $F : \mathbb{C}^s \to \mathbb{C}$ be a twice continuously differentiable function such that, for some $M > 0$,*

$$\max_{1 \leq j \leq s} \left\{ \|\partial_{z_j}^2 F\|_\infty \vee \|\partial_{\overline{z_j}}^2 F\|_\infty \vee \|\partial_{z_j} \partial_{\overline{z_j}} F\|_\infty \vee \|\partial_{\overline{z_j}} \partial_{z_j} F\|_\infty \right\} \leq M,$$

where $\|f\|_\infty \mathrel{\hat{=}} \sup_{\boldsymbol{z} \in \mathbb{C}^s} |f(\boldsymbol{z}, \overline{\boldsymbol{z}})|$. Then, for any $k \in \{1, \dots, s\}$,

$$\left| \mathbb{E}[\xi_k F(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}})] - \sum_{j=1}^s \mathbb{E}[\xi_k \overline{\xi_j}] \ \mathbb{E}[\partial_{\overline{z_j}} F(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}})] \right| \ll s^2 M \max_{1 \leq j \leq s} \mathbb{E}[|\xi_j|^3], \tag{4.5.7}$$

$$\left| \mathbb{E}[\overline{\xi_k} F(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}})] - \sum_{j=1}^s \mathbb{E}[\overline{\xi_k} \xi_j] \ \mathbb{E}[\partial_{z_j} F(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}})] \right| \ll s^2 M \max_{1 \leq j \leq s} \mathbb{E}[|\xi_j|^3], \tag{4.5.8}$$

where $f(\cdot) \ll g(\cdot)$ means that $|f(\cdot)| \leq C g(\cdot)$ for some universal constant $C > 0$ (the Vinogradov notation).

Proof. Fix $k \in \{1, \dots, s\}$. We only prove (4.5.7) because the proof of (4.5.8) is almost identical. Since $\mathbb{E}[\xi_k] = 0$ and $\mathbb{E}[\xi_k \xi_j] = 0$ for all $j \in \{1, \dots, s\}$, the left-hand side of (4.5.7) can be written as

$$\begin{aligned} &\mathbb{E}\Big[\xi_k \Big(F(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}}) - F(\mathbf{0}, \mathbf{0}) - \sum_{j=1}^s \xi_j \partial_{z_j} F(\mathbf{0}, \mathbf{0}) - \sum_{j=1}^s \overline{\xi_j} \partial_{\overline{z_j}} F(\mathbf{0}, \mathbf{0})\Big)\Big] \\ &\qquad - \sum_{j=1}^s \mathbb{E}\Big[\xi_k \overline{\xi_j}\Big] \mathbb{E}\Big[\partial_{\overline{z_j}} F(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}}) - \partial_{\overline{z_j}} F(\mathbf{0}, \mathbf{0})\Big]. \end{aligned} \tag{4.5.9}$$

By Taylor's theorem in several variables and the assumptions, the following estimates hold

$$\begin{aligned} &\left| F(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}}) - F(\mathbf{0}, \mathbf{0}) - \sum_{j=1}^s \xi_j \partial_{z_j} F(\mathbf{0}, \mathbf{0}) - \sum_{j=1}^s \overline{\xi_j} \partial_{\overline{z_j}} F(\mathbf{0}, \mathbf{0}) \right| \\ &\qquad \ll M \Big(\sum_{l=1}^s |\xi_l|\Big)^2 \leq M \, s \sum_{l=1}^s |\xi_l|^2, \end{aligned} \tag{4.5.10}$$

$$\left| \partial_{\overline{z_j}} F(\boldsymbol{\xi}, \overline{\boldsymbol{\xi}}) - \partial_{\overline{z_j}} F(\mathbf{0}, \mathbf{0}) \right| \ll M \sum_{l=1}^s |\xi_l| \quad \text{for all } j \in \{1, \dots, s\}. \tag{4.5.11}$$

Therefore,

$$
\begin{aligned}
|(4.5.9)| &\ll M \sum_{l=1}^{s} \Big(s\,\mathbb{E}\Big[|\xi_k| \cdot |\xi_l|^2\Big] + \sum_{j=1}^{s} \mathbb{E}\Big[|\xi_k| \cdot |\xi_j|\Big]\mathbb{E}\Big[|\xi_l|\Big]\Big) \\
&\leq M \sum_{l=1}^{s} \Big(s\,\mathbb{E}\Big[|\xi_k|^3\Big]^{1/3}\mathbb{E}\Big[(|\xi_l|^2)^{3/2}\Big]^{2/3} + \sum_{j=1}^{s} \mathbb{E}\Big[|\xi_k|^3\Big]^{1/3}\mathbb{E}\Big[|\xi_j|^3\Big]^{1/3}\mathbb{E}\Big[|\xi_l|^3\Big]^{1/3}\Big) \\
&\leq 2s^2 M \max_{1\leq j\leq s} \mathbb{E}[|\xi_j|^3], \qquad (4.5.12)
\end{aligned}
$$

where we used Holder's inequality to obtain the second inequality. □

Here is a generalization of Proposition 10 in Arguin and Tai (2018). It could be seen as a generalization of (4.4.2) if (4.4.2) was applied to $(W_p(h), h \in [0,1])$ instead of $(X_h(\alpha), h \in [0,1])$.

Lemma 4.5.3 (Bovier-Kurkova technique - preliminary version)**.** *Let $\beta > 0$ and $p \leq T$. For any $s \in \mathbb{N}^*$, any $k \in \{1, \ldots, s\}$, and any bounded mesurable function $\phi : [0,1]^s \to \mathbb{R}$, we have*

$$
\begin{aligned}
&\Big|\mathbb{E}G_{\beta,T}^{\times s}[W_p(h_k)\phi(\boldsymbol{h})] \\
&\quad - \beta \cdot \left\{ \begin{array}{l} \sum_{l=1}^{s} \mathbb{E}G_{\beta,T}^{\times s}\Big[\mathbb{E}[W_p(h_k)W_p(h_l)]\,\phi(\boldsymbol{h})\Big] \\ -s\,\mathbb{E}G_{\beta,T}^{\times (s+1)}\Big[\mathbb{E}[W_p(h_k)W_p(h_{s+1})]\,\phi(\boldsymbol{h})\Big] \end{array} \right\}\Bigg| \leq K p^{-3/2}, \qquad (4.5.13)
\end{aligned}
$$

where $\boldsymbol{h} \doteq (h_1, h_2, \ldots, h_s)$, $K \doteq s^2 C \beta^2 \|\phi\|_\infty$, and $C > 0$ is a universal constant.

Proof. Write for short

$$
\omega_p(h) \doteq \frac{1}{2} p^{-ih-1/2} \quad \text{and} \quad Y_p(h) \doteq \beta \sum_{\substack{q \leq T \\ q \neq p}} W_q(h). \qquad (4.5.14)
$$

Define

$$
F_p(\boldsymbol{z}, \overline{\boldsymbol{z}}) \doteq \frac{\int_{[0,1]^s} \omega_p(h_k)\phi(\boldsymbol{h}) \prod_{l=1}^{s} \exp\Big(\beta(z_l \omega_p(h_l) + \overline{z_l}\overline{\omega_p(h_l)}) + Y_p(h_l)\Big) d\boldsymbol{h}}{\int_{[0,1]^s} \prod_{l=1}^{s} \exp\Big(\beta(z_l \omega_p(h_l) + \overline{z_l}\overline{\omega_p(h_l)}) + Y_p(h_l)\Big) d\boldsymbol{h}}. \qquad (4.5.15)
$$

Then,

$$
\mathbb{E}G_{\beta,T}^{\times s}[W_p(h_k)\phi(\boldsymbol{h})] = \mathbb{E}[U_p \cdot F_p(\boldsymbol{U}_p, \overline{\boldsymbol{U}_p})] + \mathbb{E}[\overline{U_p} \cdot \overline{F_p}(\boldsymbol{U}_p, \overline{\boldsymbol{U}_p})], \qquad (4.5.16)
$$

where $\boldsymbol{U}_p \doteq (U_p, U_p, \ldots, U_p)$. Since the U_p's are i.i.d. uniform random variables on the unit circle in $\mathbb{C}$, we have $\mathbb{E}[|U_p|^3] < \infty$, $\mathbb{E}[U_p\overline{U_p}] = 1$ and $\mathbb{E}[U_p^2] = \mathbb{E}[U_p] = 0$. If we apply (4.5.7) with $F = F_p$ and $\boldsymbol{\xi} = \boldsymbol{U}_p$, and (4.5.8) with $F = \overline{F_p}$ and $\boldsymbol{\xi} = \boldsymbol{U}_p$, we get, as $T \to \infty$,

$$\begin{aligned} \mathbb{E} G_{\beta,T}^{\times s}[W_p(h_k)\phi(\boldsymbol{h})] = \sum_{j=1}^{s} \Big\{ \mathbb{E}\Big[\partial_{\overline{z_j}} F_p(\boldsymbol{U}_p, \overline{\boldsymbol{U}_p})\Big] + \mathbb{E}\Big[\partial_{z_j} \overline{F_p}(\boldsymbol{U}_p, \overline{\boldsymbol{U}_p})\Big] \Big\} \\ + s^2 \, O\Big(\max_{1 \leq j \leq s} \big\{ \|\partial_{z_j}^2 F_p\|_\infty \vee \|\partial_{\overline{z_j}}^2 F_p\|_\infty \big\} \Big). \end{aligned} \tag{4.5.17}$$

For any bounded mesurable function $H : [0,1] \to \mathbb{C}$, define

$$\langle H \rangle_{(z,\overline{z})} \mathrel{\hat{=}} \langle H(h) \rangle_{(z,\overline{z})} \mathrel{\hat{=}} \frac{\int_{[0,1]} H(h) \exp\Big(\beta(z\omega_p(h) + \overline{z}\overline{\omega_p(h)}) + Y_p(h)\Big) dh}{\int_{[0,1]} \exp\Big(\beta(z\omega_p(h) + \overline{z}\overline{\omega_p(h)}) + Y_p(h)\Big) dh}, \tag{4.5.18}$$

and for any bounded mesurable function $H : [0,1]^s \to \mathbb{C}$, define

$$\langle H \rangle_{(\boldsymbol{z},\overline{\boldsymbol{z}})}^{\phi} \mathrel{\hat{=}} \langle H(\boldsymbol{h}) \rangle_{(\boldsymbol{z},\overline{\boldsymbol{z}})}^{\phi} \mathrel{\hat{=}} \frac{\int_{[0,1]^s} H(\boldsymbol{h})\phi(\boldsymbol{h}) \prod_{l=1}^{s} \exp\Big(\beta(z_l\omega_p(h_l) + \overline{z_l}\overline{\omega_p(h_l)}) + Y_p(h_l)\Big) d\boldsymbol{h}}{\int_{[0,1]^s} \prod_{l=1}^{s} \exp\Big(\beta(z_l\omega_p(h_l) + \overline{z_l}\overline{\omega_p(h_l)}) + Y_p(h_l)\Big) d\boldsymbol{h}}. \tag{4.5.19}$$

Differentiation of the above yields

$$\begin{aligned} \partial_{\overline{z_j}} \langle H \rangle_{(\boldsymbol{z},\overline{\boldsymbol{z}})}^{\phi} &= \beta \Big\{ \langle H \overline{\omega_p(h_j)} \rangle_{(\boldsymbol{z},\overline{\boldsymbol{z}})}^{\phi} - \langle H \rangle_{(\boldsymbol{z},\overline{\boldsymbol{z}})}^{\phi} \langle \overline{\omega_p(h_{s+1})} \rangle_{(z_j,\overline{z_j})} \Big\}, \\ \partial_{z_j} \langle H \rangle_{(\boldsymbol{z},\overline{\boldsymbol{z}})}^{\phi} &= \beta \Big\{ \langle H \omega_p(h_j) \rangle_{(\boldsymbol{z},\overline{\boldsymbol{z}})}^{\phi} - \langle H \rangle_{(\boldsymbol{z},\overline{\boldsymbol{z}})}^{\phi} \langle \omega_p(h_{s+1}) \rangle_{(z_j,\overline{z_j})} \Big\}. \end{aligned} \tag{4.5.20}$$

The partial derivatives in (4.5.20) can be used to expand the summands on the right-hand side of (4.5.17). Indeed, by using the relation $F_p(\boldsymbol{z}, \overline{\boldsymbol{z}}) = \langle \omega_p(h_k) \rangle_{(\boldsymbol{z},\overline{\boldsymbol{z}})}^{\phi}$ with $\boldsymbol{z} = \boldsymbol{U}_p$,

$$\begin{aligned} &\mathbb{E}\Big[\partial_{\overline{z_j}} F_p(\boldsymbol{U}_p, \overline{\boldsymbol{U}_p})\Big] + \mathbb{E}\Big[\partial_{z_j} \overline{F_p}(\boldsymbol{U}_p, \overline{\boldsymbol{U}_p})\Big] \\ &= \mathbb{E}\Big[\partial_{\overline{z_j}} \langle \omega_p(h_k) \rangle_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})}^{\phi}\Big] + \mathbb{E}\Big[\partial_{z_j} \Big\langle \overline{\omega_p(h_k)} \Big\rangle_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})}^{\phi}\Big] \\ &\overset{(4.5.20)}{=} \beta\, \mathbb{E}\Big[\Big\langle \omega_p(h_k)\overline{\omega_p(h_j)} \Big\rangle_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})}^{\phi} - \langle \omega_p(h_k) \rangle_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})}^{\phi} \langle \overline{\omega_p(h_{s+1})} \rangle_{(U_p,\overline{U_p})} \Big] \\ &\quad + \beta\, \mathbb{E}\Big[\Big\langle \overline{\omega_p(h_k)\overline{\omega_p(h_j)}} \Big\rangle_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})}^{\phi} - \overline{\langle \omega_p(h_k) \rangle_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})}^{\phi} \langle \overline{\omega_p(h_{s+1})} \rangle_{(U_p,\overline{U_p})}} \Big] \\ &= \beta \cdot \left\{ \begin{array}{l} \mathbb{E}\Big[\Big\langle 2\mathrm{Re}\Big(\omega_p(h_k)\overline{\omega_p(h_j)}\Big) \Big\rangle_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})}^{\phi} \Big] \\ -\mathbb{E}\Big[2\mathrm{Re}\Big(\langle \omega_p(h_k) \rangle_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})}^{\phi} \langle \overline{\omega_p(h_{s+1})} \rangle_{(U_p,\overline{U_p})} \Big) \Big] \end{array} \right\}. \end{aligned} \tag{4.5.21}$$

Since, by definition,

$$\langle \cdot \rangle^{\phi}_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})} = G^{\times s}_{\beta,T}[\,\cdot\,\phi(\boldsymbol{h})], \tag{4.5.22}$$

and

$$\begin{aligned}
&2\mathrm{Re}\Big(\langle \omega_p(h_k)\rangle^{\phi}_{(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})}\langle\overline{\omega_p(h_{s+1})}\rangle_{(U_p,\overline{U_p})}\Big) \\
&\quad= 2\mathrm{Re}\left(\frac{\int_{[0,1]}\int_{[0,1]^s}\omega_p(h_k)\overline{\omega_p(h_{s+1})}\phi(\boldsymbol{h})\prod_{l=1}^{s+1}\exp\Big(\beta\sum_{p\leq T}W_p(h_l)\Big)d\boldsymbol{h}\,dh_{s+1}}{\int_{[0,1]}\int_{[0,1]^s}\prod_{l=1}^{s+1}\exp\Big(\beta\sum_{p\leq T}W_p(h_l)\Big)d\boldsymbol{h}\,dh_{s+1}}\right) \\
&\quad= G^{\times(s+1)}_{\beta,T}\Big[2\mathrm{Re}\Big(\omega_p(h_k)\overline{\omega_p(h_{s+1})}\Big)\phi(\boldsymbol{h})\Big],
\end{aligned} \tag{4.5.23}$$

and

$$2\,\mathrm{Re}(\omega_p(h)\overline{\omega_p(h')}) = \frac{1}{2p}\cos(|h-h'|\log p) \overset{(4.5.2)}{=} \mathbb{E}[W_p(h)W_p(h')], \tag{4.5.24}$$

we can rewrite (4.5.21) as

$$\begin{aligned}
&\mathbb{E}\Big[\partial_{\overline{z_j}}F_p(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})\Big] + \mathbb{E}\Big[\partial_{z_j}\overline{F_p}(\boldsymbol{U}_p,\overline{\boldsymbol{U}_p})\Big] \\
&\qquad= \beta\cdot\left\{\begin{array}{l}\mathbb{E}G^{\times s}_{\beta,T}\Big[\mathbb{E}[W_p(h_k)W_p(h_j)]\phi(\boldsymbol{h})\Big] \\ -\mathbb{E}G^{\times(s+1)}_{\beta,T}\Big[\mathbb{E}[W_p(h_k)W_p(h_{s+1})]\phi(\boldsymbol{h})\Big]\end{array}\right\}.
\end{aligned} \tag{4.5.25}$$

From (4.5.17) and (4.5.25), we conclude (4.5.13), as long as, for all $j\in\{1,\dots,s\}$,

$$\|\partial^2_{z_j}F\|_\infty \vee \|\partial^2_{\overline{z_j}}F\|_\infty \leq \widetilde{C}\beta^2\|\phi\|_\infty p^{-3/2}, \tag{4.5.26}$$

where $\widetilde{C}>0$ is a universal constant. To verify this last point, note that, by differentiating in (4.5.20),

$$\begin{aligned}
\partial^2_{z_j}\langle H\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})} &= \beta\left\{\begin{array}{l}\partial_{z_j}\langle H\omega_p(h_j)\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})} - (\partial_{z_j}\langle H\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})})\langle\omega_p(h_{s+1})\rangle_{(z_j,\overline{z_j})} \\ -\langle H\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}(\partial_{z_j}\langle\omega_p(h_{s+1})\rangle_{(z_j,\overline{z_j})})\end{array}\right\} \\
&= \beta^2\left\{\begin{array}{l}\langle H\omega_p^2(h_j)\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})} - \langle H\omega_p(h_j)\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}\langle\omega_p(h_{s+1})\rangle_{(z_j,\overline{z_j})} \\ -\Big(\langle H\omega_p(h_j)\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})} - \langle H\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}\langle\omega_p(h_{s+1})\rangle_{(z_j,\overline{z_j})}\Big)\langle\omega_p(h_{s+1})\rangle_{(z_j,\overline{z_j})} \\ -\langle H\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}\Big(\langle\omega_p^2(h_{s+1})\rangle_{(z_j,\overline{z_j})} - \langle\omega_p(h_{s+1})\rangle^2_{(z_j,\overline{z_j})}\Big)\end{array}\right\}.
\end{aligned} \tag{4.5.27}$$

Using the relation $F_p(\boldsymbol{z},\overline{\boldsymbol{z}}) = \langle \omega_p(h_k)\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}$, (4.5.27), and the triangle inequality, we obtain

$$
\begin{aligned}
|\partial^2_{z_j} F_p(\boldsymbol{z},\overline{\boldsymbol{z}})| &= \beta^2 \left|
\begin{aligned}
&\langle \omega_p(h_k)\omega_p^2(h_j)\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})} \\
&-2\langle \omega_p(h_k)\omega_p(h_j)\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}\langle \omega_p(h_{s+1})\rangle_{(z_j,\overline{z_j})} \\
&+2\langle \omega_p(h_k)\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}\langle \omega_p(h_{s+1})\rangle^2_{(z_j,\overline{z_j})} \\
&-\langle \omega_p(h_k)\rangle^{\phi}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}\langle \omega_p^2(h_{s+1})\rangle_{(z_j,\overline{z_j})}
\end{aligned}
\right| \\
&\leq \beta^2 \left\{
\begin{aligned}
&\langle |\omega_p(h_k)|\cdot|\omega_p(h_j)|^2\rangle^{|\phi|}_{(\boldsymbol{z},\overline{\boldsymbol{z}})} \\
&+2\langle |\omega_p(h_k)|\cdot|\omega_p(h_j)|\rangle^{|\phi|}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}\langle |\omega_p(h_{s+1})|\rangle_{(z_j,\overline{z_j})} \\
&+2\langle |\omega_p(h_k)|\rangle^{|\phi|}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}\langle |\omega_p(h_{s+1})|\rangle^2_{(z_j,\overline{z_j})} \\
&+\langle |\omega_p(h_k)|\rangle^{|\phi|}_{(\boldsymbol{z},\overline{\boldsymbol{z}})}\langle |\omega_p(h_{s+1})|^2\rangle_{(z_j,\overline{z_j})}
\end{aligned}
\right\}.
\end{aligned}
\tag{4.5.28}
$$

Since $|\omega_p(h)| = \frac{1}{2}p^{-1/2}$, $\langle 1\rangle_{(z_j,\overline{z_j})} = 1$ and $\langle 1\rangle^{|\phi|}_{(\boldsymbol{z},\overline{\boldsymbol{z}})} \leq \|\phi\|_\infty$, we deduce from (4.5.28) that

$$
|\partial^2_{z_j} F_p(\boldsymbol{z},\overline{\boldsymbol{z}})| \leq \frac{6}{8}\beta^2\|\phi\|_\infty p^{-3/2}. \tag{4.5.29}
$$

We obtain the bound on $\|\partial^2_{\overline{z_j}} F_p\|_\infty$ in the same manner. □

The next proposition is a consequence of the two previous lemmas. It generalizes (4.4.2), which corresponds to the special case $(k = 1, s = 1, \phi \equiv 1)$. The idea for the statement originates from Bovier and Kurkova (2004a), and the idea behind the proof generalizes the special-case application in Arguin and Zindy (2014). See Arguin and Zindy (2015); Ouimet (2017) for an application in the context of the Gaussian free field.

Proposition 4.5.4 (Bovier-Kurkova technique). *Let $\beta > 0$ and $0 \leq \alpha \leq 1$. For any $s \in \mathbb{N}^*$, any $k \in \{1,\dots,s\}$, and any bounded mesurable function $\phi : [0,1]^s \to \mathbb{R}$, we have*

$$
\begin{aligned}
&\Bigg| \frac{1}{\beta}\cdot\frac{\mathbb{E}G^{\times s}_{\beta,T}\Big[X_{h_k}(\alpha)\phi(\boldsymbol{h})\Big]}{\frac{1}{2}\log\log T} \\
&\quad - \left\{
\begin{aligned}
&\textstyle\sum_{l=1}^s \mathbb{E}G^{\times s}_{\beta,T}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_k,h_l)\}}dy\ \phi(\boldsymbol{h})\Big] \\
&-s\,\mathbb{E}G^{\times(s+1)}_{\beta,T}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_k,h_{s+1})\}}dy\ \phi(\boldsymbol{h})\Big]
\end{aligned}
\right\}\Bigg| = O\Big((\log\log T)^{-1}\Big),
\end{aligned}
\tag{4.5.30}
$$

where $\boldsymbol{h} \doteq (h_1,h_2,\dots,h_s)$.

Proof. For any $l \in \{1, \ldots, s+1\}$,

$$\mathbb{E}G_{\beta,T}^{\times(s+1)}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_k,h_l)\}} dy\ \phi(\boldsymbol{h})\Big] = \mathbb{E}G_{\beta,T}^{\times(s+1)}\Big[\rho(h_k,h_l)\,\mathbf{1}_{\{\rho(h_k,h_l)\leq\alpha\}}\,\phi(\boldsymbol{h})\Big] + \mathbb{E}G_{\beta,T}^{\times(s+1)}\Big[\alpha\,\mathbf{1}_{\{\rho(h_k,h_l)>\alpha\}}\,\phi(\boldsymbol{h})\Big]. \tag{4.5.31}$$

On the other hand, if we sum (4.5.13) over the set $\{p \text{ prime} : p \leq \exp((\log T)^\alpha)\}$ and divide by $\frac{\beta}{2}\log\log T$, we obtain

$$\left|\frac{1}{\beta}\cdot\frac{\mathbb{E}G_{\beta,T}^{\times s}[X_{h_k}(\alpha)\phi(\boldsymbol{h})]}{\frac{1}{2}\log\log T} - \left\{\begin{matrix} \sum_{l=1}^{s}\mathbb{E}G_{\beta,T}^{\times s}\left[\frac{\mathbb{E}[X_{h_k}(\alpha)X_{h_l}(\alpha)]}{\frac{1}{2}\log\log T}\,\phi(\boldsymbol{h})\right] \\ -s\,\mathbb{E}G_{\beta,T}^{\times(s+1)}\left[\frac{\mathbb{E}[X_{h_k}(\alpha)X_{h_{s+1}}(\alpha)]}{\frac{1}{2}\log\log T}\,\phi(\boldsymbol{h})\right]\end{matrix}\right\}\right| = O\Big((\log\log T)^{-1}\Big). \tag{4.5.32}$$

Now, one by one, take the difference in absolute value between each of the $s+1$ expectations inside the braces in (4.5.32) and the corresponding expectation on the left-hand side of (4.5.31). We obtain the bound (4.5.30) by using Lemma 4.5.1. □

Our goal now is to combine Proposition 4.5.4 with a concentration result (Proposition 4.5.6) in order to prove an approximate version of the GG identities (Theorem 4.5.7). We will then show that the identities must hold exactly in the limit $T \to \infty$ (Theorem 4.5.8). Before stating and proving the concentration result, we show that $f_{\alpha,\beta}(\cdot)$, the limiting perturbed free energy, is differentiable in an open interval around 0.

Lemma 4.5.5. *Let $\beta > \beta_c \doteq 2$ and $0 \leq \alpha \leq 1$. There exists $\delta = \delta(\alpha,\beta) > 0$ small enough that $f_{\alpha,\beta}(\cdot)$ from Proposition 4.4.2 is differentiable on $(-\delta,\delta)$. Also, we have $f'_{\alpha,\beta}(0) = \beta\alpha$.*

Proof. Since $\beta > 2$ and $\lim_{u\to 0} V_{\alpha,u} = 1$, there exists $\delta = \delta(\alpha,\beta) > 0$ small enough that, for all $u \in (-\delta,\delta)$,

$$f_{\alpha,\beta}(u) = \begin{cases} \beta\sqrt{V_{\alpha,u}} - 1, & \text{if } u < 0, \\ \beta(\alpha u + 1) - 1, & \text{if } u \geq 0. \end{cases} \tag{4.5.33}$$

The differentiability of $f_{\alpha,\beta}(\cdot)$ on $(-\delta,\delta)\backslash\{0\}$ is obvious. Also,

$$\frac{f_{\alpha,\beta}(u) - f_{\alpha,\beta}(0)}{u} = \begin{cases} \beta\frac{\sqrt{V_{\alpha,u}}-1}{u}, & \text{if } u < 0, \\ \beta\alpha, & \text{if } u \geq 0. \end{cases} \tag{4.5.34}$$

Take both the left and right limits at 0 to conclude. □

Here is the concentration result. It is analogous to Theorem 3.8 in Panchenko (2013b), which was proved for the mixed p-spin model. We give the proof for completeness.

Proposition 4.5.6 (Concentration)**.** *Let $\beta > \beta_c \mathrel{\hat{=}} 2$ and $0 < \alpha < 1$. For any $s \in \mathbb{N}^*$, any $k \in \{1, \dots, s\}$, and any bounded mesurable function $\phi : [0,1]^s \to \mathbb{R}$, we have*

$$\left|\frac{\mathbb{E}G_{\beta,T}^{\times s}[X_{h_k}(\alpha)\phi(\boldsymbol{h})]}{\log\log T} - \frac{\mathbb{E}G_{\beta,T}[X_{h_k}(\alpha)]}{\log\log T}\mathbb{E}G_{\beta,T}^{\times s}[\phi(\boldsymbol{h})]\right| = o_T(1), \tag{4.5.35}$$

where $\boldsymbol{h} \mathrel{\hat{=}} (h_1, h_2, \dots, h_s)$.

Proof. By applying Jensen's inequality to the expectation $\mathbb{E}G_{\beta,T}^{\times s}[\,\cdot\,]$, followed by the triangle inequality,

$$\begin{aligned}
&\left|\mathbb{E}G_{\beta,T}^{\times s}[X_{h_k}(\alpha)\phi(\boldsymbol{h})] - \mathbb{E}G_{\beta,T}[X_{h_k}(\alpha)]\mathbb{E}G_{\beta,T}^{\times s}[\phi(\boldsymbol{h})]\right| \\
&\leq \mathbb{E}G_{\beta,T}\left|X_{h_k}(\alpha) - \mathbb{E}G_{\beta,T}[X_{h_k}(\alpha)]\right| \cdot \|\phi\|_\infty \\
&\leq \left\{\begin{array}{l} \mathbb{E}G_{\beta,T}\left|X_{h_k}(\alpha) - G_{\beta,T}[X_{h_k}(\alpha)]\right| \\ +\mathbb{E}\left|G_{\beta,T}[X_{h_k}(\alpha)] - \mathbb{E}G_{\beta,T}[X_{h_k}(\alpha)]\right| \end{array}\right\} \cdot \|\phi\|_\infty \\
&\mathrel{\hat{=}} \Big\{(a) + (b)\Big\} \cdot \|\phi\|_\infty.
\end{aligned}$$

Below, we show that (a) and (b) are $o(\log\log T)$ in Step 1 and Step 2, respectively.

Step 1. Note that

$$\begin{aligned}
(a) &= \mathbb{E}G_{\beta,T}\left|\int_0^1 (X_{h_1}(\alpha) - X_{h_2}(\alpha))\frac{e^{\beta X_{h_2}}}{\int_0^1 e^{\beta X_{z_2}} dz_2} dh_2\right| \\
&\leq \mathbb{E}G_{\beta,T}^{\times 2}\left|X_{h_1}(\alpha) - X_{h_2}(\alpha)\right|.
\end{aligned} \tag{4.5.36}$$

For $u \geq 0$, we define a perturbed version of the last quantity, where the Gibbs measure $G_{\beta,T,u}$ is now defined with respect to the field $(uX_h(\alpha) + X_h, h \in [0,1])$:

$$D_{\alpha,\beta,T}(u) \mathrel{\hat{=}} \mathbb{E}G_{\beta,T,u}^{\times 2}\left|X_{h_1}(\alpha) - X_{h_2}(\alpha)\right|. \tag{4.5.37}$$

We can easily verify that

$$D'_{\alpha,\beta,T}(y) = \beta\, \mathbb{E}G_{\beta,T,y}^{\times 3}\Big[\left|X_{h_1}(\alpha) - X_{h_2}(\alpha)\right| \cdot \Big(X_{h_1}(\alpha) + X_{h_2}(\alpha) - 2X_{h_3}(\alpha)\Big)\Big]. \tag{4.5.38}$$

If we separate the expectation in (4.5.38) in two parts and apply the Cauchy-Schwarz inequality to each one of them, followed by an application of the elementary inequality $(c+d)^2 \leq 2c^2 + 2d^2$, we find, for $y \geq 0$,

$$\begin{aligned}
\left|D'_{\alpha,\beta,T}(y)\right| &\leq \beta \cdot \left\{\begin{array}{l} \mathbb{E}G^{\times 3}_{\beta,T,y}\left|X_{h_1}(\alpha) - X_{h_2}(\alpha)\right|\left|X_{h_1}(\alpha) - X_{h_3}(\alpha)\right| \\ + \mathbb{E}G^{\times 3}_{\beta,T,y}\left|X_{h_1}(\alpha) - X_{h_2}(\alpha)\right|\left|X_{h_2}(\alpha) - X_{h_3}(\alpha)\right| \end{array}\right\} \\
&\leq \beta \cdot 2\, \mathbb{E}G^{\times 2}_{\beta,T,y}[(X_{h_1}(\alpha) - X_{h_2}(\alpha))^2] \\
&\leq \beta \cdot 8\, \mathbb{E}G_{\beta,T,y}[\Big(X_h(\alpha) - G_{\beta,T,y}[X_h(\alpha)]\Big)^2]. \qquad (4.5.39)
\end{aligned}$$

Note that $\beta^{-2}(\log\log T) f''_{\alpha,\beta,T}(y) = G_{\beta,T,y}[\Big(X_h(\alpha) - G_{\beta,T,y}[X_h(\alpha)]\Big)^2]$ and apply inequality (4.5.39) in the identity $u D_{\alpha,\beta,T}(0) = \int_0^u D_{\alpha,\beta,T}(y) dy - \int_0^u \int_0^x D'_{\alpha,\beta,T}(y) dy dx$. We obtain, for $u > 0$,

$$\begin{aligned}
D_{\alpha,\beta,T}(0) &\leq \frac{1}{u}\int_0^u D_{\alpha,\beta,T}(y) dy + \int_0^u \left|D'_{\alpha,\beta,T}(y)\right| dy \\
&\leq 2\left(\frac{1}{u}\int_0^u \beta^{-2}(\log\log T)\mathbb{E}[f''_{\alpha,\beta,T}(y)] dy\right)^{1/2} \\
&\quad + 8\beta \int_0^u \beta^{-2}(\log\log T)\mathbb{E}[f''_{\alpha,\beta,T}(y)] dy. \qquad (4.5.40)
\end{aligned}$$

In order to bound $\frac{1}{u}\int_0^u D_{\alpha,\beta,T}(y) dy$, we separated $D_{\alpha,\beta,T}(y)$ in two parts (with the triangle inequality) and we applied the Cauchy-Schwarz inequality to the two resulting expectations $\frac{1}{u}\int_0^u \mathbb{E}G_{\beta,T,y}[\,\cdot\,]\, dy$. Now, on the right-hand side of (4.5.40), use the convexity of $f_{\alpha,\beta,T}(\cdot)$ and the mean convergence of $f_{\alpha,\beta,T}(z)$, $z > -1$, from Proposition 4.4.2. We get, for all $u > 0$ and all $y \in (0,1)$,

$$\begin{aligned}
\limsup_{T\to\infty} \frac{(a)}{\log\log T} &\overset{(4.5.36)}{\leq} \limsup_{T\to\infty} \frac{D_{\alpha,\beta,T}(0)}{\log\log T} \\
&\overset{(4.5.40)}{\leq} \frac{8}{\beta} \cdot \left(\frac{f_{\alpha,\beta}(u+y) - f_{\alpha,\beta}(u)}{y} - \frac{f_{\alpha,\beta}(0) - f_{\alpha,\beta}(-y)}{y}\right). \qquad (4.5.41)
\end{aligned}$$

From Lemma 4.5.5, there exists $\delta = \delta(\alpha,\beta) > 0$ such that $f_{\alpha,\beta}(\cdot)$ is differentiable on $(-\delta,\delta)$. Therefore, take $u \to 0^+$ and then $y \to 0^+$ in the above equation to conclude Step 1.

Step 2. For all $u \in (0,1)$, let

$$\begin{aligned}\eta_{\alpha,\beta,T}(u) \doteq &\left|f_{\alpha,\beta,T}(-u) - \mathbb{E}[f_{\alpha,\beta,T}(-u)]\right| + \left|f_{\alpha,\beta,T}(0) - \mathbb{E}[f_{\alpha,\beta,T}(0)]\right| \\ &+ \left|f_{\alpha,\beta,T}(u) - \mathbb{E}[f_{\alpha,\beta,T}(u)]\right|.\end{aligned} \tag{4.5.42}$$

Differentiation of the free energy gives $f'_{\alpha,\beta,T}(0) = \beta(\log\log T)^{-1} G_{\beta,T}[X_{h_k}(\alpha)]$. Then, from the convexity of $f_{\alpha,\beta,T}(\cdot)$,

$$\begin{aligned}\beta \cdot \frac{(b)}{\log\log T} &= \mathbb{E}\left|f'_{\alpha,\beta,T}(0) - \mathbb{E}[f'_{\alpha,\beta,T}(0)]\right| \\ &\leq \left|\frac{\mathbb{E}[f_{\alpha,\beta,T}(u)] - \mathbb{E}[f_{\alpha,\beta,T}(0)]}{u} - \mathbb{E}[f'_{\alpha,\beta,T}(0)]\right| \\ &+ \left|\frac{\mathbb{E}[f_{\alpha,\beta,T}(0)] - \mathbb{E}[f_{\alpha,\beta,T}(-u)]}{u} - \mathbb{E}[f'_{\alpha,\beta,T}(0)]\right| + \frac{\mathbb{E}[\eta_{\alpha,\beta,T}(u)]}{u}.\end{aligned} \tag{4.5.43}$$

Using the L^1 convergence of $f_{\alpha,\beta,T}(z)$, $z > -1$, from Proposition 4.4.2, and the mean convergence of $f'_{\alpha,\beta,T}(0)$ from Proposition 4.4.1 (the limit is $f'_{\alpha,\beta}(0)$ by Lemma 4.5.5, the convexity of $\mathbb{E}[f_{\alpha,\beta,T}(\cdot)]$ and $f_{\alpha,\beta}(\cdot)$, and by Theorem 25.7 in Rockafellar (1970)), we deduce that for all $u \in (0,1)$,

$$\limsup_{T\to\infty} \frac{(b)}{\log\log T} \leq \frac{1}{\beta} \cdot \left\{\begin{array}{l} \left|\frac{f_{\alpha,\beta}(u) - f_{\alpha,\beta}(0)}{u} - f'_{\alpha,\beta}(0)\right| \\ + \left|\frac{f_{\alpha,\beta}(0) - f_{\alpha,\beta}(-u)}{u} - f'_{\alpha,\beta}(0)\right| \end{array}\right\},$$

Take $u \to 0^+$ in the last equation, the differentiability of $f_{\alpha,\beta}(\cdot)$ at 0 (from Lemma 4.5.5) concludes Step 2. □

Theorem 4.5.7 (Approximate extended Ghirlanda-Guerra identities)**.** *Let $\beta > \beta_c \doteq 2$ and $0 < \alpha < 1$. For any $s \in \mathbb{N}^*$, any $k \in \{1, \dots, s\}$, and any bounded mesurable function $\phi : [0,1]^s \to \mathbb{R}$, we have*

$$\begin{aligned}&\left|\mathbb{E}G_{\beta,T}^{(s+1)}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_k,h_{s+1})\}} dy\, \phi(\boldsymbol{h})\Big]\right. \\ &\left. - \left\{\begin{array}{l} \frac{1}{s}\mathbb{E}G_{\beta,T}^{\times 2}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_1,h_2)\}} dy\Big] \mathbb{E}G_{\beta,T}^{\times s}[\phi(\boldsymbol{h})] \\ + \frac{1}{s}\sum_{l\neq k}^s \mathbb{E}G_{\beta,T}^{\times s}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_k,h_l)\}} dy\, \phi(\boldsymbol{h})\Big] \end{array}\right\}\right| = o_T(1),\end{aligned} \tag{4.5.44}$$

where $\boldsymbol{h} \doteq (h_1, h_2, \dots, h_s)$.

PROOF. From Proposition 4.5.4, Proposition 4.5.6 and the triangle inequality, we get

$$\left| \frac{1}{\beta} \cdot \frac{\mathbb{E}G_{\beta,T}[X_{h_k}(\alpha)]}{\frac{1}{2}\log\log T} \mathbb{E}G_{\beta,T}^{\times s}[\phi(\boldsymbol{h})] - \left\{ \begin{array}{l} \sum_{l=1}^{s} \mathbb{E}G_{\beta,T}^{\times s}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_k,h_l)\}} dy\ \phi(\boldsymbol{h})\Big] \\ -s\,\mathbb{E}G_{\beta,T}^{\times(s+1)}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_k,h_{s+1})\}} dy\ \phi(\boldsymbol{h})\Big] \end{array} \right\} \right| = o_T(1). \tag{4.5.45}$$

Furthermore, from Proposition 4.5.4 in the special case ($s = 1, k = 1, \phi \equiv 1$),

$$\left| \frac{1}{\beta} \cdot \frac{\mathbb{E}G_{\beta,T}[X_{h_k}(\alpha)]}{\frac{1}{2}\log\log T} - \left\{ \begin{array}{l} \mathbb{E}G_{\beta,T}^{\times s}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_k,h_k)\}} dy\Big] \\ -\mathbb{E}G_{\beta,T}^{\times(s+1)}\Big[\int_0^\alpha \mathbf{1}_{\{y<\rho(h_1,h_2)\}} dy\Big] \end{array} \right\} \right| = O\Big((\log\log T)^{-1}\Big). \tag{4.5.46}$$

By combining (4.5.45) and (4.5.46), we get the conclusion. □

By the representation theorem of Dovbysh and Sudakov Dovbysh and Sudakov (1982) (for an accessible proof, see Panchenko (2010)), we can show (see e.g. the reasoning on page 1459 of Arguin and Zindy (2014) or page 101 of Panchenko (2013b)) that there exists a subsequence $\{T_m\}_{m\in\mathbb{N}^*}$ converging to $+\infty$ such that for any $s \in \mathbb{N}^*$ and any continuous function $\phi : [0,1]^{s(s-1)/2} \to \mathbb{R}$, we have

$$\lim_{m\to\infty} \mathbb{E}G_{\beta,T_m}^{\times\infty}\Big[\phi((\rho(h_l,h_{l'}))_{1\le l,l'\le s})\Big] = E\mu_\beta^{\times\infty}\Big[\phi((R_{l,l'})_{1\le l,l'\le s})\Big], \tag{4.5.47}$$

where R is a random element of some probability space with measure P (and expectation E), generated by the random matrix of scalar products

$$(R_{l,l'})_{l,l'\in\mathbb{N}^*} = \Big((\rho_l,\rho_{l'})_{\mathcal{H}}\Big)_{l,l'\in\mathbb{N}^*}, \tag{4.5.48}$$

where $(\rho_l)_{l\in\mathbb{N}^*}$ is an i.i.d. sample from some random measure μ_β concentrated a.s. on the unit sphere of a separable Hilbert space $\mathcal{H}$. In particular, from Theorem 4.3.1, we have

$$E\mu_\beta^{\times 2}\Big[\mathbf{1}_{\{R_{1,2}\in A\}}\Big] = \frac{2}{\beta}\mathbf{1}_A(0) + \left(1 - \frac{2}{\beta}\right)\mathbf{1}_A(1), \quad A \in \mathcal{B}([0,1]). \tag{4.5.49}$$

Next, we show the consequence of taking the limit (4.5.47) in the statement of Theorem

4.5.7. Note that a function $\phi : \{0,1\}^{s(s-1)/2} \to \mathbb{R}$ can always be embedded in a continuous function defined on $[0,1]^{s(s-1)/2}$. Here is the main result of this section.

Theorem 4.5.8 (Extended Ghirlanda-Guerra identities in the limit)**.** *Let $\beta > \beta_c \stackrel{\circ}{=} 2$. Also, let μ_β be a subsequential limit of $\{G_{\beta,T}\}_{T\geq 2}$ in the sense of (4.5.47). For any $s \in \mathbb{N}^*$, any $k \in \{1,\dots,s\}$, and any functions $\psi : \{0,1\} \to \mathbb{R}$ and $\phi : \{0,1\}^{s(s-1)/2} \to \mathbb{R}$, we have*

$$\begin{aligned} E\mu_\beta^{(s+1)}\Big[\psi(R_{k,s+1})\phi((R_{i,i'})_{1\leq i,i'\leq s})\Big] &= \frac{1}{s} E\mu_\beta^{\times 2}\Big[\psi(R_{1,2})\Big] E\mu_\beta^{\times s}\Big[\phi((R_{i,i'})_{1\leq i,i'\leq s})\Big] \\ &\quad + \frac{1}{s}\sum_{l\neq k}^{s} E\mu_\beta^{\times s}\Big[\psi(R_{k,l})\phi((R_{i,i'})_{1\leq i,i'\leq s})\Big]. \end{aligned} \tag{4.5.50}$$

Remark 4.5.1. *The functions ψ and ϕ have $\{0,1\}$ and $\{0,1\}^{s(s-1)/2}$ as their domain, respectively, because $R_{l,l'} \in \{0,1\}$ $E\mu_\beta^{\times 2}$-almost-surely by (4.5.49) and the matrix $(R_{l,l'})_{1\leq l,l'\leq s}$ is symmetric and its diagonal elements are equal to 1 $E\mu_\beta^{\times s}$-almost-surely by (4.5.48).*

Proof of Theorem 4.5.8. From (4.5.47) and Theorem 4.5.7 (in the particular case where ϕ is a function of the overlaps), we deduce

$$\begin{aligned} &E\mu_\beta^{(s+1)}\Big[\int_0^\alpha \mathbf{1}_{\{y<R_{k,s+1}\}} dy\, \phi((R_{i,i'})_{1\leq i,i'\leq s})\Big] \\ &\quad = \frac{1}{s} E\mu_\beta^{\times 2}\Big[\int_0^\alpha \mathbf{1}_{\{y<R_{1,2}\}} dy\Big] E\mu_\beta^{\times s}\Big[\phi((R_{i,i'})_{1\leq i,i'\leq s})\Big] \\ &\quad + \frac{1}{s}\sum_{l\neq k}^{s} E\mu_\beta^{\times s}\Big[\int_0^\alpha \mathbf{1}_{\{y<R_{k,l}\}} dy\, \phi((R_{i,i'})_{1\leq i,i'\leq s})\Big]. \end{aligned} \tag{4.5.51}$$

From (4.5.49), we know that $\mathbf{1}_{\{y<R_{i,i'}\}}$ is $E\mu_\beta^{\times 2}$-a.s. constant in y on $[-1,0)$ and $[0,1)$ respectively. Therefore, for any $x \in \{-1,0\}$,

$$\begin{aligned} &E\mu_\beta^{(s+1)}\Big[\mathbf{1}_{\{x<R_{k,s+1}\}}\phi((R_{i,i'})_{1\leq i,i'\leq s})\Big] \\ &\quad = \frac{1}{s} E\mu_\beta^{\times 2}\Big[\mathbf{1}_{\{x<R_{1,2}\}}\Big] E\mu_\beta^{\times s}\Big[\phi((R_{i,i'})_{1\leq i,i'\leq s})\Big] \\ &\quad + \frac{1}{s}\sum_{l\neq k}^{s} E\mu_\beta^{\times s}\Big[\mathbf{1}_{\{x<R_{k,l}\}}\phi((R_{i,i'})_{1\leq i,i'\leq s})\Big]. \end{aligned} \tag{4.5.52}$$

But, any function $\psi : \{0,1\} \to \mathbb{R}$ can be written as a linear combination of the indicator functions $\mathbf{1}_{\{0<\cdot\}}$ and $\mathbf{1}_{\{-1<\cdot\}}$, so we get the conclusion by the linearity of (4.5.52). $\square$

4.6. Proof of Theorem 4.3.2

Once we have Theorem 4.3.1 and the Ghirlanda-Guerra identities from Theorem 4.5.8, the proof follows exactly the same steps as in the proof of Theorem 1.5 in Arguin and Zindy (2014). We can show that any subsequential limit μ_β of $\{G_{\beta,T}\}_{T\geq 2}$ in the sense of (4.5.47) must satisfy

$$\mu_\beta = \sum_{k\in\mathbb{N}^*} \xi_k \delta_{e_k}, \quad P - a.s., \tag{4.6.1}$$

where δ is the Dirac measure, $(e_k)_{k\in\mathbb{N}^*}$ is a sequence of orthonormal vectors in $\mathcal{H}$ and ξ is a Poisson-Dirichlet variable of parameter β_c/β. Since the space of probability measures on $[0,1]^{\mathbb{N}^*\times\mathbb{N}^*}$ (the space of overlap matrices) is a metric space under the weak topology, the limit in (4.5.47) must hold for the original sequence. Then, (4.3.6) is a direct consequence of (4.6.1).

Acknowledgements

I would like to thank the anonymous referees and my advisor, Louis-Pierre Arguin, for their valuable comments that led to improvements in the presentation of this paper. This work is fully supported by a NSERC Doctoral Program Alexander Graham Bell scholarship (CGS D3).

Bibliography

L.-P. Arguin and F. Ouimet. Large deviations and continuity estimates for the derivative of a random model of $\log|\zeta|$ on the critical line. *J. Math. Anal. Appl.*, 472(1):687–695, 2019. MR3906393.

L.-P. Arguin and W. Tai. Is the Riemann zeta function in a short interval a 1-RSB spin glass ? *Preprint. To appear in Sojourns in probability and statistical physics, in honor of the 70th birthday of C. Newman*, pages 1–20, 2018. arXiv:1706.08462.

L.-P. Arguin and O. Zindy. Poisson-Dirichlet statistics for the extremes of a log-correlated Gaussian field. *Ann. Appl. Probab.*, 24(4):1446–1481, 2014. MR3211001.

L.-P. Arguin and O. Zindy. Poisson-Dirichlet statistics for the extremes of the two-dimensional discrete Gaussian free field. *Electron. J. Probab.*, 20:no. 59, 19, 2015. MR3354619.

L.-P. Arguin, D. Belius, and A. J. Harper. Maxima of a randomized Riemann zeta function, and branching random walks. *Ann. Appl. Probab.*, 27(1):178–215, 2017. MR3619786.

L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł, and K. Soundararajan. Maximum of the Riemann zeta function on a short interval of the critical line. *Comm. Pure Appl. Math.*, 72(3):500–535, 2019. MR3911893.

F. Baffioni and F. Rosati. Some exact results on the ultrametric overlap distribution in mean field spin glass models. *Eur. Phys. J. B*, 17:439–447, 2000. doi:10.1007/s101050000b0002.

A. Bovier and I. Kurkova. Derrida's generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):439–480, 2004a. MR2070334.

A. Bovier and I. Kurkova. Derrida's generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):481–495, 2004b. MR2070335.

B. Derrida. Random-energy model: limit of a family of disordered models. *Phys. Rev. Lett.*, 45(2):79–82, 1980. MR575260.

B. Derrida. A generalization of the Random Energy Model which includes correlations between energies. *J. Physique Lett.*, 46(9):401–407, 1985. doi:10.1051/jphyslet:01985004609040100.

L. N. Dovbysh and V. N. Sudakov. Gram-de Finetti matrices. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 119:77–86, 1982. MR666087.

Y. V. Fyodorov and J. P. Keating. Freezing transitions and extreme values: random matrix theory, $\zeta(\frac{1}{2}+it)$ and disordered landscapes. *Philos. Trans. R. Soc. A*, 372(20120503): 1–32, 2014. MR3151088.

Y.V. Fyodorov, G.A. Hiary, and J.P. Keating. Freezing transition, characteristic polynomials of random matrices, and the Riemann zeta-function. *Phys. Rev. Lett.*, 108(170601): 1–4, 2012. doi:10.1103/PhysRevLett.108.170601.

S. Ghirlanda and F. Guerra. General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A*, 31(46):9149–9155, 1998. MR1662161.

A. J. Harper. A note on the maximum of the Riemann zeta function, and log-correlated random variables. *Preprint*, pages 1–26, 2013. arXiv:1304.0677.

A. Jagannath. On the overlap distribution of branching random walks. *Electron. J. Probab.*, 21:Paper No. 50 16, 2016. MR3539644.

A. Jagannath. Approximate ultrametricity for random measures and applications to spin glasses. *Comm. Pure Appl. Math.*, 70(4):611–664, 2017. MR3628881.

J. Najnudel. On the extreme values of the Riemann zeta function on random intervals of the critical line. *Probab. Theory Related Fields*, 172(1-2):387–452, 2018. MR3851835.

F. Ouimet. Geometry of the Gibbs measure for the discrete 2D Gaussian free field with scale-dependent variance. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(2):851–902, 2017. MR3731796.

D. Panchenko. On the Dovbysh-Sudakov representation result. *Electron. Commun. Probab.*, 15:330–338, 2010. MR2679002.

D. Panchenko. Ghirlanda-Guerra identities and ultrametricity: an elementary proof in the discrete case. *C. R. Math. Acad. Sci. Paris*, 349(13-14):813–816, 2011. MR2825947.

D. Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013a. MR2999044.

D. Panchenko. *The Sherrington-Kirkpatrick Model*. Springer Monographs in Mathematics. Springer, New York, 2013b. MR3052333.

R. T. Rockafellar. *Convex Analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970. MR0274683.

D. Ruelle. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.*, 108(2):225–239, 1987. MR875300.

M. Talagrand. *Spin Glasses: A Challenge for Mathematicians*, volume 46 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*. Springer-Verlag, Berlin, 2003. MR1993891.

M. Talagrand. *Mean Field Models for Spin Glasses. Volume II: Advanced Replica-Symmetry and Low Temperature*, volume 55. Springer, Heidelberg, 2011. MR3024566.

Article 5

Large deviations and continuity estimates for the derivative of a random model of $\log|\zeta|$ on the critical line

by

Louis-Pierre Arguin[1] and Frédéric Ouimet[2]

([1]) Baruch College and Graduate Center (CUNY), New York, NY 10010, USA.

([2]) Université de Montréal, Montréal, QC H3T 1J4, Canada.

Here is the complete citation :

Arguin, L.-P. and Ouimet, F. (2019). Large deviations and continuity estimates for the derivative of a random model of log $|\zeta|$ on the critical line. J. Math. Anal. Appl. 472, no. 1, 687-695. https://doi.org/10.1016/j.jmaa.2018.11.044.

Contributions

I wrote the paper entirely. The proofs of Propositions 5.3.1, Proposition 5.3.2, Lemma 5.4.1 and Lemma 5.4.2 are adapted from arguments presented in Arguin et al. (2017), which Louis-Pierre co-authored. The idea for the proof of Theorem 5.3.3 came to me after Louis-Pierre mentioned Lemma 2.2 in Farmer et al. (2007).

ABSTRACT. In this paper, we study the random field

$$X(h) \circeq \sum_{p \leq T} \frac{\mathrm{Re}(U_p\, p^{-ih})}{p^{1/2}}, \quad h \in [0,1],$$

where $(U_p,\, p$ primes) is an i.i.d. sequence of uniform random variables on the unit circle in $\mathbb{C}$. Harper (2013) showed that $(X(h),\, h \in (0,1))$ is a good model for the large values of $(\log|\zeta(\frac{1}{2}+i(T+h))|,\, h \in [0,1])$ when T is large, if we assume the Riemann hypothesis. The asymptotics of the maximum were found in Arguin et al. (2017) up to the second order, but the tightness of the recentered maximum is still an open problem. As a first step, we provide large deviation estimates and continuity estimates for the field's derivative $X'(h)$. The main result shows that, with probability arbitrarily close to 1,

$$\max_{h \in [0,1]} X(h) - \max_{h \in \mathcal{S}} X(h) = O(1),$$

where $\mathcal{S}$ a discrete set containing $O(\log T \sqrt{\log\log T})$ points.

Keywords: extreme value theory, large deviations, Riemann zeta function, estimates

5.1. Introduction

In Fyodorov et al. (2012) and Fyodorov and Keating (2014), it was conjectured that if τ is sampled uniformly in $[T, 2T]$ for some large T, then the law of the maximum of $(\log|\zeta(\frac{1}{2} + i(\tau + h))|, h \in [0,1])$, where ζ denotes the Riemann zeta function, should be asymptotic to $\log\log T - \frac{3}{4}\log\log\log T + \mathcal{M}_T$ where $(\mathcal{M}_T, T \geq 2)$ is a sequence of random variables converging in distribution. At present, the first order of the maximum is proved conditionally on the Riemann hypothesis in Najnudel (2018) and unconditionally in Arguin et al. (2019).

In order to study this hard problem originally, a randomized version of the Riemann zeta function was introduced in Harper (2013), see (5.2.1). The first order of the maximum was proved in Harper (2013), the second order of the maximum was proved in Arguin et al. (2017), and a related study of the Gibbs measure can be found in Arguin and Tai (2018) and Ouimet (2018). The tightness of the recentered maximum is still open.

As a first step, our main result (Theorem 5.3.3) shows that the tightness of the "continuous" maximum $\max_{h \in [0,1]} X(h)$ (once recentered) can be reduced to the tightness of a "discrete" maximum $\max_{h \in \mathcal{S}} X(h)$ (once recentered) where $\mathcal{S}$ is a discrete set containing

$O(\log T\sqrt{\log\log T})$ points. In order to prove Theorem 5.3.3, we will need continuity estimates and large deviation estimates for the field's derivative $X'(h)$, which can be found in Proposition 5.3.1 and Proposition 5.3.2, respectively.

The paper is organised as follows. In Section 5.2, we introduce the model $X(h)$. In Section 5.3, the main result is stated and proven. Proposition 5.3.1 and Proposition 5.3.2 are stated in Section 5.3 and proven in Section 5.4.

5.2. The model

Let (U_p, p primes) be an i.i.d. sequence of uniform random variables on the unit circle in $\mathbb{C}$. The random field of interest is

$$X(h) \mathrel{\hat{=}} \sum_{p\leq T} W_p(h) \mathrel{\hat{=}} \sum_{p\leq T} \frac{\mathrm{Re}(U_p\, p^{-ih})}{p^{1/2}}, \quad h\in[0,1]. \tag{5.2.1}$$

(A sum over the variable p always denotes a sum over primes.) This is a good model for the large values of $(\log|\zeta(\frac{1}{2}+i(\tau+h))|, h\in[0,1])$ for the following reason. Proposition 1 in Harper (2013) proves that, assuming the Riemann hypothesis, and for T large enough, there exists a set $B\subseteq[T,T+1]$, of Lebesgue measure at least 0.99, such that

$$\log|\zeta(\frac{1}{2}+it)| = \mathrm{Re}\left(\sum_{p\leq T}\frac{1}{p^{1/2+it}}\frac{\log(T/p)}{\log T}\right) + O(1), \quad t\in B. \tag{5.2.2}$$

If we ignore the smoothing term $\log(T/p)/\log T$ and note that the process $(p^{-i\tau}, p \text{ primes})$, where τ is sampled uniformly in $[T,2T]$, converges, as $T\to\infty$ (in the sense of convergence of its finite-dimensional distributions), to a sequence of independent random variables distributed uniformly on the unit circle (by computing the moments), then the model (5.2.1) follows. For more information, see Section 1.1 in Arguin et al. (2017).

More generally, for $-1\leq r\leq k$, denote the increments of the field by

$$X_{r,k}(h) \mathrel{\hat{=}} \sum_{2^r<\log p\leq 2^k} \frac{\mathrm{Re}(U_p\, p^{-ih})}{p^{1/2}}, \quad h\in[0,1]. \tag{5.2.3}$$

Differentiation of (5.2.3) yields

$$X'_{r,k}(h) = \sum_{2^r<\log p\leq 2^k} W'_p(h) = \sum_{2^r<\log p\leq 2^k} \frac{\mathrm{Im}(U_p\, p^{-ih})\log p}{p^{1/2}}. \tag{5.2.4}$$

5.3. Main result

Throughout the paper, we will write c, $\widetilde{c}$, c', and c'', for generic positive constants whose value may change at different occurrences. Here are the main side results of this paper.

Proposition 5.3.1 (Continuity estimates). *Let $C > 0$. For any $-1 \leq r \leq k$, $0 \leq x \leq C(2^{2k} - 2^{2r})$, $2 \leq a \leq 2^{6k} - x$ and $h \in \mathbb{R}$,*

$$\mathbb{P}\left(\max_{h':|h'-h|\leq 2^{-3k-1}} X'_{r,k}(h') \geq x + a, X'_{r,k}(h) \leq x\right) \leq c \exp\left(-2\frac{x^2}{2^{2k} - 2^{2r}} - \widetilde{c}\, a^{3/2}\right), \quad (5.3.1)$$

where the constants c and $\widetilde{c}$ only depend on C.

Proposition 5.3.2 (Large deviation estimates). *Let $C > 0$. For any $-1 \leq r \leq k$, $0 \leq x \leq C(2^{2k} - 2^{2r})$ and $h \in \mathbb{R}$,*

$$\mathbb{P}\left(\max_{h':|h'-h|\leq 2^{-3k-1}} X'_{r,k}(h') \geq x\right) \leq c \exp\left(-2\frac{x^2}{2^{2k} - 2^{2r}}\right), \quad (5.3.2)$$

where the constant c only depends on C.

From the last proposition, we obtain the following theorem.

Theorem 5.3.3 (Main result). *Let $-1 \leq r \leq k$. For all $L > 0$, let $\mathcal{S}_{r,k,L}$ be a set of equidistant points in $[0,1]$ such that $|\mathcal{S}_{r,k,L}| = \lceil L\sqrt{2^{2k} - 2^{2r}}\sqrt{k \log 2}\rceil$ and $|h' - h| \geq |\mathcal{S}_{r,k,L}|^{-1}$ for different $h, h' \in \mathcal{S}_{r,k,L}$. Then, for any $K > 0$, there exists $L \mathrel{\hat{=}} L(K) > 0$ large enough that*

$$\mathbb{P}\left(\left|\max_{h\in[0,1]} X_{r,k}(h) - \max_{h\in\mathcal{S}_{r,k,L}} X_{r,k}(h)\right| > K\right) < e^{-\frac{k}{4}(1-e^{-K})^2 L^2}. \quad (5.3.3)$$

Remark 5.3.1. *When $r = -1$ and $2^k = \log T$, $X_{r,k}(h)$ is just the original model $X(h)$. In that case, (5.3.3) shows that, with probability as close to 1 as we want, there exists a discrete set $\mathcal{S} \subseteq [0,1]$ such that*

$$\max_{h\in[0,1]} X(h) - \max_{h\in\mathcal{S}} X(h) = O(1), \quad (5.3.4)$$

where $|\mathcal{S}| = O(\log T\sqrt{\log\log T})$.

We prove Theorem 5.3.3 right away and we will prove Proposition 5.3.1 and Proposition 5.3.2 in Section 5.4.

PROOF OF THEOREM 5.3.3. For $M > 0$, define the event

$$E = \left\{ \max_{h \in [0,1]} |X'_{r,k}(h)| \geq M \sqrt{2^{2k} - 2^{2r}} \sqrt{k \log 2} \right\}. \tag{5.3.5}$$

Let $\mathcal{H}_k \triangleq 2^{-3k}\mathbb{Z}$ and note that $|\mathcal{H}_k \cap [0,1]| = 2^{3k} + 1$. By a union bound, the symmetry of $X'_{r,k}(h)$'s distribution, and Proposition 5.3.2, we obtain

$$\begin{aligned} \mathbb{P}(E) &\leq \sum_{h \in \mathcal{H}_k \cap [0,1]} 2 \cdot \mathbb{P}\left(\max_{h' : |h' - h| \leq 2^{-3k-1}} X'_{r,k}(h') \geq M \sqrt{2^{2k} - 2^{2r}} \sqrt{k \log 2} \right) \\ &\leq (2^{3k} + 1) \cdot c \, 2^{-2kM^2}. \end{aligned} \tag{5.3.6}$$

For every realisation ω of the field $\{X_{r,k}(h)\}_{h \in [0,1]}$, let $h^\star(\omega)$ be a point where the maximum is attained. When $\omega \in E^c$, the mean value theorem yields that, for any $h(\omega) \in \mathcal{S}_{r,k,L}$ such that $|h^\star(\omega) - h(\omega)| \leq 2/|\mathcal{S}_{r,k,L}|$, we have

$$e^{X_{r,k}(h^\star(\omega))} - e^{X_{r,k}(h(\omega))} = X'_{r,k}(\xi(\omega)) e^{X_{r,k}(\xi(\omega))} (h^\star(\omega) - h(\omega)) \leq \frac{2M}{L} e^{X_{r,k}(h^\star(\omega))}, \tag{5.3.7}$$

for some $\xi(\omega)$ lying between $h(\omega)$ and $h^\star(\omega)$. By taking $L \triangleq L(K) \triangleq 2M/(1 - e^{-K})$, we deduce $e^{X_{r,k}(h(\omega))} \geq e^{-K} e^{X_{r,k}(h^\star(\omega))}$. This reasoning shows that, on the event E^c,

$$\max_{h \in \mathcal{S}_{r,k,L}} X_{r,k}(h) \geq \max_{h \in [0,1]} X_{r,k}(h) - K. \tag{5.3.8}$$

The conclusion follows from (5.3.8) and (5.3.6) with $M = \frac{1}{2}(1 - e^{-K})L$. $\square$

5.4. Proof of Proposition 5.3.1 and Proposition 5.3.2

We start by controlling the tail probabilities for a single point of the field's derivative.

Lemma 5.4.1. *Let $C > 0$. For any $-1 \leq r \leq k$, $0 \leq x \leq C(2^{2k} - 2^{2r})$ and $h \in \mathbb{R}$,*

$$\mathbb{P}\Big(X'_{r,k}(h) \geq x \Big) \leq c \exp\left(-2 \frac{x^2}{2^{2k} - 2^{2r}} \right), \tag{5.4.1}$$

where the constant c only depends on C.

PROOF. Using Chernoff's inequality, the independence of the U_p's and translation invariance, we have that, for all $\lambda \geq 0$,

$$\mathbb{P}\Big(X'_{r,k}(h) \geq x \Big) \leq e^{-\lambda x} \, \mathbb{E}\Big[e^{\lambda X'_{r,k}(h)} \Big] = e^{-\lambda x} \prod_{2^r < \log p \leq 2^k} \mathbb{E}\Big[e^{\lambda W'_p(0)} \Big]. \tag{5.4.2}$$

Note that

$$\mathbb{E}\Big[e^{\lambda W_p'(0)}\Big] = \frac{1}{2\pi}\int_0^{2\pi} \exp\left(\frac{\lambda \log p}{p^{1/2}}\sin(\theta)\right) d\theta = I_0\left(\frac{\lambda \log p}{p^{1/2}}\right), \tag{5.4.3}$$

(Abramowitz and Stegun, 1964, 9.6.16, p.376), where I_0 denotes the *modified Bessel function of the first kind*. The function I_0 has the following series representation : $I_0(u) = 1 + \frac{u^2}{4} + \frac{u^4}{64} + O(u^6),\ u \in \mathbb{R}$. In turn,

$$\log(I_0(u)) = \frac{u^2}{4} - \frac{u^4}{64} + O(u^6), \quad u \in (-1,1), \tag{5.4.4}$$

because $\log(1+y) = y - \frac{y^2}{2} + O(y^3)$ for $y \in (-1,1)$, and $|I_0(u) - 1| < 1$ for $u \in (-1,1)$. Choose $\lambda = 4x/(2^{2k} - 2^{2r})$. By applying (5.4.4) in (5.4.3), the right-hand side of (5.4.2) is bounded from above by

$$c\, e^{-\lambda x} \exp\left(\sum_{2^r < \log p \leq 2^k} \frac{\lambda^2 (\log p)^2}{4p} + \widetilde{c} \sum_{2^r < \log p \leq 2^k} \frac{\lambda^6 (\log p)^6}{p^3}\right). \tag{5.4.5}$$

For the finite number of primes p for which we cannot apply (5.4.4) in (5.4.3) (note that $\lambda \log p < p^{1/2}$ holds for p large enough since $\lambda \leq 4C$ by the assumption on x), the correction terms needed for (5.4.5) to hold are absorbed in the constant c in front of the first exponential in (5.4.5). The second sum in the big exponential is bounded by a constant independent from r and k since $\lambda \leq 4C$ and $\sum_p (\log p)^6 p^{-3} < \infty$. By applying Lemma 5.5.1 with $m = 2$, $\log P = 2^r$ and $\log Q = 2^k$, the first sum in the big exponential is bounded by $2x^2/(2^{2k} - 2^{2r})$ up to an additive constant that only depends on C. The conclusion of the lemma follows. □

In the next lemma, we complement Lemma 5.4.1 by proving a large deviation estimate for $X_{r,k}'(0)$ and the difference $X_{r,k}'(h_2) - X_{r,k}'(h_1)$ jointly, where $|h_2 - h_1| \leq 2^{-3k}$.

Lemma 5.4.2. *Let $C > 0$. For any $-1 \leq r \leq k$, $0 \leq x \leq C(2^{2k} - 2^{2r})$, $0 \leq y \leq 2^{6k}$, and any distinct $h_1, h_2 \in \mathbb{R}$ such that $-2^{-3k-1} \leq h_1, h_2 \leq 2^{-3k-1}$,*

$$\begin{aligned}\mathbb{P}\Big(X_{r,k}'(0) \geq x, X_{r,k}'(h_2) - X_{r,k}'(h_1) \geq y\Big) \\ \leq c \exp\left(-2\frac{x^2}{2^{2k} - 2^{2r}} - \frac{\widetilde{c}\, y^{3/2}}{|h_2 - h_1|\, 2^{3k}}\right),\end{aligned} \tag{5.4.6}$$

where the constants c and $\widetilde{c}$ only depend on C.

Proof. Assume that $y \geq \widetilde{C}|h_2 - h_1| 2^{3k}$ for a large constant $\widetilde{C} \geq 1$ because otherwise (5.4.6) follows from (5.4.1). Since $|h_2 - h_1| 2^{3k} \leq 1$, note that this assumption also implies $y^{1/2} \geq \widetilde{C}^{1/2}|h_2 - h_1| 2^{3k}$. For all $\lambda_1, \lambda_2 \geq 0$, the left-hand side of (5.4.6) is bounded from above (using Chernoff's inequality) by

$$\mathbb{E}\Big[\exp(\lambda_1 X'_{r,k}(0) + \lambda_2(X'_{r,k}(h_2) - X'_{r,k}(h_1)))\Big] \exp(-\lambda_1 x - \lambda_2 y). \tag{5.4.7}$$

We will show that if $0 \leq \lambda_1 \leq 4C$ and $0 \leq \lambda_2 \leq |h_2 - h_1|^{-1}$, then

$$\begin{aligned} &\mathbb{E}\Big[\exp(\lambda_1 X'_{r,k}(0) + \lambda_2(X'_{r,k}(h_2) - X'_{r,k}(h_1)))\Big] \\ &\quad \leq c \exp\left(\frac{\lambda_1^2}{8}(2^{2k} - 2^{2r}) + c\,\lambda_2 |h_2 - h_1|\, 2^{3k} + c^2 \lambda_2^2 |h_2 - h_1|^2\, 2^{4k}\right). \end{aligned} \tag{5.4.8}$$

The result (5.4.6) follows by choosing $\lambda_1 = 4x/(2^{2k} - 2^{2r})$, $\lambda_2 = y^{1/2}|h_2 - h_1|^{-1}\, 2^{-3k}$ and $\widetilde{C}$ large enough (with respect to c) in (5.4.7) and (5.4.8). The assumptions on x, y, h_1 and h_2 ensure that $0 \leq \lambda_1 \leq 4C$ and $0 \leq \lambda_2 \leq |h_2 - h_1|^{-1}$. We now prove (5.4.8). For $2^r < \log p \leq 2^k$, the quantity

$$\mathbb{E}\Big[\exp(\lambda_1 W'_p(0) + \lambda_2(W'_p(h_2) - W'_p(h_1)))\Big] \tag{5.4.9}$$

(recall $W'_p(h)$ from (5.2.4)) can be written as

$$\frac{1}{2\pi}\int_0^{2\pi} \exp\left(\frac{\log p}{p^{1/2}}\Big\{\lambda_1 \sin\theta + \lambda_2(\sin(\theta - h_2 \log p) - \sin(\theta - h_1 \log p))\Big\}\right) d\theta. \tag{5.4.10}$$

Since $\sin(\theta - \eta) = \sin(\theta)\cos(\eta) - \cos(\theta)\sin(\eta)$ and

$$\frac{1}{2\pi}\int_0^{2\pi} \exp(a\cos\theta + b\sin\theta) d\theta = I_0(\sqrt{a^2 + b^2}), \tag{5.4.11}$$

(Abramowitz and Stegun, 1964, 9.6.16, p.376), then (5.4.9) is equal to

$$I_0\left(\sqrt{\frac{(\log p)^2}{p}\left\{\begin{array}{l} \Big(\lambda_1 + \lambda_2(\cos(h_2 \log p) - \cos(h_1 \log p))\Big)^2 \\ + \Big(\lambda_2(\sin(h_1 \log p) - \sin(h_2 \log p))\Big)^2 \end{array}\right\}}\right). \tag{5.4.12}$$

From (5.4.4), note that

$$\log(I_0(\sqrt{u})) = \frac{u}{4} - \frac{u^2}{64} + O(u^3), \quad u \in (-1, 1). \tag{5.4.13}$$

Also, note that

$$\begin{aligned}\sin(h_1 \log p) - \sin(h_2 \log p) &= O(|h_2 - h_1| \log p), \\ \cos(h_2 \log p) - \cos(h_1 \log p) &= O(|h_2 - h_1| \log p).\end{aligned} \tag{5.4.14}$$

If we put (5.4.9), (5.4.12), (5.4.13) and (5.4.14) together, we get, for p large enough,

$$\begin{aligned}\log\,(5.4.9) &\leq \frac{(\log p)^2}{4p} \left\{ \Big(\lambda_1 + c\,\lambda_2 |h_2 - h_1| \log p\Big)^2 + \Big(c\,\lambda_2 |h_2 - h_1| \log p)\Big)^2 \right\} + \frac{\widetilde{c}}{p^2} \\ &\leq \frac{\lambda_1^2}{4} \frac{(\log p)^2}{p} + c\,\lambda_2 |h_2 - h_1| \frac{(\log p)^3}{p} + c^2 \lambda_2^2 |h_2 - h_1|^2 \frac{(\log p)^4}{p} + \frac{\widetilde{c}}{p^2}.\end{aligned} \tag{5.4.15}$$

To obtain the last inequality, we used the fact that $\lambda_1 \leq 4C$. After summing (5.4.15) over $2^r < \log p \leq 2^k$ and using Lemma 5.5.1, we deduce

$$\begin{aligned}&\log \mathbb{E}\Big[\exp(\lambda_1 X'_{r,k}(0) + \lambda_2 (X'_{r,k}(h_2) - X'_{r,k}(h_1)))\Big] \\ &\quad\leq \widetilde{c} + \frac{\lambda_1^2}{8}(2^{2k} - 2^{2r}) + c\,\lambda_2 |h_2 - h_1|\, 2^{3k} + c^2 \lambda_2^2 |h_2 - h_1|^2\, 2^{4k},\end{aligned} \tag{5.4.16}$$

where the constants c and $\widetilde{c}$ only depend on C. This is exactly (5.4.8). $\square$

We are now ready to prove Proposition 5.3.1. For $k \in \mathbb{N}_0$, recall that $\mathcal{H}_k \mathrel{\hat{=}} 2^{-3k}\mathbb{Z}$, so that $\mathcal{H}_0 \subseteq \mathcal{H}_1 \subseteq \ldots \subseteq \mathcal{H}_k \subseteq \ldots \subseteq \mathbb{R}$ is a nested sequence of sets of equidistant points and $|\mathcal{H}_k \cap [0,1)| = 2^{3k}$.

Proof of Proposition 5.3.1. Without loss of generality, we may assume $h = 0$. We can also round x up to the nearest larger integer and decrease a so that we may assume that $x \in \mathbb{N}_0$ and $a \geq 1$. To see why this is possible, define the new values of x and a by $\widetilde{x} \mathrel{\hat{=}} \lceil x \rceil$ and $\widetilde{a} \mathrel{\hat{=}} a - \widetilde{x} + x$, respectively. Since $x + a = \widetilde{x} + \widetilde{a}$ and $x \leq \widetilde{x}$, and assuming that we can show (5.3.1) with $\widetilde{x}$ and $\widetilde{a}$, we would have

$$\begin{aligned}&\mathbb{P}\left(\max_{h' : |h' - h| \leq 2^{-3k-1}} X'_{r,k}(h') \geq x + a, X'_{r,k}(h) \leq x\right) \\ &\leq \mathbb{P}\left(\max_{h' : |h' - h| \leq 2^{-3k-1}} X'_{r,k}(h') \geq \widetilde{x} + \widetilde{a}, X'_{r,k}(h) \leq \widetilde{x}\right) \\ &\leq c\ \exp\left(-2\frac{\widetilde{x}^2}{2^{2k} - 2^{2r}} - \widetilde{c}\,\widetilde{a}^{3/2}\right) \leq c'\exp\left(-2\frac{x^2}{2^{2k} - 2^{2r}} - c'' a^{3/2}\right),\end{aligned} \tag{5.4.17}$$

where the constants c' and c'' only depend on C.

It remains to show (5.3.1) when $x \in \mathbb{N}_0$ and $a \geq 1$. We choose to adapt the chaining argument found in (Arguin et al., 2017, Proposition 2.5). Define the events

$$B_x \mathrel{\hat=} \{X'_{r,k}(0) \leq 0\}$$

and

$$B_q \mathrel{\hat=} \{X'_{r,k}(0) \in [x-q-1, x-q]\}, \quad q \in \{0,1,\dots,x-1\}. \tag{5.4.18}$$

Note that the left-hand side of (5.3.1) is at most

$$\sum_{q=0}^{x} \mathbb{P}\left(B_q \cap \left\{\max_{h'\in A}\{X'_{r,k}(h') - X'_{r,k}(0)\} \geq a+q\right\}\right), \tag{5.4.19}$$

where $A = [-2^{-3k-1}, 2^{-3k-1}]$. Let $(h_i, i \in \mathbb{N}_0)$ be a sequence such that $h_0 = 0$, $h_i \in \mathcal{H}_{k+i} \cap A$, $\lim_{i\to\infty} h_i = h'$ and $|h_{i+1} - h_i| \in \{0, \frac{1}{8}2^{-3(k+i)}, \frac{2}{8}2^{-3(k+i)}, \frac{3}{8}2^{-3(k+i)}, \frac{4}{8}2^{-3(k+i)}\}$ for all i. Because the map $h \mapsto X'_{r,k}(h)$ is almost-surely continuous,

$$X'_{r,k}(h') - X'_{r,k}(0) = \sum_{i=0}^{\infty}(X'_{r,k}(h_{i+1}) - X'_{r,k}(h_i)). \tag{5.4.20}$$

Since $\sum_{i=0}^{\infty} \frac{1}{2(i+1)^2} \leq 1$, we have the inclusion of events,

$$\left\{X'_{r,k}(h') - X'_{r,k}(0) \geq a+q\right\} \subseteq \bigcup_{i=0}^{\infty}\left\{X'_{r,k}(h_{i+1}) - X'_{r,k}(h_i) \geq \frac{a+q}{2(i+1)^2}\right\}. \tag{5.4.21}$$

This implies that $\{\max_{h'\in A} X'_{r,k}(h') - X'_{r,k}(0) \geq a+q\}$ is included in

$$\bigcup_{i=0}^{\infty} \bigcup_{\substack{h_1\in\mathcal{H}_{k+i}\cap A \\ |h_2-h_1| = \frac{j}{8}2^{-3(k+i)} \\ \text{for some } j\in\{1,2,3,4\}}} \left\{X'_{r,k}(h_2) - X'_{r,k}(h_1) \geq \frac{a+q}{2(i+1)^2}\right\}, \tag{5.4.22}$$

where we have ignored the case $h_1 = h_2$ since the event

$$\left\{X'_{r,k}(h_2) - X'_{r,k}(h_1) \geq \frac{a+q}{2(i+1)^2}\right\}$$

is the empty set. Because $|\mathcal{H}_{k+i} \cap A| \leq c\, 2^{3i}$, the q-th summand in (5.4.19) is at most

$$\sum_{i=0}^{\infty} c\, 2^{3i} \sup_{\substack{h_1\in\mathcal{H}_{k+i}\cap A \\ |h_2-h_1| = \frac{j}{8}2^{-3(k+i)} \\ \text{for some } j\in\{1,2,3,4\}}} \mathbb{P}\left(B_q \cap \left\{X'_{r,k}(h_2) - X'_{r,k}(h_1) \geq \frac{a+q}{2(i+1)^2}\right\}\right). \tag{5.4.23}$$

Note that $a+q \leq a+x \leq 2^{6k}$ by assumption. Lemma 5.4.2 can thus be applied to get that

(5.4.23) is at most

$$c\sum_{i=0}^{\infty} 2^{3i}\exp\left(-2\frac{(x-q-1)^2}{2^{2k}-2^{2r}} - \widetilde{c}\, 2^{3i}\frac{(a+q)^{3/2}}{(i+1)^3}\right) \le c' e^{-2\frac{(x-q-1)^2}{2^{2k}-2^{2r}} - \widetilde{c}(a+q)^{3/2}}. \tag{5.4.24}$$

Since $e^{-\widetilde{c}(a+q)^{3/2}} \le e^{-\widetilde{c}a^{3/2} - \widetilde{c}q^{3/2}}$, (5.4.19) is at most

$$\begin{aligned} c'\, e^{-\widetilde{c}a^{3/2}} \sum_{q=0}^{x} e^{-2\frac{(x-q-1)^2}{2^{2k}-2^{2r}} - \widetilde{c}q^{3/2}} &\le c'\, e^{-\frac{2x^2}{2^{2k}-2^{2r}} - \widetilde{c}a^{3/2}} \sum_{q=0}^{x} e^{4C(q+1) - \widetilde{c}q^{3/2}} \\ &\le c'' e^{-\frac{2x^2}{2^{2k}-2^{2r}} - \widetilde{c}a^{3/2}}, \end{aligned} \tag{5.4.25}$$

where we used the assumption $x \le C(2^{2k} - 2^{2r})$ to obtain the first inequality in (5.4.25). This proves (5.3.1). □

Proof of Proposition 5.3.2. The left-hand side of (5.3.2) is at most

$$\begin{aligned} &\mathbb{P}\Big(X'_{r,k}(h) \ge x-2\Big) \\ &\quad + \mathbb{P}\begin{pmatrix} \max_{h':|h'-h|\le 2^{-3k-1}} X'_{r,k}(h') \ge (x-2)+2, \\ X'_{r,k}(h) \le x-2 \end{pmatrix} \end{aligned} \tag{5.4.26}$$

The conclusion follows from Lemma 5.4.1 and Proposition 5.3.1 with $x-2$ in place of x and $a = 2$. □

5.5. Appendix : Technical lemma

Lemma 5.5.1. *Let $m \ge 1$ and $1 \le P < Q$, then*

$$\left|\sum_{P<p\le Q} \frac{(\log p)^m}{p} - \left(\frac{(\log Q)^m}{m} - \frac{(\log P)^m}{m}\right)\right| \le D, \tag{5.5.1}$$

where $D > 0$ is a constant that only depends on m.

Proof. Without loss of generality, assume that $P \ge 2$. We use a standard form of the prime number theorem (Montgomery and Vaughan, 2007, Theorem 6.9) which states that

$$\#\{p \text{ prime} : p \le x\} = \int_2^x \frac{1}{\log u}du + R(x), \tag{5.5.2}$$

where $R(x) = O(xe^{-c\sqrt{\log x}})$, uniformly for $x \ge 2$. Using (5.5.2) and integration by parts, we have

$$\sum_{P<p\leq Q} \frac{(\log p)^m}{p} = \int_P^Q \frac{(\log u)^{m-1}}{u} du + \int_P^Q \frac{(\log u)^m}{u} dR(u)$$
$$= \frac{(\log Q)^m}{m} - \frac{(\log P)^m}{m} + \frac{(\log Q)^m}{Q} R(Q) - \frac{(\log P)^m}{P} R(P) \qquad (5.5.3)$$
$$- \int_P^Q \frac{(m - \log u)(\log u)^{m-1}}{u^2} R(u) du.$$

By making the change of variable $z = c\sqrt{\log u}$ on the right-hand side of (5.5.3), note that

$$\left| \int_P^Q \frac{(m - \log u)(\log u)^{m-1}}{u^2} R(u) du \right| \leq \widetilde{D} \int_0^\infty z^{2m+1} e^{-z} dz = \widetilde{D}\, \Gamma(2m+2), \qquad (5.5.4)$$

where $\widetilde{D} > 0$ is a constant that only depends on m. This ends the proof. □

Acknowledgements

We would like to thank the anonymous referee for his valuable comments that led to improvements in the presentation of this paper. L.-P. Arguin is supported in part by NSF Grant DMS-1513441 and by NSF CAREER DMS-1653602. F. Ouimet is supported by a NSERC Doctoral Program Alexander Graham Bell scholarship (CGS D3).

Bibliography

M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964. MR0167642.

L.-P. Arguin and W. Tai. Is the Riemann zeta function in a short interval a 1-RSB spin glass ? *Preprint. To appear in Sojourns in probability and statistical physics, in honor of the 70th birthday of C. Newman*, pages 1–20, 2018. arXiv:1706.08462.

L.-P. Arguin, D. Belius, and A. J. Harper. Maxima of a randomized Riemann zeta function, and branching random walks. *Ann. Appl. Probab.*, 27(1):178–215, 2017. MR3619786.

L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł, and K. Soundararajan. Maximum of the Riemann zeta function on a short interval of the critical line. *Comm. Pure Appl. Math.*, 72(3):500–535, 2019. MR3911893.

D. W. Farmer, S. M. Gonek, and C. P. Hughes. The maximum size of L-functions. *J. Reine Angew. Math.*, 609:215–236, 2007. MR2350784.

Y. V. Fyodorov and J. P. Keating. Freezing transitions and extreme values: random matrix theory, $\zeta(\frac{1}{2}+it)$ and disordered landscapes. *Philos. Trans. R. Soc. A*, 372(20120503): 1–32, 2014. MR3151088.

Y.V. Fyodorov, G.A. Hiary, and J.P. Keating. Freezing transition, characteristic polynomials of random matrices, and the Riemann zeta-function. *Phys. Rev. Lett.*, 108(170601): 1–4, 2012. doi:10.1103/PhysRevLett.108.170601.

A. J. Harper. A note on the maximum of the Riemann zeta function, and log-correlated random variables. *Preprint*, pages 1–26, 2013. arXiv:1304.0677.

H. L. Montgomery and R. C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007. MR2378655.

J. Najnudel. On the extreme values of the Riemann zeta function on random intervals of the critical line. *Probab. Theory Related Fields*, 172(1-2):387–452, 2018. MR3851835.

F. Ouimet. Poisson-Dirichlet statistics for the extremes of a randomized Riemann zeta function. *Electron. Commun. Probab.*, 23:Paper No. 46, 15, 2018. MR3841407.

Article 6

Moments of the Riemann zeta function on short intervals of the critical line

by

Louis-Pierre Arguin[1], Frédéric Ouimet[2] and Maksym Radziwiłł[3]

([1]) Baruch College and Graduate Center (CUNY), New York, NY 10010, USA.

([2]) Université de Montréal, Montréal, QC H3T 1J4, Canada.

([3]) California Institute of Technology, Los Angeles (Pasadena), CA 91125, USA.

Here is the complete citation :

Arguin, L.-P., Ouimet, F. and Radziwill, M. (2019). Moments of the Riemann zeta function on short intervals of the critical line. https://arxiv.org/abs/1901.04061v2.

Contributions

In an earlier version (conditional on RH), Louis-Pierre sketched the proof structure of the entire article, wrote the introduction and wrote the majority of the proofs. Some parts were missing, and that's where my biggest contribution was : details for the analogue of Section 6.2.3, the non-contribution of the primes up to scales $|\theta|$ when $\theta < 0$ (Corollary 6.2.11), some small details concerning the localization argument, a good portion of Section 6.3.4 and another portion analogous to Lemmas 4.4 and 4.5 in Arguin et al. (2019) that is now cut out. Maksym is the main contributor for the transition to the unconditional version of the article. His discretization in Section 6.2.2 changed most proofs significantly.

Abstract. We show that as $T \to \infty$, for all $t \in [T, 2T]$ outside of a set of measure $o(T)$,

$$\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\tfrac{1}{2} + it + ih)|^\beta dh = (\log T)^{f_\theta(\beta)+o(1)},$$

for some explicit exponent $f_\theta(\beta)$, where $\theta > -1$ and $\beta > 0$. This proves an extended version of a conjecture of Fyodorov and Keating (2014). In particular, it shows that, for all $\theta > -1$, the moments exhibit a phase transition at a critical exponent $\beta_c(\theta)$, below which $f_\theta(\beta)$ is quadratic and above which $f_\theta(\beta)$ is linear. The form of the exponent f_θ also differs between mesoscopic intervals ($-1 < \theta < 0$) and macroscopic intervals ($\theta > 0$), a phenomenon that stems from an approximate tree structure for the correlations of zeta. We also prove that, for all $t \in [T, 2T]$ outside a set of measure $o(T)$,

$$\max_{|h| \le \log^\theta T} |\zeta(\tfrac{1}{2} + it + ih)| = (\log T)^{m(\theta)+o(1)},$$

for some explicit $m(\theta)$. This generalizes earlier results of Najnudel (2018) and Arguin et al. (2019) for $\theta = 0$. The proofs are unconditional, except for the upper bounds when $\theta > 3$, where the Riemann hypothesis is assumed.

Keywords: Extreme value theory, Riemann zeta function, maximum, moments

6.1. Introduction

6.1.1. Maxima and moments over large intervals

Understanding the growth of the Riemann zeta function $\zeta(s)$ on the critical line $\operatorname{Re} s = \frac{1}{2}$ is a central problem in number theory due, among other things, to its relationship with the distribution of the zeros of $\zeta(s)$, see e.g. Theorem 9.3 in Titchmarsh (1986), and the more general subconvexity problem, see e.g. Michel and Venkatesh (2010); Venkatesh (2010), and see Iwaniec and Sarnak (2000) for a general discussion.

The Lindelöf hypothesis predicts that, for any $\varepsilon > 0$ and all $t \in \mathbb{R}$, we have $|\zeta(\frac{1}{2} + it)| = O((1 + |t|)^\varepsilon)$, whereas it follows from the Riemann hypothesis that

$$|\zeta(\tfrac{1}{2} + it)| = O\left(\exp\left(\left(\frac{\log 2}{2} + o(1)\right)\frac{\log t}{\log\log t}\right)\right), \tag{6.1.1}$$

as $t \to \infty$; see Chandee and Soundararajan (2011).

Unfortunately, there is a large gap between these conditional results and the best unconditional upper bounds, such as Bourgain (2017), which shows that $|\zeta(\frac{1}{2}+it)| = O\left((1+|t|)^{13/84+\varepsilon}\right)$ for any given $\varepsilon > 0$ and all $t \in \mathbb{R}$. Currently, the best unconditional lower bound,

$$\max_{t\in[0,T]} |\zeta(\tfrac{1}{2}+it)| \geq \exp\left((\sqrt{2}+o(1))\sqrt{\frac{\log T \log\log\log T}{\log\log T}}\right), \tag{6.1.2}$$

as $T \to \infty$, is established in de la Bretèche and Tenenbaum (2019) building on a method from Bondarenko and Seip (2018).

The true order of the maximum of $|\zeta(\frac{1}{2}+it)|$ remains a subject of dispute to this day. A conjecture that we find plausible is stated in Farmer et al. (2007), where it is conjectured based on probabilistic models that

$$\max_{t\in[0,T]} |\zeta(\tfrac{1}{2}+it)| = \exp\left(\left(\frac{1}{\sqrt{2}}+o(1)\right)\sqrt{\log T \cdot \log\log T}\right), \quad \text{as } T \to \infty. \tag{6.1.3}$$

Another set of central objects in the theory of the Riemann zeta function are the *moments*

$$\frac{1}{T}\int_T^{2T} |\zeta(\tfrac{1}{2}+it)|^\beta dt, \quad \beta > 0. \tag{6.1.4}$$

Their importance comes from their relationship to the size and zero-distribution of $\zeta(s)$. However, unlike the problem of understanding the size of the global maximum of $|\zeta(\frac{1}{2}+it)|$, we are in possession of widely believed conjectures as to the behavior of moments. Following the work Keating and Snaith (2000), it is expected that, for all $\beta > 0$,

$$\frac{1}{T}\int_T^{2T} |\zeta(\tfrac{1}{2}+it)|^\beta dt \sim C_\beta (\log T)^{\beta^2/4}, \tag{6.1.5}$$

as $T \to \infty$, and that the constant $C_\beta > 0$ factors into a product of two constants: one is computed from the moments of the characteristic polynomial of random unitary matrices, and the other is an arithmetic factor coming from the small primes.

There are a few results supporting (6.1.5). First, the conjecture (6.1.5) is known for $\beta = 2$ and $\beta = 4$ following the classical work of Hardy-Littlewood and Ingham. Upper bounds of the correct order of magnitude are established in Heap et al. (2019) for $0 < \beta \leq 4$.

Meanwhile, lower bounds of the correct order of magnitude have been established for all $\beta \geq 2$ in Radziwiłł and Soundararajan (2013). Conditionally on the Riemann hypothesis, the correct order of magnitude of (6.1.5) is known for all $\beta > 0$ (see Soundararajan (2009); Harper (2013b) for the upper bounds and Heath-Brown (1981) for the lower bounds).

6.1.2. Maxima and moments over short intervals

Motivated by the problem of understanding the global maximum, Fyodorov et al. (2012); Fyodorov and Keating (2014) initiated the question of understanding the true size of *the local maximum of* $\zeta(\frac{1}{2}+it)$ by establishing a connection with log-correlated processes. If τ is sampled uniformly on $[T, 2T]$ under $\mathbb{P}$, they conjectured that for any $0 < \delta < 1$, there exists $C = C(\delta) > 0$ large enough and independent of T, such that with $\mathbb{P}$-probability $1-\delta$,

$$\max_{h \in [-1,1]} \log |\zeta(\tfrac{1}{2}+i\tau+ih)| - \Big(\log\log T - \frac{3}{4}\log\log\log T\Big) \in [-C, C]. \tag{6.1.6}$$

They also conjectured the type of fluctuations around the recentering term.

The leading order $\log\log T$ was proved in Najnudel (2018) (conditionally on the Riemann hypothesis for the lower bound) and in Arguin et al. (2019) unconditionally. Around Equation (14) in Fyodorov et al. (2012), it is also argued that the moments in a short interval undergo a *freezing phase transition*, that is, as $T \to \infty$, the event,

$$\int_{[-1,1]} |\zeta(\tfrac{1}{2}+i\tau+ih)|^{\beta} dh = \begin{cases} (\log T)^{\beta^2/4+o(1)}, & \text{if } \beta \leq 2, \\ (\log T)^{\beta-1+o(1)}, & \text{if } \beta > 2, \end{cases} \tag{6.1.7}$$

has probability $1 - o(1)$ as $T \to \infty$. Fyodorov and Keating (2014) state corresponding conjectures for mesoscopic intervals of length $\log^{\theta} T$ when $\theta \in (-1, 0)$, as well as finer asymptotics for the moments.

In view of Equations (6.1.5) and (6.1.7), an obvious question is to determine up to which interval size the freezing phase transition persists. In this paper, we establish that freezing transitions occur exactly for interval sizes of order $\log^{\theta} T$ with $\theta > -1$, including large intervals with $\theta > 0$. We also obtain the corresponding results for local maxima over such intervals. The following functions will be crucial to our analysis :

$$\begin{aligned}
&\theta \le 0: \qquad m(\theta) := 1+\theta, \qquad f_\theta(\beta) := \begin{cases} \frac{\beta^2}{4}(1+\theta)+\theta, & \text{if } \beta \le \beta_c(\theta) = 2, \\ \beta m(\theta) - 1, & \text{if } \beta > \beta_c(\theta), \end{cases} \\
&\theta > 0: \qquad m(\theta) := \sqrt{1+\theta}, \qquad f_\theta(\beta) := \begin{cases} \frac{\beta^2}{4}+\theta, & \text{if } \beta \le \beta_c(\theta) = 2\sqrt{1+\theta}, \\ \beta m(\theta) - 1, & \text{if } \beta > \beta_c(\theta). \end{cases}
\end{aligned} \tag{6.1.8}$$

Theorem 6.1.1 (Moments)**.** *Let $\theta > -1$, $\beta > 0$ and $\varepsilon > 0$ be given. Let τ be a random variable uniformly distributed on $[T, 2T]$ under the probability measure $\mathbb{P}$. Then, as $T \to \infty$, we have*

$$\mathbb{P}\Big(\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)|^\beta dh < (\log T)^{f_\theta(\beta) - \varepsilon} \Big) = o(1). \tag{6.1.9}$$

Moreover, if $\theta \le 3$ or if the Riemann hypothesis holds, then as $T \to \infty$,

$$\mathbb{P}\Big(\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)|^\beta dh > (\log T)^{f_\theta(\beta) + \varepsilon} \Big) = o(1). \tag{6.1.10}$$

Proof. For the upper bound, see Section 6.2.3, and for the lower bound, see Proposition 6.3.2. ∎

When $\beta > \beta_c(\theta)$, the moments exhibit *freezing*, i.e. they are dominated by just one large value corresponding to the local maximum of $|\zeta(\frac{1}{2} + i\tau + ih)|$, $|h| \le \log^\theta T$. Theorem 6.1.1 also suggests that freezing does not occur for intervals larger than any fixed power of $\log T$, since $\beta_c(\theta) \to \infty$ as $\theta \to \infty$.

Theorem 6.1.2 (Local maximum)**.** *Let $\theta > -1$ and $\varepsilon > 0$ be given. Let τ be a random variable uniformly distributed on $[T, 2T]$ under the probability measure $\mathbb{P}$. Then, as $T \to \infty$, we have*

$$\mathbb{P}\Big(\max_{|h| \le \log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)| < (\log T)^{m(\theta) - \varepsilon} \Big) = o(1). \tag{6.1.11}$$

Moreover, if $\theta \le 3$ or if the Riemann hypothesis holds, then as $T \to \infty$,

$$\mathbb{P}\Big(\max_{|h| \le \log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)| > (\log T)^{m(\theta) + \varepsilon} \Big) = o(1). \tag{6.1.12}$$

PROOF. For the upper bound, see Section 6.2.3, and for the lower bound, see Proposition 6.3.1. □

It is instructive to put these results in the context of two well-known facts on ζ. First, Selberg's central limit theorem, see for example Radziwiłł and Soundararajan (2017), states that, for any given $a < b$,

$$\mathbb{P}\left(\frac{\log|\zeta(\frac{1}{2}+i\tau)|}{\sqrt{\frac{1}{2}\log\log T}} \in (a,b)\right) \xrightarrow{T\to\infty} \int_a^b \frac{e^{-u^2/2}}{\sqrt{2\pi}}\,du. \tag{6.1.13}$$

In other words, a typical value of $\log|\zeta(\frac{1}{2}+i\tau)|$ is a Gaussian random variable of variance $\frac{1}{2}\log\log T$. This is consistent with the moment conjecture (6.1.5) which gives a precise expression for the Laplace transform of $\log|\zeta(\frac{1}{2}+i\tau)|$. Second, since $\zeta(\frac{1}{2}+it)$ with $T \le t \le 2T$ varies on the scale of $(\log T)^{-1}$, the statistics of extreme values of $\log|\zeta(\frac{1}{2}+i\tau+ih)|$, $|h| \le \log^\theta T$, should be similar to the ones of $(\log T)^{1+\theta}$ Gaussian random variables of variance $\frac{1}{2}\log\log T$. If the random variables were independent, this is the so-called *Random Energy Model* (REM) in statistical mechanics introduced in Derrida (1981). For $\theta \ge 0$, it is not hard to check, using basic Gaussian tail estimates, that the expression (6.1.8) corresponds to the free energy of the model, and the results of Theorem 6.1.2, to the maximum of the REM. For more on this, we refer to Kistler (2015), where many techniques from REM were introduced to analyze log-correlated processes.

The REM heuristic is of course limited as the values of $\log|\zeta(\frac{1}{2}+i\tau+ih)|$, $|h| \le \log^\theta T$, are correlated. In fact, they are log-correlated as first noticed Bourgade (2010). This is explained in Section 6.1.4. For $\theta < 0$, the correct model is a branching random walk which accurately predicts the changes in $m(\theta)$ and $f_\theta(\beta)$. For $\theta > 0$, our results show that the correlations do not affect large values at leading order (though the proofs must take them into account). As argued in Section 6.1.4, we believe that the correct probabilistic model for large values in this case is $\log^\theta T$ independent branching random walks. One implication is that, unlike the case $\theta \le 0$, the REM heuristic should persist to subleading order (but fail at the level of fluctuations). In view of this, we believe that conjecture (6.1.6) needs to be expanded as follows to include large intervals:

Conjecture 6.1.3. *Let $\theta > -1$ be given and let $m(\theta)$ be as in (6.1.8). Let τ be sampled uniformly on $[T, 2T]$ under $\mathbb{P}$. For any $0 < \delta < 1$, there exists $C = C(\delta) > 0$ large enough and independent of T, such that with probability $1-\delta$,*

$$\max_{|h| \le \log^\theta T} \log|\zeta(\tfrac{1}{2} + i\tau + ih)| - \Big(m(\theta)\log\log T - r(\theta)\log\log\log T\Big) \in [-C, C], \quad (6.1.14)$$

where

$$r(\theta) = \frac{3}{4} \quad \text{if } \theta \le 0 \qquad \text{and} \qquad r(\theta) = \frac{1}{4\sqrt{1+\theta}} \quad \text{if } \theta > 0.$$

6.1.3. Relations to other models

When $-1 < \theta \le 0$, Conjecture 6.1.3 is based on modelling ζ by the characteristic polynomial of a random unitary matrix (CUE). More precisely, if M_N is a random matrix sampled from the Haar measure on the unitary group $\mathcal{U}(N)$, one can consider the moments

$$\mathbb{E}\Bigg[\bigg(\frac{1}{2\pi}\int_0^{2\pi} |\det(\mathbb{I} - e^{-ih}M_N)|^{2\beta} dh\bigg)^k\Bigg], \quad k > 0,\ \beta > 0. \quad (6.1.15)$$

These can be computed in the limit $N \to \infty$, at least heuristically, using Selberg integrals and the Fisher-Hartwig formula, cf. Fyodorov and Keating (2014). Exact expressions were recently obtained in Bailey and Keating (2019) in the regime $k, \beta \in \mathbb{N}$. The statistics of $\log \int_0^{2\pi} |\det(\mathbb{I} - e^{-ih}M_N)|^{2\beta} dh$ and of $\max_{h\in[0,2\pi]} |\det(\mathbb{I} - e^{-ih}M_N)|$ in the limit $N \to \infty$ can be inferred from the asymptotics of the moments by comparison with log-correlated processes, cf. Fyodorov et al. (2018) for a numerical study. In the CUE setting, the freezing analogue of (6.1.7) and the leading order as in (6.1.6) were proved in Arguin et al. (2017a). The subleading order of the maximum was proved in Paquette and Zeitouni (2018), and up to constant C in Chhaibi et al. (2018).

In the subcritical regime $\beta < \frac{1}{2}$, it is expected from the analysis of log-correlated processes, cf. Fyodorov and Bouchaud (2008), that the fluctuations of the maximum can be captured by a sum of two Gumbel variables. This was proved in Rémy (2018) for a specific log-correlated model by computing the moments in the range $k < \frac{1}{4\beta^2}$ of a random measure related to the theory of Gaussian multiplicative chaos, cf. Rhodes and Vargas (2014). In the CUE setting, this measure is the limit of

$$\frac{|\det(\mathbb{I}-e^{-ih}M_N)|^{2\beta}}{\mathbb{E}[|\det(\mathbb{I}-e^{-ih}M_N)|^{2\beta}]}\frac{dh}{2\pi}. \tag{6.1.16}$$

The limit of the above was shown to be non-degenerate for $\beta < 1$ in Webb (2015); Nikula et al. (2018). Such a random measure can also be considered in the context of the Riemann zeta function for mesoscopic intervals of length $\log^\theta T$, $-1 < \theta \leq 0$, with $|\zeta(\frac{1}{2}+i\tau+ih)|$ in place of $|\det(\mathbb{I}-e^{-ih}M_N)|$. (There does not seem to be any obvious equivalent for macroscopic intervals, $\theta > 0$, in the CUE model.) A step in this direction was made in Saksman and Webb (2018) where $\zeta(\frac{1}{2}+i\tau+ih)$, $h\in\mathbb{R}$, was shown to converge as $T\to\infty$ when considered as a random variable on the space of tempered distributions.

Another model for the large values of $\log|\zeta(\frac{1}{2}+i\tau+ih)|$, $h\in[-1,1]$, is to consider a random Dirichlet polynomial $X_h = \mathrm{Re}\sum_{p\leq T} p^{-1/2-ih}U_p$, where ($U_p$, p primes) are i.i.d. uniform random variables on the unit circle, cf. Harper (2013a); Arguin et al. (2017b); Arguin and Ouimet (2019). The analogue of conjecture (6.1.6) for this model was proved up to second-order corrections in Arguin et al. (2017b), and large deviations and continuity estimates for the derivative were found in Arguin and Ouimet (2019). The limit of the corresponding multiplicative chaos measure was obtained in Saksman and Webb (2018). A proof of the freezing phase transition was given in Arguin and Tai (2018). In the latter, the limit of the Gibbs measure $\exp(\beta X_h)dh$ is also studied in the supercritical regime $\beta > 2$, showing that it is supported on h's that are at a relative distance of order one or order $(\log T)^{-1}$ of each other. This result was used in Ouimet (2018) to prove that the normalized Gibbs weights converge to a Poisson-Dirichlet distribution.

Notation. Throughout the article, the notation τ will denote a random variable uniformly distributed on $[T, 2T]$ under $\mathbb{P}$. Expectations under $\mathbb{P}$ are denoted by $\mathbb{E}$. We write $f(T) = o(g(T))$ if $|f(T)/g(T)|$ tends to 0 as $T\to\infty$ when the parameters θ, β and ε are fixed. Similarly, we write $f(T) = O(g(T))$ if $\limsup|f(T)/g(T)|$ is bounded for θ, β and ε fixed. We will sometimes write for conciseness $f(T) \ll g(T)$ if $f(T) = O(g(T))$, and also $f(T) \asymp g(T)$ if both $f(T) \ll g(T)$ and $g(T) \ll f(T)$ hold. Finally, in some of the proofs, we use the common convention in analytic number theory that ε denotes an arbitrarily small positive quantity that may vary from line to line.

6.1.4. Outline of the proof

For $\theta > 0$, the upper bound part of Theorem 6.1.1 and Theorem 6.1.2 follows from the moment estimates

$$\mathbb{E}\Big[|\zeta(\tfrac{1}{2}+i\tau)|^{\beta}\Big] \ll (\log T)^{\beta^2/4+\varepsilon}, \tag{6.1.17}$$

and from a discretization result which roughly shows that for a Dirichlet polynomial D that approximates zeta, and for $\beta \geq 1$, we have

$$\max_{|h|\leq \log^{\theta} T} |D(\tfrac{1}{2}+i\tau+ih)|^{\beta} \ll \sum_{|k|\leq \log^{1+\theta} T} \Big|D\Big(\tfrac{1}{2}+i\tau+\tfrac{2\pi i k}{\log T}\Big)\Big|^{\beta}. \tag{6.1.18}$$

Equation (6.1.18) tells us that the process $(\zeta(\frac{1}{2}+i\tau+ih),\ |h| \leq \log^{\theta} T)$ varies on a $(\log T)^{-1}$ scale, so that the maximum and moments on an interval of length $O(\log^{\theta} T)$ behave as those of $O(\log^{1+\theta} T)$ i.i.d. Gaussian random variables of variance $\frac{1}{2}\log\log T$. The limitation to $\theta \leq 3$ comes from the fact that the upper bounds (6.1.17) are not known unconditionally for $\beta > 4$.

When $\theta < 0$, the upper bounds in Theorem 6.1.1 and Theorem 6.1.2 are a bit more delicate. We follow essentially the same strategy, but we apply it to the function

$$(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau), \quad \text{where } \mathcal{P}_{\alpha}(s) = \sum_{\log p \leq \log^{\alpha} T} \frac{1}{p^{s}}, \tag{6.1.19}$$

instead of $\zeta(\frac{1}{2}+i\tau)$. As discussed in more detail below, the reason for this is that when $\theta < 0$, the contribution of the primes up to scale $|\theta|$ is negligible with high probability, namely, with probability $1-o(1)$,

$$\max_{|h|\leq \log^{\theta} T} \Big|\mathcal{P}_{|\theta|}(\tfrac{1}{2}+i\tau+ih)\Big| = o(\log\log T). \tag{6.1.20}$$

When τ is restricted to a specific event $\mathcal{A}(T)$ on which (6.1.19) can be discretized as in (6.1.18), we can show that

$$\mathbb{E}\Big[\Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau)\Big|^{\beta}\Big] \ll (\log T)^{(\beta^2/4)\cdot(1+\theta)+\varepsilon}, \tag{6.1.21}$$

for $\beta \leq 2$. This explains the additional factor $(\beta^2/4)\theta$ in $f_{\theta}(\beta)$ when $-1 < \theta < 0$ and $\beta \leq 2$.

We then turn to the lower bound part of Theorem 6.1.1 and Theorem 6.1.2. The lower bounds in Theorem 6.1.2 follow directly from Theorem 6.1.1 (see (6.3.74)), so it is enough to discuss Theorem 6.1.1.

The problem is first reduced to obtaining lower bounds for moments off the critical line. In particular, it is shown, uniformly in $\frac{1}{2} \le \sigma \le \frac{1}{2} + (\log T)^{\theta - 3\varepsilon}$ and for any given $\varepsilon > 0$, that with probability $1 - o(1)$,

$$\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\sigma + i\tau + ih)|^\beta dh \ll \int_{-2\log^\theta T}^{2\log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)|^\beta dh + \frac{1}{(\log T)^7}. \tag{6.1.22}$$

This is accomplished by using a result of Gabriel (1927) for subharmonic functions, and the construction of an explicit entire function which is a good approximation to the indicator function of the rectangle $\mathcal{R} = \{\sigma + iu : |u| \le (\log T)^\theta, \frac{1}{2} \le \sigma \le \frac{1}{2} + (\log T)^{\theta - 3\varepsilon}\}$ in the whole strip $\frac{1}{2} \le \operatorname{Re} s$. The fact that the interval can be very small when $\theta < 0$ makes this part rather technical. We believe that this result might be useful in other applications as well.

The problem is therefore reduced to obtaining a good lower bound for

$$\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\sigma_0 + i\tau + ih)|^\beta dh, \quad \text{with } \sigma_0 = \frac{1}{2} + \frac{1}{(\log T)^{1-\delta}}, \tag{6.1.23}$$

for some sufficiently small $\delta > 0$. We adapt mollification results from Arguin et al. (2019) to show that, outside of an event of probability $o(1)$, the problem can be reduced to understanding

$$\int_{-\log^\theta T}^{\log^\theta T} \exp\Big(\beta \operatorname{Re} \mathcal{P}_{1-\delta}(\sigma_0 + i\tau + ih)\Big) dh. \tag{6.1.24}$$

The proof of the lower bound is now restricted to the problem of understanding the correlation structure of the process

$$\Big(\operatorname{Re} \mathcal{P}_{1-\delta}(\sigma_0 + i\tau + ih),\ |h| \le \log^\theta T\Big). \tag{6.1.25}$$

The remaining part of the argument is done in Section 6.3.4 by a multiscale second moment method introduced in Kistler (2015). The covariance of the process (6.1.25) can be computed using Lemma 6.4.3 with $a(p) = p^{-\sigma_0}(p^{-ih} + p^{-ih'})$:

$$\mathbb{E}\Big[\operatorname{Re}\mathcal{P}_{1-\delta}(\sigma_0 + i\tau + ih) \cdot \operatorname{Re}\mathcal{P}_{1-\delta}(\sigma_0 + i\tau + ih')\Big] \\ = \frac{1}{2}\sum_{\log p \le (\log T)^{1-\delta}} \frac{\cos(|h-h'|\log p)}{p^{2\sigma_0}} + O(1). \tag{6.1.26}$$

The cosine factor implies that primes smaller than $\exp(|h-h'|^{-1})$ are almost perfectly correlated, whereas primes greater than $\exp(|h-h'|^{-1})$ decorrelate quickly. In fact, the covariance can be evaluated precisely using the prime number theorem and equals $\frac{1}{2}\log|h-h'|^{-1} + O(1)$. This shows that the process is approximatively a log-correlated Gaussian process. (This is also true for ζ in the sense of finite-dimensional distributions as shown in Bourgade (2010).)

The identification with a log-correlated process is useful as it suggests that the Dirichlet polynomials have an underlying tree structure. To see this, consider the *increments*

$$P_k(h) = \sum_{e^{k-1} < \log p \le e^k} \operatorname{Re}\frac{1}{p^{\sigma_0 + i\tau + ih}}, \quad 1 \le k \le \log\log T. \tag{6.1.27}$$

The range of primes is chosen so that each P_k has variance $\frac{1}{2} + o(1)$. In this framework, the Dirichlet polynomial at h can be seen as a random walk with independent and identically distributed increments. However, the random walks for different h's are not independent by (6.1.26). In fact, the walks are almost perfectly correlated until they *branch out* around the prime $p \approx \exp(|h-h'|^{-1})$, corresponding to the increment $k(h,h') = \log|h-h'|^{-1}$. Since k goes to essentially $\log\log T$, the analysis can be restricted to h's at a distance $(\log T)^{-1}$ of each other. Furthermore, the h's in an interval of size $(\log T)^{-\alpha}$ for $0 < \alpha < 1$ will share the same increments up to $k \approx \alpha\log\log T$.

The above observations have important consequences for the probabilistic analysis. For $\theta = 0$, this means that the process (6.1.25) on an interval of order one is well approximated by a Gaussian process indexed by a tree of average degree $e = 2.718\ldots$, where the independent increments $P_k(h)$ are identified with the edges of the tree. Note that the number of leaves on the interval $[-1,1]$ is then $\approx e^{\log\log T} = \log T$. Equivalently, the walks $\sum_k P_k(h)$, $h \in [-1,1]$, can be seen as a branching random walk on a Galton-Watson tree with an average number of offspring e, cf. Figure 6.1.1.

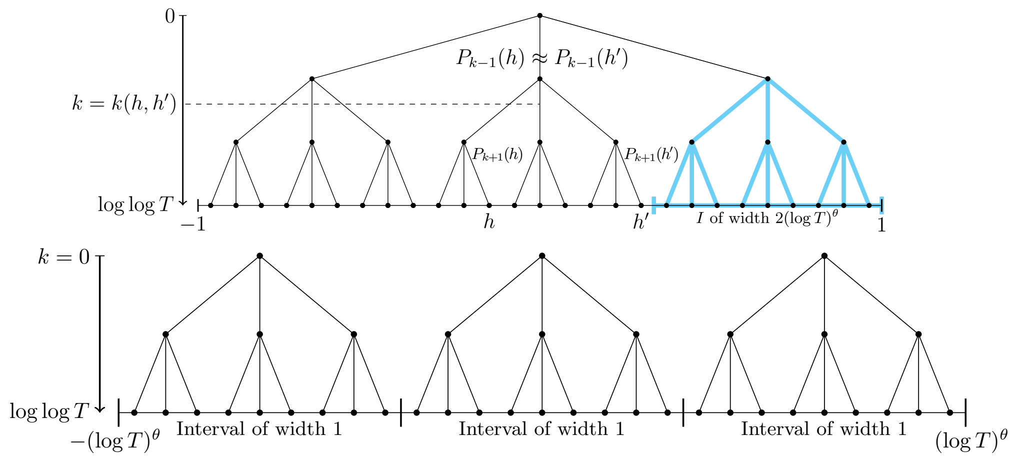


FIGURE 6.1.1. **(Top)** An illustration of the branching random walk $\sum_k P_k$ for the interval I with $\theta = 0$. The one for a subinterval with $\theta < 0$ is depicted in blue. **(Bottom)** An illustration of the independent branching random walks $\sum_k P_k$ for disjoint intervals of width 1 inside I of length $2\log^\theta T$ with $\theta > 0$.

For $\theta < 0$, the tree structure suggests that the primes up to $\exp(\log^{|\theta|} T)$ do not contribute to large values, since they should be essentially the same for all h's in the interval . Therefore these primes can be cutoff at a low cost, cf. Corollary 6.2.11. This is equivalent to restricting to a subtree of the one on $[-1, 1]$ with $(1+\theta)\log\log T$ increments and $\log^{1+\theta} T$ leaves, yielding a maximum at leading order of $(1 + \theta)\log\log T$ by the REM heuristic.

The case $\theta > 0$ stands out as the analogy with branching random walks fails. This is because the random walks for h and h' are essentially independent for $|h - h'| > 1$. Therefore the right probabilistic model seems to consist of $\log^\theta T$ independent branching random walks corresponding to different intervals of order one, see Figure 6.1.1. A large class of similar models (called CREM's for *Continuous Random Energy Models*) have been studied in Bovier and Kurkova (2004), see Bovier (2006, 2017) for a review. It turns out that the large values at leading order correspond to the ones of a REM with $\log^{1+\theta} T$ variables with variance $\frac{1}{2}\log\log T$. This yields a maximum of $\sqrt{1+\theta}\log\log T$ at leading order. In fact, in view of the extreme value statistics of CREM's, we expect that the REM heuristic holds for subleading corrections. This is the motivation for Conjecture 6.1.3.

6.2. Upper bounds

6.2.1. Moment estimates

We will need a number of moment estimates which we state below.

Proposition 6.2.1. *Assume the Riemann hypothesis. Let $\beta > 0$ and $\varepsilon > 0$ be given. Then,*

$$\mathbb{E}\Big[|\zeta(\tfrac{1}{2} + i\tau)|^{\beta}\Big] \ll (\log T)^{\beta^2/4+\varepsilon}. \tag{6.2.1}$$

Proof. See Corollary A in Soundararajan (2009). □

Proposition 6.2.2. *Let $0 < \beta \leq 4$ be given. Then,*

$$\mathbb{E}\Big[|\zeta(\tfrac{1}{2} + i\tau)|^{\beta}\Big] \ll (\log T)^{\beta^2/4}. \tag{6.2.2}$$

Proof. See Theorem 1 in Heap et al. (2019). □

The proof of Proposition 6.2.1 is based on the following deterministic upper bound for ζ: Suppose that T is large. Let $T \leq t \leq 2T$, and let $2 \leq x \leq T^2$. Then, as $T \to \infty$, we have

$$\log|\zeta(\frac{1}{2} + it)| \leq \operatorname{Re} \sum_{p \leq x} \frac{1}{p^{\frac{1}{2}+\frac{1}{\log x}+it}} \frac{\log(x/p)}{\log x} + \frac{\log T}{\log x} + O(\log\log\log T), \tag{6.2.3}$$

see Proposition and Lemma 2 in Soundararajan (2009). On the Riemann hypothesis, the upper bounds to Theorems 6.1.1 and 6.1.2 could be proved in a simpler way by using this deterministic bound, and by proving the corresponding results for the Dirichlet polynomials. For unconditional results, such a deterministic upper bound is not available. We need to work *on average* to discard the contribution of large primes. This is the purpose of Lemmas 6.2.3, 6.2.4, 6.2.5 and Proposition 6.2.6 below.

Everywhere in Section 6.2, we will denote, for $\alpha > 0$ and $s \in \mathbb{C}$,

$$\mathcal{P}_{\alpha}(s) = \sum_{\log p \leq \log^{\alpha} T} p^{-s}. \tag{6.2.4}$$

To compute the moments of $\zeta \cdot e^{-\mathcal{P}_{|\theta|}}$, we will need to express $e^{-\mathcal{P}_{|\theta|}}$ as a finite Dirichlet polynomial. To this aim, notice that if $|z| \leq \nu/10$ for some ν, we have $\Big|e^z - \sum_{j=0}^{\nu} z^j/j!\Big| \leq e^{-\nu}$.

Consider more generally $e^{\lambda\mathcal{P}(s)}$ with $\lambda \in \mathbb{C}$ and $\mathcal{P}(s) = \sum_{p\le X} a(p)p^{-s}$ for some bounded multiplicative function a. We have by the above, assuming $|\lambda\mathcal{P}(s)| \le \nu/10$, and by the multinomial formula, that

$$\Big|e^{\lambda\mathcal{P}(s)} - \sum_{k=0}^{\nu}\frac{\lambda^k}{k!}\Big(\sum_{p\le X}\frac{a(p)}{p^s}\Big)^k\Big| = \Big|e^{\lambda\mathcal{P}(s)} - \sum_{\substack{\Omega(n)\le\nu\\ p|n\implies p\le X}}\frac{\lambda^{\Omega(n)}a(n)\mathfrak{g}(n)}{n^s}\Big| \le e^{-\nu}, \tag{6.2.5}$$

where $\Omega(n)$ is the number of prime factors of n with multiplicity. Here, $\mathfrak{g}$ is the multiplicative function defined by $\mathfrak{g}(p^k) = 1/k!$ for all integers k and primes p.

The relevant multiplicative function a for $e^{-\mathcal{P}_{|\theta|}}$ will be of the following form: Given $\alpha, \beta \in \mathbb{R}$ and $\theta > -1$, let $\mathfrak{F}_{\alpha,\beta,\theta}(n)$ denote a completely multiplicative function such that

$$\mathfrak{F}_{\alpha,\beta,\theta}(p) := \begin{cases} \alpha, & \text{if } \log p \le \log^{|\theta|} T, \\ \beta, & \text{if } \log^{|\theta|} T \le \log p. \end{cases} \tag{6.2.6}$$

In the next three lemmas, we control various terms with the aim of proving the moment estimate in Proposition 6.2.6, which we will need in the case of short intervals.

Lemma 6.2.3. *Let $-1 < \theta < 0$, $\beta > 0$ and $\varepsilon > 0$ be given. Then,*

$$\mathbb{E}\Big[\Big|\sum_{\substack{\Omega(n)\le 100\lfloor\log\log T\rfloor\\ p|n\implies \log p\le\log^{1-\varepsilon}T}}\frac{\mathfrak{F}_{0,\beta/2,\theta}(n)\mathfrak{g}(n)}{n^{1/2+i\tau}}\Big|^2\Big] \ll (\log T)^{\beta^2(1+\theta)/4}. \tag{6.2.7}$$

Proof. Notice that the Dirichlet polynomial in (6.2.7) has length $\ll T^\delta$ for any fixed $\delta > 0$. In particular, by the mean-value formula (Lemma 6.4.2),

$$\mathbb{E}\Big[\Big|\sum_{\substack{\Omega(n)\le 100\lfloor\log\log T\rfloor\\ p|n\implies \log p\le\log^{1-\varepsilon}T}}\frac{\mathfrak{F}_{0,\beta/2,\theta}(n)\mathfrak{g}(n)}{n^{1/2+i\tau}}\Big|^2\Big] \ll \sum_{\substack{\Omega(n)\le 100\lfloor\log\log T\rfloor\\ p|n\implies \log p\le\log^{1-\varepsilon}T}}\frac{\mathfrak{F}_{0,\beta/2,\theta}(n)^2\mathfrak{g}(n)^2}{n}.$$

Dropping the restriction over $\Omega(n)$ and expressing the sum as an Euler product yield

$$\sum_{p|n\implies \log p\le\log^{1-\varepsilon}T}\frac{\mathfrak{F}_{0,\beta/2,\theta}(n)^2\mathfrak{g}(n)^2}{n} = \prod_{\log p\le\log^{1-\varepsilon}T}\Big(1 + \frac{\mathfrak{F}_{0,\beta/2,\theta}(p)^2}{p} + O(p^{-2})\Big). \tag{6.2.8}$$

The logarithm of the right-hand side is easily evaluated using the prime number theorem (see Lemma 6.4.1) and is $(\beta^2(1+\theta)/4)\log\log T + O(1)$. This proves the claimed bound. $\square$

Lemma 6.2.4. *Let $-1 < \theta < 0$, $0 < \beta \leq 2$ and $\varepsilon > 0$ be given. Then,*

$$\mathbb{E}\Bigg[|\zeta(\tfrac{1}{2}+i\tau)|^2 \cdot \Big| \sum_{\substack{\Omega(n)\leq 100\lfloor \log\log T\rfloor \\ p|n \implies \log p \leq \log^{1-\varepsilon} T}} \frac{\mathfrak{F}_{-1,\beta/2-1,\theta}(n)\mathfrak{g}(n)}{n^{1/2+i\tau}}\Big|^2\Bigg] \ll (\log T)^{\beta^2(1+\theta)/4+\varepsilon}. \tag{6.2.9}$$

Proof. By Theorem 1 in Bettin et al. (2017), the left-hand side of (6.2.9) is

$$\leq \sum_{\substack{\Omega(n)\leq 100\lfloor \log\log T\rfloor \\ \Omega(m)\leq 100\lfloor \log\log T\rfloor \\ p|n \implies \log p\leq \log^{1-\varepsilon} T \\ p|m \implies \log p\leq \log^{1-\varepsilon} T}} \frac{\mathfrak{F}_{-1,\beta/2-1,\theta}(nm)\mathfrak{g}(n)\mathfrak{g}(m)}{[n,m]} \cdot \frac{1}{T}\int_{\mathbb{R}} \Big(\log\Big(\frac{t(n,m)^2}{2\pi nm}\Big)+2\gamma\Big)\Phi\Big(\frac{t}{T}\Big)dt + O(T^{-\varepsilon}), \tag{6.2.10}$$

where Φ is a smooth non-negative function such that $\Phi(x) \geq 1$ for $1 \leq x \leq 2$, and (n,m) and $[n,m]$ stand for the greatest common divisor and the least common multiple, respectively.

We first note that if n, m are square-free then $[n,m]$ is the product over the distinct prime factors of n and m. This means that if $a(n)$ and $b(m)$ are two bounded multiplicative functions, we have

$$\sum_{\substack{p|n \implies p\leq X \\ p|m \implies p \leq X}} \frac{a(n)b(m)}{[n,m]} = \prod_{p\leq X}\Bigg(1+\frac{a(p)}{p}+\frac{b(p)}{p}+\frac{a(p)\cdot b(p)}{p}+O(p^{-2})\Bigg). \tag{6.2.11}$$

This holds simply by enumerating the ordered pairs of integers in terms of the prime factors considering the four possibilities: p does not divide n nor m, p divides n, p divides m and p divides both n and m.

Using Chernoff's bound, we can get rid of the restriction $\Omega(n) \leq 100\lfloor \log\log T\rfloor$ in (6.2.10). It suffices to notice that the contribution of the sum over n with $\Omega(n) > 100\lfloor \log\log T\rfloor$ is

$$\begin{aligned} &\ll \log T \sum_{\substack{p|n \implies p\leq T \\ p|m \implies p\leq T}} \frac{|\mathfrak{F}_{-1,\beta/2-1,\theta}(nm)|}{[n,m]} e^{\Omega(n)-100\log\log T} \\ &\ll (\log T)^{-99}\prod_{p\leq T}\Big(1+\frac{(1+e)|\mathfrak{F}_{-1,\beta/2-1,\theta}(p)|}{p}+\frac{e\cdot|\mathfrak{F}_{-1,\beta/2-1,\theta}(p)|^2}{p}\Big) \\ &\ll (\log T)^{-99}\cdot(\log T)^{1+2e} = o(1), \end{aligned} \tag{6.2.12}$$

where we used (6.2.11) and the fact that $|\mathfrak{F}_{-1,\beta/2-1,\theta}(p)| \leq 1$ for $0 < \beta \leq 2$. The contribution of the sum over m with $\Omega(m) > 100\lfloor \log\log T \rfloor$ can be removed in the same manner.

Considering the sums in (6.2.10) without the restriction on $\Omega(n)$ and $\Omega(m)$, we get by (6.2.11) and Lemma 6.4.1,

$$\begin{aligned}\sum_{\substack{p|n \implies \log p \leq \log^{1-\varepsilon} T \\ p|m \implies \log p \leq \log^{1-\varepsilon} T}} \frac{\mathfrak{F}_{-1,\beta/2-1,\theta}(nm)\mathfrak{g}(n)\mathfrak{g}(m)}{[n,m]} &\asymp \prod_{\log p \leq \log^{1-\varepsilon} T} \Big(1 + \frac{2\mathfrak{F}_{-1,\beta/2-1,\theta}(p) + \mathfrak{F}_{-1,\beta/2-1,\theta}(p)^2}{p}\Big) \\ &\asymp (\log T)^{-|\theta|} \cdot (\log T)^{(\beta^2/4-1)\cdot(1+\theta-\varepsilon)} \\ &\ll (\log T)^{\beta^2(1+\theta)/4-1+\varepsilon}. \end{aligned} \tag{6.2.13}$$

To evaluate the remaining part of the sum, write

$$\log\Big(\frac{(m,n)^2}{mn}\Big) = \frac{1}{2\pi i}\oint_{|z|=1/\log T} \Big(\frac{(m,n)^2}{mn}\Big)^z \cdot \frac{dz}{z^2}. \tag{6.2.14}$$

Then, we end up having to evaluate

$$\frac{1}{2\pi i}\oint_{|z|=1/\log T} \sum_{\substack{p|n \implies \log p \leq \log^{1-\varepsilon} T \\ p|m \implies \log p \leq \log^{1-\varepsilon} T}} \frac{\mathfrak{F}_{-1,\beta/2-1,\theta}(mn)\mathfrak{g}(m)\mathfrak{g}(n)}{[m,n]} \cdot \Big(\frac{(m,n)^2}{mn}\Big)^z \cdot \frac{dz}{z^2}. \tag{6.2.15}$$

As above, the sum over m and n factors into an Euler product which is

$$\asymp \prod_{\log p \leq \log^{1-\varepsilon} T} \Big(1 + \frac{2\mathfrak{F}_{-1,\beta/2-1,\theta}(p)}{p^{1+z}} + \frac{\mathfrak{F}_{-1,\beta/2-1,\theta}(p)^2}{p} + O(p^{-2+2|z|})\Big). \tag{6.2.16}$$

For $|z| = 1/\log T$, note that a Taylor expansion yields

$$\sum_{\log p \leq \log^{1-\varepsilon} T} \frac{\mathfrak{F}_{-1,\beta/2-1,\theta}(p)}{p^{1+z}} = \sum_{\log p \leq \log^{1-\varepsilon} T} \frac{\mathfrak{F}_{-1,\beta/2-1,\theta}(p)}{p} + O\Big(\frac{1}{\log T}\sum_{\log p \leq \log^{1-\varepsilon} T} \frac{\log p}{p}\Big),$$

and since the above error term is $o(1)$ by Lemma 6.4.1, the Euler product in (6.2.16) is

$$\asymp \prod_{\log p \leq \log^{1-\varepsilon} T} \Big(1 + \frac{2\mathfrak{F}_{-1,\beta/2-1,\theta}(p) + \mathfrak{F}_{-1,\beta/2-1,\theta}(p)^2}{p}\Big) \ll (\log T)^{\beta^2(1+\theta)/4-1+\varepsilon}. \tag{6.2.17}$$

Therefore, by putting this back in the contour integral and using a trivial bound on z^2, (6.2.15) is $\ll (\log T)^{\beta^2(1+\theta)/4+\varepsilon}$ as required. □

Lemma 6.2.5. *Let $\varepsilon > 0$ be given. For $\ell = 50\lfloor \log\log T \rfloor$, we have*

$$\mathbb{E}\Big[|\zeta(\tfrac{1}{2}+i\tau)|^2 \cdot \Big|\frac{\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)}{100\log\log T}\Big|^{2\ell}\Big] \ll (\log T)^{-21}, \tag{6.2.18}$$

and

$$\mathbb{E}\Big[\Big|\frac{\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)}{100\log\log T}\Big|^{2\ell}\Big] \ll (\log T)^{-21}. \tag{6.2.19}$$

Proof. For (6.2.18), we apply the Cauchy-Schwarz inequality, a fourth moment bound on zeta, and a moment estimate (Lemma 6.4.4) followed by a prime number theorem estimate (Lemma 6.4.1) on the remaining term to conclude that the expectation is

$$\ll (\log T)^2 \cdot \mathbb{E}\Big[\Big|\frac{\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)}{100\log\log T}\Big|^{4\ell}\Big]^{1/2} \ll (\log T)^2 \cdot (\log T)^{-50}. \tag{6.2.20}$$

The proof of (6.2.19) is even more straightforward. □

The last three lemmas show a moment bound of the right order for $\zeta \cdot e^{-\mathcal{P}_{|\theta|}}$.

Proposition 6.2.6. *Let $-1 < \theta < 0$, $0 < \beta \leq 2$ and $\varepsilon > 0$ be given. Then, as $T \to \infty$,*

$$\mathbb{E}\Big[\Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau)\Big|^{\beta} \mathbf{1}_{\mathcal{A}(T)}\Big] \ll (\log T)^{\beta^2(1+\theta)/4+\varepsilon}, \tag{6.2.21}$$

with the event

$$\mathcal{A}(T) = \Big\{|\mathcal{P}_{|\theta|}(\tfrac{1}{2}+i\tau)| \leq 2\log\log T\Big\}. \tag{6.2.22}$$

Proof. Let $0 < \beta < 2$. By Young's inequality with $p = 2/\beta$ and $q = 2/(2-\beta)$,

$$\begin{aligned} \Big|\zeta(\tfrac{1}{2}+i\tau)\Big|^{\beta} &\leq \tfrac{1}{p} \cdot |\zeta(\tfrac{1}{2}+i\tau)|^2 \cdot e^{-\frac{2}{q}\mathrm{Re}\,\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)} + \tfrac{1}{q} \cdot e^{\frac{2}{p}\mathrm{Re}\,\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)} \\ &= \tfrac{\beta}{2} \cdot |\zeta(\tfrac{1}{2}+i\tau)|^2 \cdot e^{-(2-\beta)\mathrm{Re}\,\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)} + \tfrac{2-\beta}{2} \cdot e^{\beta\mathrm{Re}\,\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)}. \end{aligned} \tag{6.2.23}$$

Note that (6.2.23) holds trivially for $\beta = 2$. Hence, for $0 < \beta \leq 2$,

$$\begin{aligned} \Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau)\Big|^{\beta} \leq{}& \tfrac{\beta}{2}\, |\zeta(\tfrac{1}{2}+i\tau)|^2 \cdot e^{-(2-\beta)\mathrm{Re}\,\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)-\beta\mathrm{Re}\,\mathcal{P}_{|\theta|}(\frac{1}{2}+i\tau)} \\ &+ \tfrac{2-\beta}{2}\, e^{\beta\mathrm{Re}\,\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)-\beta\mathrm{Re}\,\mathcal{P}_{|\theta|}(\frac{1}{2}+i\tau)}. \end{aligned} \tag{6.2.24}$$

On the event $\mathcal{A}(T) \cap \{|\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)| \leq 100 \log\log T\}$, we get by (6.2.5) that

$$e^{-(2-\beta)\mathrm{Re}\,\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)-\beta \mathrm{Re}\,\mathcal{P}_{|\theta|}(\frac{1}{2}+i\tau)} \ll \Bigg| \sum_{\substack{\Omega(n)\leq 100\lfloor \log\log T\rfloor \\ p|n \implies \log p \leq \log^{1-\varepsilon} T}} \frac{\mathfrak{F}_{-1,\beta/2-1,\theta}(n)\mathfrak{g}(n)}{n^{1/2+i\tau}} \Bigg|^2 \tag{6.2.25}$$

where $\mathfrak{F}_{\alpha,\beta,\theta}(n)$ is the completely multiplicative function defined in (6.2.6). Likewise, on the same event, we have

$$e^{\beta\mathrm{Re}\,\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)-\beta \mathrm{Re}\,\mathcal{P}_{|\theta|}(\frac{1}{2}+i\tau)} \ll \Bigg| \sum_{\substack{\Omega(n)\leq 100\lfloor \log\log T\rfloor \\ p|n \implies \log p \leq \log^{1-\varepsilon} T}} \frac{\mathfrak{F}_{0,\beta/2,\theta}(n)\mathfrak{g}(n)}{n^{1/2+i\tau}} \Bigg|^2 \tag{6.2.26}$$

Finally, on the event $\mathcal{A}(T) \cap \{|\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)| > 100 \log\log T\}$, we get, for any $\ell \geq 1$,

$$\Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau)\Big|^\beta \leq (\log T)^4 \cdot (1+|\zeta(\tfrac{1}{2}+i\tau)|^2) \cdot \left|\frac{\mathcal{P}_{1-\varepsilon}(\frac{1}{2}+i\tau)}{100\log\log T}\right|^{2\ell}, \tag{6.2.27}$$

since for $\beta \leq 2$, $|\zeta|^\beta$ is bounded by $(1+|\zeta|^2)$ and $e^{-\beta \mathrm{Re}\,\mathcal{P}_{|\theta|}}$ is bounded by $(\log T)^4$ on $\mathcal{A}(T)$. We choose $\ell = 50\lfloor \log\log T\rfloor$. Now, take the expectation with τ restricted to $\mathcal{A}(T)$ in (6.2.24), then split the terms on the right-hand side over the associated events in (6.2.25), (6.2.26) and (6.2.27). We use Lemmas 6.2.3, 6.2.4 and 6.2.5 to bound the expectations. $\square$

6.2.2. Discretization

The analysis of the maximum of zeta on an interval can often be reduced to the analysis on a discrete set of points at a distance of roughly $(\log T)^{-1}$ of each other. This can be proved for the maximum using the functional equation for zeta, see for example Lemma 2.2 in Farmer et al. (2007). We will need a more elaborate variant for general Dirichlet polynomials.

Proposition 6.2.7. *Let $\theta > -1$, $\beta \geq 1$, and $\varepsilon > 0$ be given. Let D be a Dirichlet polynomial of length $T^{1+\varepsilon}$. Then, for all $A > 0$, $T \leq t \leq 2T$, and $\sigma \geq 1/2$,*

$$\begin{aligned} \sup_{|h|\leq \log^\theta T} |D(\sigma+it+ih)|^\beta \ll_A &\sum_{|k|\leq 2\log^{1+\theta} T} \Big|D\Big(\sigma+it+\tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\Big|^\beta \\ &+ \sum_{|k|> 2\log^{1+\theta} T} \Big|D\Big(\sigma+it+\tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\Big|^\beta \cdot \frac{1}{1+|k|^A}. \end{aligned} \tag{6.2.28}$$

Proof. Let V be a smooth compactly supported function with $V(x) = 1$ for $0 \le x \le 1+\varepsilon$ and compactly supported in $[-\varepsilon, 1 + 2\varepsilon]$. We show

$$|D(\sigma + it + ih)|^{\beta} \ll \sum_{k \in \mathbb{Z}} \Big| D\Big(\sigma + it + \tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\Big|^{\beta} \cdot \Big| \widehat{V}\Big(\tfrac{k}{2+3\varepsilon} - \tfrac{h \log T}{2\pi}\Big)\Big|. \tag{6.2.29}$$

Taking a supremum over $|h| \le \log^{\theta} T$, and using the rapid decay of $\widehat{V}$, we get (6.2.28).

Let $G(x) = V(2\pi x/\log T)$, so that $G\big(\frac{1}{2\pi}\log n\big) = 1$ for $1 \le n \le T^{1+\varepsilon}$. We have $\widehat{G}(x) = \frac{\log T}{2\pi} \cdot \widehat{V}\big(\frac{x \log T}{2\pi}\big)$. By the Paley-Wiener theorem (see for example Theorem IX.11 in Reed and Simon (1972)), uniformly in $T \le t \le 2T$ and $|h| \le \log^{\theta} T$, we have

$$|D(\sigma + it + ih + ix)\widehat{G}(x)| \ll \exp((2 + 3\varepsilon)\log T \cdot |x|), \quad x \in \mathbb{C}. \tag{6.2.30}$$

Now, consider

$$\sum_{k \in \mathbb{Z}} D\Big(\sigma + it + \tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\widehat{G}\Big(\tfrac{2\pi k}{(2+3\varepsilon)\log T} - h\Big). \tag{6.2.31}$$

By the Poisson summation formula, the above is equal to

$$\sum_{\ell \in \mathbb{Z}} \int_{\mathbb{R}} D\Big(\sigma + it + \tfrac{2\pi i x}{(2+3\varepsilon)\log T}\Big)\widehat{G}\Big(\tfrac{2\pi x}{(2+3\varepsilon)\log T} - h\Big)e^{-2\pi i \ell x}dx. \tag{6.2.32}$$

By a change of variable, this is equal to

$$(2 + 3\varepsilon) \cdot \frac{\log T}{2\pi} \sum_{\ell \in \mathbb{Z}} \int_{\mathbb{R}} D\Big(\sigma + it + ix\Big)\widehat{G}(x - h)e^{-i\ell x(2+3\varepsilon)\log T}dx. \tag{6.2.33}$$

Using (6.2.30), all the terms with $\ell \neq 0$ in (6.2.33) are equal to zero. The term $\ell = 0$ is equal to $D(\sigma + it + ih)$ since $G(\frac{1}{2\pi}\log n) = 1$ for $1 \le n \le T^{1+\varepsilon}$. It follows that

$$D(\sigma + it + ih) = \frac{1}{2 + 3\varepsilon} \sum_{k \in \mathbb{Z}} D\Big(\sigma + it + \tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\widehat{V}\Big(\tfrac{k}{2+3\varepsilon} - \tfrac{h \log T}{2\pi}\Big). \tag{6.2.34}$$

Taking absolute values and applying Hölder's inequality with $\beta \ge 1$, we obtain

$$\begin{aligned} |D(\sigma + it + ih)| \le {} & \Big(\frac{1}{2 + 3\varepsilon} \sum_{k \in \mathbb{Z}} \Big| D\Big(\sigma + it + \tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\Big|^{\beta} \cdot \Big| \widehat{V}\Big(\tfrac{k}{2+3\varepsilon} - \tfrac{h \log T}{2\pi}\Big)\Big|\Big)^{1/\beta} \\ & \times \Big(\frac{1}{2 + 3\varepsilon} \sum_{k \in \mathbb{Z}} \Big| \widehat{V}\Big(\tfrac{k}{2+3\varepsilon} - \tfrac{h \log T}{2\pi}\Big)\Big|\Big)^{1-1/\beta}. \end{aligned} \tag{6.2.35}$$

This proves (6.2.29) using the rapid decay of $\widehat{V}$. $\square$

Proposition 6.2.7 implies five important corollaries to tackle the maximum of ζ and of Dirichlet polynomials. We first observe that the discretization applies to ζ.

Corollary 6.2.8. *Let $\theta > -1$, $\beta \geq 1$ and $\varepsilon > 0$ be given. Then, for any $A, B > 0$ and all $T \leq t \leq 2T$,*

$$\begin{aligned}\max_{|h| \leq \log^\theta T} &|\zeta(\tfrac{1}{2} + it + ih)|^\beta \\ &\ll_{A,B} \sum_{|k| \leq 2\log^{1+\theta} T} \Big|\zeta\Big(\tfrac{1}{2} + it + \tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\Big|^\beta \\ &+ \sum_{|k| > 2\log^{1+\theta} T} \Big|\zeta\Big(\tfrac{1}{2} + it + \tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\Big|^\beta \cdot \frac{1}{1+|k|^A} + T^{-B}.\end{aligned} \tag{6.2.36}$$

Proof. From Proposition 2 in Bombieri and Friedlander (1995), we have, for any $A > 0$,

$$\sum_{n \leq T} \frac{1}{n^{1/2+it}} \cdot \Big(1 - \frac{n}{T}\Big)^A = \zeta(\tfrac{1}{2} + it) + O_A(T^{-A/2}). \tag{6.2.37}$$

We apply Proposition 6.2.7 to conclude. $\square$

As a consequence, we get a suboptimal upper bound using the second moment.

Corollary 6.2.9. *For any $A \geq 0$,*

$$\mathbb{P}\Big(\max_{|h| \leq \log^A T} |\zeta(\tfrac{1}{2} + i\tau + ih)| > 2^A (\log T)^{2+A}\Big) \ll \frac{1}{\log T} \cdot 2^{-A}. \tag{6.2.38}$$

Proof. Using the integral representation for ζ on the critical strip, we certainly know that $\zeta(\frac{1}{2}+it) = O(1+|t|)$ for all t (see for example (2.12.2) in Titchmarsh (1986)), which means that (6.2.38) is trivially satisfied when $A > \log T/\log\log T$. Therefore, assume

$$A \leq \log T/\log\log T.$$

By applying Chebyshev's inequality and Corollary 6.2.8, the probability in (6.2.38) is bounded above by

$$
\begin{aligned}
&2^{-2A}(\log T)^{-4-2A}\,\mathbb{E}\Big[\max_{|h|\le \log^A T}|\zeta(\tfrac12+i\tau+ih)|^2\Big]\\
&\ll 2^{-2A}(\log T)^{-4-2A}\sum_{|k|\le 2\log^{1+A}T}\mathbb{E}\Big[\Big|\zeta\Big(\tfrac12+i\tau+\tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\Big|^2\Big] \qquad (6.2.39)\\
&+2^{-2A}(\log T)^{-4-2A}\sum_{|k|>2\log^{1+A}T}\mathbb{E}\Big[\Big|\zeta\Big(\tfrac12+i\tau+\tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\Big|^2\Big]\cdot\frac{1}{1+|k|^{100}}+T^{-101},
\end{aligned}
$$

for any fixed $\varepsilon > 0$. The first expectation is $\ll \log T$ by using a standard second moment bound. We bound the second expectation by enlarging the integration to $|t| \le T|k|$ and then applying the second moment bound, i.e.

$$
\mathbb{E}\Big[\Big|\zeta\Big(\tfrac12+i\tau+\tfrac{2\pi i k}{(2+3\varepsilon)\log T}\Big)\Big|^2\Big] \le |k|\cdot\frac{1}{T|k|}\int_{|t|\le T|k|}\Big|\zeta(\tfrac12+it)\Big|^2 dt \ll |k|\cdot\log(T|k|). \qquad (6.2.40)
$$

We conclude that the right-hand side of (6.2.39) is, by the assumption on A,

$$
\ll 2^{-2A}(\log T)^{-2-A} \ll \frac{1}{\log T}\cdot 2^{-A}. \qquad (6.2.41)
$$

□

A similar reasoning using Markov's inequality can be applied to get a suboptimal upper bound for the maximum of $\mathcal{P}_\alpha$, $0<\alpha<1$.

Corollary 6.2.10. *For any $\theta > -1$, $\varepsilon > 0$ and $\sigma \ge 1/2$,*

$$
\mathbb{P}\Big(\max_{|h|\le \log^\theta T}|\mathcal{P}_\alpha(\sigma+i\tau+ih)| > (\sqrt{\alpha(1+\theta)}+\varepsilon)\log\log T\Big) = o(1). \qquad (6.2.42)
$$

Proof. Apply Markov's inequality with exponent $2k$, discretize as in (6.2.39) using Proposition 6.2.7, and then use Lemma 6.4.4 with $k = \lfloor (1+\theta)\log\log T\rfloor$ to bound the expectations. □

The same bound holds trivially for the maximum of $\operatorname{Re}\mathcal{P}_\alpha$ since $|\operatorname{Re}\mathcal{P}_\alpha| \le |\mathcal{P}_\alpha|$. This is sharp for $\theta > 0$. For $\theta < 0$ and $\alpha > |\theta|$, this bound (and the bound for ζ) needs to be refined by discarding the contribution of small primes. The result below directly implies that for $\theta < 0$ and $\alpha > |\theta|$, the sharp upper bound for $\operatorname{Re}\mathcal{P}_\alpha$ is $\sqrt{(\alpha+\theta)(1+\theta)}\log\log T$ since the effective variance is $\frac{(\alpha+\theta)}{2}\log\log T$.

Corollary 6.2.11. *Let $-1 < \theta < 0$ and $\sigma \geq 1/2$. For any $0 < \varepsilon < C$ and $V = V(T)$ that satisfies $\varepsilon \log\log T \leq V \leq C \log\log T$, we have*

$$\mathbb{P}\Big(\max_{|h| \leq \log^\theta T} \Big| \mathcal{P}_{|\theta|}(\sigma + i\tau + ih) \Big| > V \Big) \ll e^{-cV}, \tag{6.2.43}$$

for some constant $c = c(\varepsilon, C) > 0$.

Proof. For a lighter notation, write $S(h) = \mathcal{P}_{|\theta|}(\sigma + i\tau + ih)$. (We keep the dependence on τ implicit, consistent with the probabilistic notation for random variables.) We have

$$\begin{aligned} \mathbb{P}\Big(\max_{|h| \leq \log^\theta T} |S(h)| > V \Big) \leq \mathbb{P}\Big(&\max_{|h| \leq \log^\theta T} |S(h) - S(0)| > V/2 \Big) \\ &+ \mathbb{P}\Big(|S(0)| > V/2 \Big). \end{aligned} \tag{6.2.44}$$

Let ℓ denote a generic natural integer. By Chebyshev's inequality, a moment estimate (Lemma 6.4.4) and a prime number theorem estimate (Lemma 6.4.1), we have

$$\mathbb{P}\Big(|S(0)| > V/2 \Big) \leq \frac{\mathbb{E}\Big[|S(0)|^{2\ell} \Big]}{(V/2)^{2\ell}} \ll \ell! \left(\frac{\sum_{p \leq T} p^{-2\sigma}}{(V/2)^2} \right)^\ell \ll \left(\frac{4\ell \log\log T}{\varepsilon^2 (\log\log T)^2} \right)^\ell. \tag{6.2.45}$$

With the choice $\ell = \lfloor \frac{\varepsilon^2}{8} \log\log T \rfloor$, this probability is $\ll \exp(-aV)$ for some constant $a = a(\varepsilon, C) > 0$.

It remains to control the first probability on the right-hand side of (6.2.44). Let ℓ denote another natural integer to be chosen later. By applying Proposition 6.2.7, we get

$$\mathbb{E}\Big[\max_{|h| \leq \log^\theta T} |S(h) - S(0)|^{2\ell} \Big] \ll \log^{1+\theta} T \cdot \mathbb{E}\Big[|S(h) - S(0)|^{2\ell} \Big] \tag{6.2.46}$$

A short calculation, using moment estimates (Lemma 6.4.4) followed by prime number theorem estimates (Lemma 6.4.1), gives, for all $|h| \leq \log^\theta T$,

$$\mathbb{E}\Big[|S(h) - S(0)|^{2\ell} \Big] \ll \ell! \left(\sum_{\log p \leq \log^{|\theta|} T} \frac{2 - 2\cos(|h| \log p)}{p} \right)^\ell \ll (\ell\, c)^\ell, \tag{6.2.47}$$

for some constant $c > 0$ (to obtain the last inequality, note that $|h| \cdot \log^{|\theta|} T \leq 1$).

Then, by Chebyshev's inequality and the choice $\ell = \lfloor \frac{\varepsilon^2}{8c} \log\log T \rfloor$, we deduce

$$\mathbb{P}\Big(\max_{|h| \le \log^\theta T} |S(h) - S(0)| > V/2 \Big) \ll \left(\frac{4\,\ell\, c}{V^2} \right)^{\ell} \ll e^{-bV}, \tag{6.2.48}$$

for some constant $b = b(\varepsilon, C) > 0$. □

As before the maximum of $\zeta \cdot e^{-\mathcal{P}_{|\theta|}}$ can be discretized by truncating the exponential.

Corollary 6.2.12. *Let $0 \ge \theta > -1$ and $\varepsilon > 0$ be given. Then, the event*

$$\begin{aligned} &\max_{|h| \le \log^\theta T} \Big| (\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2} + i\tau + ih) \Big|^2 \\ &\quad \ll \sum_{|k| \le 2\log^{1+\theta} T} \Big| (\zeta \cdot e^{-\mathcal{P}_{|\theta|}}) \Big(\tfrac{1}{2} + i\tau + \tfrac{2\pi i k}{(2+3\varepsilon)\log T} \Big) \Big|^2 + o(1) \end{aligned} \tag{6.2.49}$$

has probability $1 - o(1)$.

Proof. Define the event

$$\widetilde{\mathcal{A}}(T) = \Big\{ \max_{|h| \le \log^\theta T} \Big| \mathcal{P}_{|\theta|}(\tfrac{1}{2} + i\tau + ih) \Big| \le 2 \log\log T \Big\}. \tag{6.2.50}$$

By Corollary 6.2.11, we have $\mathbb{P}(\widetilde{\mathcal{A}}(T)) = 1 - o(1)$. By (6.2.5), for all $\tau \in \widetilde{\mathcal{A}}(T)$,

$$\begin{aligned} \Big| \sum_{\substack{\Omega(n) \le 20\lfloor \log\log T \rfloor \\ p | n \implies \log p \le \log^{|\theta|} T}} \frac{(-1)^{\Omega(n)} \mathfrak{g}(n)}{n^{1/2 + i\tau + ih}} \Big| &= \Big| e^{-\mathcal{P}_{|\theta|}(\frac{1}{2} + i\tau + ih)} \Big| + O\Big((\log T)^{-20} \Big) \\ &\asymp \Big| e^{-\mathcal{P}_{|\theta|}(\frac{1}{2} + i\tau + ih)} \Big|. \end{aligned} \tag{6.2.51}$$

Combining this with (6.2.37), we conclude that, for all $\tau \in \widetilde{\mathcal{A}}(T)$,

$$\Big| (\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2} + i\tau + ih) \Big| \asymp |D(\tfrac{1}{2} + i\tau + ih)| + o(1), \tag{6.2.52}$$

with D a Dirichlet polynomial of length $\ll T^{1+\varepsilon}$ for every fixed $\varepsilon > 0$. Proposition 6.2.7 implies

$$\max_{|h| \le \log^\theta T} |D(\tfrac{1}{2} + i\tau + ih)|^2 \ll \sum_{|k| \le 2\log^{1+\theta} T} \Big| D\Big(\tfrac{1}{2} + i\tau + \tfrac{2\pi i k}{(2+3\varepsilon)\log T} \Big) \Big|^2 + o(1). \tag{6.2.53}$$

Together with (6.2.52), this concludes the proof. □

6.2.3. Proofs of the upper bounds

6.2.3.1. *The case of $\theta \geq 0$*

Proof of Theorem 6.1.2 for $\theta \geq 0$. By Markov's inequality, for any $\beta > 0$, we have

$$\begin{aligned} &\mathbb{P}\Big(\max_{|h|\leq \log^\theta T} |\zeta(\tfrac{1}{2}+i\tau+ih)| > (\log T)^{m(\theta)+\varepsilon}\Big) \\ &\qquad \ll (\log T)^{-\beta m(\theta)-\beta\varepsilon} \cdot \mathbb{E}\Big[\max_{|h|\leq \log^\theta T} |\zeta(\tfrac{1}{2}+i\tau+ih)|^\beta\Big]. \end{aligned} \tag{6.2.54}$$

For $\beta > 1$, we get, by picking A large enough in Corollary 6.2.8, that the right-hand side of the above equation is

$$\ll (\log T)^{-\beta m(\theta)-\beta\varepsilon+1+\theta} \cdot \mathbb{E}\Big[|\zeta(\tfrac{1}{2}+i\tau)|^\beta\Big]. \tag{6.2.55}$$

(The sum on large k's is handled as in (6.2.40).) By applying Proposition 6.2.2 if $\theta \leq 3$ and Proposition 6.2.1 if $\theta > 3$, the expectation is bounded by $(\log T)^{\beta^2/4+\beta\varepsilon/2}$. The optimal bound is at $\beta = 2m(\theta) > 1$. Therefore, the claim follows. □

Proof of Theorem 6.1.1 for $\theta \geq 0$. For all $\beta > 0$, Markov's inequality yields the bound

$$\begin{aligned} &\mathbb{P}\Big(\int_{|h|\leq \log^\theta T} |\zeta(\tfrac{1}{2}+i\tau+ih)|^\beta dt \geq (\log T)^{f_\theta(\beta)+\varepsilon}\Big) \\ &\qquad \ll (\log T)^{-f_\theta(\beta)-\varepsilon} \log^\theta T \cdot \mathbb{E}\Big[|\zeta(\tfrac{1}{2}+i\tau)|^\beta\Big]. \end{aligned} \tag{6.2.56}$$

When $\beta \leq 2\sqrt{1+\theta}$, we have $f_\theta(\beta) = \beta^2/4+\theta$, so the right-hand side of (6.2.56) is $\ll (\log T)^{-\varepsilon/2}$ by Proposition 6.2.2 for $\theta \leq 3$ and by Proposition 6.2.1 for $\theta > 3$.

It remains to sharpen the bound in the case $\beta > 2\sqrt{1+\theta}$. We use the Lebesgue measure of high points. Let $a, b > 0$. Two successive applications of Markov's inequality yield

$$\begin{aligned} &\mathbb{P}\Big(\mathrm{Leb}\Big\{|h| \leq \log^\theta T : |\zeta(\tfrac{1}{2}+i\tau+ih)| > (\log T)^a\Big\} \geq (\log T)^{-a^2+\theta+\varepsilon}\Big) \\ &\ll (\log T)^{a^2-\varepsilon} \cdot (\log T)^{-ba} \cdot \mathbb{E}\Big[|\zeta(\tfrac{1}{2}+i\tau)|^b\Big]. \end{aligned} \tag{6.2.57}$$

Again, the optimal bound is at $b = 2a$. Using Proposition 6.2.2 for $\theta \leq 3$ and Proposition 6.2.1 for $\theta > 3$ and choosing $b = 2a$, we conclude that this is $\ll (\log T)^{-\varepsilon/2}$ for $0 < a \leq m(\theta)$.

We now partition the integral according to the value of the integrand. Let $M \geq 1$ be an integer and $0 \leq j \leq M$. Theorem 6.1.2 and the above imply that, with probability $1-o(1)$,

$$\int_{|h|\leq \log^\theta T} |\zeta(\tfrac{1}{2}+i\tau+ih)|^\beta dh \ll \sum_{0\leq j\leq M} (\log T)^{\beta((j+1)/M)m(\theta)} \cdot (\log T)^{-(j/M)^2 m(\theta)^2+\theta+\varepsilon}. \quad (6.2.58)$$

For $\beta > 2\sqrt{1+\theta} \geq 2m(\theta)$, the last term $j = M$ dominates and, in particular, the above is bounded by

$$\ll (\log T)^{\beta m(\theta)-m(\theta)^2+\theta+2\varepsilon} = (\log T)^{\beta m(\theta)-1+2\varepsilon}, \quad (6.2.59)$$

provided that M is chosen sufficiently large. $\square$

6.2.3.2. *The case of $\theta < 0$*

Proof of Theorem 6.1.2 for $\theta < 0$. We notice that

$$\begin{aligned} &\mathbb{P}\Big(\max_{|h|\leq \log^\theta T} |\zeta(\tfrac{1}{2}+i\tau+ih)| > (\log T)^{m(\theta)+\varepsilon}\Big) \\ &\quad \leq \mathbb{P}\Big(\max_{|h|\leq \log^\theta T} \Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau+ih)\Big| > (\log T)^{m(\theta)+\varepsilon/2}\Big) \\ &\qquad + \mathbb{P}\Big(\max_{|h|\leq \log^\theta T} \Big|e^{\mathcal{P}_{|\theta|}(\frac{1}{2}+i\tau+ih)}\Big| > (\log T)^{\varepsilon/2}\Big). \end{aligned} \quad (6.2.60)$$

By Corollary 6.2.11, the last term is $o(1)$ as $T\to\infty$. As in (6.2.50), let

$$\widetilde{\mathcal{A}}(T) = \Big\{\max_{|h|\leq \log^\theta T} \Big|\mathcal{P}_{|\theta|}(\tfrac{1}{2}+i\tau+ih)\Big| \leq 2\log\log T\Big\}. \quad (6.2.61)$$

By Corollary 6.2.11 again, the probability of $\widetilde{\mathcal{A}}(T)$ is $1-o(1)$. We let $\mathcal{A}_0(T)$ denote the subset of $\widetilde{\mathcal{A}}(T)$ for which the conclusion of Corollary 6.2.12 holds. The probability of $\mathcal{A}_0(T)$ is $1-o(1)$. Then, by Chebyshev's inequality, we have

$$\begin{aligned} &\mathbb{P}\Big(\Big\{\max_{|h|\leq \log^\theta T} \Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau+ih)\Big| > (\log T)^{m(\theta)+\varepsilon/2}\Big\} \cap \mathcal{A}_0(T)\Big) \\ &\quad \leq (\log T)^{-2m(\theta)-\varepsilon} \cdot \mathbb{E}\Big[\max_{|h|\leq \log^\theta T} \Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau+ih)\Big|^2 \mathbf{1}_{\mathcal{A}_0(T)}\Big]. \end{aligned} \quad (6.2.62)$$

By Corollary 6.2.12, and since $m(\theta) = 1+\theta$, this is

$$\ll (\log T)^{-(1+\theta)-\varepsilon} \cdot \mathbb{E}\Big[\Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2}+i\tau)\Big|^2 \mathbf{1}_{\widetilde{\mathcal{A}}(T)}\Big]. \quad (6.2.63)$$

By Proposition 6.2.6, this is

$$\ll (\log T)^{-(1+\theta)-\varepsilon} \cdot (\log T)^{(1+\theta)+\varepsilon/2} \ll (\log T)^{-\varepsilon/2}, \tag{6.2.64}$$

as needed. □

Proof of Theorem 6.1.1 for $\theta < 0$. Similarly to (6.2.60), we can restrict to $\zeta \cdot e^{-\mathcal{P}_{|\theta|}}$ as follows

$$\begin{aligned} &\mathbb{P}\Big(\int_{|h| \le \log^\theta T} |\zeta(\tfrac{1}{2} + i\tau + ih)|^\beta dh > (\log T)^{f_\theta(\beta)+\varepsilon} \Big) \\ &\quad \le \mathbb{P}\Big(\int_{|h| \le \log^\theta T} \Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2} + i\tau + ih)\Big|^\beta dh > (\log T)^{f_\theta(\beta)+\varepsilon/2} \Big) + o(1). \end{aligned} \tag{6.2.65}$$

As in (6.2.61), $\mathbb{P}(\widetilde{\mathcal{A}}(T)) = 1 - o(1)$, and by Markov's inequality, we have

$$\begin{aligned} &\mathbb{P}\Big(\Big\{ \int_{|h| \le \log^\theta T} \Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2} + i\tau + ih)\Big|^\beta dh > (\log T)^{f_\theta(\beta)+\varepsilon/2} \Big\} \cap \widetilde{\mathcal{A}}(T) \Big) \\ &\quad \ll (\log T)^{-f_\theta(\beta)-\varepsilon/2} \cdot \log^\theta T \cdot \mathbb{E}\Big[\Big|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2} + i\tau)\Big|^\beta \mathbf{1}_{\widetilde{\mathcal{A}}(T)} \Big]. \end{aligned} \tag{6.2.66}$$

By Proposition 6.2.6, the above is

$$\ll (\log T)^{-(\beta^2/4)(1+\theta)-\varepsilon/2} \cdot (\log T)^{(\beta^2/4)\cdot(1+\theta)+\varepsilon/4} \ll (\log T)^{-\varepsilon/4}. \tag{6.2.67}$$

This bound proves the claim for $\beta \le 2$.

It remains to refine the bound for the case $\beta > 2$. This proceeds in the same way as in the proof of Theorem 6.1.1 in the case of $\theta \ge 0$, with ζ replaced by $\zeta \cdot e^{-\mathcal{P}_{|\theta|}}$ restricted on the event $\widetilde{\mathcal{A}}(T)$. Namely, we have, for $0 < a \le m(\theta)$,

$$\begin{aligned} &\mathbb{P}\Big(\Big\{ \mathrm{Leb}\Big\{ |h| \le \log^\theta T : |(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2} + i\tau + ih)| > (\log T)^a \Big\} \ge (\log T)^{-a^2+\theta+\varepsilon} \Big\} \cap \widetilde{\mathcal{A}}(T) \Big) \\ &\ll (\log T)^{a^2-\varepsilon} \cdot (\log T)^{-ba} \cdot \mathbb{E}\Big[|(\zeta \cdot e^{-\mathcal{P}_{|\theta|}})(\tfrac{1}{2} + i\tau)|^b \mathbf{1}_{\widetilde{\mathcal{A}}(T)} \Big]. \end{aligned} \tag{6.2.68}$$

This is $o(1)$ by Proposition 6.2.6 with the optimal choice $b = 2a/(1+\theta) \le 2$. The remainder is done exactly as in the proof of Theorem 6.1.1 in the case of $\theta \ge 0$, by partitioning the integral over values of the integrand on the range $[0, m(\theta) + \varepsilon]$. □

6.3. Lower bounds

In this section, we prove:

Proposition 6.3.1. *Let $\theta > -1$ and $\varepsilon > 0$ be given. Then,*

$$\mathbb{P}\Big(\max_{|h|\leq \log^\theta T} |\zeta(1/2 + i\tau + ih)| > (\log T)^{m(\theta)-\varepsilon}\Big) = 1 - o(1). \tag{6.3.1}$$

Proposition 6.3.2. *Let $\theta > -1$, $\beta > 0$ and $\varepsilon > 0$ be given. Then,*

$$\mathbb{P}\Big(\int_{-\log^\theta T}^{\log^\theta T} |\zeta(1/2 + i\tau + ih)|^\beta dh > (\log T)^{f_\theta(\beta)-\varepsilon}\Big) = 1 - o(1). \tag{6.3.2}$$

The lower bound for the maximum will be an easy consequence of the lower bound for the moments. The idea is to approximate zeta by an appropriate Dirichlet polynomial. This can be done with good precision off-axis, cf. Section 6.3.1. The approximation to a Dirichlet polynomial is then shown in Section 6.3.2. The lower bound for the moments of the Dirichlet polynomials is proved in Section 6.3.3 using Kistler's multiscale second moment method. Finally, the two propositions above are proved in Section 6.3.4.

6.3.1. Reduction off-axis

In Arguin et al. (2019), the maximum on a short interval of the critical line was compared to the one on a short interval away from the critical line by exploiting the analyticity of ζ away from its pole. More precisely, a value off-axis can be seen as an average of zeta over the critical line weighed by the corresponding Poisson kernel. This approach could also be used in the case of the moments by using the subharmonicity of the function $z \mapsto |z|^\beta$. We choose to apply a different method based on the following convexity theorem of Gabriel, which handles error terms more efficiently.

Proposition 6.3.3 (Theorem 2 of Gabriel (1927) in the special case $a = b = 1$)**.** *Let F be a complex valued function which is regular in the strip $\alpha \leq \operatorname{Re} z \leq \beta$. Suppose that $|F(z)|$ tends to zero as $|\operatorname{Im} z| \to \infty$, uniformly for $\alpha \leq \operatorname{Re} z \leq \beta$. Then, for any $\gamma \in [\alpha, \beta]$ and any fixed $k > 0$,*

$$I(\gamma) \leq I(\alpha)^{(\beta-\gamma)/(\beta-\alpha)} \cdot I(\beta)^{(\gamma-\alpha)/(\beta-\alpha)} \tag{6.3.3}$$

where

$$I(\sigma) = \int_{\mathbb{R}} |F(\sigma + it)|^k dt. \tag{6.3.4}$$

This theorem has the following useful consequence.

Corollary 6.3.4. *Let F be a complex valued function which is regular in the strip $\frac{1}{2} \le \operatorname{Re} z$. Suppose that $|F(z)|$ tends to zero uniformly as $|\operatorname{Im} z| \to \infty$. Suppose that $I(\sigma) \to 0$ as $\sigma \to \infty$. Then, for any $\sigma > \frac{1}{2}$ and any fixed $k > 0$,*

$$I(\sigma) \le I(\tfrac{1}{2}). \tag{6.3.5}$$

Proof. Let $\sigma^\star$ be such that

$$I(\sigma^\star) = \sup_{\sigma \ge 1/2} I(\sigma). \tag{6.3.6}$$

Note that because of the assumption that $I(\sigma) \to 0$ as $\sigma \to \infty$, the above $\sigma^\star$ has a finite value. Let $\varepsilon > 0$ be given. If $\sigma^\star = \frac{1}{2}$, then we are done. If $\sigma^\star \neq \frac{1}{2}$, then by Proposition 6.3.3 applied with $\gamma = \sigma^\star$, $\alpha = \frac{1}{2}$ and $\beta = \sigma^\star + \varepsilon$, we get

$$I(\sigma^\star) \le I(\tfrac{1}{2})^\lambda \cdot I(\sigma^\star + \varepsilon)^\mu, \tag{6.3.7}$$

for some appropriate $\lambda, \mu > 0$ that satisfy $\lambda + \mu = 1$.

Therefore, by definition of $\sigma^\star$ in (6.3.6),

$$I(\sigma^\star) \le I(\tfrac{1}{2})^\lambda \cdot I(\sigma^\star)^\mu, \tag{6.3.8}$$

and hence $I(\sigma^\star)^\lambda \le I(\frac{1}{2})^\lambda$. Since $\lambda > 0$, we get $I(\sigma^\star) \le I(\frac{1}{2})$. By (6.3.6), the claim follows. □

We now construct a special analytic approximation for the indicator function of a rectangle.

Lemma 6.3.5. *Let $0 \le \Delta \le L$ and $\varepsilon > 0$ be given. There exists an entire function $\Phi_{\Delta,L}(z)$ such that, for $z = \sigma + iv$ with $\sigma \ge \frac{1}{2}$ and $v \in \mathbb{R}$,*

(1) For $K > 1 + \varepsilon$ and $|v| > KL$, uniformly in $\sigma \ge \frac{1}{2}$, $\Phi_{\Delta,L}(z) \ll_A ((K-1)\Delta)^{-A}$.

(2) For any $|v| \le (1-\varepsilon)L$, $|\Phi_{\Delta,L}(z)| = 1 + O_A(\Delta^{-A}) + O((\sigma - \frac{1}{2})\Delta^2/L)$.

(3) For any $(1-\varepsilon)L \le |v| \le (1+\varepsilon)L$, $|\Phi_{\Delta,L}(z)| \ll 1 + (\sigma - \frac{1}{2})\Delta^2/L$.

(4) $\Phi_{\Delta,L}(z) \to 0$ *uniformly in* v *as* $\sigma \to \infty$.

PROOF. Let V be a smooth function, compactly supported in $[0,\infty)$ and such that $V(1) = 1$. Given a parameter $\eta > 0$ and given $z \in \mathbb{C}$ with $\operatorname{Re} z \ge \frac{1}{2}$ and $u \in \mathbb{R}$, consider the following function :

$$\delta_\eta(z) = \eta \int_0^\infty e^{-2\pi(z-\frac{1}{2})x} \cdot V\Big(\eta x\Big)dx. \tag{6.3.9}$$

Then $\delta_\eta(z)$ defines an entire function of exponential type. By integration by parts, we see that

$$\delta_\eta(z) \ll_A (1 + |z - \tfrac{1}{2}|\eta^{-1})^{-A}, \tag{6.3.10}$$

for any $A > 0$ and uniformly in $\operatorname{Re} z \ge \frac{1}{2}$. Therefore, we may think of $\delta_\eta(z)$ as localizing to $z = \frac{1}{2} + O(\eta)$. Furthermore, notice that if $z = \frac{1}{2} + iv$ and $u \in \mathbb{R}$, then

$$\delta_\eta(z - iu) = \widehat{V}(\eta^{-1}(v-u)), \tag{6.3.11}$$

and for $z = \sigma + iv$, we have by a Taylor expansion of the exponential,

$$\begin{aligned}
\delta_\eta(z - iu) &= \eta \int_0^\infty e^{-2\pi(\sigma - \frac{1}{2} + i(v-u))x} \cdot V\Big(\eta x\Big)dx \\
&= \eta \int_0^\infty e^{-2\pi i(v-u)x} \cdot \Big(1 + O\Big((\sigma - \tfrac{1}{2})\eta^{-1}\Big)\Big) \cdot V\Big(\eta x\Big)dx \\
&= \widehat{V}(\eta^{-1}(v-u)) + O\Big((\sigma - \tfrac{1}{2})\eta^{-1}\Big).
\end{aligned} \tag{6.3.12}$$

Finally, for $z = \sigma + iv$ with $\sigma \ge \frac{1}{2}$, we have from (6.3.10) that

$$|\delta_\eta(z - iu)| \ll_A \frac{1}{1 + (\eta^{-1}|u-v|)^A}. \tag{6.3.13}$$

The candidate function is

$$\Phi_{\Delta,L}(z) = \frac{\Delta}{L}\int_{-L}^{L} e^{-2\pi iu(\Delta/L)} \cdot \delta_{L/\Delta}(z - iu)du. \tag{6.3.14}$$

We will now describe some of the features of this function. Write $z = \sigma + iv$ with $\sigma \ge \frac{1}{2}$. Using the bound (6.3.13), we see that, if $|v| > KL$ with $K > 1 + \varepsilon$ and $\sigma \ge \frac{1}{2}$, then

$$\Phi_{\Delta,L}(z) \ll_A \frac{\Delta}{L}\int_{-L}^{L} \frac{1}{1+(\frac{\Delta}{L}\cdot|u-v|)^A}\,du \ll_A (K-1)^{-A}\Delta^{1-A}. \tag{6.3.15}$$

This gives the first claim.

If $|v| \leq (1+\varepsilon)L$, then by (6.3.14) and (6.3.12), we have

$$\Phi_{\Delta,L}(z) = \frac{\Delta}{L}\int_{-L}^{L} e^{-2\pi i u(\Delta/L)}\cdot \widehat{V}\Big(\tfrac{\Delta}{L}\cdot(v-u)\Big)du + O\Big((\sigma-\tfrac{1}{2})\Delta^2/L\Big). \tag{6.3.16}$$

In particular, it follows that if $\frac{1}{2} \leq \sigma$ and $|v| \leq (1-\varepsilon)L$, then due to the rapid decay of $\widehat{V}$,

$$\begin{aligned}\Phi_{\Delta,L}(z) &= e^{-2\pi i v(\Delta/L)}\int_{v\Delta/L-\Delta}^{v\Delta/L+\Delta} e^{2\pi i u}\cdot \widehat{V}(u)du + O\Big((\sigma-\tfrac{1}{2})\Delta^2/L\Big)\\ &= e^{-2\pi i v(\Delta/L)} + O_A(\Delta^{-A}) + O\Big((\sigma-\tfrac{1}{2})\Delta^2/L\Big),\end{aligned} \tag{6.3.17}$$

by Fourier inversion and the assumption that $V(1)=1$. This proves the second claim. If $\frac{1}{2} \leq \sigma \ll 1$ and $|v| \leq (1+\varepsilon)L$, then we have the bound

$$|\Phi_{\Delta,L}(z)| \ll \int_{\mathbb{R}} |\widehat{V}(u)|du + O\Big((\sigma-\tfrac{1}{2})\Delta^2/L\Big), \tag{6.3.18}$$

which proves the third claim. Finally, notice that $\delta_{L/\Delta}(z-iu) \to 0$ uniformly as $\sigma \to \infty$ by (6.3.10), which implies the last claim that $\Phi_{\Delta,L}(z) \to 0$ uniformly in $v \in \mathbb{R}$ as $\sigma \to \infty$. ☐

The following proposition relates the moments off and on axis.

Proposition 6.3.6. *Let $\theta > -1$, $\beta > 0$, $\varepsilon > 0$, $T \geq 10^9$. Then, for all $\frac{1}{2} \leq \sigma \leq \frac{1}{2} + (\log T)^{\theta-3\varepsilon}$, the event*

$$\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\sigma+i\tau+iu)|^\beta du \ll_{\theta,\beta,\varepsilon} \int_{-2\log^\theta T}^{2\log^\theta T} |\zeta(\tfrac{1}{2}+i\tau+iu)|^\beta du + \frac{1}{(\log T)^7} \tag{6.3.19}$$

has probability $1-o(1)$.

Proof. Let

$$D(\sigma+i\tau) = \sum_{n\leq T}\frac{1}{n^{\sigma+i\tau}}\cdot\Big(1-\frac{n}{T}\Big)^A, \tag{6.3.20}$$

with $A > 100$ fixed. From Proposition 2 in Bombieri and Friedlander (1995), we have, for

$T \le \tau \le 2T$ and $\frac{1}{2} \le \sigma \le \frac{1}{2} + (\log T)^{\theta - 3\varepsilon}$,

$$\zeta(\sigma + i\tau) = D(\sigma + i\tau) + O_A(T^{-A/2}). \tag{6.3.21}$$

Consider

$$I(\sigma) = \int_{\mathbb{R}} |D(\sigma + i\tau + iu)|^{\beta} \cdot |\Phi_{\Delta,L}(\sigma + iu)|^{\beta} du, \tag{6.3.22}$$

with $\Delta = \log^{\varepsilon} T$ and $L = 2\log^{\theta} T$. Then, by Lemma 6.3.5 and Corollary 6.3.4, we have

$$\begin{aligned} &\int_{\mathbb{R}} |D(\sigma + i\tau + iu)|^{\beta} \cdot |\Phi_{\Delta,L}(\sigma + iu)|^{\beta} du \\ &\qquad \ll \int_{\mathbb{R}} |D(\tfrac{1}{2} + i\tau + iu)|^{\beta} \cdot |\Phi_{\Delta,L}(\tfrac{1}{2} + iu)|^{\beta} du. \end{aligned} \tag{6.3.23}$$

Now, it remains to unsmooth the expression. By Lemma 6.3.5, provided that $\sigma - \frac{1}{2} \le (\log T)^{\theta - 3\varepsilon}$, we have

$$\int_{-\log^{\theta} T}^{\log^{\theta} T} |D(\sigma + i\tau + iu)|^{\beta} du \ll \int_{\mathbb{R}} |D(\sigma + i\tau + iu)|^{\beta} \cdot |\Phi_{\Delta,L}(\sigma + iu)|^{\beta} du. \tag{6.3.24}$$

On the other hand, by Lemma 6.3.5, we have

$$\begin{aligned} &\int_{\mathbb{R}} |D(\tfrac{1}{2} + i\tau + iu)|^{\beta} \cdot |\Phi_{\Delta,L}(\tfrac{1}{2} + iu)|^{\beta} du \\ &\qquad \ll \int_{-2\log^{\theta} T}^{2\log^{\theta} T} |D(\tfrac{1}{2} + i\tau + iu)|^{\beta} du \\ &\qquad + \sum_{A=0}^{\infty} \int_{\mathcal{U}_A} |D(\tfrac{1}{2} + i\tau + iu)|^{\beta} \cdot |\Phi_{\Delta,L}(\tfrac{1}{2} + iu)|^{\beta} du, \end{aligned} \tag{6.3.25}$$

where $\mathcal{U}_A = \{2(\log T)^{\theta + A} \le |u| \le 2(\log T)^{\theta + A + 1}\}$. By Corollary 6.2.9, the approximation in (6.3.21), and a union bound, the event

$$\mathcal{S}(T) = \left\{ \max_{A \in \mathbb{N} \cup \{0\}} \max_{|u| \le \log^{A} T} |D(\tfrac{1}{2} + i\tau + iu)| \le 2^{A} (\log T)^{2+A} \right\} \tag{6.3.26}$$

has probability $1 - o(1)$. Moreover, by Lemma 6.3.5, for all $2(\log T)^{\theta + A} \le |u|$, we have

$$|\Phi_{\Delta,L}(\tfrac{1}{2} + iu)| \ll_{\theta,\beta,\varepsilon} (\log T)^{-4A(1+1/\beta)} \cdot (\log T)^{-(10\lceil\theta\rceil + 10)\cdot(1+1/\beta)}. \tag{6.3.27}$$

Therefore, on the event $\mathcal{S}(T)$, and for each $A \geq 0$,

$$\int_{\mathcal{U}_A} |D(\tfrac{1}{2}+i\tau+iu)|^\beta \cdot |\Phi_{\Delta,L}(\tfrac{1}{2}+iu)|^\beta du$$

$$\ll_{\theta,\beta,\varepsilon} (\log T)^{(\beta+1)(\lceil\theta\rceil+A+3)} \cdot 2^{A\beta} \cdot (\log T)^{-(\beta+1)\cdot(10\lceil\theta\rceil+4A+10)}$$

$$\ll_{\theta,\beta,\varepsilon} (\log T)^{-(\beta+1)\cdot(A+7)}. \tag{6.3.28}$$

Thus, on $\mathcal{S}(T)$, the contribution of the sum on the right-hand side of (6.3.25) is negligible. By combining (6.3.23), (6.3.24) and (6.3.25), the claim follows. □

6.3.2. Mollification

This step is an adaptation of Section 4.2 of Arguin et al. (2019), which is itself based on the work of Radziwiłł and Soundararajan (2017). The treatment is slightly different as the width of the interval needs to be taken into account. Also, we choose to use the discretization in Proposition 6.2.7 to obtain a uniform control on the interval as opposed to a Sobolev inequality.

The main idea is to define a mollifier for the zeta function

$$M(s) = \sum_n \frac{\mu(n)a(n)}{n^s}. \tag{6.3.29}$$

Here μ denotes the Möbius function $\mu(n) = (-1)^{\omega(n)}$ if n is square-free, where $\omega(n)$ is the number of distinct prime factors, and $\mu(n) = 0$ if n is non-square free. The term $a(n)$ equals 1 if all prime factors of n are smaller than

$$X = \exp((\log T)^{1-K^{-1}}), \quad K \geq 2, \tag{6.3.30}$$

and if

$$\Omega(n) \leq 100Ke^{\theta\vee 0}\log\log T =: \nu_\theta, \tag{6.3.31}$$

with $a(n) = 0$ otherwise. The estimate will be done slightly off-axis:

$$\sigma_0 = \frac{1}{2} + \frac{(\log T)^{3/(2K)}}{\log T}. \tag{6.3.32}$$

The parameter K will eventually be assumed to be large enough depending on θ, β and ε.

The goal of this section is to prove that M is an approximate inverse of ζ:

Lemma 6.3.7. *Let $\theta > -1$ and $\varepsilon > 0$ be given. Then,*

$$\mathbb{P}\left(\max_{|h|\leq \log^\theta T} |(\zeta \cdot M)(\sigma_0 + i(\tau + h)) - 1| > \varepsilon\right) = o(1). \tag{6.3.33}$$

This was proved in the case $\theta = 0$ in Lemma 4.2 of Arguin et al. (2019). In particular, it also holds *verbatim* for $-1 < \theta < 0$ since the interval is just smaller. The proof of Lemma 6.3.7 also holds in the case $\theta > 0$ with slight modifications that we highlight. The key idea is the following L^2-control:

Lemma 6.3.8. *Let $\theta > -1$ be given. Then,*

$$\mathbb{E}\left[\left|(\zeta \cdot M)(\sigma_0 + i\tau) - 1\right|^2\right] \ll (\log T)^{-100\, e^\theta}. \tag{6.3.34}$$

PROOF. We only have to prove the case $\theta > 0$. The proof is exactly as in Arguin et al. (2019) with a new error term due to the choice of ν_θ. (The manipulations are very similar to the ones in Lemma 6.2.4.) The error appears after Equation (4.10) in Arguin et al. (2019) and is given by

$$(\log T)\, e^{-\nu_\theta} \prod_{p \leq X} (1 + 7p^{-1}). \tag{6.3.35}$$

The Euler product is bounded by $\ll (\log T)^7$ using Lemma 6.4.1. Using this and the definition of ν_θ in (6.3.31) yields

$$(\log T)\, e^{-\nu_\theta} \prod_{p \leq X} (1 + 7p^{-1}) \ll (\log T)^8 \cdot (\log T)^{-100 K e^\theta}. \tag{6.3.36}$$

Since $K \geq 2$, this gives the correct estimate. Note that the expression $\sum_{p>X} \log(1-p^{-2\sigma_0})^{-1}$ entering in the remainder of the proof of Lemma 4.2 is

$$\ll \sum_{p>X} p^{-2\sigma_0} \ll X^{-(\sigma_0 - 1/2)} = \exp(-(\log T)^{\frac{1}{2K}}) \ll (\log T)^{-100\, e^\theta}. \tag{6.3.37}$$

This ends the proof. $\square$

PROOF OF LEMMA 6.3.7. By (6.2.37), ζ is well approximated by a Dirichlet polynomial of length T. Moreover, M is a Dirichlet polynomial of length less than T^δ for any fixed

$\delta > 0$. Therefore, an application of Chebyshev's inequality and Proposition 6.2.7 yield that the probability is

$$\ll \log^{1+\theta} T \cdot \mathbb{E}\Big[\Big|(\zeta \cdot M)(\sigma_0 + i\tau) - 1\Big|^2\Big]. \tag{6.3.38}$$

The conclusion follows from Lemma 6.3.8. □

6.3.3. Bounds for Dirichlet polynomials

We now approximate the mollifier M by the exponential of a Dirichlet polynomial. We first note that, on the region of absolute convergence, we have the following exact identity by expanding the log

$$\sum_n \mu(n) n^{-s} = \exp\Big(\log \prod_p (1 - p^{-s})\Big) = \exp\Big(-\sum_p p^{-s} - \sum_{k\geq 2}\sum_p \frac{p^{-ks}}{k}\Big). \tag{6.3.39}$$

Write

$$\widetilde{\mathcal{P}}_{1-K^{-1}}(s) := \sum_{p^k \leq X} \frac{1}{k p^{ks}}. \tag{6.3.40}$$

Note that $\exp(-\widetilde{\mathcal{P}}_{1-K^{-1}}(s))$ corresponds to a Dirichlet polynomial with coefficients $\mu(n)$ supported on integers n such that all the prime factors of n are $\leq X$.

Lemma 6.3.9. *Let $\theta > -1$ be given. Then, for any $K > 2$, we have*

$$\mathbb{P}\Big(\max_{|h|\leq \log^\theta T}\Big|M(\sigma_0 + i\tau + ih) - e^{-\widetilde{\mathcal{P}}_{1-K^{-1}}}(\sigma_0 + i\tau + ih)\Big)\Big| > (\log T)^{-10}\Big) = o(1). \tag{6.3.41}$$

Proof. Proceeding as in (6.2.42), it follows that for any $\varepsilon > 0$,

$$\mathbb{P}\Big(\max_{|h|\leq \log^\theta T} |\widetilde{\mathcal{P}}_{1-K^{-1}}(\sigma_0 + i\tau + ih)| > \frac{\nu_\theta}{10}\Big) = o(1). \tag{6.3.42}$$

This is done by noticing that the sum for $k > 2$ is of order one, and that the sum for $k = 2$ is of negligible order:

$$\mathbb{P}\Big(\max_{|h|\leq \log^\theta T}\Big|\sum_{p^2\leq X} p^{-2(\sigma_0+i\tau+ih)}/2\Big| > A\Big) \ll A^{-2\ell}(\log^{1+\theta} T) \cdot \ell!\,, \tag{6.3.43}$$

where we use Lemma 6.4.4. This is $o(1)$ for the choice $A = \sqrt{\nu_\theta}$ and $\ell = \lfloor (1+\theta)\log\log T \rfloor$.

Equations (6.3.42), (6.2.5) and (6.3.39) imply that, on a set of probability $1-o(1)$, the Dirichlet polynomial $\exp(-\widetilde{\mathcal{P}}_{1-K^{-1}}(s))$ is well approximated (with an error $e^{-\nu_\theta} \ll (\log T)^{-100}$) by a Dirichlet polynomial with the same coefficients as M on the set of integers with at most ν_θ prime factors. Denote this truncation by $\mathcal{M}$. In particular, Proposition 6.2.7 and Lemma 6.4.2 yield

$$\mathbb{E}\Big[\max_{|h|\le \log^\theta T} |M-\mathcal{M}|^2(\sigma_0+i\tau+ih)\Big] \ll \log^{1+\theta} T \cdot \sum_{\substack{p|n \implies p\le X\\ \Omega(n)>\nu_\theta}} n^{-1}. \tag{6.3.44}$$

The right-hand side is $\ll (\log T)^{-100}$ since

$$\log^{1+\theta} T \cdot \sum_{\substack{p|n \implies p\le X\\ \Omega(n)>\nu_\theta}} n^{-1} \ll \log^{1+\theta} T \cdot e^{-\nu_\theta} \sum_{p|n \implies p\le X} e^{\Omega(n)} n^{-1} \ll (\log T)^{-100}. \tag{6.3.45}$$

The result follows by Chebyshev's inequality. □

6.3.4. Proofs of the lower bounds

Consider, for $0 \le j \le K-2$, the Dirichlet polynomials

$$P_j(h) = \mathrm{Re} \sum_{p\in J_j} \frac{1}{p^{\sigma_0+i\tau+ih}}, \qquad J_j = (\exp((\log T)^{\frac{j}{K}}, \exp((\log T)^{\frac{j+1}{K}})]. \tag{6.3.46}$$

We choose a probabilistic notation for the increments P_j's seen as random variable, omitting the dependence on the random τ. We first prove a lower bound for the moments of Dirichlet polynomials.

Proposition 6.3.10. *Let $\theta > -1$ and $\varepsilon > 0$ be given. Then,*

$$\mathbb{P}\Big(\int_{-\log^\theta T}^{\log^\theta T} \exp\Big(\beta \sum_{j=1}^{K-3} P_j(h)\Big) dh > (\log T)^{f_\theta(\beta)-\varepsilon}\Big) = 1-o(1). \tag{6.3.47}$$

The polynomial P_{K-2} is not included in the sum to ensure that the variances of the P_j's are almost equal. Indeed, for all $|h| \le \log^\theta T$ and $j \le K-3$, an application of (6.4.6) yields

$$s_j^2 = \mathbb{E}[P_j(h)^2] = \frac{1}{2K}\log\log T + O((\log T)^{-\frac{1}{2K}}), \tag{6.3.48}$$

since $\sigma_0 - \frac{1}{2} = (\log T)^{-1+3/(2K)}$. The polynomial P_0 is ignored to ensure that the polynomials $\sum_{j=1}^{K-3} P_j(h)$ are almost independent for h's that are far apart, which will be crucial for the second-moment method to go through; see below (6.3.65) in the proof of Proposition 6.3.10.

Proof of Proposition 6.3.10. This is similar to the upper bound proof of Theorem 6.1.1. We first relate the moments to the measure of high points. Let $\varepsilon > 0$ and $M \in \mathbb{N}$, and set

$$\mathcal{E}_\theta(\gamma) := \begin{cases} \theta - \frac{\gamma^2}{1+\theta}, & \text{if } \theta \leq 0, \\ \theta - \gamma^2, & \text{if } \theta > 0. \end{cases} \tag{6.3.49}$$

Consider $\gamma_j = \frac{j}{M} m(\theta) + \varepsilon$ for $1 \leq j \leq M$, and the good event

$$\begin{aligned} E = \bigcap_{j=1}^{M} &\left\{ \mathrm{Leb}\Big\{ |h| \leq \log^\theta T : \exp\Big(\sum_{j=1}^{K-3} P_j(h) \Big) > (\log T)^{\gamma_{j-1}} \Big\} \geq (\log T)^{\mathcal{E}_\theta(\gamma_{j-1}) - \varepsilon/2} \right\} \\ &\bigcap \left\{ \max_{|h| \leq \log^\theta T} \exp\Big(\sum_{j=1}^{K-3} P_j(h) \Big) \leq (\log T)^{m(\theta)+\varepsilon} \right\}. \end{aligned} \tag{6.3.50}$$

We will show below that $\mathbb{P}(E)$ is $1 - o(1)$. First, we prove the lower bound on the moments on the event E. We have

$$\frac{\log \int_{-\log^\theta T}^{\log^\theta T} \exp\Big(\beta \sum_{j=1}^{K-3} P_j(h)\Big) dh}{\log \log T} \geq \max_{1 \leq j \leq M} \{\beta \gamma_{j-1} + \mathcal{E}_\theta(\gamma_{j-1})\} - \varepsilon/2. \tag{6.3.51}$$

By the continuity of the function $\gamma \mapsto \beta\gamma + \mathcal{E}_\theta(\gamma)$, Equation (6.3.51) implies that, on the event E and for M large enough with respect to ε and β,

$$\frac{\log \int_{-\log^\theta T}^{\log^\theta T} \exp\Big(\beta \sum_{j=1}^{K-3} P_j(h)\Big) dh}{\log \log T} > \max_{\gamma \in [\varepsilon, m(\theta)]} \Big\{ \beta\gamma + \mathcal{E}_\theta(\gamma) \Big\} - \varepsilon. \tag{6.3.52}$$

When $0 < \beta \leq 2m(\theta)/(1 + (\theta \wedge 0))$, take $\varepsilon > 0$ small enough so that $\beta > 2\varepsilon/(1 + (\theta \wedge 0))$. The maximum is attained at $\gamma = \frac{\beta}{2}(1 + (\theta \wedge 0))$, in which case the right-hand side of (6.3.52) is equal to $\frac{\beta^2}{4}(1 + (\theta \wedge 0)) + \theta - \varepsilon$. When $\beta > 2m(\theta)/(1 + (\theta \wedge 0))$, the maximum is attained at $\gamma = m(\theta)$, in which case the right-hand side of (6.3.52) is equal to $(\beta m(\theta) - 1) - \varepsilon$. Thus, on the event E and for M large enough, the lower bound in (6.3.47) is satisfied.

To conclude the proof of the proposition, it remains to show that $\mathbb{P}(E) \to 1$ as $T \to \infty$. By the upper bound on the maximum of $\sum_{j=1}^{K-3} P_j(h)$ in (6.2.42) (and the remark below it for $\theta < 0$), it is sufficient to prove that, for all $\eta > 0$ and all $0 < \gamma < m(\theta)$, the event

$$\left\{ \mathrm{Leb}\left\{ |h| \leq \log^{\theta} T : \sum_{j=1}^{K-3} P_j(h) > \gamma \log\log T \right\} \geq (\log T)^{\mathcal{E}_\theta(\gamma)-\eta} \right\} \tag{6.3.53}$$

has probability $1 - o(1)$.

Consider

$$\mathcal{J}(\theta) = \begin{cases} 1, & \text{if } \theta \geq 0, \\ \lfloor K|\theta| \rfloor + 1, & \text{if } \theta < 0. \end{cases} \tag{6.3.54}$$

For $\theta < 0$, Corollary 6.2.11 ensures that the primes up to $\exp(\log^{|\theta|} T)$ only make a very small contribution, namely the event

$$\left\{ \left| \textstyle\sum_{j=1}^{\mathcal{J}(\theta)-1} P_j(h) \right| \leq \tfrac{\gamma}{(1+(\theta\wedge 0))\, K} \log\log T \right\} \tag{6.3.55}$$

has probability $1 - o(1)$. We consider the random variable

$$\mathcal{N} = \mathrm{Leb}\left\{ |h| \leq \log^{\theta} T : P_j(h) > x_j, \text{ for } \mathcal{J}(\theta) \leq j \leq K-3 \right\}, \tag{6.3.56}$$

where

$$x_j = \left(1 + \frac{100}{(1+(\theta\wedge 0))\, K}\right) \cdot \frac{\gamma}{(1+(\theta\wedge 0))\, K} \log\log T. \tag{6.3.57}$$

By summing the x_j's, it is not hard to check that the intersection of the events $\{\mathcal{N} \geq (\log T)^{\mathcal{E}_\theta(\gamma)-\eta}\}$ and the one in (6.3.55) is included in the event in (6.3.53). Therefore, the proof of the proposition is reduced to show

$$\mathbb{P}\Big(\mathcal{N} \geq (\log T)^{\mathcal{E}_\theta(\gamma)-\eta}\Big) = 1 - o(1). \tag{6.3.58}$$

This is established by the Paley-Zygmund inequality.

To this aim, we shall need one-point and two-point large deviation estimates for the event

$$A(h) = \left\{ P_j(h) > x_j, \text{ for } \mathcal{J}(\theta) \leq j \leq K-3 \right\}. \tag{6.3.59}$$

The next two propositions are stated as Propositions 5.4 and 5.5 in Arguin et al. (2019). They are consequences of the Gaussian moments in Lemma 6.4.3.

Proposition 6.3.11 (One-point large deviation estimates)**.** *Let $\theta > -1$ be given, and let $h \in [-\log^\theta T, \log^\theta T]$. For any choices of $0 < x_j \leq \log\log T$, where $1 \leq j \leq K-3$, we have*

$$\mathbb{P}(A(h)) = (1+o(1)) \prod_{j=\mathcal{J}(\theta)}^{K-3} \int_{x_j/s_j}^{\infty} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy \asymp \prod_{j=\mathcal{J}(\theta)}^{K-3} \frac{s_j}{x_j} \cdot e^{-x_j^2/(2s_j^2)}. \tag{6.3.60}$$

In the case of two points h, h', the primes are essentially correlated up to $\exp(|h-h'|^{-1})$ and quickly decorrelate afterwards. For $\theta \geq 0$, this means that the P_j's are essentially independent whenever $|h-h'| > (\log T)^{-\frac{1}{2K}}$, since $j = 0$ is excluded. For $\theta < 0$, we must exclude the j's up to $\mathcal{J}(\theta) - 1$. Therefore, the P_j's are essentially independent whenever $|h-h'| > (\log T)^{\theta - \frac{1}{2K}}$. We get:

Proposition 6.3.12 (Two-point large deviation estimates)**.** *Let $\theta > -1$ be given, and let $h, h' \in [-\log^\theta T, \log^\theta T]$ be such that $|h-h'| > (\log T)^{-\frac{\mathcal{J}(\theta)}{K} + \frac{1}{2K}}$. Then,*

$$\mathbb{P}(A(h) \cap A(h')) = (1+o(1))\, \mathbb{P}(A(h))\, \mathbb{P}(A(h')). \tag{6.3.61}$$

If $|h-h'| \leq 1$, let $0 \leq \ell \leq K-3$ denote the largest integer in this range with $|h-h'| \leq (\log T)^{-\ell/K}$. Then, for any choices of $\sqrt{\log\log T} \ll x_j \leq \log\log T$, we have

$$\mathbb{P}(A(h) \cap A(h')) \ll \exp\Bigg(- \sum_{j=\mathcal{J}(\theta)}^{\ell} \frac{x_j^2}{2s_j^2} - \sum_{j=(\ell+1)\vee\mathcal{J}(\theta)}^{K-3} \frac{x_j^2}{s_j^2} \Bigg). \tag{6.3.62}$$

Now, in order to prove (6.3.58), we start by finding a lower bound on $\mathbb{E}[\mathcal{N}]$. By (6.3.60), the x_j's in (6.3.57) and the s_j's in (6.3.48), we have

$$\mathbb{E}[\mathcal{N}] = \int_{-\log^\theta T}^{\log^\theta T} \mathbb{P}(A(h)) dh \gg \log^\theta T \prod_{j=\mathcal{J}(\theta)}^{K-3} \frac{s_j}{x_j} \cdot e^{-x_j^2/(2s_j^2)} \gg (\log T)^{\mathcal{E}_\theta(\gamma) - \eta/3}, \tag{6.3.63}$$

assuming that K is large enough with respect to θ, γ and η. By the Paley-Zygmund inequality, this implies

$$\begin{aligned} \mathbb{P}\Big(\mathcal{N} \geq (\log T)^{\mathcal{E}_\theta(\gamma) - \eta}\Big) &\geq \mathbb{P}\Big(\mathcal{N} \geq (\log T)^{-\eta/3} \mathbb{E}[\mathcal{N}]\Big) \\ &\geq \Big(1 - (\log T)^{-\eta/3}\Big) (\mathbb{E}[\mathcal{N}])^2 / \mathbb{E}[\mathcal{N}^2]. \end{aligned} \tag{6.3.64}$$

It remains to show $\mathbb{E}[\mathcal{N}^2] = (1+o(1))(\mathbb{E}[\mathcal{N}])^2$. With $I = [-\log^\theta T, \log^\theta T]$, linearity yields

$$\mathbb{E}[\mathcal{N}^2] = \int_{I\times I} \mathbb{P}(A(h)\cap A(h'))\, dhdh'. \tag{6.3.65}$$

The integral can be divided into $(K - \mathcal{J}(\theta)+1)$ parts:

$$\begin{aligned}
&B = \{(h,h') : |h-h'| > (\log T)^{-\frac{\mathcal{J}(\theta)}{K}+\frac{1}{2K}}\};\\
&B_0 = \{(h,h') : (\log T)^{-\frac{\mathcal{J}(\theta)}{K}} < |h-h'| \le (\log T)^{-\frac{\mathcal{J}(\theta)}{K}+\frac{1}{2K}}\};\\
&B_\ell = \{(h,h') : (\log T)^{-(\ell+1)/K} < |h-h'| \le (\log T)^{-\ell/K}\}, \quad \text{for } \ell = \mathcal{J}(\theta),\dots,K-3;\\
&B_{K-2} = \{(h,h') : |h-h'| \le (\log T)^{-(K-2)/K}\}.
\end{aligned} \tag{6.3.66}$$

The dominant term will be the one on B. Note that $\mathrm{Leb}(B) = \mathrm{Leb}(I)^2(1+o(1))$. Hence, by (6.3.61), we have

$$\int_B \mathbb{P}(A(h)\cap A(h'))dhdh' = (1+o(1))(\mathbb{E}[\mathcal{N}])^2. \tag{6.3.67}$$

By (6.3.62) and the estimate (6.3.63), the integral on B_0 is

$$\begin{aligned}
&\ll (\log T)^{\theta-\frac{\mathcal{J}(\theta)}{K}+\frac{1}{2K}} \exp\left(\sum_{j=\mathcal{J}(\theta)}^{K-3} -\frac{x_j^2}{s_j^2}\right)\\
&\ll (\log T)^{-(\theta\vee 0)-\frac{1}{3K}}(\mathbb{E}[\mathcal{N}])^2,
\end{aligned} \tag{6.3.68}$$

assuming that K is large enough with respect to θ and γ. For $\ell = \mathcal{J}(\theta),\dots,K-3$, the integral on B_ℓ is, by (6.3.62) and the estimate (6.3.63),

$$\begin{aligned}
&\ll (\log T)^{\theta-\ell/K} \exp\left(-\sum_{j=\mathcal{J}(\theta)}^{\ell} \frac{x_j^2}{2s_j^2} - \sum_{j=\ell+1}^{K-3}\frac{x_j^2}{s_j^2}\right)\\
&= (\log T)^{-\theta-\ell/K} \exp\left(\sum_{j=\mathcal{J}(\theta)}^{\ell} \frac{x_j^2}{2s_j^2}\right)\cdot (\log T)^{2\theta} \exp\left(-\sum_{j=\mathcal{J}(\theta)}^{K-3}\frac{x_j^2}{s_j^2}\right)\\
&\ll (\log T)^{-\theta-\ell/K+(\ell/K+(\theta\wedge 0))\frac{\gamma^2}{(1+(\theta\wedge 0))^2}+\eta}(\mathbb{E}[\mathcal{N}])^2,
\end{aligned} \tag{6.3.69}$$

assuming again that K is large enough with respect to θ, γ and η. Since $\gamma^2 < m(\theta)^2 = (1+\theta)(1+(\theta\wedge 0))$, the right-hand side of (6.3.69) is $o\Big((\mathbb{E}[\mathcal{N}])^2\Big)$ if we fix $\eta > 0$ small enough with respect to θ. Similarly, by (6.3.60) and the estimate (6.3.63), the integral on B_{K-2} is

$$\leq \int_{B_{K-2}} \mathbb{P}(A(h))dhdh' \ll (\log T)^{-\theta-1+2/K+\eta/3} \cdot \mathbb{E}[\mathcal{N}] = o\Big((\mathbb{E}[\mathcal{N}])^2\Big). \tag{6.3.70}$$

This concludes the proof of Proposition 6.3.10. $\square$

Putting all the work of Section 3 together, we can prove the lower bound in Theorem 6.1.1.

Proof of Proposition 6.3.2. By Proposition 6.3.6, the probability in (6.3.2) is

$$\geq \mathbb{P}\Big(\int_{-\frac{1}{2}\log^\theta T}^{\frac{1}{2}\log^\theta T} |\zeta(\sigma_0+i\tau+ih)|^\beta dh > (\log T)^{f_\theta(\beta)-\varepsilon}\Big) - o(1). \tag{6.3.71}$$

By Lemma 6.3.7 and Lemma 6.3.9, the above is

$$\geq \mathbb{P}\Big(\int_{-\frac{1}{2}\log^\theta T}^{\frac{1}{2}\log^\theta T} \exp\Big(\beta \mathrm{Re}\, \widetilde{\mathcal{P}}_{1-K^{-1}}(h)\Big) dh > (\log T)^{f_\theta(\beta)-\varepsilon}\Big) - o(1). \tag{6.3.72}$$

By Equation (6.3.43), $\widetilde{\mathcal{P}}_{1-K^{-1}}$ can be replaced by $\mathcal{P}_{1-K^{-1}}$ with an error less than $\log^\varepsilon T$. By (6.2.42), we may discard the terms with $j=0$ and $j=K-2$ with a similar error. For K large enough with respect to ε, β and θ, the probability in (6.3.72) is therefore

$$\geq \mathbb{P}\Big(\int_{-\frac{1}{2}\log^\theta T}^{\frac{1}{2}\log^\theta T} \exp\Big(\beta \sum_{j=1}^{K-3} P_j(h)\Big) dh > (\log T)^{f_\theta(\beta)-\varepsilon}\Big) - o(1). \tag{6.3.73}$$

Finally, the probability in (6.3.73) tends to 1 as $T\to\infty$ by Proposition 6.3.10. $\square$

We now prove the lower bound in Theorem 6.1.2.

Proof of Proposition 6.3.1. From (6.1.8), we have that $f_\theta(\beta) = \beta m(\theta) - 1$ when $\beta > \beta_c(\theta) = 2\sqrt{1+(\theta\wedge 0)}$. Thus, on the event in the statement of Proposition 6.3.2 (which has probability $1-o(1)$), and for β large enough with respect to ε and θ, we have

$$\max_{|h|\leq \log^\theta T} |\zeta(\tfrac{1}{2}+i\tau+ih)| \geq \Big(\frac{1}{2\log^\theta T}\int_{-\log^\theta T}^{\log^\theta T} |\zeta(\tfrac{1}{2}+i\tau+ih)|^\beta dh\Big)^{1/\beta}$$
$$\gg (\log T)^{m(\theta)-\frac{(1+\varepsilon+\theta)}{\beta}} \geq (\log T)^{m(\theta)-\varepsilon}. \tag{6.3.74}$$

This ends the proof. $\square$

6.4. Useful estimates

The prime number theorem yields estimates on the sum of primes with a good error.

Lemma 6.4.1. *Let $1 \le P \le Q$, then*

$$\sum_{P<p\le Q} \frac{(\log p)^m}{p} = \begin{cases} \frac{(\log Q)^m}{m} - \frac{(\log P)^m}{m} + O_m(1), & \text{if } m \ge 1, \\ \log\log Q - \log\log P + O(e^{-c\sqrt{\log P}}), & \text{if } m = 0. \end{cases} \tag{6.4.1}$$

Also, for $|\eta \log Q| \le 1$,

$$\sum_{P<p\le Q} \frac{\cos(\eta \log p)}{p} = \log\log Q - \log\log P + O(1). \tag{6.4.2}$$

Proof. For (6.4.1), see Lemma A.1 of Arguin and Ouimet (2019) and Lemma 2.1 of Arguin et al. (2017b). For (6.4.2), see p.20 in Harper (2013a). □

The next three results yield moment estimates for Dirichlet polynomials. The first one is an elementary bound. The second ensures that moments of Dirichlet polynomials that are not too high are Gaussian.

Lemma 6.4.2 (Lemma 3.3 in Arguin et al. (2019))**.** *For any complex numbers $a(n)$ and $b(n)$, and for $N \le T$, we have*

$$\begin{aligned} &\mathbb{E}\Big[\Big(\sum_{m\le N} a(m) m^{-i\tau}\Big)\Big(\sum_{n\le N} b(n) n^{i\tau}\Big)\Big] \\ &\quad = \sum_{n\le N} a(n)b(n) + O\Big(\frac{N\log N}{T}\sum_{n\le N}(|a(n)|^2 + |b(n)|^2)\Big). \end{aligned} \tag{6.4.3}$$

Lemma 6.4.3 (Lemma 3.4 in Arguin et al. (2019))**.** *Let $x \in [2,\infty)$, and suppose that for primes $p \le x$, $a(p) \in \mathbb{C}$ such that $|a(p)| \le 1$. Then, for any $k \in \mathbb{N}$,*

$$\mathbb{E}\Big[\Big(\frac{1}{2}\sum_{p\le x}(a(p)p^{-i\tau} + a^\star(p)p^{i\tau})\Big)^k\Big] = \frac{\partial^k}{\partial z^k}\Big(\prod_{p\le x} I_0(\sqrt{a(p)a^\star(p)}z)\Big)\Big|_{z=0} + O\Big(\frac{x^{2k}}{T}\Big), \tag{6.4.4}$$

where $I_0(z) = \sum_{n\ge 0} z^{2n}/(2^{2n}(n!)^2)$ denotes the modified Bessel function of the first kind. In particular, the expression is $O\left(x^{2k}/T\right)$ for odd k.

The relations with Gaussian moments in the case where $a(p) = p^{-\sigma-ih}$ is obtained by expanding the product to get

$$\prod_{p\le x} I_0(|a(p)|z) = F(z)\cdot \exp\left(\frac{z^2}{2}\cdot\frac{1}{2}\sum_{p\le x} p^{-2\sigma}\right) \tag{6.4.5}$$

where $F(z)$ is analytic in a neighborhood of 0 with $F(0) = 1$ and its derivatives uniformly bounded by $\sum_{p\le x} p^{-4\sigma}$. In particular, this implies that, for $\sigma \ge 1/2$ and k small enough so that $x^{2k}/T = o(1)$,

$$\mathbb{E}\left[\Big(\sum_{p\le x} \operatorname{Re} p^{-\sigma-i\tau-ih}\Big)^{2k}\right] = (1+o(1))\,\frac{(2k)!}{2^k\cdot k!}\left(\frac{1}{2}\sum_{p\le x} p^{-2\sigma}\right)^k. \tag{6.4.6}$$

The above also holds if $a(p) = 0$ for $p \le y$ (say) with the sum over primes restricted to $y < p \le x$. In particular, the error $\sum_{y<p\le x} p^{-4\sigma}$ can be made $o(1)$ by taking y large. We note that the moments yield a Gaussian tail

$$\mathbb{P}\Big(\sum_{p\le x} \operatorname{Re} p^{-\sigma-i\tau-ih} > V\Big) \ll \exp(-V^2/(2\sigma^2)), \tag{6.4.7}$$

by picking the moment $k = \lfloor V^2/2\sigma^2\rfloor$ with $\sigma^2 = \frac{1}{2}\sum_{p\le x} p^{-2\sigma}$, for V not too large.

Finally the third estimate is a cruder version of the Gaussian moment estimates that yields quick upper bounds on moments.

Lemma 6.4.4 (Lemma 3 in Soundararajan (2009))**.** *Let T be large, and let $2 \le x \le T$. Let k be a natural number such that $x^k \ll T/\log T$. For any complex numbers $a(p)$, we have*

$$\mathbb{E}\left[\Big|\sum_{p\le x} \frac{a(p)}{p^{1/2+i\tau}}\Big|^{2k}\right] \ll k!\left(\sum_{p\le x}\frac{|a(p)|^2}{p}\right)^k. \tag{6.4.8}$$

Acknowledgements

L.-P. Arguin is supported in part by NSF Grant DMS-1513441 and by NSF CAREER DMS-1653602. F. Ouimet is supported by a NSERC Doctoral Program Alexander Graham Bell scholarship (CGS D3). M. Radziwiłł acknowledges support of a Sloan fellowship.

Bibliography

L.-P. Arguin and F. Ouimet. Large deviations and continuity estimates for the derivative of a random model of $\log|\zeta|$ on the critical line. *J. Math. Anal. Appl.*, 472(1):687–695, 2019. MR3906393.

L.-P. Arguin and W. Tai. Is the Riemann zeta function in a short interval a 1-RSB spin glass ? *Preprint. To appear in Sojourns in probability and statistical physics, in honor of the 70th birthday of C. Newman*, pages 1–20, 2018. arXiv:1706.08462.

L.-P. Arguin, D. Belius, and P. Bourgade. Maximum of the characteristic polynomial of random unitary matrices. *Comm. Math. Phys.*, 349(2):703–751, 2017a. MR3594368.

L.-P. Arguin, D. Belius, and A. J. Harper. Maxima of a randomized Riemann zeta function, and branching random walks. *Ann. Appl. Probab.*, 27(1):178–215, 2017b. MR3619786.

L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwiłł, and K. Soundararajan. Maximum of the Riemann zeta function on a short interval of the critical line. *Comm. Pure Appl. Math.*, 72(3):500–535, 2019. MR3911893.

E. C. Bailey and J. P. Keating. On the moments of the moments of the characteristic polynomials of random unitary matrices. *Preprint, To appear in Communications in Mathematical Physics*, pages 1–35, 2019. arXiv:1807.06605.

S. Bettin, V. Chandee, and M. Radziwiłł. The mean square of the product of the Riemann zeta-function with Dirichlet polynomials. *J. Reine Angew. Math.*, 729:51–79, 2017. MR3680370.

E. Bombieri and J. B. Friedlander. Dirichlet polynomial approximations to zeta functions. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 22(3):517–544, 1995. MR1360548.

A. Bondarenko and K. Seip. Extreme values of the Riemann zeta function and its argument. *Math. Ann.*, 372(3-4):999–1015, 2018. MR3880290.

P. Bourgade. Mesoscopic fluctuations of the zeta zeros. *Probab. Theory Related Fields*, 148 (3-4):479–500, 2010. MR2678896.

J. Bourgain. Decoupling, exponential sums and the Riemann zeta function. *J. Amer. Math. Soc.*, 30(1):205–224, 2017. MR3556291.

A. Bovier. *Statistical mechanics of disordered systems*, volume 18 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge,

2006. A mathematical perspective. MR2252929.

A. Bovier. *Gaussian processes on trees*, volume 163 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2017. From spin glasses to branching Brownian motion. MR3618123.

A. Bovier and I. Kurkova. Derrida's generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):481–495, 2004. MR2070335.

V. Chandee and K. Soundararajan. Bounding $|\zeta(\frac{1}{2}+it)|$ on the Riemann hypothesis. *Bull. Lond. Math. Soc.*, 43(2):243–250, 2011. MR2781205.

R. Chhaibi, T. Madaule, and J. Najnudel. On the maximum of the CβE field. *Duke Math. J.*, 167(12):2243–2345, 2018. MR3848391.

R. de la Bretèche and G. Tenenbaum. Sommes de Gál et applications. *Proc. London Math. Soc.*, 119(3):104–134, 2019. doi:10.1112/plms.12224.

B. Derrida. Random-energy model: an exactly solvable model of disordered systems. *Phys. Rev. B (3)*, 24(5):2613–2626, 1981. MR627810.

D. W. Farmer, S. M. Gonek, and C. P. Hughes. The maximum size of L-functions. *J. Reine Angew. Math.*, 609:215–236, 2007. MR2350784.

Y. V. Fyodorov and J.-P. Bouchaud. Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential. *J. Phys. A*, 41(37):372001, 12, 2008. MR2430565.

Y. V. Fyodorov and J. P. Keating. Freezing transitions and extreme values: random matrix theory, $\zeta(\frac{1}{2}+it)$ and disordered landscapes. *Philos. Trans. R. Soc. A*, 372(20120503): 1–32, 2014. MR3151088.

Y. V. Fyodorov, S. Gnutzmann, and J. P. Keating. Extreme values of CUE characteristic polynomials: a numerical study. *J. Phys. A*, 51(46):464001, 22, 2018. MR3876698.

Y.V. Fyodorov, G.A. Hiary, and J.P. Keating. Freezing transition, characteristic polynomials of random matrices, and the Riemann zeta-function. *Phys. Rev. Lett.*, 108(170601): 1–4, 2012. doi:10.1103/PhysRevLett.108.170601.

R. M. Gabriel. Some Results Concerning the Integrals of Moduli of Regular Functions Along Certain Curves. *J. London Math. Soc.*, 2(2):112–117, 1927. MR1574583.

A. J. Harper. A note on the maximum of the Riemann zeta function, and log-correlated random variables. *Preprint*, pages 1–26, 2013a. arXiv:1304.0677.

A. J. Harper. Sharp conditional bounds for moments of the Riemann zeta function. *Preprint*, pages 1–20, 2013b. arXiv:1305.4618.

W. Heap, M. Radziwiłł, and K. Soundararajan. Sharp upper bounds for fractional moments of the Riemann zeta function. *Preprint*, pages 1–10, 2019. arXiv:1901.08423.

D. R. Heath-Brown. Fractional moments of the Riemann zeta function. *J. London Math. Soc. (2)*, 24(1):65–78, 1981. MR623671.

H. Iwaniec and P. Sarnak. Perspectives on the analytic theory of L-functions. *Geom. Funct. Anal.*, Special Volume(Part II):705–741, 2000. MR1826269.

J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2+it)$. *Comm. Math. Phys.*, 214(1):57–89, 2000. MR1794265.

N. Kistler. Derrida's random energy models. From spin glasses to the extremes of correlated random fields. In *Correlated Random Systems: Five Different Methods*, volume 2143 of *Lecture Notes in Mathematics*. Springer, 2015. MR3380419.

P. Michel and A. Venkatesh. The subconvexity problem for GL_2. *Publ. Math. Inst. Hautes Études Sci.*, 111:171–271, 2010. MR2653249.

J. Najnudel. On the extreme values of the Riemann zeta function on random intervals of the critical line. *Probab. Theory Related Fields*, 172(1-2):387–452, 2018. MR3851835.

M. Nikula, E. Saksman, and C. Webb. Multiplicative chaos and the characteristic polynomial of the CUE: the L^1-phase. *Preprint*, pages 1–49, 2018. arXiv:1806.01831.

F. Ouimet. Poisson-Dirichlet statistics for the extremes of a randomized Riemann zeta function. *Electron. Commun. Probab.*, 23:Paper No. 46, 15, 2018. MR3841407.

E. Paquette and O. Zeitouni. The maximum of the CUE field. *Int. Math. Res. Not. IMRN*, 2018(16):5028–5119, 2018. MR3848227.

M. Radziwiłł and K. Soundararajan. Continuous lower bounds for moments of zeta and L-functions. *Mathematika*, 59(1):119–128, 2013. MR3028174.

M. Radziwiłł and K. Soundararajan. Selberg's central limit theorem for $\log|\zeta(1/2+it)|$. *Enseign. Math.*, 63(1-2):1–19, 2017. MR3832861.

M. Reed and B. Simon. *Methods of modern mathematical physics. I. Functional analysis.*

Academic Press, New York-London, 1972. MR0493419.

G. Rémy. The Fyodorov-Bouchaud formula and Liouville conformal field theory. *Preprint*, pages 1–24, 2018. arXiv:1710.06897.

R. Rhodes and V. Vargas. Gaussian multiplicative chaos and applications: a review. *Probab. Surv.*, 11:315–392, 2014. MR3274356.

E. Saksman and C. Webb. The Riemann zeta function and Gaussian multiplicative chaos: statistics on the critical line. *Preprint*, pages 1–85, 2018. arXiv:1609.00027.

K. Soundararajan. Moments of the Riemann zeta function. *Ann. of Math. (2)*, 170(2): 981–993, 2009. MR2552116.

E. C. Titchmarsh. *The theory of the Riemann zeta-function*. The Clarendon Press, Oxford University Press, New York, second edition, 1986. MR882550.

A. Venkatesh. Sparse equidistribution problems, period bounds and subconvexity. *Ann. of Math. (2)*, 172(2):989–1094, 2010. MR2680486.

C. Webb. The characteristic polynomial of a random unitary matrix and Gaussian multiplicative chaos—the L^2-phase. *Electron. J. Probab.*, 20:no. 104, 21, 2015. MR3407221.

Part 3 : Asymptotic statistics

Article 7

Complete monotonicity of multinomial probabilities and its application to Bernstein estimators on the simplex

by

Frédéric Ouimet[1]

([1]) Université de Montréal, Montréal, QC H3T 1J4, Canada.

Here is the complete citation :

Ouimet, F. (2018). Complete monotonicity of multinomial probabilities and its application to Bernstein estimators on the simplex. J. Math. Anal. Appl. 466, no. 2, 1609-1617. https://doi.org/10.1016/j.jmaa.2018.06.049.

Abstract. Let $d \in \mathbb{N}$ and let $\gamma_i \in [0, \infty)$, $x_i \in (0, 1)$ be such that $\sum_{i=1}^{d+1} \gamma_i = M \in (0, \infty)$ and $\sum_{i=1}^{d+1} x_i = 1$. We prove that

$$a \mapsto \frac{\Gamma(aM + 1)}{\prod_{i=1}^{d+1} \Gamma(a\gamma_i + 1)} \prod_{i=1}^{d+1} x_i^{a\gamma_i}$$

is completely monotonic on $(0, \infty)$. This result generalizes the one found by Alzer (2018) for binomial probabilities ($d = 1$). As a consequence of the log-convexity, we obtain some combinatorial inequalities for multinomial coefficients. We also show how the main result can be used to derive asymptotic formulas for quantities of interest in the context of statistical density estimation based on Bernstein polynomials on the d-dimensional simplex.

Keywords: multinomial probability, complete monotonicity, Gamma function, combinatorial inequalities, Bernstein polynomials, simplex

7.1. Introduction

For any $d \in \mathbb{N}$, let $[d] \mathrel{\hat{=}} \{1, 2, \dots, d\}$. For any $\boldsymbol{v} \mathrel{\hat{=}} (v_1, v_2, \dots, v_d) \in \mathbb{R}^d$, write $\|\boldsymbol{v}\| \mathrel{\hat{=}} \sum_{i=1}^d |v_i|$. Denote the d-dimensional simplex and its interior by

$$\mathcal{S} \mathrel{\hat{=}} \Big\{\boldsymbol{x} \in [0,1]^d : \|\boldsymbol{x}\| \leq 1\Big\} \quad \text{and} \quad \mathrm{Int}(\mathcal{S}) \mathrel{\hat{=}} \Big\{\boldsymbol{x} \in (0,1)^d : \|\boldsymbol{x}\| < 1\Big\}.$$

Given a random sample $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_n$ on $\mathcal{S}$ from some unknown distribution F, define the Bernstein estimator on the simplex

$$\hat{F}_{m,n}(\boldsymbol{x}) \mathrel{\hat{=}} \sum_{\boldsymbol{k} \in \mathbb{N}_0^d : \|\boldsymbol{k}\| \leq m} F_n(\boldsymbol{k}/m) P_{\boldsymbol{k},m}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathcal{S}, \tag{7.1.1}$$

where $m, n \in \mathbb{N}$, $F_n(\boldsymbol{y}) \mathrel{\hat{=}} \frac{1}{n} \sum_{j=1}^n 1_{\{\boldsymbol{y} \leq \boldsymbol{y}_j\}}$ is the empirical cumulative distribution function, $x_{d+1} \mathrel{\hat{=}} 1 - \|\boldsymbol{x}\|$, $k_{d+1} \mathrel{\hat{=}} m - \|\boldsymbol{k}\|$, and

$$P_{\boldsymbol{k},m}(\boldsymbol{x}) \mathrel{\hat{=}} \frac{m!}{\prod_{i=1}^{d+1} k_i!} \prod_{i=1}^{d+1} x_i^{k_i}. \tag{7.1.2}$$

Our first goal is to prove that $a \mapsto P_{a\boldsymbol{k},am}(\boldsymbol{x})$ is completely monotonic on $(0, \infty)$, see Definition 7.1.1 below. In fact, we prove a slightly more general statement in Theorem 7.2.1. From the log-convexity, we deduce some combinatorial inequalities for multinomial coefficients in Section 7.3. The proof of the theorem and the combinatorial inequalities follow very closely, and generalize, the work of Alzer (2018). In Section 7.4, we show how

Theorem 7.2.1 can be used to prove asymptotic formulas for quantities of interest related to (7.1.1). To our knowledge, the statistical properties (bias, variance, mean integrated squared error, etc.) of the estimator in (7.1.1) (and the associated density estimator, see e.g. Babu and Chaubey (2006); Leblanc (2010)) have never been studied when $d > 1$, except for the pointwise mean squared error of the density estimator in Tenbusch (1994) when $d = 2$. This was our motivation for this article.

Definition 7.1.1 (Complete monotonicity)**.** *A non-constant function $a \mapsto g(a)$ is said to be completely monotonic on $(0, \infty)$, if g has derivatives of all orders and satisfies*

$$(-1)^n g^{(n)}(a) > 0, \quad \text{for all } n \in \mathbb{N}_0,\ a \in (0, \infty). \tag{7.1.3}$$

Remark 7.1.1. *Inequality (7.1.3) is usually not strict when defining complete monotonicity, but non-constant functions that satisfy the non-strict version of (7.1.3) automatically satisfy the strict version, see (Dubourdieu, 1939, p.98) for the original proof or (van Haeringen, 1996, p.395) for a simpler proof.*

We will need the two following lemmas during the proof of Theorem 7.2.1.

Lemma 7.1.2. *Let $g : (0, \infty) \to (0, 1)$. If $(-\log g)'$ is completely monotonic on $(0, \infty)$, then g is completely monotonic on $(0, \infty)$.*

Proof. Take $f : (0, \infty) \to (0, 1) : y \mapsto e^{-y}$ and $h : (0, \infty) \to (0, \infty) : x \mapsto -\log g(x)$. Since h is positive and $h' = (-\log g)'$ is completely monotonic by hypothesis, then $g = f \circ h$ is completely monotonic by Theorem 2 in Miller and Samko (2001). □

Lemma 7.1.3. *If $\boldsymbol{u} \mathrel{\hat=} (u_1, u_2, \dots, u_d) \in Int(\mathcal{S})$, $u_{d+1} \mathrel{\hat=} 1 - \|\boldsymbol{u}\| > 0$ and $y > 1$, then*

$$J_{\boldsymbol{u}}(y) \mathrel{\hat=} \frac{1}{y-1} - \sum_{i=1}^{d+1} \frac{1}{y^{1/u_i} - 1} > 0. \tag{7.1.4}$$

Proof. Lemma 1 in Alzer (2018) proves (7.1.4) in the case $d = 1$. Fix $d \geq 2$ and assume that (7.1.4) is true for any smaller integer. Let $y > 1$. By Lemma 1 in Alzer (2018),

$$\frac{1}{y-1} - \frac{1}{y^{1/\|\boldsymbol{u}\|} - 1} - \frac{1}{y^{1/(1-\|\boldsymbol{u}\|)} - 1} > 0. \tag{7.1.5}$$

Therefore, (7.1.4) will follow if we can show that

$$\frac{1}{y^{1/\|\boldsymbol{u}\|}-1}-\sum_{i=1}^{d}\frac{1}{y^{1/u_i}-1}>0. \tag{7.1.6}$$

Simply define $z \mathrel{\hat{=}} y^{1/\|\boldsymbol{u}\|}$ and $v_i \mathrel{\hat{=}} u_i/\|\boldsymbol{u}\|$, then (7.1.6) is equivalent to

$$\frac{1}{z-1}-\sum_{i=1}^{d}\frac{1}{z^{1/v_i}-1}>0, \tag{7.1.7}$$

which is true by the induction hypothesis. □

7.2. Main result

Below is a generalization of the theorem in Alzer (2018).

Theorem 7.2.1. *For any $d \in \mathbb{N}$, $M \in (0,\infty)$, $\boldsymbol{x} \in Int(\mathcal{S})$, $x_{d+1} \mathrel{\hat{=}} 1-\|\boldsymbol{x}\| > 0$, and any $\boldsymbol{\gamma} \in [0,\infty)^d$ such that $\|\boldsymbol{\gamma}\| \le M$ and $\gamma_{d+1} \mathrel{\hat{=}} M-\|\boldsymbol{\gamma}\| \ge 0$, the function*

$$g(a) \mathrel{\hat{=}} \frac{\Gamma(aM+1)}{\prod_{i=1}^{d+1}\Gamma(a\gamma_i+1)}\prod_{i=1}^{d+1}x_i^{a\gamma_i} \tag{7.2.1}$$

is completely monotonic on $(0,\infty)$.

Remark 7.2.1. *In the proof of Theorem 7.2.1, we will show that $(-\log g)'$ is completely monotonic on $(0,\infty)$, which is a stronger statement by Lemma 7.1.2.*

Remark 7.2.2. *Soon after the first version of the present paper was posted on arXiv.org, Qi et al. (2018) gave an alternative proof of the complete monotonicity of $(-\log g)'$ and rewrote the combinatorial inequalities of Section 7.3 in terms of multivariate beta functions.*

Proof. Let $M \in (0,\infty)$, $\boldsymbol{x} \in \mathrm{Int}(\mathcal{S})$ and $a > 0$. The theorem in Alzer (2018) proves our statement in the case $d = 1$ (when the components of $\boldsymbol{\gamma}$ are integers, but the adjustment is trivial). Therefore, fix $d \ge 2$ and assume that the theorem is true for any smaller integer. If there exists $i \in [d+1]$ such that $\gamma_i = 0$, the theorem reduces to proving that (7.2.1) is completely monotonic for a d that is smaller then the one that we previously fixed, which is true by the induction hypothesis. Thus, assume for the remainder of the proof that

$$\gamma_i > 0, \quad \text{for all } i \in [d+1]. \tag{7.2.2}$$

Define

$$h(a) \doteq -\log g(a) = -\log\Gamma(aM+1) + \sum_{i=1}^{d+1}\log\Gamma(a\gamma_i+1) - a\sum_{i=1}^{d+1}\gamma_i\log x_i. \tag{7.2.3}$$

Then,

$$h'(a) = -M\psi(aM+1) + \sum_{i=1}^{d+1}\gamma_i\psi(a\gamma_i+1) - \sum_{i=1}^{d+1}\gamma_i\log x_i, \tag{7.2.4}$$

where $\psi \doteq (\log\Gamma)' = \Gamma'/\Gamma$. Using the integral representation

$$\psi'(z) = \int_0^\infty \frac{te^{-(z-1)t}}{e^t-1}dt, \quad z>0, \tag{7.2.5}$$

see (Abramowitz and Stegun, 1964, p.260), we obtain (take $t = s/M$ and $t = s/\gamma_i$)

$$\begin{aligned} h''(a) &= -M^2\psi'(aM+1) + \sum_{i=1}^{d+1}\gamma_i^2\psi'(a\gamma_i+1) \\ &= -M^2\int_0^\infty \frac{te^{-aMt}}{e^t-1}dt + \sum_{i=1}^{d+1}\gamma_i^2\int_0^\infty \frac{te^{-a\gamma_i t}}{e^t-1}dt \\ &= -\int_0^\infty se^{-as}J_{\boldsymbol{\gamma}/M}(e^{s/M})ds, \end{aligned} \tag{7.2.6}$$

where $J_{\boldsymbol{u}}(y)$ is defined in (7.1.4). Applying Lemma 7.1.3 gives

$$(-1)^n h^{(n+1)}(a) = \int_0^\infty s^n e^{-as}J_{\boldsymbol{\gamma}/M}(e^{s/M})ds > 0, \quad n\in\mathbb{N},\ a>0. \tag{7.2.7}$$

If we show that $h'(a) > 0$ for $a > 0$, then h' will be completely monotonic under Definition 7.1.1 and we will be able to conclude that g is completely monotonic by Lemma 7.1.2. Since h' is decreasing (see (7.2.7) when $n = 1$), we show that $\lim_{a\to\infty} h'(a) \geq 0$ to conclude the proof.

If we apply the recurrence formula

$$\psi(z+1) = \psi(z) + \frac{1}{z}, \quad z>0, \tag{7.2.8}$$

see (Abramowitz and Stegun, 1964, p.258), we obtain from (7.2.4) the representation

$$h'(a) = \frac{d}{a} - MR(aM) + \sum_{i=1}^{d+1}\gamma_i R(a\gamma_i) + \sum_{i=1}^{d+1}\gamma_i\log\left(\frac{\gamma_i/M}{x_i}\right), \tag{7.2.9}$$

where $R(z) \doteq \psi(z) - \log z$.

Using the asymptotic formula

$$\psi(z) \sim \log z - \frac{1}{2z} - \dots \quad (\text{as } z \to \infty) \tag{7.2.10}$$

see (Abramowitz and Stegun, 1964, p.259), we conclude from (7.2.9) and Jensen's inequality (for the convex function $-\log(\cdot)$ and the probability weights $P_i \circeq \gamma_i/M$ and $Q_i \circeq x_i$) that

$$\lim_{a\to\infty} h'(a) = M \sum_{i=1}^{d+1} \frac{\gamma_i}{M} \log\left(\frac{\gamma_i/M}{x_i}\right) \geq -M \log\left(\sum_{i=1}^{d+1} x_i\right) = 0. \tag{7.2.11}$$

This ends the proof. $\square$

Remark 7.2.3. *Interestingly, the sum on the left-end side of the inequality in (7.2.11) is the Kullback-Leibler divergence $D_{KL}(P\|Q)$. It is well defined because of (7.2.2) and the fact that $\boldsymbol{x} \in \mathrm{Int}(\mathcal{S})$ by hypothesis (which implies $0 < x_i < 1$ for all $i \in [d+1]$).*

7.3. Some combinatorial inequalities

In the context of Theorem 7.2.1, define

$$C(a) \circeq \frac{\Gamma(aM+1)}{\prod_{i=1}^{d+1} \Gamma(a\gamma_i + 1)}, \quad a \in (0,\infty). \tag{7.3.1}$$

Below are three simple combinatorial inequalities for the multinomial coefficients in (7.3.1). They generalize the ones proved in Alzer (2018) for binomial coefficients.

Corollary 7.3.1. *Let $k \in \mathbb{N}$ and let $a_j \in (0,\infty)$, $\lambda_j \in (0,1)$, $j \in \{1,2,\dots,k\}$, be such that $\sum_{j=1}^k \lambda_j = 1$. The following inequalities hold :*

(a) $C(\sum_{j=1}^k \lambda_j a_j) \leq \prod_{j=1}^k C(a_j)^{\lambda_j}$, where equality holds if and only if all the a_j's are the same.

(b) $\prod_{j=1}^k C(a_j) < C(\sum_{j=1}^k a_j)$.

(c) If $a_1 \leq a_3$, then $C(a_1+a_2)C(a_3) \leq C(a_1)C(a_2+a_3)$, where equality holds if and only if $a_1 = a_3$.

Proof. By (7.2.7) in the case $n = 1$, we know that g in the statement of Theorem 7.2.1 is strictly log-convex, which implies (a) by definition. Point (b) follows from Lemma 3 in Alzer (2018) because g is differentiable on $[0,\infty)$, $g(0) = 1$ and g is (strictly) positive, (strictly) decreasing and strictly log-convex on $(0, \infty)$. Point (c) follows from a trivial adaptation of the proof of Corollary 3 in Alzer (2018) using (7.2.7). $\square$

7.4. Application to Bernstein estimators on the simplex

In recent years, there has been a sustained interest in the study of statistical properties of Bernstein estimators on the unit hypercube, whether we talk about the cumulative distribution function (cdf) estimators

$$\hat{F}_{m,n}(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^d \cap [0,m]^d} F_n(\boldsymbol{k}/m) \prod_{i=1}^d \binom{m}{k_i} x_i^{k_i} (1 - x_i)^{k_i}, \quad \boldsymbol{x} \in [0,1]^d, \tag{7.4.1}$$

where F_n denotes the empirical cdf (given a random sample $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_n$ from an unknown cdf F), or the density estimators

$$\hat{f}_{m,n}(\boldsymbol{x}) = m^d \sum_{\boldsymbol{k} \in \mathbb{N}_0^d \cap [0,m-1]^d} \mathbb{P}_n\left(\left(\frac{\boldsymbol{k}}{m}, \frac{\boldsymbol{k}+\boldsymbol{1}}{m}\right]\right) \prod_{i=1}^d \binom{m-1}{k_i} x_i^{k_i} (1 - x_i)^{k_i}, \quad \boldsymbol{x} \in [0,1]^d, \tag{7.4.2}$$

where $\mathbb{P}_n$ denotes the empirical measure. For more information, the reader is referred to Babu et al. (2002), Babu and Chaubey (2006), Belalia (2016), Belalia et al. (2017), Ghosal (2001), Igarashi and Kakizawa (2014), Kakizawa (2011), Janssen et al. (2012, 2014, 2017), Leblanc and Johnson (2007), Leblanc (2009, 2010, 2012b,a), Lu (2015), Petrone (1999), Prakasa Rao (2005), Tenbusch (1994) and Vitale (1975).

One clear advantage of Bernstein estimators over kernel estimators (for example) is that they generally perform better near the boundary, see e.g. Leblanc (2012a). To our knowledge, the statistical properties of Bernstein estimators on the simplex (see (7.1.1)), and the associated density estimators, have never been studied in the literature, except in the univariate case where they coincide with (7.4.1) and (7.4.2) above, and except for the pointwise mean squared error of the density estimator in Tenbusch (1994) when $d = 2$. This subject is worth investigating because there are instances in practice where the distribution that we would like to estimate lives naturally on the d-dimensional simplex. One such example is the Dirichlet distribution, which is the conjugate prior of the multinomial distribution in Bayesian estimation, see e.g. Lange (1995) for an application in the context of allele frequency estimation in genetics. In those instances, we would expect that the estimators defined on the simplex perform better than the ones defined on the unit hypercube, especially near the boundary $\|\boldsymbol{x}\| = 1$.

Following Leblanc and Johnson (2007) and Leblanc (2010), define

$$S_{r,s,m}(\boldsymbol{x}) \mathrel{\hat{=}} \sum_{\boldsymbol{k}\in\mathbb{N}_0^d : \|\boldsymbol{k}\|\leq m} P_{r\boldsymbol{k},rm}(\boldsymbol{x}) P_{s\boldsymbol{k},sm}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathcal{S},$$

for $r, s, m \in \mathbb{N}$. This family of polynomials would arise in the context of statistical density estimation based on the Bernstein estimators in (7.1.1) (see e.g. the appendix in Leblanc (2010)). Theorem 7.2.1 will be used to prove Proposition 7.4.2 below.

The following lemma generalizes Theorem 1.1 (iii) in Leblanc and Johnson (2007), and Lemma 3 (ii) and (iv) in Leblanc (2010) when $j = 0$.

Lemma 7.4.1. *Let $d, r, s, m \in \mathbb{N}$, $\boldsymbol{x} \in Int(\mathcal{S})$, and define the covariance matrix*

$$\boldsymbol{\Sigma} \mathrel{\hat{=}} rs(r+s)(diag(\boldsymbol{x}) - \boldsymbol{x}\boldsymbol{x}^T). \tag{7.4.3}$$

We have

$$m^{d/2} S_{r,s,m}(\boldsymbol{x}) = \phi_{r,s}(\boldsymbol{x}) + o_{\boldsymbol{x}}(1), \quad \text{as } m \to \infty,$$

where

$$\phi_{r,s}(\boldsymbol{x}) \mathrel{\hat{=}} \frac{(gcd(r,s))^d}{(2\pi)^{d/2}(\det(\boldsymbol{\Sigma}))^{1/2}}. \tag{7.4.4}$$

Proof. Let $\boldsymbol{U}_1, \dots, \boldsymbol{U}_m$ and $\boldsymbol{V}_1, \dots, \boldsymbol{V}_m$ be two (independent) sequences of independent random vectors such that $\boldsymbol{U}_i \sim \text{Multinomial}(r, \boldsymbol{x})$ and $\boldsymbol{V}_i \sim \text{Multinomial}(s, \boldsymbol{x})$ for each $i \in [d]$. Now, let $\boldsymbol{H} \mathrel{\hat{=}} \text{gcd}(r,s)\boldsymbol{I}_d$ where $\boldsymbol{I}_d$ is the $d \times d$ identity matrix, and define $\boldsymbol{W}_i \mathrel{\hat{=}} s\boldsymbol{U}_i - r\boldsymbol{V}_i$ so that the j-th component of $\boldsymbol{W}_i$ has a lattice distribution with span $H_{jj} = \text{gcd}(r,s)$. Note that $\boldsymbol{W}_i^\star \mathrel{\hat{=}} \boldsymbol{H}^{-1}\boldsymbol{W}_i$ has span 1 in all d directions. The covariance matrix of $\boldsymbol{W}_i$ is given by $\boldsymbol{\Sigma}$ in (7.4.3). We can write $S_{r,s,m}(\boldsymbol{x})$ in terms of the $\boldsymbol{W}_i^\star$'s as

$$S_{r,s,m}(\boldsymbol{x}) = \mathbb{P}\left(\sum_{i=1}^m s\boldsymbol{U}_i = \sum_{i=1}^m r\boldsymbol{V}_i\right) = \mathbb{P}\left(\sum_{i=1}^m \boldsymbol{W}_i^\star = \boldsymbol{0}\right).$$

Therefore, using Theorem 3.1 of Athreya and Janicki (2016) (a local central limit theorem for random vectors with lattice distributions), $\det(\boldsymbol{H}) = (\text{gcd}(r,s))^d$ and the fact that the covariance matrix of $\boldsymbol{W}_i^\star$ is equal to $\boldsymbol{H}^{-1}\boldsymbol{\Sigma}\boldsymbol{H}^{-1}$, we obtain the conclusion. $\square$

The following proposition generalizes Lemma 4 in Leblanc (2010) when $j = 0$.

Proposition 7.4.2. *Let $r, s, m \in \mathbb{N}$ and let $h : \mathcal{S} \to \mathbb{R}$ be any bounded measurable function. As $m \to \infty$,*

(a) $m^{d/2} \int_{\mathcal{S}} S_{r,s,m}(\boldsymbol{x}) d\boldsymbol{x} = \frac{2^{-d}\sqrt{\pi}}{\Gamma(d/2+1/2)} + O(m^{-1}) = \int_{\mathcal{S}} \phi_{r,s}(\boldsymbol{x}) d\boldsymbol{x} + O(m^{-1})$,

(b) $\int_{\mathcal{S}} h(\boldsymbol{x})(m^{d/2} S_{r,s,m}(\boldsymbol{x}) - \phi_{r,s}(\boldsymbol{x})) d\boldsymbol{x} = o(1)$.

Proof. Assume for now that $r = s = 1$. We have

$$
\begin{aligned}
\int_{\mathcal{S}} S_{1,1,m}(\boldsymbol{x}) d\boldsymbol{x} &= \sum_{\|\boldsymbol{k}\| \leq m} \int_{\mathcal{S}} (P_{\boldsymbol{k},m}(\boldsymbol{x}))^2 d\boldsymbol{x} = \sum_{\|\boldsymbol{k}\| \leq m} \left(\frac{\Gamma(m+1)}{\prod_{i=1}^{d+1} \Gamma(k_i+1)} \right)^2 \int_{\mathcal{S}} \prod_{i=1}^{d+1} x_i^{2k_i} d\boldsymbol{x} \\
&= \sum_{\|\boldsymbol{k}\| \leq m} \left(\frac{\Gamma(m+1)}{\prod_{i=1}^{d+1} \Gamma(k_i+1)} \right)^2 \frac{\prod_{i=1}^{d+1} \Gamma(2k_i+1)}{\Gamma(2m+d+1)} \\
&= \frac{(\Gamma(m+1))^2}{\Gamma(2m+d+1)} \sum_{\|\boldsymbol{k}\| \leq m} \prod_{i=1}^{d+1} \binom{2k_i}{k_i}. \qquad (7.4.5)
\end{aligned}
$$

To obtain the third equality, we used the normalization constant for the Dirichlet distribution. Note that

$$
\begin{aligned}
\sum_{\|\boldsymbol{k}\| \leq m} \prod_{i=1}^{d+1} \binom{2k_i}{k_i} &= (-4)^m \sum_{\|\boldsymbol{k}\| \leq m} \prod_{i=1}^{d+1} \frac{1}{(-4)^m} \binom{2k_i}{k_i} = (-4)^m \sum_{\|\boldsymbol{k}\| \leq m} \prod_{i=1}^{d+1} \binom{-1/2}{k_i} \\
&= (-4)^m \binom{-(d+1)/2}{m} \\
&= \binom{m + \frac{d-1}{2}}{m} 4^m, \qquad (7.4.6)
\end{aligned}
$$

where the last three equalities follow, respectively, from (5.37), the Chu-Vandermonde convolution (p. 248), and (5.14) in Graham et al. (1994). By applying (7.4.6) and the duplication formula

$$
4^y = \frac{2\sqrt{\pi}\,\Gamma(2y)}{\Gamma(y)\Gamma(y+1/2)}, \quad y \in (0, \infty), \qquad (7.4.7)
$$

see (Abramowitz and Stegun, 1964, p.256), in (7.4.5), we get

$$
\begin{aligned}
\int_{\mathcal{S}} S_{1,1,m}(\boldsymbol{x}) d\boldsymbol{x} &= \frac{(\Gamma(m+1))^2}{\Gamma(2m+d+1)} \cdot \frac{\Gamma(m+d/2+1/2)}{\Gamma(m+1)\Gamma(d/2+1/2)} \cdot 4^m \\
&= \frac{2\sqrt{\pi}\,\Gamma(m+1)}{\Gamma(d/2+1/2)\Gamma(m+d/2+1)} \cdot \frac{\Gamma(m+d/2+1/2)\Gamma(m+d/2+1)}{2\sqrt{\pi}\,\Gamma(2m+d+1)} \cdot 4^m
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{2\sqrt{\pi}\,\Gamma(m+1)}{\Gamma(d/2+1/2)\Gamma(m+d/2+1)} \cdot \frac{4^m}{4^{m+d/2+1/2}} \\
&= \frac{2^{-d}\sqrt{\pi}\,\Gamma(m+1)}{\Gamma(d/2+1/2)\Gamma(m+d/2+1)} \\
&= \begin{cases} \frac{2^{-d}\sqrt{\pi}}{\Gamma(d/2+1/2)} \prod_{i=1}^{d/2} (m+i)^{-1}, & \text{if } d \text{ is even}, \\ \frac{2^{-d}\sqrt{\pi}}{\Gamma(d/2+1/2)} \prod_{i=1}^{d/2+1/2} (m+d/2+1-i)^{-1} \cdot \frac{\Gamma(m+1)}{\Gamma(m+1/2)}, & \text{if } d \text{ is odd}. \end{cases}
\end{aligned}
$$

Using the fact that

$$\frac{\Gamma(m+1)}{m^{1/2}\Gamma(m+1/2)} = 1 + \frac{1}{8m} + O(m^{-2}), \tag{7.4.8}$$

see (Abramowitz and Stegun, 1964, p.257), we obtain

$$m^{d/2} \int_{\mathcal{S}} S_{1,1,m}(\boldsymbol{x}) d\boldsymbol{x} = \frac{2^{-d}\sqrt{\pi}}{\Gamma(d/2+1/2)} + O(m^{-1}). \tag{7.4.9}$$

In the case $r = s = 1$, the expression for $\boldsymbol{\Sigma}$ in (7.4.3) is equal to $2(\mathrm{diag}(\boldsymbol{x}) - \boldsymbol{x}\boldsymbol{x}^T)$. Using the square-root-free symbolic Cholesky decomposition for covariance matrices of multinomial distributions (see Theorem 1 in Tanabe and Sagae (1992)), we deduce that $\det(\boldsymbol{\Sigma}) = 2^d \det(\mathrm{diag}(\boldsymbol{x}) - \boldsymbol{x}\boldsymbol{x}^T) = 2^d \prod_{i=1}^{d+1} x_i$. Therefore,

$$
\begin{aligned}
\int_{\mathcal{S}} \frac{1}{(2\pi)^{d/2}(\det(\boldsymbol{\Sigma}))^{1/2}} d\boldsymbol{x} &= \frac{1}{2^d \pi^{d/2}} \int_{\mathcal{S}} \prod_{i=1}^{d+1} x_i^{1/2-1} d\boldsymbol{x} = \frac{1}{2^d \pi^{d/2}} \cdot \frac{(\Gamma(1/2))^{d+1}}{\Gamma(d/2+1/2)} \\
&= \frac{2^{-d}\sqrt{\pi}}{\Gamma(d/2+1/2)}.
\end{aligned} \tag{7.4.10}
$$

Together with (7.4.9) and (7.4.4), this proves (a) for $r = s = 1$.

Now, the almost-everywhere convergence from Lemma 7.4.1 and the mean convergence from (a) imply that $\{S_{1,1,m}(\cdot)\}_{m\in\mathbb{N}}$ is uniformly integrable, see (Shiryaev, 1996, p.189). By Theorem 7.2.1, $a \mapsto P_{a\boldsymbol{k},am}$ is decreasing on $(0,\infty)$, so

$$S_{r,s,m}(\boldsymbol{x}) \leq \sum_{\|\boldsymbol{k}\|\leq m} (P_{\boldsymbol{k},m}(\boldsymbol{x}))^2 = S_{1,1,m}(\boldsymbol{x}), \tag{7.4.11}$$

which implies that $\{S_{r,s,m}(\cdot)\}_{m\in\mathbb{N}}$ is also uniformly integrable. Hence, by Lemma 7.4.1, we must have (a) in the general case $r, s \in \mathbb{N}$. Finally, the almost-everywhere convergence and the uniform integrability imply the L^1 convergence, so (b) follows immediately from Jensen's inequality and the fact that h is bounded. $\square$

Acknowledgements

I would like to thank Alexandre Leblanc for promptly giving me access to Leblanc (2006). This work is supported by a NSERC Doctoral Program Alexander Graham Bell scholarship (CGS D3).

Bibliography

M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964. MR0167642.

H. Alzer. Complete monotonicity of a function related to the binomial probability. *J. Math. Anal. Appl.*, 459(1):10–15, 2018. MR3730425.

K. B. Athreya and R. Janicki. Asymptotics of powers of binomial and multinomial probabilities. *Statist. Probab. Lett.*, 112:58–62, 2016. MR3475488.

G. J. Babu and Y. P. Chaubey. Smooth estimation of a distribution and density function on a hypercube using Bernstein polynomials for dependent random vectors. *Statist. Probab. Lett.*, 76(9):959–969, 2006. MR2270097.

G. J. Babu, A. J. Canty, and Y. P. Chaubey. Application of Bernstein polynomials for smooth estimation of a distribution and density function. *J. Statist. Plann. Inference*, 105(2):377–392, 2002. MR1910059.

M. Belalia. On the asymptotic properties of the Bernstein estimator of the multivariate distribution function. *Statist. Probab. Lett.*, 110:249–256, 2016. MR3474765.

M. Belalia, T. Bouezmarni, and A. Leblanc. Smooth conditional distribution estimators using Bernstein polynomials. *Comput. Statist. Data Anal.*, 111:166–182, 2017. MR3630225.

M. J. Dubourdieu. Sur un théorème de M. S. Bernstein relatif à la transformation de Laplace-Stieltjes. *Compositio Math.*, 7:96–111, 1939. MR0000436.

S. Ghosal. Convergence rates for density estimation with Bernstein polynomials. *Ann. Statist.*, 29(5):1264–1280, 2001. MR1873330.

R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete mathematics*. Addison-Wesley Publishing Company, Reading, MA, second edition, 1994. MR1397498.

G. Igarashi and Y. Kakizawa. On improving convergence rate of Bernstein polynomial density estimator. *J. Nonparametr. Stat.*, 26(1):61–84, 2014. MR3174309.

P. Janssen, J. Swanepoel, and N. Veraverbeke. Large sample behavior of the Bernstein copula estimator. *J. Statist. Plann. Inference*, 142(5):1189–1197, 2012. MR2879763.

P. Janssen, J. Swanepoel, and N. Veraverbeke. A note on the asymptotic behavior of the Bernstein estimator of the copula density. *J. Multivariate Anal.*, 124:480–487, 2014. MR3147339.

P. Janssen, J. Swanepoel, and N. Veraverbeke. Smooth copula-based estimation of the conditional density function with a single covariate. *J. Multivariate Anal.*, 159:39–48, 2017. MR3668546.

Y. Kakizawa. A note on generalized Bernstein polynomial density estimators. *Stat. Methodol.*, 8(2):136–153, 2011. MR2769276.

K. Lange. Applications of the Dirichlet distribution to forensic match probabilities. *Genetica*, 96(1-2):107–117, 1995. doi:10.1007/BF01441156.

A. Leblanc. A Bias-Corrected Approach to Density Estimation Using Bernstein Polynomials. *Technical Report*, pages 1–24, 2006. University of Manitoba, Dept. of Statistics.

A. Leblanc. Chung-Smirnov property for Bernstein estimators of distribution functions. *J. Nonparametr. Stat.*, 21(2):133–142, 2009. MR2488150.

A. Leblanc. A bias-reduced approach to density estimation using Bernstein polynomials. *J. Nonparametr. Stat.*, 22(3-4):459–475, 2010. MR2662607.

A. Leblanc. On the boundary properties of Bernstein polynomial estimators of density and distribution functions. *J. Statist. Plann. Inference*, 142(10):2762–2778, 2012a. MR2925964.

A. Leblanc. On estimating distribution functions using Bernstein polynomials. *Ann. Inst. Statist. Math.*, 64(5):919–943, 2012b. MR2960952.

A. Leblanc and B. C. Johnson. On a uniformly integrable family of polynomials defined on the unit interval. *JIPAM. J. Inequal. Pure Appl. Math.*, 8(3): Article 67, 5pp., 2007. MR2345922.

L. Lu. On the uniform consistency of the Bernstein density estimator. *Statist. Probab. Lett.*, 107:52–61, 2015. MR3412755.

K. S. Miller and S. G. Samko. Completely monotonic functions. *Integral Transform. Spec. Funct.*, 12(4):389–402, 2001. MR1872377.

S. Petrone. Bayesian density estimation using Bernstein polynomials. *Canad. J. Statist.*, 27(1):105–126, 1999. MR1703623.

B. L. S. Prakasa Rao. Estimation of distribution and density functions by generalized Bernstein polynomials. *Indian J. Pure Appl. Math.*, 36(2):63–88, 2005. MR2153833.

F. Qi, D.-W. Niu, D. Lim, and B.-N. Guo. Some logarithmically completely monotonic functions and inequalities for multinomial coefficients and multivariate beta functions. *Preprint, HAL archives-ouvertes*, pages 1–13, 2018. hal-01769288.

A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996. MR1368405.

K. Tanabe and M. Sagae. An exact Cholesky decomposition and the generalized inverse of the variance-covariance matrix of the multinomial distribution, with applications. *J. Roy. Statist. Soc. Ser. B*, 54(1):211–219, 1992. MR1157720.

A. Tenbusch. Two-dimensional Bernstein polynomial density estimators. *Metrika*, 41(3-4): 233–253, 1994. MR1293514.

H. van Haeringen. Completely monotonic and related functions. *J. Math. Anal. Appl.*, 204 (2):389–408, 1996. MR1421454.

R. A. Vitale. Bernstein polynomial approach to density function estimation. In *Statistical Inference and Related Topics (Proc. Summer Res. Inst. Statist. Inference for Stochastic Processes, Indiana Univ., Bloomington, Ind., 1974, Vol. 2; dedicated to Z. W. Birnbaum)*, pages 87–99. Academic Press, New York, 1975. MR0397977.

Article 8

A uniform L^1 law of large numbers for functions of i.i.d. random variables that are translated by a consistent estimator

by

Pierre Lafaye de Micheaux[1] and Frédéric Ouimet[2]

([1]) UNSW Sydney, Sydney (Kensington), NSW 2052 Australia.

([2]) Université de Montréal, Montréal, QC H3T 1J4, Canada.

Here is the complete citation :

Lafaye de Micheaux, P. and Ouimet, F. (2018). A uniform L^1 law of large numbers for functions of i.i.d. random variables that are translated by a consistent estimator. Statist. Probab. Lett. 142, 109-117. https://doi.org/10.1016/j.spl.2018.06.006.

Contributions

I wrote the article and the proofs alone. Pierre helped me refine the introduction. Pierre was the one to originally raise the question about the convergence of

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \neq \mathrm{Median}(\boldsymbol{X})\}} \log |X_i - \mathrm{Median}(\boldsymbol{X})|$$

when the X_i's are i.i.d. random variables with Laplace(μ) distribution. It was the motivation behind the article.

Abstract. We develop a new L^1 law of large numbers where the i-th summand is given by a function $h(\cdot)$ evaluated at $X_i - \theta_n$, and where $\theta_n \doteq \theta_n(X_1, X_2, \ldots, X_n)$ is an estimator converging in probability to some parameter $\theta \in \mathbb{R}$. Under broad technical conditions, the convergence is shown to hold uniformly in the set of estimators interpolating between θ and another consistent estimator $\theta_n^\star$. Our main contribution is the treatment of the case where $|h|$ blows up at 0, which is not covered by standard uniform laws of large numbers.

Keywords: uniform law of large numbers, Taylor expansion, M-estimators, score function

8.1. Introduction

Let $X_1, X_2, X_3, \ldots$ be a sequence of i.i.d. random variables and consider the statistic $T_n(\theta_n^\star)$ where the random variable $T_n(\theta) \doteq T_n(X_1, X_2, \ldots, X_n; \theta) : \Omega \to \mathbb{R}$ depends on an unknown parameter $\theta \in \mathbb{R}$ for which we have a consistent sequence of estimators $\theta_n^\star \doteq \theta_n^\star(X_1, X_2, \ldots, X_n)$. Assume further that the following first-order Taylor expansion is valid

$$T_n(\theta_n^\star) = T_n(\theta) + (\theta_n^\star - \theta) \int_0^1 T_n'(\theta + v(\theta_n^\star - \theta)) dv, \tag{8.1.1}$$

where

$$T_n'(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \neq t\}} h(X_i - t), \tag{8.1.2}$$

and where $h : \mathbb{R} \backslash \{0\} \to \mathbb{R}$ is a measurable function (possibly nonlinear). In statistics, one is often interested in knowing if estimating a parameter (θ here) has an impact on the asymptotic law of a given statistic. See for example the interesting results of de Wet and Randles (1987) in the context of limiting χ^2 U and V statistics. Equations (8.1.1) and (8.1.2) provide a natural setting for studying the question of whether or not $T_n(\theta_n^\star) - T_n(\theta) \to 0$ whenever $\theta_n^\star \to \theta$, as $n \to \infty$.

Given some regularity conditions on the behavior of $h(\cdot)$ around the origin and in its tails, proving the convergence to $\mathbb{E}[h(X_1 - \theta)]$, in probability say, of the integral on the right-hand side of (8.1.1) is often possible under weak assumptions by adapting standard uniform laws of large numbers. For instance, one can use (Ferguson, 1996, Theorem 16 (a)), which was introduced by LeCam (1953) and Rubin (1956). One can also use entropy conditions:

see, e.g., (van de Geer, 2000, Chapter 3) and (van der Vaart and Wellner, 1996, Section 2.4). Some of these theorems go back to or evolved from the works of Blum (1955), Dehardt (1971), Vapnik and Červonenkis (1971); Vapnik and Chervonenkis (1981), Giné and Zinn (1984), Pollard (1984) and Talagrand (1987). For extensive notes on the origins of the entropy conditions, we refer the interested reader to (van de Geer, 2000, Section 3.8) and (Pollard, 1984, pp. 36–38).

However, when $|h|$ blows up at 0, namely when $\limsup_{x\to 0} |h(x)| = \infty$, these results are not applicable because the enveloppe function $h^{\sup}(x) \mathrel{\hat{=}} \sup_{t:|t-\theta|<\delta} \mathbf{1}_{\{x\neq t\}}|h(x-t)|$ is infinite in any small enough neighborhood of θ and, in particular, $h^{\sup}(X_1)$ is not integrable for the outer measure.

We faced such a problem when analysing the convergence of score functions in the context of testing the goodness-of-fit of the Laplace distribution with unknown location and scale parameters (μ, σ). If the family of alternatives is taken to be the asymmetric power distribution (Komunjer, 2007) or the skewness exponential power distribution (Fernández et al., 1995), a score function evaluated at the maximum likelihood estimator $(\mu_n^\star, \sigma_n^\star)$ can be used, in the spirit of (Desgagné et al., 2013; Desgagné and Lafaye de Micheaux, 2018). If the score function is expanded around (μ, σ), then a multivariate version of (8.1.1) is obtained. One of the integrals in the expansion will have an integrand (8.1.2) where $h(\cdot)$ contains a logarithmic term. Standard uniform laws of large numbers cannot be applied to show the convergence of such integrals because the enveloppe function of the class of functions $\{\log(\,\cdot\, - t)\}_{t:|t-\mu|<\delta}$ is infinite in any small enough neighborhood of μ. In section 8.3, we show how the main result of this paper (Theorem 8.2.4) can be used to prove a crucial part of the problem described above.

More generally, the main result is that, under broad conditions, one obtains

$$\lim_{n\to\infty} \sup_{v\in[0,1]} \mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i\neq\theta+v(\theta_n^\star-\theta)\}} h(X_i-\theta-v(\theta_n^\star-\theta)) - \mathbb{E}\Big[h(X_1-\theta)\Big]\right| = 0. \tag{8.1.3}$$

From (8.1.3) and the setting above, one can conclude that $T_n(\theta_n^\star) - T_n(\theta) \to 0$ in probability as $n \to \infty$.

8.2. A new uniform L^1 law of large numbers

Throughout the paper, the labels $(X.k)$, $(H.k)$ and $(E.k)$ denote, respectively, assumptions that we will make on X_1, $h(\cdot)$ and θ_n. Figure 8.2.1 at the end of the current section illustrates the logical structure of these assumptions and their implications. We start by proving a non-uniform version of Theorem 8.2.4.

Proposition 8.2.1. *Let $\theta \in \mathbb{R}$ and let $X_1, X_2, X_3, \ldots$ be a sequence of i.i.d. random variables such that*

(X.1): $\mathbb{P}(X_1 = \theta) = 0$.

Let $h : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ be a mesurable function that satisfies

(H.1): $\mathbb{P}(X_1 - \theta \in \mathcal{D}_h) = 0$, *where $\mathcal{D}_h$ is the set of discontinuity points of $h(\cdot)$,*

(H.2): $\mathbb{E}\,|h(X_1 - \theta)| < \infty$.

Let $\theta_n \mathrel{\hat=} \theta_n(X_1, X_2, \ldots, X_n)$ be an estimator that satisfies

(E.1): $\theta_n \stackrel{\mathbb{P}}{\longrightarrow} \theta$,

(E.2): *For all $n \in \mathbb{N}$ and all $i \in \{1, 2, \ldots, n\}$, $(X_i - \theta_n, X_i - \theta) \stackrel{law}{=} (X_1 - \theta_n, X_1 - \theta)$,*

(E.3): *There exists $N_0 \in \mathbb{N}$ such that $\left\{\mathbf{1}_{\{X_1 \neq \theta_n\}} h(X_1 - \theta_n)\right\}_{n \geq N_0}$ is uniformly integrable.*

Then,

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \neq \theta_n\}} h(X_i - \theta_n) - \mathbb{E}\Big[h(X_1 - \theta)\Big]\right| \longrightarrow 0. \tag{8.2.1}$$

Remark 8.2.1. *Condition $(E.2)$ is satisfied for any estimator that is symmetric with respect to its n variables. For example, this is the case for any maximum likelihood estimator that is based on i.i.d. observations.*

Proof of Proposition 8.2.1. From $(X.1)$ and $(E.1)$, we know that $\mathbf{1}_{\{X_1 = \theta_n\}} \stackrel{\mathbb{P}}{\longrightarrow} 0$. Indeed, for any $\varepsilon > 0$,

- take $\delta \mathrel{\hat=} \delta_\varepsilon > 0$ such that $\mathbb{P}(|X_1 - \theta| < \delta) < \varepsilon/2$, and
- take $N \mathrel{\hat=} N_{\delta,\varepsilon}$ such that for all $n \geq N$, we have $\mathbb{P}(|\theta_n - \theta| \geq \delta) < \varepsilon/2$.

We get, for all $n \geq N$,

$$\mathbb{P}(X_1 = \theta_n) \leq \mathbb{P}(X_1 = \theta_n, |\theta_n - \theta| < \delta) + \mathbb{P}(|\theta_n - \theta| \geq \delta) < \varepsilon.$$

In particular, this shows $\mathbf{1}_{\{X_1=\theta_n\}}|h(X_1-\theta)| \stackrel{\mathbb{P}}{\longrightarrow} 0$. Since this sequence is uniformly integrable by $(H.2)$, we also have the L^1 convergence. By using Jensen's inequality and $(E.2)$, we deduce

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\{X_i=\theta_n\}}h(X_i-\theta)\right| \leq \mathbb{E}\Big[\mathbf{1}_{\{X_1=\theta_n\}}|h(X_1-\theta)|\Big] \longrightarrow 0. \tag{8.2.2}$$

By $(H.2)$ and the law of large numbers in L^1 (see, e.g., Theorem 1.2.6 in Stroock (2011)), we also know that

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n}h(X_i-\theta)-\mathbb{E}\Big[h(X_1-\theta)\Big]\right| \longrightarrow 0. \tag{8.2.3}$$

By combining (8.2.2) and (8.2.3), we have shown

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\{X_i\neq\theta_n\}}h(X_i-\theta)-\mathbb{E}\Big[h(X_1-\theta)\Big]\right| \longrightarrow 0. \tag{8.2.4}$$

To conclude the proof, we show that

$$Y_n \mathrel{\hat=} \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\{X_i\neq\theta_n\}}h(X_i-\theta_n)-\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\{X_i\neq\theta_n\}}h(X_i-\theta) \stackrel{L^1}{\longrightarrow} 0.$$

From Jensen's inequality and $(E.2)$, we have

$$\mathbb{E}|Y_n| \leq \mathbb{E}\Big[\mathbf{1}_{\{X_1\neq\theta_n\}}\Big|h(X_1-\theta_n)-h(X_1-\theta)\Big|\Big]. \tag{8.2.5}$$

The sequence $\{\mathbf{1}_{\{X_1\neq\theta_n\}}|h(X_1-\theta_n)-h(X_1-\theta)|\}_{n\in\mathbb{N}}$ converges to 0 in probability by $(H.1)$, $(E.1)$ and the continuous mapping theorem (van der Vaart, 1998, Theorem 2.3). Furthermore, the sequence is uniformly integrable for $n \geq N_0$ by $(H.2)$, $(E.3)$ and the fact that the sums of random variables coming (respectively) from two uniformly integrable sequences form a uniformly integrable sequence. Hence, $Y_n \to 0$ in L^1. $\square$

Since the distribution of $X_1-\theta_n$ is rarely known, condition $(E.3)$ in Proposition 8.2.1 is impractical to verify. The next lemma fix this problem.

Lemma 8.2.2. *Let $\theta \in \mathbb{R}$. Let $X_1, X_2, X_3, \ldots$ be a sequence of i.i.d. random variables. Let $h : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ be a mesurable function. Let $\theta_n \mathrel{\hat=} \theta_n(X_1, X_2, \ldots, X_n)$ be an estimator that satisfies*

- **(E.4):** *If $\limsup_{x\to 0}|h(x)| < \infty$, we impose no condition. Otherwise, assume that there exist $N_1 \in \mathbb{N}$, $\alpha_0 > 0$ and a constant $C_{\alpha_0} > 0$ such that*

$$\sup_{n\geq N_1}\sup_{A\in\mathcal{B}_{>0}([-\alpha_0,\alpha_0])}\frac{\mathbb{P}(X_1-\theta_n\in A)}{\textit{Lebesgue}(A)}\leq C_{\alpha_0}<\infty,$$

where $\mathcal{B}_{>0}([-\alpha_0,\alpha_0])$ denotes the Borel sets of positive Lebesgue measure on the interval $[-\alpha_0,\alpha_0]$.

(E.5): *There exist $N_2\geq 2$, $C,\gamma,p>0$ and $\beta_0>\gamma$ such that, for $\mathbb{P}(X_1-\theta\in\cdot)$-almost-all $x\in\mathbb{R}$, we have*

— *For all $u\geq(x+\gamma)\vee\beta_0$ and for all $n\geq N_2$,*

$$\mathbb{P}(\theta_n-\theta\leq x-u\,|\,X_1-\theta=x)\leq Ce^{-|x-u|^p}.$$

— *For all $u\leq(x-\gamma)\wedge(-\beta_0)$ and for all $n\geq N_2$,*

$$\mathbb{P}(\theta_n-\theta\geq x-u\,|\,X_1-\theta=x)\leq Ce^{-|x-u|^p}.$$

(E.6): *There exists $N_3\in\mathbb{N}$ such that for all $n\geq N_3$, there exists $A_n\in\mathcal{B}(\mathbb{R})$ such that $\mathbb{P}(X_1-\theta\in A_n)=1$ and, for all $x\in A_n$, the conditional measure $\mathbb{P}(x-(\theta_n-\theta)\in\cdot\,|\,X_1-\theta=x)$, when restricted to $\{u\in\mathbb{R}:|u|\geq\beta_0,|x-u|>\gamma\}$, is absolutely continuous with respect to the Lebesgue measure.*

Assume that $h(\cdot)$ satisfies

(H.3): *For all $x_0\in\mathbb{R}\backslash\{0\}$, $\limsup_{x\to x_0}|h(x)|<\infty$,*

(H.4): $\int_{|u|\leq\alpha_0}|h(u)|du<\infty$,

(H.5): **(1):** *$h(\cdot)$ is absolutely continuous on bounded sub-intervals of $(-\infty,-\beta_0)\cup(\beta_0,+\infty)$;*

(2): *There exists an integrable random variable M such that $\sup_{|t|\leq\gamma}|h(X_1-\theta-t)|\,\mathbf{1}_{\{|X_1-\theta-t|\geq\beta_0\}}\leq M$ $\mathbb{P}$-almost-surely;*

(3): *$\lim_{|\beta|\to\infty}|h(\beta)|e^{-|x-\beta|^p}=0$ for $\mathbb{P}(X_1-\theta\in\cdot)$-almost-all $x\in\mathbb{R}$, and $\{|h(\beta)|e^{-|X_1-\theta-\beta|^p}\}_{|\beta|\geq\beta_0}$ is uniformly integrable;*

(4): $\int_{|u|\geq\beta_0}\mathbb{E}\Big[|h'(u)|\,e^{-|X_1-\theta-u|^p}\Big]du<\infty$*;*

(5): *For almost-all $|u|\geq\beta_0$, we have $-sign(u)sign(h(u))h'(u)\leq 0$.*

Then, $(E.3)$ from Proposition 8.2.1 is satisfied, namely $\Big\{\mathbf{1}_{\{X_1\neq\theta_n\}}h(X_1-\theta_n)\Big\}_{n\geq N_0}$ is uniformly integrable, where $N_0\doteq N_1\vee N_2\vee N_3$.

Remark 8.2.2. *If $X_1 - \theta_n$ has a density for n large enough and, in a neighborhood of 0, those densities are uniformly bounded from above by the same positive constant, then $(E.4)$ is satisfied. In general, when θ_n is even only slightly non-trivial, we rarely know the distribution of $X_1 - \theta_n$. However, if θ_n concentrates more and more around θ as $n \to \infty$ (like most maximum likelihood estimators for instance), then we expect the weight of the distribution of X_1 around θ to dominate the weight of the distribution of $X_1 - \theta_n$ around 0. In that case, we can expect $(E.4)$ to be satisfied when X_1 has a regular enough distribution around θ. Condition $(E.5)$ is a way to control the tail behavior of θ_n's distribution for the above heuristic to work. Since the lemma is intended to be used when $|h|$ blows up at 0, condition $(E.4)$ is there to control the distribution of $X_1 - \theta_n$ around 0.*

Proof. We want to prove that for $N_0 \doteq N_1 \vee N_2 \vee N_3$, we have

$$\lim_{K \to \infty} \sup_{n \geq N_0} \mathbb{E}\Big[\Big|h(X_1 - \theta_n)\Big| \mathbf{1}_{\{X_1 \neq \theta_n\} \cap \{|h(X_1 - \theta_n)| \geq K\}}\Big] = 0.$$

By $(H.3)$, $h(\cdot)$ is uniformly bounded on compact subsets of $\mathbb{R}\backslash\{0\}$. It is therefore sufficient to show both

$$\lim_{\alpha \to 0} \sup_{n \geq N_0} \mathbb{E}\Big[\Big|h(X_1 - \theta_n)\Big| \mathbf{1}_{\{X_1 \neq \theta_n\} \cap \{|X_1 - \theta_n| \leq \alpha\}}\Big] = 0, \tag{8.2.6}$$

$$\lim_{\beta \to \infty} \sup_{n \geq N_0} \mathbb{E}\Big[\Big|h(X_1 - \theta_n)\Big| \mathbf{1}_{\{|X_1 - \theta_n| \geq \beta\}}\Big] = 0. \tag{8.2.7}$$

When $\limsup_{x \to 0} |h(x)| < \infty$, then (8.2.6) is trivially satisfied because $h(\cdot)$ is uniformly bounded on compact subsets of $\mathbb{R}$ by $(H.3)$. When $\limsup_{x \to 0} |h(x)| = \infty$, then (8.2.6) follows directly from $(E.4)$, $(H.4)$ and the dominated convergence theorem (DCT).

Assume for the remaining of the proof that

$$n \geq N_0 \quad \text{and} \quad \beta > \beta_0 > \gamma,$$

where γ and β_0 are fixed in $(E.5)$. Separate the expectation in (8.2.7) in two parts :

$$\begin{aligned}(a) + (b) \doteq \mathbb{E}\Big[\Big|h(X_1 - \theta_n)\Big| \mathbf{1}_{\{|X_1 - \theta_n| \geq \beta\} \cap \{|\theta_n - \theta| \leq \gamma\}}\Big] \\ + \mathbb{E}\Big[\Big|h(X_1 - \theta_n)\Big| \mathbf{1}_{\{|X_1 - \theta_n| \geq \beta\} \cap \{|\theta_n - \theta| > \gamma\}}\Big].\end{aligned}$$

By $(H.5).2$ and the DCT, we have $(a) \to 0$ as $\beta \to \infty$, uniformly in n. For the term (b), condition on the value of $X_1 - \theta$, integrate by parts (see $(E.6)$ and $(H.5).1$) and then use

$(E.5)$ and $(H.5).5$. We obtain

$$
\begin{aligned}
(b) &= \int_{\{(u,x)\,:\,|u|\geq\beta,\ |x-u|>\gamma\}} |h(u)|\,\mathbb{P}((X_1-\theta_n, X_1-\theta)\in d(u,x)) \\
&= \int_{-\infty}^{\infty}\left(\int_{\{u\,:\,|u|\geq\beta,\ |x-u|>\gamma\}} |h(u)|\,\mathbb{P}(x-(\theta_n-\theta)\in du\,|\,X_1-\theta=x)\right)\mathbb{P}(X_1-\theta\in dx) \\
&= \int_{-\infty}^{-(\beta+\gamma)} \left\{\begin{array}{l} \Big[-|h(u)|\,\mathbb{P}(\theta_n-\theta\leq x-u\,|\,X_1-\theta=x)\Big]\Big|_{u=x+\gamma}^{-\beta} \\ +\int_{x+\gamma}^{-\beta}\operatorname{sign}(h(u))\,h'(u)\,\mathbb{P}(\theta_n-\theta\leq x-u\,|\,X_1-\theta=x)\,du \end{array}\right\}\mathbb{P}(X_1-\theta\in dx) \\
&\quad+\int_{-\infty}^{\infty}\lim_{t\to\infty}\left\{\begin{array}{l} \Big[-|h(u)|\,\mathbb{P}(\theta_n-\theta\leq x-u\,|\,X_1-\theta=x)\Big]\Big|_{u=(x+\gamma)\vee\beta}^{t} \\ +\int_{(x+\gamma)\vee\beta}^{t}\operatorname{sign}(h(u))\,h'(u)\,\mathbb{P}(\theta_n-\theta\leq x-u\,|\,X_1-\theta=x)\,du \\ +\Big[|h(u)|\,\mathbb{P}(\theta_n-\theta\geq x-u\,|\,X_1-\theta=x)\Big]\Big|_{u=-t}^{(x-\gamma)\wedge(-\beta)} \\ -\int_{-t}^{(x-\gamma)\wedge(-\beta)}\operatorname{sign}(h(u))\,h'(u)\,\mathbb{P}(\theta_n-\theta\geq x-u\,|\,X_1-\theta=x)\,du \end{array}\right\}\mathbb{P}(X_1-\theta\in dx) \\
&\quad+\int_{\beta+\gamma}^{\infty}\left\{\begin{array}{l} \Big[|h(u)|\,\mathbb{P}(\theta_n-\theta\geq x-u\,|\,X_1-\theta=x)\Big]\Big|_{u=\beta}^{x-\gamma} \\ -\int_{\beta}^{x-\gamma}\operatorname{sign}(h(u))\,h'(u)\,\mathbb{P}(\theta_n-\theta\geq x-u\,|\,X_1-\theta=x)\,du \end{array}\right\}\mathbb{P}(X_1-\theta\in dx) \\
&\leq \int_{-\infty}^{-(\beta+\gamma)}\Big\{|h(x+\gamma)|+0\Big\}\,\mathbb{P}(X_1-\theta\in dx) \\
&\quad+C\int_{-\infty}^{\infty}\left\{\begin{array}{l} |h((x+\gamma)\vee\beta)|\,e^{-|x-((x+\gamma)\vee\beta)|^p}+\int_{\beta}^{\infty}|h'(u)|\,e^{-|x-u|^p}du \\ |h((x-\gamma)\wedge(-\beta))|\,e^{-|x-((x-\gamma)\wedge(-\beta))|^p}+\int_{-\infty}^{-\beta}|h'(u)|\,e^{-|x-u|^p}du \end{array}\right\}\mathbb{P}(X_1-\theta\in dx) \\
&\quad+\int_{\beta+\gamma}^{\infty}\Big\{|h(x-\gamma)|+0\Big\}\,\mathbb{P}(X_1-\theta\in dx) \\
&\lesssim \mathbb{E}\Big[|h(X_1-\theta+\gamma)|\mathbf{1}_{\{|X_1-\theta+\gamma|\geq\beta\}}\Big]+\mathbb{E}\Big[|h(\beta)|\,e^{-|X_1-\theta-\beta|^p}\Big]+\int_{\beta}^{\infty}\mathbb{E}\Big[|h'(u)|\,e^{-|X_1-\theta-u|^p}\Big]du \\
&\quad+\mathbb{E}\Big[|h(X_1-\theta-\gamma)|\mathbf{1}_{\{|X_1-\theta-\gamma|\geq\beta\}}\Big]+\mathbb{E}\Big[|h(-\beta)|\,e^{-|X_1-\theta+\beta|^p}\Big]+\int_{-\infty}^{-\beta}\mathbb{E}\Big[|h'(u)|\,e^{-|X_1-\theta-u|^p}\Big]du,
\end{aligned}
$$

where $y \lesssim z$ means $y \leq (1\vee C)z$. As $\beta\to\infty$, the first and fourth terms go to 0 by $(H.5).2$ and the DCT, the second and fifth terms go to 0 by $(H.5).3$ and the DCT, the third and sixth terms go to 0 by $(H.5).4$ and the DCT. None of the terms depended on n, so the convergence is uniform in $n\geq N_0$. $\square$

If $\{\theta_n^\star\}_{n\in\mathbb{N}}$ is a sequence of M-estimators, then the next lemma proposes an easy-to-verify condition on the tail probabilities of $\theta_n^\star$ for $(E.5)$ in Lemma 8.2.2 to hold uniformly in the set of estimators

$$\mathcal{E}_{n,\theta} \mathrel{\hat=} \{\theta + v(\theta_n^\star - \theta)\}_{v\in[0,1]}, \quad \text{for some } \theta \in \mathbb{R}. \tag{8.2.8}$$

Lemma 8.2.3. *Let $\theta \in \mathbb{R}$ and let $X_1, X_2, X_3, \dots$ be a sequence of i.i.d. random variables. Let $\{\theta_n^\star\}_{n\in\mathbb{N}}$ be a sequence of estimators satisfying*

$$\sum_{i=1}^n \psi(X_i - \theta_n^\star) = 0, \tag{8.2.9}$$

where $\psi : \mathbb{R} \to \mathbb{R}$ is measurable, non-decreasing and $\psi(0) = 0$. Assume that there exist $N \geq 1$ and $C, \gamma, p > 0$ such that

$$\sup_{n\geq N} \mathbb{P}\Big(|\theta_n^\star - \theta| \geq |t|\Big) \leq C e^{-|t|^p}, \quad \textit{for all } |t| \geq \gamma. \tag{8.2.10}$$

Then, condition $(E.5)$ from Lemma 8.2.2 is satisfied uniformly on $\mathcal{E}_{n,\theta}$, namely :

(E.5.unif): *There exist $N_2 \geq 2$, $C, \gamma, p > 0$ and $\beta_0 > \gamma$ such that, for $\mathbb{P}(X_1 - \theta \in \,\cdot\,)$-almost-all $x \in \mathbb{R}$, we have*

- *For all $u \geq (x + \gamma) \vee \beta_0$ and for all $n \geq N_2$,*

$$\sup_{\theta_n \in \mathcal{E}_{n,\theta}} \mathbb{P}(\theta_n - \theta \leq x - u \,|\, X_1 - \theta = x) \leq C e^{-|x-u|^p}.$$

- *For all $u \leq (x - \gamma) \wedge (-\beta_0)$ and for all $n \geq N_2$,*

$$\sup_{\theta_n \in \mathcal{E}_{n,\theta}} \mathbb{P}(\theta_n - \theta \geq x - u \,|\, X_1 - \theta = x) \leq C e^{-|x-u|^p}.$$

Proof. For all $n \geq 2$, let $\theta_{2:n}^\star \mathrel{\hat=} \theta_{2:n}^\star(X_2, X_3, \dots, X_n)$ be an estimator that satisfies

$$\sum_{i=2}^n \psi(X_i - \theta_{2:n}^\star) = 0 \quad \text{and} \quad \theta_{2:n}^\star \overset{\text{law}}{=} \theta_{n-1}^\star. \tag{8.2.11}$$

Since ψ is non-decreasing and $\psi(0) = 0$,

- $$\theta_n^\star \leq X_1 \implies \psi(X_1 - \theta_n^\star) \geq 0$$
$$\overset{(8.2.9)}{\implies} \sum_{i=2}^n \psi(X_i - \theta_n^\star) \leq 0 \overset{(8.2.11)}{\implies} \theta_{2:n}^\star \leq \theta_n^\star \leq X_1, \tag{8.2.12}$$

- $\theta_n^\star \geq X_1 \implies \psi(X_1 - \theta_n^\star) \leq 0$
$$\overset{(8.2.9)}{\implies} \sum_{i=2}^{n} \psi(X_i - \theta_n^\star) \geq 0 \overset{(8.2.11)}{\implies} \theta_{2:n}^\star \geq \theta_n^\star \geq X_1. \tag{8.2.13}$$

Let $\theta_n \in \mathcal{E}_{n,\theta}$ for all $n \in \mathbb{N}$. In order to prove (8.2.14) (respectively (8.2.15)) below, we use the following facts in succession : $\theta_n - \theta \leq 0 \implies \theta_n^\star - \theta \leq \theta_n - \theta$ (respectively $\theta_n - \theta \geq 0 \implies \theta_n^\star - \theta \geq \theta_n - \theta$), (8.2.12) (respectively (8.2.13)), the independence between X_1 and $\theta_{2:n}^\star$, (8.2.11), and (8.2.10).

- For all $u \geq (x + \gamma) \vee \beta_0 > 0$ (note that $x - u \leq -\gamma < 0$) and for all $n \geq N + 1$,

$$\begin{aligned}
\mathbb{P}(\theta_n - \theta \leq x - u \,|\, X_1 - \theta = x) &\leq \mathbb{P}(\theta_n^\star - \theta \leq x - u \,|\, X_1 - \theta = x) \\
&\leq \mathbb{P}(\theta_{2:n}^\star - \theta \leq x - u) \\
&= \mathbb{P}(\theta_{n-1}^\star - \theta \leq x - u) \\
&\leq Ce^{-|x-u|^p}.
\end{aligned} \tag{8.2.14}$$

- For all $u \leq (x - \gamma) \wedge (-\beta_0) < 0$ (note that $x - u \geq \gamma > 0$) and for all $n \geq N + 1$,

$$\begin{aligned}
\mathbb{P}(\theta_n - \theta \geq x - u \,|\, X_1 - \theta = x) &\leq \mathbb{P}(\theta_n^\star - \theta \geq x - u \,|\, X_1 - \theta = x) \\
&\leq \mathbb{P}(\theta_{2:n}^\star - \theta \geq x - u) \\
&= \mathbb{P}(\theta_{n-1}^\star - \theta \geq x - u) \\
&\leq Ce^{-|x-u|^p}.
\end{aligned} \tag{8.2.15}$$

Simply choose $N_2 \mathrel{\hat{=}} N + 1$ in (E.5.unif). This ends the proof. □

We can now state the main result. The structure of the assumptions is illustrated in Figure 8.2.1.

Theorem 8.2.4. *Let $\theta \in \mathbb{R}$ and let $X_1, X_2, X_3, \dots$ be a sequence of i.i.d. random variables satisfying*

(X.1): $\mathbb{P}(X_1 = \theta) = 0$.

Let $\{\theta_n^\star\}_{n\in\mathbb{N}}$ be a sequence of estimators satisfying **(E.5.unif)** *directly or the conditions in Lemma 8.2.3. Denote $\mathcal{E}_{n,\theta} \mathrel{\hat{=}} \{\theta + v(\theta_n^\star - \theta)\}_{v\in[0,1]}$, and assume that*

(E.1.unif): $\theta_n^\star \xrightarrow{\mathbb{P}} \theta$;

(E.2.unif): *For all* $n \in \mathbb{N}$, *all* $i \in \{1, 2, \ldots, n\}$ *and all* $\theta_n \in \mathcal{E}_{n,\theta}$, $(X_i - \theta_n, X_i - \theta) \stackrel{law}{=} (X_1 - \theta_n, X_1 - \theta)$;

(E.4.unif): *If* $\limsup_{x \to 0} |h(x)| < \infty$, *we impose no condition. Otherwise, assume that there exist* $N_1 \in \mathbb{N}$, $\alpha_0 > 0$ *and a constant* $C_{\alpha_0} > 0$ *such that*

$$\sup_{n \geq N_1} \sup_{\theta_n \in \mathcal{E}_{n,\theta}} \sup_{A \in \mathcal{B}_{>0}([-\alpha_0, \alpha_0])} \frac{\mathbb{P}(X_1 - \theta_n \in A)}{\textit{Lebesgue}(A)} \leq C_{\alpha_0} < \infty.$$

(E.6.unif): *There exists* $N_3 \in \mathbb{N}$ *such that for all* $n \geq N_3$ *and for all* $\theta_n \in \mathcal{E}_{n,\theta}$, *there exists* $A_{n,\theta_n} \in \mathcal{B}(\mathbb{R})$ *such that* $\mathbb{P}(X_1 - \theta \in A_{n,\theta_n}) = 1$ *and, for all* $x \in A_{n,\theta_n}$, *the measure* $\mathbb{P}(x - (\theta_n - \theta) \in \cdot \mid X_1 - \theta = x)$, *when restricted to* $\{u \in \mathbb{R} : |u| \geq \beta_0, |x - u| > \gamma\}$, *is absolutely continuous with respect to the Lebesgue measure.*

Finally, assume

(H.1), (H.2): *from Proposition 8.2.1,*

(H.3), (H.4), (H.5): *from Lemma 8.2.2.*

Then, the conclusion in Proposition 8.2.1 holds uniformly for $\theta_n \in \mathcal{E}_{n,\theta}$, *namely*

$$\lim_{n \to \infty} \sup_{\theta_n \in \mathcal{E}_{n,\theta}} \mathbb{E}\left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i \neq \theta_n\}} h(X_i - \theta_n) - \mathbb{E}\Big[h(X_1 - \theta)\Big] \right| = 0. \tag{8.2.16}$$

Proof. We know that $(E.5.\text{unif})$ holds, either directly or via the conditions in Lemma 8.2.3. By combining $(E.4.\text{unif})$ to $(E.6.\text{unif})$ and $(H.3)$ to $(H.5)$, a proof along the lines of Lemma 8.2.2 shows

(E.3.unif):

$$\lim_{K \to \infty} \sup_{n \geq N_0} \sup_{\theta_n \in \mathcal{E}_{n,\theta}} \mathbb{E}\left[\Big|h(X_1 - \theta_n)\Big| \mathbf{1}_{\{X_1 \neq \theta_n\} \cap \{|h(X_1 - \theta_n)| \geq K\}} \right] = 0.$$

By $(E.3.\text{unif})$, the identity $|U_n + V_n| \mathbf{1}_{\{|U_n + V_n| \geq 2K\}} \leq 2|U_n| \mathbf{1}_{\{|U_n| \geq K\}} + 2|V_n| \mathbf{1}_{\{|V_n| \geq K\}}$, and $(H.2)$, we deduce

$$\lim_{K \to \infty} \sup_{n \geq N_0} \sup_{\theta_n \in \mathcal{E}_{n,\theta}} \mathbb{E}\left[|h(X_1 - \theta_n) - h(X_1 - \theta)| \mathbf{1}_{\{X_1 \neq \theta_n\} \cap \{|h(X_1 - \theta_n) - h(X_1 - \theta)| \geq K\}} \right] = 0. \tag{8.2.17}$$

To conclude, we rerun the proof of Proposition 8.2.1 with our new assumptions. By $(X.1)$, $(H.2)$, $(E.1.\text{unif})$ and $(E.2.\text{unif})$, the convergence in (8.2.2) is valid for $\sup_{\theta_n \in \mathcal{E}_{n,\theta}}$ of the expectation. This implies that the convergence in (8.2.4) is also valid for $\sup_{\theta_n \in \mathcal{E}_{n,\theta}}$ of the expectation. Furthermore, by $(H.1)$, $(E.1.\text{unif})$ and the continuous mapping theorem, we have, for all $\varepsilon > 0$,

$$\lim_{n\to\infty} \sup_{\theta_n \in \mathcal{E}_{n,\theta}} \mathbb{P}\Big(\mathbf{1}_{\{X_1 \neq \theta_n\}} \big|h(X_1 - \theta_n) - h(X_1 - \theta)\big| > \varepsilon\Big) = 0. \tag{8.2.18}$$

By combining (8.2.17) and (8.2.18), the $\sup_{\theta_n \in \mathcal{E}_{n,\theta}}$ of the expectation on the right-hand side of (8.2.5) converges to 0. In summary, we have shown that $\sup_{\theta_n \in \mathcal{E}_{n,\theta}}$ of the expectations in (8.2.2), (8.2.4) and (8.2.5) all converge (respectively) to 0. Hence, the conclusion of Proposition 8.2.1 holds for $\sup_{\theta_n \in \mathcal{E}_{n,\theta}}$ of the expectation, which is exactly the claim made in (8.2.16). □

Remark 8.2.3. *By following the proof of Theorem 8.2.4, we see that $(X.1)$, $(H.1)$, $(H.2)$, $(E.1.\mathit{unif})$, $(E.2.\mathit{unif})$ and $(E.3.\mathit{unif})$ alone imply the conclusion in (8.2.16). The other assumptions in the statement of the theorem are simply there to give a more practical way to verify $(E.3.\mathit{unif})$.*

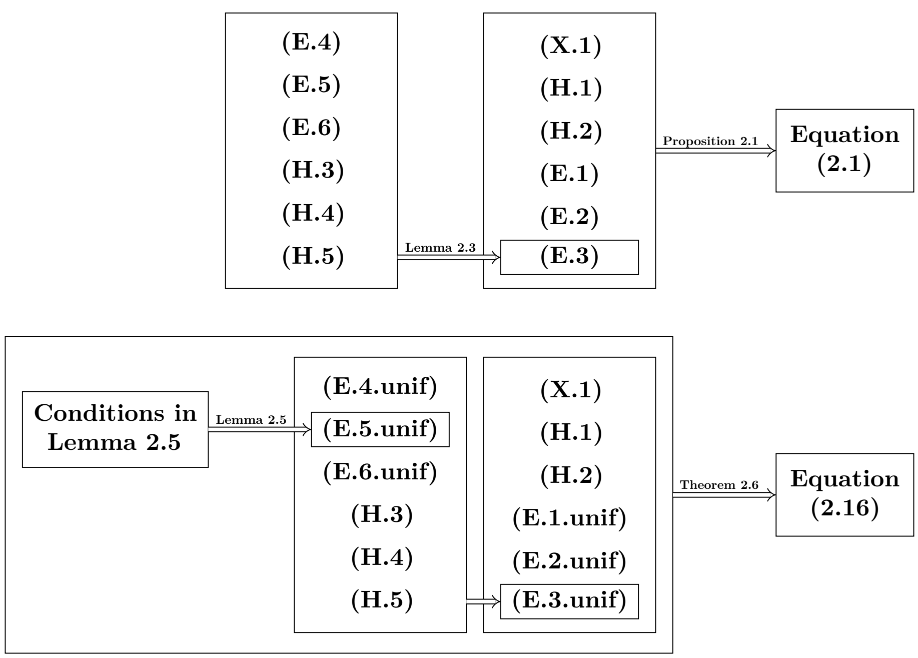


FIGURE 8.2.1. Logical structure of the assumptions and their implications.

8.3. Example

We now give an application of the previous theorem. The context of the problem is described at the end of Section 8.1.

Lemma 8.3.1. *Let $X_1, X_2, X_3, \dots$ be a sequence of i.i.d. random variables with density function*

$$f_{X_1}(x) \mathrel{\hat{=}} \frac{1}{4\sigma} e^{-\frac{1}{2}\left|\frac{x-\mu}{\sigma}\right|}, \quad x \in \mathbb{R},$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$. Define $h : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ by

$$h(y) \mathrel{\hat{=}} \mathit{sign}(y) \log |y|.$$

Let

$$\mu_n^\star \mathrel{\hat{=}} \mathit{median}(X_1, X_2, \dots, X_n) \mathrel{\hat{=}} \begin{cases} X_{((n+1)/2)}, & \text{if } n \text{ is odd}, \\ \frac{1}{2}(X_{(n/2)} + X_{(n/2+1)}), & \text{if } n \text{ is even}. \end{cases} \tag{8.3.1}$$

For $v \in [0,1]$, define $\mu_{n,v}^\star \mathrel{\hat{=}} \mu + v(\mu_n^\star - \mu)$, and let $\mathcal{E}_{n,\mu} \mathrel{\hat{=}} \{\mu_{n,v}^\star\}_{v\in[0,1]}$. Then,

$$\lim_{n\to\infty} \sup_{v\in[0,1]} \mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n \mathbf{1}_{\{X_i \neq \mu_{n,v}^\star\}} h(X_i - \mu_{n,v}^\star) - \mathbb{E}\left[h(X_1 - \mu)\right]\right| = 0. \tag{8.3.2}$$

Proof. Without loss of generality, assume that $\mu = 0$. Below, we verify the conditions of Theorem 8.2.4.

(X.1): $\mathbb{P}(X_1 = 0) = 0$. This is obvious.

(Conditions in Lemma 8.2.3): We show that the conditions are satisfied with $\psi(y) \mathrel{\hat{=}} \mathrm{sign}(y)$ and $\psi(0) \mathrel{\hat{=}} 0$. Indeed, by (8.3.1), we know that $\sum_{i=1}^n \psi(X_i - \mu_n^\star) = 0$. Furthermore, for $N \in \mathbb{N}$ and $\gamma > 0$ both large enough (depending on σ), we have, for all $n \geq N$ and all $t \geq \gamma$,

$$\begin{aligned} \mathbb{P}(\mu_n^\star \geq t) &\leq \sum_{k=\lceil n/2\rceil}^n \binom{n}{k} \mathbb{P}(X_1 \geq t)^k \, \mathbb{P}(X_1 \leq t)^{n-k} \\ &\leq (n - \lceil n/2\rceil) \cdot \binom{n}{\lceil n/2\rceil} \cdot \mathbb{P}(X_1 \geq t)^{\lceil n/2\rceil} \\ &\leq \lfloor n/2\rfloor \cdot 2\frac{2^n}{\sqrt{n}} \cdot \left(\frac{1}{2}e^{-\frac{t}{2\sigma}}\right)^{\lceil n/2\rceil} \leq \frac{\sqrt{n}}{2} 2^n e^{-\frac{nt}{8\sigma}} \cdot e^{-\frac{nt}{8\sigma}} \leq \frac{1}{2}e^{-t}. \end{aligned} \tag{8.3.3}$$

To obtain the third inequality, we use Stirling's formula and assume that N is large enough. To obtain the last inequality, assume that $N \geq 8\sigma$ and $\gamma \geq 8\sigma$. This proves (8.2.10) with $C = 1$ and $p = 1$.

(E.1.unif): $\mu_n^\star \xrightarrow{\mathbb{P}} 0$. This is explained in Example 5.11 of van der Vaart (1998).

(E.2.unif): For any $v \in [0,1]$, the estimator $\mu_{n,v}^\star = v\mu_n^\star$ is symmetric with respect to its n variables because the median, $\mu_n^\star$, is symmetric with respect to its n variables. Since the X_i's are i.i.d., the condition is satisfied.

(E.4.unif): We have $\limsup_{x\to 0} |h(x)| = \infty$, so we need to verify the condition. For any $n \geq 2$ and any $v \in [0,1]$, note that $X_1 - v\mu_n^\star$ has a density function. It suffices to show that the densities are bounded, uniformly in n and v, by a positive constant. Since the density $u \mapsto f_{X_1 - v\mu_n^\star}(u)$ is symmetric around 0, we will assume, without loss of generality, that $u > 0$. For $v \in (0,1]$, denote $z \circeq (x-u)/v$ and notice that $z < x$.

When $v \in (0,1]$ and $n \geq 3$ is odd, we have

$$
\begin{aligned}
f_{X_1 - v\mu_n^\star}(u) &= \int_{-\infty}^{\infty} f_{X_1 - v\mu_n^\star | X_1}(u \,|\, x) f_{X_1}(x) dx = \int_{-\infty}^{\infty} \frac{1}{v} f_{\mu_n^\star | X_1}(z \,|\, x) f_{X_1}(x) dx \\
&= \int_{-\infty}^{\infty} \frac{1}{v} \binom{n}{\lfloor n/2 \rfloor} (F_{X_1}(z))^{\lfloor n/2 \rfloor} f_{X_1}(z)(1 - F_{X_1}(z))^{\lfloor n/2 - 1 \rfloor} f_{X_1}(x) dx \\
&\leq C \, \|f_{X_1}\|_\infty \underbrace{\int_{-\infty}^{\infty} \frac{1}{v} f_{\mu_{n-2}^\star}(z) dx}_{=\,1} \\
&= C \, \|f_{X_1}\|_\infty < \infty.
\end{aligned}
$$

In the inequality above, we took $C \circeq \sup_{n\geq 3} \binom{n}{\lfloor n/2 \rfloor} / \binom{n-2}{\lfloor (n-2)/2 \rfloor}$, which is finite by Stirling's formula. When $v \in (0,1]$ and $n \geq 4$ is even, we can apply a similar argument and also obtain a uniform bound. Finally, when $v = 0$ and $n \in \mathbb{N}$, $f_{X_1 - v\mu_n^\star}(u) = f_{X_1}(u) \leq 1/(4\sigma)$.

In summary, $f_{X_1 - v\mu_n^\star}(u)$ is uniformly bounded in $u \in \mathbb{R}$, $n \geq 3$ and $v \in [0,1]$, which proves (E.4.unif) with any $\alpha_0 > 0$ and any $N_1 \geq 3$.

(E.6.unif): In our case, this is trivial because the conditional density $f_{X_1 - v\mu_n^\star | X_1}(\cdot \,|\, x)$ exists for all $x \in \mathbb{R}$, all $n \geq 2$ and all $v \in (0,1]$.

(H.1): The function h is continuous on $\mathbb{R}\backslash\{0\}$, so $\mathcal{D}_h = \emptyset$ and thus $\mathbb{P}(X_1 \in \mathcal{D}_h) = 0$.

(H.2): $\mathbb{E}\Big|h(X_1)\Big| \leq \int_{|x|\leq 1} |\log|x|| \frac{1}{4\sigma} dx + \int_{|x|\geq 1} |x| f_{X_1}(x) dx \leq \frac{2}{4\sigma} + 2\sigma < \infty$.

(H.3): For all $x_0 \in \mathbb{R}\backslash\{0\}$, $\limsup_{x\to x_0} |h(x)| < \infty$. This is obvious.

(H.4): $\int_{|u|\leq\alpha_0} |\log|u|| du < \infty$ is true for any $\alpha_0 > 0$ since $\int_{|u|\leq 1} |\log|u|| du = 2$.

(H.5): (1) This is obviously true for any $\beta_0 > 0$ (use the fundamental theorem of calculus).

(2) For any $\gamma > 0$ and any $\beta_0 > \gamma$, the supremum $\sup_{|t|\leq\gamma} |h(X_1 - t)| \mathbf{1}_{\{|X_1 - t|\geq\beta_0\}}$ is attained at the boundary with probability 1 (not necessarily the same end of the boundary for different ω's). Therefore, take $M = |h(X_1 - \gamma)| \mathbf{1}_{\{|X_1-\gamma|\geq\beta_0\}} + |h(X_1 + \gamma)| \mathbf{1}_{\{|X_1+\gamma|\geq\beta_0\}}$. It is easy to show that $\mathbb{E}[M] < \infty$ because $|\log|x|| \leq |x|$ for $|x| \geq 1$ and $\int_{|x|\geq(1\vee\beta_0)} |x| f_{X_1\pm\gamma}(x) dx < \infty$.

(3) We need to verify this condition for $p = 1$ since this is the p that we used above to verify the conditions of Lemma 8.2.3. First, $\lim_{|\beta|\to\infty} |h(\beta)| e^{-|x-\beta|^p} = 0$ is true for all $x \in \mathbb{R}$ and all $p > 0$ (true in particular for $p = 1$). For the second part, assume that $\beta \geq 1$. We have

$$\begin{aligned}
\mathbb{E}[e^{-|X_1-\beta|}] &= \int_{(-\infty,0)\cup(0,\beta)\cup(\beta,\infty)} e^{-|x-\beta|} \cdot \frac{1}{4\sigma} e^{-\frac{1}{2\sigma}|x|} dx \\
&\leq \frac{1}{2} e^{-|\beta|} \underbrace{\int_{-\infty}^{0} \frac{1}{2\sigma} e^{-\frac{1}{2\sigma}|x|} dx}_{=1} + \frac{|\beta|}{4\sigma} e^{-(1\wedge\frac{1}{2\sigma})|\beta|} + \frac{1}{4\sigma} e^{-\frac{1}{2\sigma}|\beta|} \underbrace{\int_{\beta}^{\infty} e^{-|x-\beta|} dx}_{=1} \\
&\leq \frac{|\beta|}{2} \left(1 \vee \frac{1}{2\sigma}\right) e^{-(1\wedge\frac{1}{2\sigma})|\beta|}.
\end{aligned} \tag{8.3.4}$$

By the symmetry of f_{X_1}, we also have (8.3.4) for $\beta \leq -1$. Hence, for any $\beta_0 \geq 1$,

$$\sup_{|\beta|\geq\beta_0} \mathbb{E}\left[\left(|h(\beta)| e^{-|X_1-\beta|}\right)^2\right] < \infty,$$

which is a well-known sufficient condition for the uniform integrability of

$$\{|h(\beta)| e^{-|X_1-\beta|}\}_{|\beta|\geq\beta_0},$$

see e.g. (Klenke, 2014, Corollary 6.21).

(4) Take any $\beta_0 \geq 1$, then (8.3.4) implies

$$\int_{|u|\geq\beta_0} \mathbb{E}\Big[|h'(u)|e^{-|X_1-u|}\Big]du \leq \frac{1}{\beta_0}\int_{|u|\geq\beta_0} \mathbb{E}\Big[e^{-|X_1-u|}\Big]du < \infty.$$

(5) Take any $\beta_0 \geq 1$, then, for all $|u| \geq \beta_0$,

$$-\mathrm{sign}(u)\cdot\mathrm{sign}(h(u))\cdot h'(u) = -\mathrm{sign}(u)\cdot\mathrm{sign}(u)\cdot\frac{1}{|u|} \leq 0.$$

This ends the proof. $\square$

Acknowledgements

We would like to thank the anonymous reviewer for helpful comments that contributed to improving the final version of the paper. F. Ouimet is supported by a NSERC Doctoral Program Alexander Graham Bell scholarship (CGS D3).

Bibliography

J. R. Blum. On the convergence of empiric distribution functions. *Ann. Math. Statist.*, 26: 527–529, 1955. MR0070871.

T. de Wet and R. H. Randles. On the effect of substituting parameter estimators in limiting χ^2 U and V statistics. *Ann. Statist.*, 15(1):398–412, 1987. MR885745.

J. Dehardt. Generalizations of the Glivenko-Cantelli theorem. *Ann. Math. Statist.*, 42: 2050–2055, 1971. MR0297000.

A. Desgagné and P. Lafaye de Micheaux. A powerful and interpretable alternative to the Jarque-Bera test of normality based on 2nd-power skewness and kurtosis, using the Rao's score test on the APD family. *J. Appl. Stat.*, 45(13):2307–2327, 2018. MR3841677.

A. Desgagné, P. Lafaye de Micheaux, and A. Leblanc. Test of normality against generalized exponential power alternatives. *Comm. Statist. Theory Methods*, 42(1):164–190, 2013. MR3004652.

T. S. Ferguson. *A course in large sample theory.* Texts in Statistical Science Series. Chapman & Hall, London, 1996. MR1699953.

C. Fernández, J. Osiewalski, and M. F. J. Steel. Modeling and inference with v-spherical distributions. *J. Amer. Statist. Assoc.*, 90(432):1331–1340, 1995. MR1379475.

E. Giné and J. Zinn. Some limit theorems for empirical processes. *Ann. Probab.*, 12(4): 929–998, 1984. MR757767.

A. Klenke. *Probability theory.* Universitext. Springer, London, second edition, 2014. MR3112259.

I. Komunjer. Asymmetric power distribution: theory and applications to risk measurement. *J. Appl. Econometrics*, 22(5):891–921, 2007. MR2395888.

L. LeCam. On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates. *Univ. California Publ. Statist.*, 1:277–329, 1953. MR0054913.

D. Pollard. *Convergence of stochastic processes.* Springer Series in Statistics. Springer-Verlag, New York, 1984. MR762984.

H. Rubin. Uniform convergence of random functions with applications to statistics. *Ann. Math. Statist.*, 27:200–203, 1956. MR0076246.

D. W. Stroock. *Probability theory.* Cambridge University Press, Cambridge, second edition, 2011. MR2760872.

M. Talagrand. The Glivenko-Cantelli problem. *Ann. Probab.*, 15(3):837–870, 1987. MR893902.

S. A. van de Geer. *Applications of empirical process theory*, volume 6 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, Cambridge, 2000. MR1739079.

A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, Cambridge, 1998. MR1652247.

A. W. van der Vaart and J. A. Wellner. *Weak convergence and empirical processes.* Springer Series in Statistics. Springer-Verlag, New York, 1996. MR1385671.

V. N. Vapnik and A. Ya. Chervonenkis. Necessary and sufficient conditions for the uniform convergence of empirical means to their true values. *Teor. Veroyatnost. i Primenen.*, 26 (3):543–563, 1981. MR627861.

V. N. Vapnik and A. Ya. Červonenkis. The uniform convergence of frequencies of the appearance of events to their probabilities. *Teor. Verojatnost. i Primenen.*, 16:264–279, 1971. MR0288823.

Article 9

Asymptotic law of a modified score statistic for the asymmetric power distribution with unknown location and scale parameters

by

Frédéric Ouimet[1]

([1]) Université de Montréal, Montréal, QC H3T 1J4, Canada.

Here is the complete citation :

Desgagné, A., Lafaye de Micheaux, P. and Ouimet, F. (2018). Asymptotic law of a modified score statistic for the asymmetric power distribution with unknown location and scale parameters. https://arxiv.org/abs/1809.02852v1.

Contributions

The version of the article that is presented here is entirely written by me (the statements of the results are inspired by the work in Desgagné and Lafaye de Micheaux (2018), which the present article generalizes). Before we submit to a journal, Pierre Lafaye de Micheaux and Alain Desgagné will expand this version with further motivations in the introduction and some simulations.

Abstract. For an i.i.d. sample of observations, we study a modified score statistic that tests the goodness-of-fit of a given exponential power distribution against a family of alternatives, called the asymmetric power distribution. The family of alternatives was introduced in Komunjer (2007) and is a reparametrization of the skewed exponential power distribution from Fernández et al. (1995) and Kotz et al. (2001). The score is modified in the sense that the location and scale parameters (assumed to be unknown) are replaced by their maximum likelihood estimators. We find the asymptotic law of the modified score statistic under the null hypothesis (H_0) and under local alternatives, using the notion of contiguity. Our work generalizes and extends the findings of Desgagné and Lafaye de Micheaux (2018), where the data points were normally distributed under H_0. The special case where each data point has a Laplace distribution under H_0 is the hardest to treat and requires a recent result from Lafaye de Micheaux and Ouimet (2018) on a uniform law of large numbers for summands that blow up.

Keywords: asymptotic statistics, exponential power distribution, asymmetric power distribution, skewed exponential power distribution, Lagrange multiplier test, score test, uniform law of large numbers

9.1. The asymmetric power distribution (APD)

The asymmetric power distribution (APD), proposed by Komunjer (2007), can be viewed as a generalization of the exponential power distribution (EPD) – also known as the generalized error distribution or the generalized normal distribution (Nadarajah (2005)) – to a broader family that includes asymmetric densities. The APD family combines the large range of exponential tail behaviors provided by the EPD family with various levels of asymmetry. The probability density function $f(u)$ of the standard APD is defined in Section 2 of Komunjer (2007). In order to relate it more easily to the skewed exponential power distribution of Fernández et al. (1995) and Kotz et al. (2001) (see Remark 9.1.1 below), we modify its scaling with the change of variable $u = 2^{-1/\theta_2} y$ and we obtain

$$f(y \mid \boldsymbol{\theta}) \mathrel{\hat=} \frac{\delta_{\boldsymbol{\theta}}^{1/\theta_2}}{2^{1/\theta_2}\Gamma(1 + 1/\theta_2)} \exp\left(-\frac{1}{2}\frac{\delta_{\boldsymbol{\theta}}}{A_{\boldsymbol{\theta}}(y)}|y|^{\theta_2}\right), \quad y \in \mathbb{R}, \tag{9.1.1}$$

where $\boldsymbol{\theta} \mathrel{\hat=} (\theta_1, \theta_2)^\top$, $\theta_1 \in (0,1)$, $\theta_2 \in (0,\infty)$,

$$\delta_{\boldsymbol{\theta}} \mathrel{\hat=} \frac{2\theta_1^{\theta_2}(1-\theta_1)^{\theta_2}}{\theta_1^{\theta_2} + (1-\theta_1)^{\theta_2}} \quad \text{and} \quad A_{\boldsymbol{\theta}}(y) \mathrel{\hat=} \Big[1/2 + \mathrm{sign}(y)(1/2 - \theta_1)\Big]^{\theta_2}. \tag{9.1.2}$$

More generally, we can add location and scale parameters $(\mu, \sigma) \in \mathbb{R} \times (0, \infty)$. We define

$$g(x \,|\, \boldsymbol{\theta}, \boldsymbol{\kappa}) \mathrel{\hat{=}} \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma} \,\middle|\, \boldsymbol{\theta}\right), \quad x \in \mathbb{R}, \tag{9.1.3}$$

where

$$\boldsymbol{\kappa} \mathrel{\hat{=}} (\mu, \sigma)^\top. \tag{9.1.4}$$

When X has density (9.1.3), we denote $X \sim \mathrm{APD}(\boldsymbol{\theta}, \boldsymbol{\kappa})$.

Remark 9.1.1. *In Equation* (8) *of Fernández et al. (1995) and page 271 of Kotz et al. (2001), the skewed exponential power distribution (where the location and scale parameters m and s are added as μ and σ were added in* (9.1.3)*) is defined by the density function*

$$\widetilde{g}(x \,|\, \gamma, q, m, s) \mathrel{\hat{=}} \begin{cases} c_{\gamma,q} \, \frac{1}{s} \exp\left(-\frac{1}{2}\left|\frac{\gamma(x-m)}{s}\right|^q\right), & \text{if } x \leq m, \\ c_{\gamma,q} \, \frac{1}{s} \exp\left(-\frac{1}{2}\left|\frac{(x-m)}{\gamma s}\right|^q\right), & \text{if } x \geq m, \end{cases} \tag{9.1.5}$$

where $\gamma, q \in (0, \infty)$ and $c_{\gamma,q}^{-1} \mathrel{\hat{=}} 2^{1/q}\Gamma(1+1/q)(\gamma+1/\gamma)$. The reader can verify that (9.1.3) *is a reparametrization of* (9.1.5) *where*

$$\theta_1 \mathrel{\hat{=}} 1/(1+\gamma^2), \quad \theta_2 \mathrel{\hat{=}} q, \quad \mu \mathrel{\hat{=}} m \quad \text{and} \quad \sigma \mathrel{\hat{=}} \delta_{\boldsymbol{\theta}}^{1/\theta_2}(\gamma+1/\gamma)s. \tag{9.1.6}$$

Remark 9.1.2. *One interesting property of the parametrization* (9.1.3) *is that θ_1 represents the proportion of the density that is left of the mode μ. It can be useful for modelling purposes.*

9.2. Preliminaries

Throughout this paper, we assume that $\boldsymbol{\kappa} = (\mu, \sigma)^\top$ is *unknown*. Additionally, fix a constant $\lambda \geq 1$ and let $\boldsymbol{\theta}_0 \mathrel{\hat{=}} (1/2, \lambda)^\top$. For an i.i.d. sample $X_1, X_2, \dots, X_n$, we want to test the hypotheses

$$\begin{aligned} &H_0 : X_i \sim \mathrm{APD}(\boldsymbol{\theta}_0, \boldsymbol{\kappa}); \\ &H_1 : X_i \sim \mathrm{APD}(\boldsymbol{\theta}, \boldsymbol{\kappa}), \ \boldsymbol{\theta} \neq \boldsymbol{\theta}_0. \end{aligned} \tag{9.2.1}$$

If $\boldsymbol{\kappa}$ were known, this could be achieved with the score statistic

$$\boldsymbol{r}_n(\boldsymbol{\kappa}) \mathrel{\hat{=}} \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \log g(X_i \,|\, \boldsymbol{\theta}_0, \boldsymbol{\kappa}). \tag{9.2.2}$$

Indeed, we can show (see Proposition 9.3.1 below) that, under H_0, $\boldsymbol{r}_n(\boldsymbol{\kappa})^\top J_{\boldsymbol{\theta\theta}}^{-1}\, \boldsymbol{r}_n(\boldsymbol{\kappa}) \rightsquigarrow \chi^2_2$, where $J_{\boldsymbol{\theta\theta}}$ denotes the asymptotic covariance matrix of $\boldsymbol{r}_n(\boldsymbol{\kappa})$. Since we assumed that $\boldsymbol{\kappa}$ is unknown, we propose to test (9.2.1) by replacing $\boldsymbol{\kappa}$ in (9.2.2) by its maximum likelihood estimator

$$\hat{\boldsymbol{\kappa}}_n \triangleq (\hat{\mu}_n, \hat{\sigma}_n)^\top. \tag{9.2.3}$$

We are thus interested in determining the asymptotic law of the *modified score statistic*

$$\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) \triangleq \frac{1}{n}\sum_{i=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \log g(X_i \,|\, \boldsymbol{\theta}_0, \hat{\boldsymbol{\kappa}}_n). \tag{9.2.4}$$

Remark 9.2.1. *Our first main result (Theorem 9.3.4) gives the asymptotic law of $\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)$ under H_0, and our second main result (Theorem 9.3.8) gives it under local alternatives (which are defined in (9.3.13)). Falk et al. (2008) did a similar study in the context of Pareto distributions.*

Remark 9.2.2. *Two special cases are of particular interest in (9.2.1). When $\lambda = 1$, the X_i's have a Laplace distribution under H_0, and when $\lambda = 2$, the X_i's are normally distributed under H_0. The case $\lambda = 2$ was previously treated in Desgagné and Lafaye de Micheaux (2018), but not under local alternatives. In this paper, we treat all the cases $\lambda \geq 1$ under H_0 and under local alternatives. The case $\lambda = 1$ is the hardest to handle and will require a recent result from Lafaye de Micheaux and Ouimet (2018) on a uniform law of large numbers for summands that blow up (see the proof of Proposition 9.3.2).*

Below, we introduce some notations (see also the Notation section at the end of the paper). Define

$$\boldsymbol{d}_{\boldsymbol{\theta}}(y) \triangleq \frac{\partial}{\partial \boldsymbol{\theta}} \log g(x \,|\, \boldsymbol{\theta}_0, \boldsymbol{\kappa})\Big|_{x=\mu+\sigma y} = \frac{\partial}{\partial \boldsymbol{\theta}} \log f(y \,|\, \boldsymbol{\theta}_0), \tag{9.2.5}$$

$$\boldsymbol{d}_{\boldsymbol{\kappa}}(y) \triangleq \sigma \frac{\partial}{\partial \boldsymbol{\kappa}} \log g(x \,|\, \boldsymbol{\theta}_0, \boldsymbol{\kappa})\Big|_{x=\mu+\sigma y} = \begin{pmatrix} -\frac{\partial}{\partial y} \log f(y \,|\, \boldsymbol{\theta}_0) \\ -1 - y \frac{\partial}{\partial y} \log f(y \,|\, \boldsymbol{\theta}_0) \end{pmatrix}. \tag{9.2.6}$$

We can easily verify (using Wolfram Mathematica) that

$$\boldsymbol{d}_{\boldsymbol{\theta}}(y) = \begin{pmatrix} -\lambda |y|^\lambda \mathrm{sign}(y) \\ -\frac{1}{2}\Big\{|y|^\lambda \log|y| - \frac{2}{\lambda^2}\big[\log 2 + \psi(1 + 1/\lambda)\big]\Big\} \end{pmatrix}, \quad \boldsymbol{d}_{\boldsymbol{\kappa}}(y) = \begin{pmatrix} \frac{\lambda}{2}|y|^{\lambda-1}\mathrm{sign}(y), \\ \frac{\lambda}{2}|y|^\lambda - 1 \end{pmatrix}, \tag{9.2.7}$$

where $\psi(z) \mathrel{\hat{=}} \frac{d}{dz}\log\Gamma(z)$ is the digamma function and $\Gamma(z) \mathrel{\hat{=}} \int_0^\infty t^{z-1}e^{-t}dt$ is the gamma function. Using the notation in (9.2.5), we can write the score statistic (9.2.2) as

$$\boldsymbol{r}_n(\boldsymbol{\kappa}) = \frac{1}{n}\sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i), \quad \text{where } \; Y_i \mathrel{\hat{=}} \sigma^{-1}(X_i - \mu). \tag{9.2.8}$$

Under the null hypothesis, $X_i \sim APD(\boldsymbol{\theta}_0, \boldsymbol{\kappa})$, we find the maximum likelihood estimator $\hat{\boldsymbol{\kappa}}_n = (\hat{\mu}_n, \hat{\sigma}_n)^\top$ by solving

$$(\hat{\mu}_n, \hat{\sigma}_n) \in \operatorname{argmax}_{\boldsymbol{\kappa}\in\mathbb{R}\times(0,\infty)} \sum_{i=1}^n \left\{\frac{1}{2}\left|\frac{X_i - \kappa_1}{\kappa_2}\right|^\lambda - \log\kappa_2\right\}, \tag{9.2.9}$$

or equivalently, by finding the values who jointly satisfy the equations

$$\sum_{i=1}^n d_\mu\left(\frac{X_i - \hat{\mu}_n}{\hat{\sigma}_n}\right) = 0 \quad \text{and} \quad \sum_{i=1}^n d_\sigma\left(\frac{X_i - \hat{\mu}_n}{\hat{\sigma}_n}\right) = 0. \tag{9.2.10}$$

We obtain the estimators

$$\hat{\mu}_n = \begin{cases} \text{median}(X_1, X_2, \ldots, X_n), & \text{if } \lambda = 1, \\ \frac{1}{n}\sum_{i=1}^n X_i, & \text{if } \lambda = 2, \\ \text{the unique solution to } \sum_{i=1}^n |X_i - \hat{\mu}_n|^{\lambda-1}\text{sign}(X_i - \hat{\mu}_n) = 0, & \text{if } \lambda > 1, \end{cases} \tag{9.2.11}$$

$$\hat{\sigma}_n = \left(\frac{1}{n}\sum_{i=1}^n \frac{\lambda}{2}|X_i - \hat{\mu}_n|^\lambda\right)^{1/\lambda}.$$

Remark 9.2.3. *When $\lambda \notin \{1, 2\}$, $\hat{\mu}_n$ doesn't have an explicit expression.*

Remark 9.2.4. *The median is not well-defined when n is even. If the values in the sample are all different, then any real number inside the interval $(X_{(n/2)}, X_{(n/2+1)})$, where $X_{(k)}$ denotes the k-th smallest value of the sample, satisfies the definition of a median with respect to the empirical distribution. To avoid ambiguity, assume for the remainder of this article that the median is uniquely defined by*

$$\textit{median}(X_1, X_2, \ldots, X_n) \mathrel{\hat{=}} \begin{cases} X_{((n+1)/2)}, & \textit{if } n \textit{ is odd}, \\ \frac{1}{2}(X_{(n/2)} + X_{(n/2+1)}), & \textit{if } n \textit{ is even}. \end{cases} \tag{9.2.12}$$

Below, we state a small adaptation of a well-known uniform law of large numbers due to Lucien Le Cam. We will use it several times in this article. The proof, which is deferred to Section 9.4.1, follows the strategy described in Section 16 of Ferguson (1996). A small adaptation is needed to treat the case where the parameter space is not compact.

Lemma 9.2.1. *Let $X_1, X_2, X_3, \dots$ be a sequence of i.i.d. random variables, and let $\hat{\boldsymbol{\xi}}_n \mathrel{\hat=} \hat{\boldsymbol{\xi}}_n(X_1, X_2, \dots, X_n)$ be an estimator such that $\hat{\boldsymbol{\xi}}_n \xrightarrow{a.s.} \boldsymbol{\xi} \in \mathbb{R}^d$. For $\delta \geq 0$, let $B_\delta[\boldsymbol{\xi}] \mathrel{\hat=} \{\boldsymbol{t} \in \mathbb{R}^d : \|\boldsymbol{t} - \boldsymbol{\xi}\|_2 \leq \delta\}$. Assume that $U : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is a measurable function and there exists $\delta > 0$ such that*

(A.1): *For all $x \in \mathbb{R}$, $\boldsymbol{t} \mapsto U(x, \boldsymbol{t})$ is continuous on $B_\delta[\boldsymbol{\xi}]$;*

(A.2): *There exists $K : \mathbb{R} \to \mathbb{R}$ such that $|U(x, \boldsymbol{t})| \leq K(x)$ for all $(x, \boldsymbol{t}) \in \mathbb{R} \times B_\delta[\boldsymbol{\xi}]$ and $\mathbb{E}\Big[|K(X_1)|\Big] < \infty$.*

If $\rho_n \mathrel{\hat=} \|\hat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}\|_2$ and $\overline{U}(\boldsymbol{t}) \mathrel{\hat=} \mathbb{E}[U(X_1, \boldsymbol{t})]$, then

$$\mathbb{P}\left(\limsup_{n\to\infty} \sup_{\boldsymbol{t} \in B_{\rho_n}[\boldsymbol{\xi}]} \left|\frac{1}{n}\sum_{i=1}^n U(X_i, \boldsymbol{t}) - \overline{U}(\boldsymbol{\xi})\right| > 0\right) = 0. \tag{9.2.13}$$

By combining Lemma 9.2.1 and a result of from Rubin and Rukhin (1983) on the convergence rates of M-estimators, we can show (see Section 9.4.1) that the maximum likelihood estimators in (9.2.11) are strongly consistent.

Lemma 9.2.2. *Under H_0 and under H_1,*

$$\hat{\boldsymbol{\kappa}}_n \mathrel{\hat=} \begin{pmatrix} \hat{\mu}_n \\ \hat{\sigma}_n \end{pmatrix} \xrightarrow{a.s.} \begin{pmatrix} \mu \\ \sigma \end{pmatrix} \mathrel{\hat=} \boldsymbol{\kappa}, \quad \text{as } n \to \infty. \tag{9.2.14}$$

9.3. Asymptotic law of the modified score statistic

Using the notation in (9.2.5), we can write the modified score statistic (9.2.4) as

$$\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) = \frac{1}{n}\sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\theta}}(Z_i), \quad \text{where } Z_i \mathrel{\hat=} \hat{\sigma}_n^{-1}(X_i - \hat{\mu}_n). \tag{9.3.1}$$

Below, we establish the asymptotic law of $\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)$ under the null hypothesis (Section 9.3.1) and under local alternatives (Section 9.3.2). The proofs are deferred to Section 9.4.2 and Section 9.4.3, respectively.

9.3.1. Under the null hypothesis (H_0)

The strategy consists first in determining the asymptotic law of the vector

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \begin{pmatrix} \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) \end{pmatrix} \tag{9.3.2}$$

under H_0. The second step consists in writing $n^{1/2}\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)$ as a linear combination of the components of this vector plus a negligible term (via a first-order Taylor expansion). We will then be able to deduce the asymptotic distribution of $n^{1/2}\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)$ under H_0. Recall that H_0 means that for all $i \in \mathbb{N}$, $X_i \sim \mathrm{APD}(\boldsymbol{\theta}_0, \boldsymbol{\kappa})$, or equivalently,

$$Y_i \mathrel{\hat{=}} \sigma^{-1}(X_i - \mu) \sim \mathrm{APD}(\boldsymbol{\theta}_0, (0,1)^\top). \tag{9.3.3}$$

The following proposition is a direct application of the central limit theorem. The computations for the entries of the asymptotic covariance matrix J are given in Section 9.4.2.

Proposition 9.3.1. *We have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \begin{pmatrix} \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) \end{pmatrix} \overset{\mathbb{P}_{H_0}}{\rightsquigarrow} \mathcal{N}_4\left(\boldsymbol{0}\,; J \mathrel{\hat{=}} \begin{pmatrix} J_{\boldsymbol{\theta}\boldsymbol{\theta}} & J_{\boldsymbol{\theta}\boldsymbol{\kappa}} \\ J_{\boldsymbol{\theta}\boldsymbol{\kappa}} & J_{\boldsymbol{\kappa}\boldsymbol{\kappa}} \end{pmatrix}\right), \tag{9.3.4}$$

where $\boldsymbol{d}_{\boldsymbol{\theta}}$ and $\boldsymbol{d}_{\boldsymbol{\kappa}}$ are given in (9.2.7), and

$$J = \begin{pmatrix} 4(1+\lambda) & 0 & -\frac{2^{1-1/\lambda}\lambda}{\Gamma(\beta)} & 0 \\ 0 & \lambda^{-3}[\phi^2 + \beta\psi'(\beta) - 1] & 0 & -\frac{\phi}{\lambda} \\ -\frac{2^{1-1/\lambda}\lambda}{\Gamma(\beta)} & 0 & \frac{\lambda\Gamma(3-\beta)}{2^{2/\lambda}\Gamma(\beta)} & 0 \\ 0 & -\frac{\phi}{\lambda} & 0 & \lambda \end{pmatrix}, \tag{9.3.5}$$

where $\phi \mathrel{\hat{=}} 1 + \log 2 + \psi(\beta)$, $\beta \mathrel{\hat{=}} 1 + 1/\lambda$ and ψ denotes the digamma function.

In the next proposition, we use a first-order Taylor expansion with the aim of writing $n^{1/2}\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)$ as a linear combination of the components of the vector on the left-hand side of (9.3.4), plus a negligible term.

Proposition 9.3.2. *Let $\boldsymbol{1}_2 \mathrel{\hat{=}} (1,1)^\top$. We have*

$$n^{1/2}\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) = n^{1/2}\boldsymbol{r}_n(\boldsymbol{\kappa}) + \boldsymbol{r}_n'(\boldsymbol{\kappa})\, n^{1/2}(\hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa}) + o_{\mathbb{P}_{H_0}}(1)\boldsymbol{1}_2. \tag{9.3.6}$$

Now, we study the term $\boldsymbol{r}_n'(\boldsymbol{\kappa})\, n^{1/2}(\hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa})$ and rewrite (9.3.6).

Proposition 9.3.3. *Recall $J_{\boldsymbol{\theta\kappa}}$ and $J_{\boldsymbol{\kappa\kappa}}$ from Proposition 9.3.1. Then,*

$$\boldsymbol{r}_n'(\boldsymbol{\kappa}) = -\sigma^{-1} J_{\boldsymbol{\theta\kappa}} + o_{\mathbb{P}_{H_0}}(1) I_2, \tag{9.3.7}$$

$$n^{1/2}(\hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa}) = \sigma J_{\boldsymbol{\kappa\kappa}}^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) + o_{\mathbb{P}_{H_0}}(1) \mathbf{1}_2, \tag{9.3.8}$$

where I_2 is the identity matrix of size 2. Furthermore,

$$n^{1/2} \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) = \left(I_2\,;\, -J_{\boldsymbol{\theta\kappa}} J_{\boldsymbol{\kappa\kappa}}^{-1}\right) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \begin{pmatrix} \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) \end{pmatrix} + o_{\mathbb{P}_{H_0}}(1) \mathbf{1}_2. \tag{9.3.9}$$

By combining Proposition 9.3.1 and Proposition 9.3.3, we obtain the asymptotic distribution of $n^{1/2} \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)$ under the null hypothesis H_0.

Theorem 9.3.4 (First main result)**.** *We have*

$$n^{1/2} \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) \overset{\mathbb{P}_{H_0}}{\rightsquigarrow} \mathcal{N}_2(\mathbf{0}, \Sigma), \quad \text{as } n \to \infty, \tag{9.3.10}$$

where

$$\Sigma = J_{\boldsymbol{\theta\theta}} - J_{\boldsymbol{\kappa\kappa}}^{-1} J_{\boldsymbol{\theta\kappa}}^2 = \begin{pmatrix} 4(1+\lambda) - \frac{4\lambda}{\Gamma(3-\beta)\Gamma(\beta)} & 0 \\ 0 & \frac{\beta\psi'(\beta)-1}{\lambda^3} \end{pmatrix}. \tag{9.3.11}$$

In particular,

$$n\, \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)^\top \Sigma^{-1} \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) \overset{\mathbb{P}_{H_0}}{\rightsquigarrow} \chi_2^2, \quad \text{as } n \to \infty. \tag{9.3.12}$$

Remark 9.3.1. *In Desgagné and Lafaye de Micheaux (2018), the case $\lambda = 2$ was treated.*

9.3.2. Under local alternatives $(H_{1,n})$

The local alternatives are defined by

$$H_{1,n} : X_i \sim \mathrm{APD}(\boldsymbol{\theta}_n, \boldsymbol{\kappa}), \quad \boldsymbol{\theta}_n = \boldsymbol{\theta}_0 + \frac{\boldsymbol{\delta}}{\sqrt{n}}(1 + o(1)), \tag{9.3.13}$$

where $\boldsymbol{\delta} \in \mathbb{R}^2 \backslash \{\mathbf{0}\}$ is fixed. The vector $\boldsymbol{\delta}$ indicates the direction of the alternative.

The following proposition will be a crucial tool to prove the weak convergence of our modified score statistic under local alternatives. It is a consequence of the concept of *contiguity*, see e.g. Section 6.2 in van der Vaart (1998).

Proposition 9.3.5. *For any statistics* $\boldsymbol{T}_n \mathrel{\hat{=}} \boldsymbol{T}_n(X_1, X_2, \dots, X_n; \boldsymbol{\kappa})$ *taking values in* $\mathbb{R}^d$,

$$\boldsymbol{T}_n \xrightarrow{\mathbb{P}_{H_0}} 0 \quad \text{if and only if} \quad \boldsymbol{T}_n \xrightarrow{\mathbb{P}_{H_{1,n}}} 0, \tag{9.3.14}$$

as $n \to \infty$.

As an immediate consequence, we obtain the same decomposition under $H_{1,n}$ that we found for the modified score statistic under H_0 in Proposition 9.3.3.

Corollary 9.3.6. *Let* $\boldsymbol{\delta} \in \mathbb{R}^2 \backslash \{\mathbf{0}\}$. *Then, as* $n \to \infty$,

$$n^{1/2} \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) = \left(I_2\,;\, -J_{\boldsymbol{\theta\kappa}} J_{\boldsymbol{\kappa\kappa}}^{-1}\right) \frac{1}{\sqrt{n}} \sum_{i=1}^n \begin{pmatrix} \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) \end{pmatrix} + o_{\mathbb{P}_{H_{1,n}}}(1) \mathbf{1}_2. \tag{9.3.15}$$

We now use Le Cam's third lemma to prove the analogue of Proposition 9.3.1 under $H_{1,n}$. Our aim is to obtain the asymptotic distribution of the right-hand side of (9.3.15).

Proposition 9.3.7. *Let* $\boldsymbol{\delta} \in \mathbb{R}^2 \backslash \{\mathbf{0}\}$. *Then, as* $n \to \infty$,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \begin{pmatrix} \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) \end{pmatrix} \overset{\mathbb{P}_{H_{1,n}}}{\rightsquigarrow} \mathcal{N}_4\left(\begin{pmatrix} J_{\boldsymbol{\theta\theta}} \boldsymbol{\delta} \\ J_{\boldsymbol{\theta\kappa}} \boldsymbol{\delta} \end{pmatrix} ; J \mathrel{\hat{=}} \begin{pmatrix} J_{\boldsymbol{\theta\theta}} & J_{\boldsymbol{\theta\kappa}} \\ J_{\boldsymbol{\theta\kappa}} & J_{\boldsymbol{\kappa\kappa}} \end{pmatrix} \right), \tag{9.3.16}$$

where J *is given in* (9.3.5).

Finally, by combining Corollary 9.3.6 and Proposition 9.3.7, we obtain the asymptotic distribution of $n^{1/2} \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)$ under the local alternatives $H_{1,n}$.

Theorem 9.3.8 (Second main result)**.** *Let* $\boldsymbol{\delta} \in \mathbb{R}^2 \backslash \{\mathbf{0}\}$. *Then,*

$$n^{1/2} \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) \overset{\mathbb{P}_{H_{1,n}}}{\rightsquigarrow} \mathcal{N}_2(\Sigma \boldsymbol{\delta}\,;\, \Sigma), \quad \text{as } n \to \infty, \tag{9.3.17}$$

where Σ *is given in* (9.3.11). *In particular,*

$$n\, \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)^\top \Sigma^{-1} \boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) \overset{\mathbb{P}_{H_{1,n}}}{\rightsquigarrow} \chi_2^2(\boldsymbol{\delta}^\top \Sigma \boldsymbol{\delta}), \quad \text{as } n \to \infty, \tag{9.3.18}$$

where $\boldsymbol{\delta}^\top \Sigma \boldsymbol{\delta}$ *represents the noncentrality parameter of the* χ_2^2 *distribution.*

9.4. Proofs

9.4.1. Proof of the results stated in Section 9.2

Proof of Lemma 9.2.1. Fix $\delta > 0$ to a value for which $(A.1)$ and $(A.2)$ hold. By the triangle inequality, and since $\rho_n \xrightarrow{\text{a.s.}} 0$ by hypothesis, we have

$$
\begin{aligned}
&\mathbb{P}\bigg(\limsup_{n\to\infty} \sup_{\boldsymbol{t}\in B_{\rho_n}[\boldsymbol{\xi}]} \bigg|\frac{1}{n}\sum_{i=1}^{n} U(X_i,\boldsymbol{t}) - \overline{U}(\boldsymbol{\xi})\bigg| > 0\bigg) \\
&\qquad \leq \mathbb{P}\bigg(\limsup_{n\to\infty} \sup_{\boldsymbol{t}\in B_{\rho_n}[\boldsymbol{\xi}]} \bigg|\frac{1}{n}\sum_{i=1}^{n} U(X_i,\boldsymbol{t}) - \overline{U}(\boldsymbol{t})\bigg| > 0\bigg) \\
&\qquad\qquad + \mathbb{P}\bigg(\limsup_{n\to\infty} \sup_{\boldsymbol{t}\in B_{\rho_n}[\boldsymbol{\xi}]} \Big|\overline{U}(\boldsymbol{t}) - \overline{U}(\boldsymbol{\xi})\Big| > 0\bigg) \qquad (9.4.1)\\
&\qquad \leq \mathbb{P}\bigg(\limsup_{n\to\infty} \sup_{\boldsymbol{t}\in B_{\delta}[\boldsymbol{\xi}]} \bigg|\frac{1}{n}\sum_{i=1}^{n} U(X_i,\boldsymbol{t}) - \overline{U}(\boldsymbol{t})\bigg| > 0\bigg) \\
&\qquad\qquad + \mathbb{P}\bigg(\limsup_{n\to\infty} \sup_{\boldsymbol{t}\in B_{\rho_n}[\boldsymbol{\xi}]} \Big|\overline{U}(\boldsymbol{t}) - \overline{U}(\boldsymbol{\xi})\Big| > 0\bigg).
\end{aligned}
$$

By applying a uniform law of large numbers on the compact set $B_\delta[\boldsymbol{\xi}]$ (Theorem 16 (a) in Ferguson (1996) with our assumptions $(A.1)$ and $(A.2)$), the first probability on the right-hand side of (9.4.1) is zero. By $(A.1)$, $(A.2)$ and the dominated convergence theorem, we know that $\overline{U}(\boldsymbol{t}) \mathrel{\hat{=}} \mathbb{E}[U(X_1,\boldsymbol{t})]$ is continuous on $B_\delta[\boldsymbol{\xi}]$. Since $\rho_n \xrightarrow{\text{a.s.}} 0$ by hypothesis, the second probability on the right-hand side of (9.4.1) is also zero. $\square$

Proof of Lemma 9.2.2. By (9.2.10), the estimator $\hat{\mu}_n$ is determined by the equation

$$\sum_{i=1}^{n} w(X_i,\hat{\mu}_n) = 0, \quad \text{where } w(x,\mu) \mathrel{\hat{=}} |x-\mu|^{\lambda-1}\mathrm{sign}(x-\mu). \qquad (9.4.2)$$

For any $x \in \mathbb{R}$, $w(x,\cdot)$ is non-increasing when $\lambda \geq 1$. From Theorem 2 and Remark 1 in Rubin and Rukhin (1983) (the proof is a simple application of Chernoff's theorem), we get that, for any $\varepsilon > 0$, the probabilities $\mathbb{P}(|\hat{\mu}_n - \mu| > \varepsilon)$ decay exponentially fast in n (using the fact that $\mathbb{E}[w(X_1,\mu+\varepsilon)] < 0$ and $\mathbb{E}[w(X_1,\mu-\varepsilon)] > 0$ both hold under H_0 and under H_1). In particular, for any $\varepsilon > 0$, the probabilities are summable in n. Hence, by the Borel-Cantelli lemma, we have $\hat{\mu}_n \to \mu$ a.s.

Also, from (9.2.11), we have

$$\frac{2}{\lambda}\hat{\sigma}_n^\lambda = \frac{1}{n}\sum_{i=1}^n |X_i - \hat{\mu}_n|^\lambda. \tag{9.4.3}$$

If we denote $U(x,t) \doteq |x-t|^\lambda$ and $\overline{U}(t) \doteq \mathbb{E}\big[U(X_1,t)\big]$, then it is easily verified that $\overline{U}(\mu) = (2/\lambda)\sigma^\lambda$. From Lemma 9.2.1, we deduce

$$\mathbb{P}\left(\lim_{n\to\infty}\left|\frac{1}{n}\sum_{i=1}^n U(X_i,\hat{\mu}_n) - \overline{U}(\mu)\right| = 0\right) = 1. \tag{9.4.4}$$

This implies $\hat{\sigma}_n \to \sigma$ a.s. □

9.4.2. Proof of the results stated in Section 9.3.1

Proof of Proposition 9.3.1. The proposition is a direct application of the central limit theorem. Let $X \sim \mathrm{APD}(\boldsymbol{\theta}_0, \boldsymbol{\kappa})$ and $Y \doteq \sigma^{-1}(X-\mu)$. Below, we show the computations for the covariances between $d_{\theta_1}(Y)$, $d_{\theta_2}(Y)$, $d_\mu(Y)$ and $d_\sigma(Y)$. Before that, we gather some facts. The density of Y is

$$f(y\,|\,1/2,\lambda) = \frac{e^{-\frac{1}{2}|y|^\lambda}}{2^{1+1/\lambda}\Gamma(1+1/\lambda)}. \tag{9.4.5}$$

Recall the definition of the gamma and digamma functions (where $x>0$):

$$\Gamma(x) \doteq \int_0^\infty t^{x-1}e^{-t}dt, \quad \text{and} \quad \psi(x) \doteq \frac{d}{dx}\log\Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}, \tag{9.4.6}$$

and some well-known properties they satisfy (see, e.g., (Abramowitz and Stegun, 1964, Chapter 6)):

$$\Gamma(1+x) = x\Gamma(x), \tag{9.4.7}$$

$$\psi(1+x) = \psi(x) + \frac{1}{x}, \tag{9.4.8}$$

$$\psi'(1+x) = \psi'(x) - \frac{1}{x^2}, \tag{9.4.9}$$

$$\int_0^\infty t^{x-1}(\log t)e^{-t}dt = \Gamma(x)\psi(x), \tag{9.4.10}$$

$$\int_0^\infty t^{x-1}(\log t)^2 e^{-t}dt = \Gamma(x)(\psi'(x) + \psi^2(x)). \tag{9.4.11}$$

The computations below are valid for all $\lambda > 0$, except for $J_{\mu\mu}$, which only exists when $\lambda > 1/2$. Since we assume $\lambda \geq 1$ in this article, there are no limitations.

By the symmetry of the density f and the anti-symmetry of the integrands, we have

$$J_{\theta_1\theta_2} = J_{\theta_1\sigma} = J_{\theta_2\mu} = J_{\mu\sigma} = \boxed{0}. \tag{9.4.12}$$

Here are the other cases:

$$\begin{aligned}
J_{\theta_1\theta_1} &= \mathbb{E}[d_{\theta_1}(Y)d_{\theta_1}(Y)] \overset{(9.2.7)}{=} \lambda^2\mathbb{E}[|Y|^{2\lambda}] \\
&\overset{(9.4.5)}{=} \frac{2\lambda^2}{2^{1+1/\lambda}\Gamma(1+1/\lambda)} \cdot 2^2 \int_0^\infty \left(\frac{1}{2}y^\lambda\right)^2 e^{-\frac{1}{2}y^\lambda} dy \\
&= \frac{4\lambda^2}{2^{1/\lambda}\Gamma(1+1/\lambda)} \int_0^\infty t^2 e^{-t} \left[\frac{2}{\lambda}(2t)^{1/\lambda-1}\right] dt \qquad (\text{with } t = \frac{1}{2}y^\lambda) \\
&\overset{(9.4.6)}{=} \frac{4\lambda}{\Gamma(1+1/\lambda)}\Gamma(2+1/\lambda) \\
&\overset{(9.4.7)}{=} \boxed{4(\lambda+1)},
\end{aligned} \tag{9.4.13}$$

$$\begin{aligned}
J_{\theta_1\mu} &= \mathbb{E}[d_{\theta_1}(Y)d_{\mu}(Y)] \overset{(9.2.7)}{=} \frac{-\lambda^2}{2}\mathbb{E}[|Y|^{2\lambda-1}] \\
&\overset{(9.4.5)}{=} \frac{-\lambda^2}{2^{1+1/\lambda}\Gamma(1+1/\lambda)} \cdot 2^{2-1/\lambda} \int_0^\infty \left(\frac{1}{2}y^\lambda\right)^{2-1/\lambda} e^{-\frac{1}{2}y^\lambda} dy \\
&= \frac{-2^{1-2/\lambda}\lambda^2}{\Gamma(1+1/\lambda)} \int_0^\infty t^{2-1/\lambda} e^{-t} \left[\frac{2}{\lambda}(2t)^{1/\lambda-1}\right] dt \qquad (\text{with } t = \frac{1}{2}y^\lambda) \\
&\overset{(9.4.6)}{=} \boxed{\frac{-2^{1-1/\lambda}\lambda}{\Gamma(1+1/\lambda)}},
\end{aligned} \tag{9.4.14}$$

$$\begin{aligned}
J_{\mu\mu} &= \mathbb{E}[d_{\mu}(Y)d_{\mu}(Y)] \overset{(9.2.7)}{=} \frac{\lambda^2}{4}\mathbb{E}\Big[|Y|^{2\lambda-2}\Big] \\
&\overset{(9.4.5)}{=} \frac{\lambda^2}{2^{2+1/\lambda}\Gamma(1+1/\lambda)} \cdot 2^{2-2/\lambda} \int_0^\infty \left(\frac{1}{2}y^\lambda\right)^{2-2/\lambda} e^{-\frac{1}{2}y^\lambda} dy \\
&= \frac{\lambda^2}{2^{3/\lambda}\Gamma(1+1/\lambda)} \int_0^\infty t^{2-2/\lambda} e^{-t} \left[\frac{2}{\lambda}(2t)^{1/\lambda-1}\right] dt \qquad (\text{with } t = \frac{1}{2}y^\lambda) \\
&\overset{(9.4.6)}{=} \boxed{\frac{\lambda\Gamma(2-1/\lambda)}{2^{2/\lambda}\Gamma(1+1/\lambda)}}.
\end{aligned} \tag{9.4.15}$$

Denote $\nu \doteq \log 2 + \psi(1+1/\lambda)$. We have

$$J_{\theta_2\theta_2} = \mathbb{E}[d_{\theta_2}(Y)d_{\theta_2}(Y)] \overset{(9.2.7)}{=} 2^{-2}\,\mathbb{E}\left[\left\{|Y|^\lambda \log|Y| - \frac{2}{\lambda^2}\nu\right\}^2\right]$$

$$\begin{aligned}
&\overset{(9.4.5)}{=} \frac{2^{-1}}{2^{1+1/\lambda}\Gamma(1+1/\lambda)} \cdot \frac{2^2}{\lambda^2}\int_0^\infty \left(\frac{1}{2}y^\lambda\right)^2 (\log y^\lambda)^2 e^{-\frac{1}{2}y^\lambda} dy \\
&\quad - \frac{\frac{2}{\lambda^2}\nu}{2^{1+1/\lambda}\Gamma(1+1/\lambda)} \cdot \frac{2}{\lambda}\int_0^\infty \left(\frac{1}{2}y^\lambda\right)(\log y^\lambda) e^{-\frac{1}{2}y^\lambda} dy + \frac{1}{\lambda^4}\nu^2 \\
&= \frac{1}{2^{1/\lambda}\lambda^2\Gamma(1+1/\lambda)}\int_0^\infty t^2 (\log 2 + \log t)^2 e^{-t}\left[\frac{2}{\lambda}(2t)^{1/\lambda-1}\right] dt \qquad (\text{with } t = \frac{1}{2}y^\lambda) \\
&\quad - \frac{2\nu}{2^{1/\lambda}\lambda^3\Gamma(1+1/\lambda)}\int_0^\infty t\,(\log 2 + \log t) e^{-t}\left[\frac{2}{\lambda}(2t)^{1/\lambda-1}\right] dt + \frac{1}{\lambda^4}\nu^2 \\
&= \frac{1}{\lambda^3\Gamma(1+1/\lambda)}\left\{\begin{array}{l}(\log 2)^2\Gamma(2+1/\lambda) + (2\log 2)\Gamma(2+1/\lambda)\psi(2+1/\lambda) \\ +\Gamma(2+1/\lambda)(\psi'(2+1/\lambda)+\psi^2(2+1/\lambda))\end{array}\right\} \\
&\quad - \frac{2\nu}{\lambda^4\Gamma(1+1/\lambda)}\left\{\begin{array}{l}(\log 2)\Gamma(1+1/\lambda) \\ +\Gamma(1+1/\lambda)\psi(1+1/\lambda)\end{array}\right\} + \frac{1}{\lambda^4}\nu^2 \\
&\quad\ \text{by (9.4.6), (9.4.10) and (9.4.11),} \\
&\overset{(9.4.7)}{=} \frac{1}{\lambda^4}\left[(\lambda+1)\left\{\begin{array}{l}(\log 2)^2 + (2\log 2)\psi(2+1/\lambda) \\ +\psi'(2+1/\lambda)+\psi^2(2+1/\lambda)\end{array}\right\} - \nu^2\right] \\
&\overset{(9.4.8)}{=} \frac{1}{\lambda^4}\left[(\lambda+1)\left\{\begin{array}{l}(2\log 2)\frac{\lambda}{\lambda+1} + \psi'(2+1/\lambda) \\ +\left(\frac{\lambda}{\lambda+1}\right)^2 + \frac{2\lambda}{\lambda+1}\psi(1+1/\lambda)\end{array}\right\} + \lambda\nu^2\right] \\
&\overset{(9.4.9)}{=} \frac{1}{\lambda^4}\left[2\lambda\nu + (\lambda+1)\psi'(1+1/\lambda) + \lambda\nu^2\right] \\
&= \boxed{\frac{1}{\lambda^3}\left[\nu(2+\nu) + (1+1/\lambda)\psi'(1+1/\lambda)\right]}, \qquad (9.4.16)
\end{aligned}$$

$$\begin{aligned}
J_{\theta_2\sigma} &= \mathbb{E}[d_{\theta_2}(Y)d_\sigma(Y)] \\
&\overset{(9.2.7)}{=} \frac{-\lambda}{4}\mathbb{E}\Big[|Y|^{2\lambda}\log|Y|\Big] + \frac{\nu}{2\lambda}\mathbb{E}\Big[|Y|^\lambda\Big] - \frac{\nu}{\lambda^2} + \frac{1}{2}\mathbb{E}\Big[|Y|^\lambda\log|Y|\Big] \\
&\overset{(9.4.5)}{=} \frac{-\lambda}{2^{2+1/\lambda}\Gamma(1+1/\lambda)} \cdot \frac{2^2}{\lambda}\int_0^\infty \left(\frac{1}{2}y^\lambda\right)^2 (\log y^\lambda) e^{-\frac{1}{2}y^\lambda} dy \\
&\quad + \frac{\nu}{2^{1+1/\lambda}\lambda\Gamma(1+1/\lambda)} \cdot 2\int_0^\infty \left(\frac{1}{2}y^\lambda\right) e^{-\frac{1}{2}y^\lambda} dy - \frac{\nu}{\lambda^2} \\
&\quad + \frac{1}{2^{1+1/\lambda}\Gamma(1+1/\lambda)} \cdot \frac{2}{\lambda}\int_0^\infty \left(\frac{1}{2}y^\lambda\right)(\log y^\lambda) e^{-\frac{1}{2}y^\lambda} dy \\
&= \frac{-1}{2^{1/\lambda}\Gamma(1+1/\lambda)}\int_0^\infty t^2 (\log 2 + \log t) e^{-t}\left[\frac{2}{\lambda}(2t)^{1/\lambda-1}\right] dt
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{\nu}{2^{1/\lambda}\lambda\Gamma(1+1/\lambda)} \int_0^\infty t\, e^{-t} \left[\frac{2}{\lambda}(2t)^{1/\lambda-1}\right] dt - \frac{\nu}{\lambda^2} \\
&+ \frac{1}{2^{1/\lambda}\lambda\Gamma(1+1/\lambda)} \int_0^\infty t\,(\log 2 + \log t) e^{-t} \left[\frac{2}{\lambda}(2t)^{1/\lambda-1}\right] dt \qquad (\text{with } t = \frac{1}{2}y^\lambda) \\
= \;& \frac{-1}{\lambda\Gamma(1+1/\lambda)} \Big\{(\log 2)\Gamma(2+1/\lambda) + \Gamma(2+1/\lambda)\psi(2+1/\lambda)\Big\} \\
&+ \frac{1}{\lambda^2\Gamma(1+1/\lambda)} \Big\{(\log 2)\Gamma(1+1/\lambda) + \Gamma(1+1/\lambda)\psi(1+1/\lambda)\Big\} \\
&\quad \text{by (9.4.6) and (9.4.10),} \\
\overset{(9.4.7)}{=} \;& \frac{-(\lambda+1)(\log 2 + \psi(2+1/\lambda))}{\lambda^2} + \frac{\nu}{\lambda^2} \overset{(9.4.8)}{=} \frac{-\lambda - (\lambda+1)\nu + \nu}{\lambda^2} \\
= \;& \boxed{-(1+\nu)/\lambda}, \qquad (9.4.17)
\end{aligned}
$$

$$
\begin{aligned}
J_{\sigma\sigma} \;=\; & \mathbb{E}[d_\sigma(Y) d_\sigma(Y)] \\
\overset{(9.2.7)}{=} \;& \frac{\lambda^2}{4}\mathbb{E}\Big[|Y|^{2\lambda}\Big] - \lambda\mathbb{E}\Big[|Y|^\lambda\Big] + 1 \\
\overset{(9.4.13)}{=} \;& (\lambda+2) - \lambda\mathbb{E}\Big[|Y|^\lambda\Big] \\
\overset{(9.4.5)}{=} \;& (\lambda+2) - \frac{\lambda}{2^{1/\lambda}\Gamma(1+1/\lambda)} \cdot 2\int_0^\infty \left(\frac{1}{2}y^\lambda\right) e^{-\frac{1}{2}y^\lambda} dy \\
= \;& (\lambda+2) - \frac{2\lambda}{2^{1/\lambda}\Gamma(1+1/\lambda)} \int_0^\infty t e^{-t} \left[\frac{2}{\lambda}(2t)^{1/\lambda-1}\right] dy \qquad (\text{with } t = \frac{1}{2}y^\lambda) \\
\overset{(9.4.6)}{=} \;& \boxed{\lambda}. \qquad (9.4.18)
\end{aligned}
$$

This ends the proof. □

Proof of Proposition 9.3.2. Assume H_0 throughout this proof. Use the fundamental theorem of calculus to expand $\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)$ around $\boldsymbol{\kappa}$:

$$
\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) = \boldsymbol{r}_n(\boldsymbol{\kappa}) + \int_0^1 \boldsymbol{r}_n'(\boldsymbol{\kappa}_{n,v}^\star) dv\, (\hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa}), \qquad (9.4.19)
$$

where $\boldsymbol{\kappa}_{n,v}^\star \stackrel{\circ}{=} \boldsymbol{\kappa} + v(\hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa})$.

From (9.2.8) and (9.2.7), we know that for all $\boldsymbol{t} \in \mathbb{R} \times (0,\infty)$,

$$
\boldsymbol{r}_n'(\boldsymbol{t}) = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n U_1(X_i,\boldsymbol{t}) & \frac{1}{n}\sum_{i=1}^n U_2(X_i,\boldsymbol{t}) \\ \frac{1}{n}\sum_{i=1}^n U_3(X_i,\boldsymbol{t}) & \frac{1}{n}\sum_{i=1}^n U_4(X_i,\boldsymbol{t}) \end{pmatrix} \qquad (9.4.20)
$$

where $y \doteq (x - t_1)/t_2$ and

$$U_1(x, \boldsymbol{t}) \doteq \frac{\lambda^2}{\sigma}|y|^{\lambda-1}; \qquad U_2(x, \boldsymbol{t}) \doteq \frac{\lambda^2}{\sigma}y|y|^{\lambda-1};$$
$$U_3(x, \boldsymbol{t}) \doteq \frac{1}{2\sigma}|y|^{\lambda-1}\mathrm{sign}\,(y)\,\{\lambda \log|y| + 1\}\,; \quad U_4(x, \boldsymbol{t}) \doteq \frac{1}{2\sigma}|y|^{\lambda}\,\{\lambda \log|y| + 1\}\,.$$

By the triangle inequality and Lemma 9.2.1, we can verify that for all $(k, \lambda) \in \{1, 2, 3, 4\} \times [1, \infty)\backslash\{(3, 1)\}$,

$$\mathbb{P}\bigg(\limsup_{n\to\infty}\sup_{v\in[0,1]}\bigg|\frac{1}{n}\sum_{i=1}^{n} U_k(X_i, \boldsymbol{\kappa}^{\star}_{n,v}) - U_k(X_i, \boldsymbol{\kappa})\bigg| > 0\bigg) = 0. \tag{9.4.21}$$

Since we already know from Proposition 9.3.3 that

$$\hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa} = O_{\mathbb{P}}(n^{-1/2})\mathbf{1}_2, \tag{9.4.22}$$

we deduce from (9.4.19), (9.4.20), (9.4.21) and (9.4.22) that, for all $(k, \lambda) \in \{1, 2, 3, 4\} \times [1, \infty)\backslash\{(3, 1)\}$,

$$\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n) = \boldsymbol{r}_n(\boldsymbol{\kappa}) + \boldsymbol{r}'_n(\boldsymbol{\kappa})(\hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa}) + o_{\mathbb{P}}(n^{-1/2})\mathbf{1}_2, \tag{9.4.23}$$

which is the statement we wanted to prove.

When $(k, \lambda) = (3, 1)$, we have to be a bit more careful. Indeed, Lemma 9.2.1 cannot be applied to U_3 in this case because the log term implies that, for any $\delta > 0$, $\sup_{\boldsymbol{t}\in B_\delta[\boldsymbol{\kappa}]}|U_3(x, \boldsymbol{t})| = \infty$ for all $x \in B_\delta[\mu]$, and thus (A.2) cannot be satisfied. Instead, we use the result from 9.5, which is a consequence of a uniform law of large numbers developed in Lafaye de Micheaux and Ouimet (2018) for summands that blow up. By using successively Jensen's inequality, Fubini's theorem, the triangle inequality and Lemma 9.5.1, we have

$$\begin{aligned}
&\mathbb{E}\bigg|\int_0^1 \frac{1}{n}\sum_{i=1}^{n} U_3(X_i, \boldsymbol{\kappa}^{\star}_{n,v})dv - \int_0^1 \frac{1}{n}\sum_{i=1}^{n} U_3(X_i, \boldsymbol{\kappa})dv\bigg| \\
&\qquad \leq \int_0^1 \mathbb{E}\bigg|\frac{1}{n}\sum_{i=1}^{n} U_3(X_i, \boldsymbol{\kappa}^{\star}_{n,v}) - \frac{1}{n}\sum_{i=1}^{n} U_3(X_i, \boldsymbol{\kappa})\bigg|dv \\
&\qquad \leq 2\sup_{v\in[0,1]}\mathbb{E}\bigg|\frac{1}{n}\sum_{i=1}^{n} U_3(X_i, \boldsymbol{\kappa}^{\star}_{n,v}) - \mathbb{E}\Big[U_3(X_1, \boldsymbol{\kappa})\Big]\bigg| \stackrel{n\to\infty}{\longrightarrow} 0.
\end{aligned} \tag{9.4.24}$$

By Markov's inequality, this yields, for $\lambda = 1$,

$$\left| \int_0^1 \frac{1}{n} \sum_{i=1}^n U_3(X_i, \boldsymbol{\kappa}_{n,v}^{\star}) dv - \int_0^1 \frac{1}{n} \sum_{i=1}^n U_3(X_i, \boldsymbol{\kappa}) dv \right| \xrightarrow{\mathbb{P}} 0. \tag{9.4.25}$$

Combining (9.4.22) and (9.4.25) into (9.4.19) proves the statement of the proposition when $(k, \lambda) = (3, 1)$. □

Proof of Proposition 9.3.3. Let $X \sim \mathrm{APD}(\boldsymbol{\theta}_0, \boldsymbol{\kappa})$ and $Y \mathrel{\hat=} (X - \mu)/\sigma$. By the weak law of large numbers, the chain rule and integration by parts,

$$\begin{aligned} \boldsymbol{r}_n'(\boldsymbol{\kappa}) &= \mathbb{E}\Big[\frac{\partial}{\partial \boldsymbol{\kappa}^\top} \boldsymbol{d}_{\boldsymbol{\theta}}(Y)\Big] + o_{\mathbb{P}}(1)\mathbf{1}_2 \\ &= \mathbb{E}\Big[\boldsymbol{d}_{\boldsymbol{\theta}}'(Y) \frac{\partial Y}{\partial \boldsymbol{\kappa}^\top}\Big] + o_{\mathbb{P}}(1)\mathbf{1}_2 \\ &= \Big[\boldsymbol{d}_{\boldsymbol{\theta}}(y) \frac{\partial y}{\partial \boldsymbol{\kappa}^\top} f(y \,|\, \boldsymbol{\theta}_0)\Big]\Big|_{-\infty}^{\infty} \\ &\qquad - \int_{-\infty}^{\infty} \boldsymbol{d}_{\boldsymbol{\theta}}(y) \frac{\partial}{\partial y}\Big[\frac{\partial y}{\partial \boldsymbol{\kappa}^\top} f(y \,|\, \boldsymbol{\theta}_0)\Big] dy + o_{\mathbb{P}}(1)\mathbf{1}_2 \\ &\overset{(9.2.6)}{=} [0] - \mathbb{E}\Big[\boldsymbol{d}_{\boldsymbol{\theta}}(Y) \sigma^{-1} \boldsymbol{d}_{\boldsymbol{\kappa}}(Y)^\top\Big] + o_{\mathbb{P}}(1)\mathbf{1}_2 \\ &\overset{(9.3.4)}{=} -\sigma^{-1} J_{\boldsymbol{\theta}\boldsymbol{\kappa}} + o_{\mathbb{P}}(1) I_2. \end{aligned} \tag{9.4.26}$$

This proves (9.3.7). Now, we show the asymptotics of $n^{1/2}(\hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa})$. From (9.2.6), note that

$$\frac{\partial}{\partial \boldsymbol{\kappa}} \log g(X \,|\, \boldsymbol{\theta}_0, \boldsymbol{\kappa}) = \sigma^{-1} d_{\boldsymbol{\kappa}}(Y). \tag{9.4.27}$$

A direct application of Theorem 5.23 in van der Vaart (1998) with $m_{\boldsymbol{\kappa}}(x) \mathrel{\hat=} |(x - \kappa_1)/\kappa_2|^{\lambda} - \log \kappa_2$ (by definition, $\hat{\boldsymbol{\kappa}}_n \in \mathrm{argmax}_{\boldsymbol{\kappa} \in \mathbb{R} \times (0,\infty)} \sum_{i=1}^n m_{\boldsymbol{\kappa}}(X_i)$, recall (9.2.9)), combined with the almost-sure convergence $\hat{\boldsymbol{\kappa}}_n \to \boldsymbol{\kappa}$ from Lemma 9.2.2, yields

$$\begin{aligned} n^{1/2}(\hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa}) &= \mathbb{E}\Big[\sigma^{-1} \boldsymbol{d}_{\boldsymbol{\kappa}}(Y) \sigma^{-1} \boldsymbol{d}_{\boldsymbol{\kappa}}(Y)^\top\Big]^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \sigma^{-1} \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) + o_{\mathbb{P}}(1)\mathbf{1}_2 \\ &\overset{(9.3.4)}{=} \sigma J_{\boldsymbol{\kappa}\boldsymbol{\kappa}}^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) + o_{\mathbb{P}}(1)\mathbf{1}_2. \end{aligned} \tag{9.4.28}$$

This proves (9.3.8). Finally, since $\frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i)$ is $O_{\mathbb{P}}(1)$ by Proposition 9.3.1, Equation (9.3.9) follows directly from Proposition 9.3.2, (9.3.7) and (9.3.8). □

Proof of Theorem 9.3.4. The asymptotic normality of $n^{1/2}\boldsymbol{r}_n(\hat{\boldsymbol{\kappa}}_n)$ follows directly from Proposition 9.3.3 and Proposition 9.3.1. The asymptotic covariance matrix Σ is given by (note that $J_{\boldsymbol{\theta\kappa}}$ and $J_{\boldsymbol{\kappa\kappa}}$ are diagonal):

$$\begin{aligned}
\Sigma \quad &= \quad \left(I_2\,;\,-J_{\boldsymbol{\theta\kappa}}J_{\boldsymbol{\kappa\kappa}}^{-1}\right)\begin{pmatrix} J_{\boldsymbol{\theta\theta}} & J_{\boldsymbol{\theta\kappa}} \\ J_{\boldsymbol{\theta\kappa}} & J_{\boldsymbol{\kappa\kappa}} \end{pmatrix}\begin{pmatrix} I_2 \\ -J_{\boldsymbol{\kappa\kappa}}^{-1}J_{\boldsymbol{\theta\kappa}} \end{pmatrix} \quad = \quad J_{\boldsymbol{\theta\theta}} - J_{\boldsymbol{\kappa\kappa}}^{-1}J_{\boldsymbol{\theta\kappa}}^2 \\
&\overset{(9.3.4)}{=} \begin{pmatrix} 4(1+\lambda) & 0 \\ 0 & \frac{\phi^2+\beta\psi'(\beta)-1}{\lambda^3} \end{pmatrix} - \begin{pmatrix} \frac{2^{2/\lambda}\Gamma(\beta)}{\lambda\Gamma(3-\beta)} & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix}\begin{pmatrix} \frac{2^{2-2/\lambda}\lambda^2}{(\Gamma(\beta))^2} & 0 \\ 0 & \frac{\phi^2}{\lambda^2} \end{pmatrix} \qquad (9.4.29) \\
&= \begin{pmatrix} 4(1+\lambda) - \frac{4\lambda}{\Gamma(3-\beta)\Gamma(\beta)} & 0 \\ 0 & \frac{\beta\psi'(\beta)-1}{\lambda^3} \end{pmatrix}.
\end{aligned}$$

This ends the proof. □

9.4.3. Proof of the results stated in Section 9.3.2

In order to establish our results under the local alternatives $H_{1,n}$, we use Le Cam's first and third lemma (see Lemma 6.4 and Example 6.7 in van der Vaart (1998)). The proof structure in this section is inspired by the one presented in Falk et al. (2008).

Lemma 9.4.1 (Le Cam's first lemma)**.** *Let $(P_n, n\in\mathbb{N})$ and $(Q_n, n\in\mathbb{N})$ be sequences of probability measures on the measurable spaces $(\Omega_n,\mathcal{A}_n)$. Then, the following statements are equivalent:*

- *(i) $Q_n \lhd P_n$, i.e. $(Q_n, n\in\mathbb{N})$ is contiguous with respect to $(P_n, n\in\mathbb{N})$.*
- *(ii) If $\frac{dP_n}{dQ_n} \overset{Q_n}{\rightsquigarrow} U$ along a subsequence, then $\mathbb{P}(U>0)=1$.*
- *(iii) If $\frac{dQ_n}{dP_n} \overset{P_n}{\rightsquigarrow} V$ along a subsequence, then $\mathbb{E}[V]=1$.*
- *(iv) For any statistics $\boldsymbol{T}_n:\Omega_n\to\mathbb{R}^k$: If $\boldsymbol{T}_n \xrightarrow{P_n} 0$, then $\boldsymbol{T}_n \xrightarrow{Q_n} 0$.*

Lemma 9.4.2 (Le Cam's third lemma)**.** *Let $(P_n, n\in\mathbb{N})$ and $(Q_n, n\in\mathbb{N})$ be sequences of probability measures on the measurable spaces $(\Omega_n,\mathcal{A}_n)$, and let $\boldsymbol{W}_n:\Omega_n\to\mathbb{R}^k$ be a sequence of random vectors. Suppose that $Q_n \lhd P_n$ and*

$$\begin{pmatrix} \boldsymbol{W}_n \\ \log\frac{dQ_n}{dP_n} \end{pmatrix} \overset{P_n}{\rightsquigarrow} \mathcal{N}_{k+1}\left(\begin{pmatrix} m \\ -\frac{1}{2}s^2 \end{pmatrix}, \begin{pmatrix} M & \tau \\ \tau^\top & s^2 \end{pmatrix}\right), \qquad (9.4.30)$$

then $\boldsymbol{W}_n \overset{Q_n}{\rightsquigarrow} \mathcal{N}_k(m+\tau, M)$.

Proof of Proposition 9.3.5. To prove this result, we use Le Cam's first lemma. Assume that our vector of observations is the identity function

$$\boldsymbol{X} \mathrel{\hat=} (X_1, X_2, \dots, X_n) \mathrel{\hat=} \mathrm{Id} : (\Omega_n \mathrel{\hat=} \mathbb{R}^n, \mathcal{A}_n \mathrel{\hat=} \mathcal{L}(\mathbb{R}^n), \lambda) \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda), \tag{9.4.31}$$

where $\mathcal{L}(\mathbb{R}^n)$ denotes the completion of the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$, and where λ denotes the Lebesgue measure. On $(\Omega_n, \mathcal{A}_n)$, define the probability measures

$$\begin{aligned} \mathbb{P}_{H_{0,n}}(A) &\mathrel{\hat=} \int_A \prod_{i=1}^n g(X_i(\omega) \,|\, \boldsymbol{\theta}_0, \boldsymbol{\kappa}) \, d\lambda(\omega), \quad A \in \mathcal{A}_n, \\ \mathbb{P}_{H_{1,n}}(A) &\mathrel{\hat=} \int_A \prod_{i=1}^n g(X_i(\omega) \,|\, \boldsymbol{\theta}_n, \boldsymbol{\kappa}) \, d\lambda(\omega), \quad A \in \mathcal{A}_n, \end{aligned} \tag{9.4.32}$$

where $\boldsymbol{\theta}_n \mathrel{\hat=} \boldsymbol{\theta}_0 + (1 + o(1)) n^{-1/2} \boldsymbol{\delta}$. By construction, the law of $\boldsymbol{X}$ under $\mathbb{P}_{H_{0,n}}$ corresponds to the null hypothesis H_0 and the law $\boldsymbol{X}$ under $\mathbb{P}_{H_{1,n}}$ corresponds the alternative hypothesis $H_{1,n}$. Since g is positive on $\mathbb{R}$, the measures $\mathbb{P}_{H_{0,n}}$, $\mathbb{P}_{H_{1,n}}$ and λ are equivalent on $(\Omega_n, \mathcal{A}_n)$. From (9.4.32), we deduce that

$$\frac{d\mathbb{P}_{H_{1,n}}}{d\mathbb{P}_{H_{0,n}}} = \frac{d\mathbb{P}_{H_{1,n}}/d\lambda}{d\mathbb{P}_{H_{0,n}}/d\lambda} = \frac{\prod_{i=1}^n g(X_i \,|\, \boldsymbol{\theta}_n, \boldsymbol{\kappa})}{\prod_{i=1}^n g(X_i \,|\, \boldsymbol{\theta}_0, \boldsymbol{\kappa})} = \frac{\prod_{i=1}^n f(Y_i \,|\, \boldsymbol{\theta}_n)}{\prod_{i=1}^n f(Y_i \,|\, \boldsymbol{\theta}_0)}, \tag{9.4.33}$$

where $Y_i \mathrel{\hat=} (X_i - \mu)/\sigma$.

Using a second-order Taylor expansion around $\boldsymbol{\theta}_0$, we have, under $H_0 : X_i \sim \mathrm{APD}(\boldsymbol{\theta}_0, \boldsymbol{\kappa})$,

$$\begin{aligned} \log\left(\frac{d\mathbb{P}_{H_{1,n}}}{d\mathbb{P}_{H_{0,n}}}\right) &= \sum_{i=1}^n (\log f(Y_i \,|\, \boldsymbol{\theta}_n) - \log f(Y_i \,|\, \boldsymbol{\theta}_0)) \\ &= (1 + o(1)) \, \boldsymbol{\delta}^\top \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ &\quad + (1 + o(1))^2 \, \boldsymbol{\delta}^\top \int_0^1 \int_0^1 v \frac{1}{n} \sum_{i=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \log f(Y_i \,|\, \boldsymbol{t}_{n,u,v}) du dv \, \boldsymbol{\delta}, \end{aligned} \tag{9.4.34}$$

where $\boldsymbol{t}_{n,u,v} \mathrel{\hat=} \boldsymbol{\theta}_0 + uv(\boldsymbol{\theta}_n - \boldsymbol{\theta}_0)$. From the convergence of the first two components in (9.3.4), we know that, as $n \to \infty$,

$$(1 + o(1)) \, \boldsymbol{\delta}^\top \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \overset{\mathbb{P}_{H_{0,n}}}{\rightsquigarrow} \mathcal{N}(0, \boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}} \, \boldsymbol{\delta}). \tag{9.4.35}$$

For the second term on the right-hand side of (9.4.34), we want to apply a standard uniform law of large numbers (Lemma 9.2.1). From the expression of $f(y \,|\, \boldsymbol{t})$ in (9.1.1), we see that for each $(j,k) \in \{1,2\}^2$, the function $U_{j,k}(y, \boldsymbol{t}) \mathrel{\hat{=}} \frac{\partial^2}{\partial\theta_j \partial\theta_k} \log f(y \,|\, \boldsymbol{t})$ satisfies:

(A.1): For all $y \in \mathbb{R}$, $\boldsymbol{t} \mapsto U_{j,k}(y, \boldsymbol{t})$ is continuous on the compact $\mathcal{C} \mathrel{\hat{=}} [\frac{1}{4}, \frac{3}{4}] \times [\frac{\lambda}{2}, \frac{3\lambda}{2}]$;

(A.2): There exists a finite polynomial $K : \mathbb{R} \to \mathbb{R}$ such that $|U_{j,k}(y, \boldsymbol{t})| \leq K(|y|)$ for all $(y, \boldsymbol{t}) \in \mathbb{R} \times \mathcal{C}$ (which implies that $K(|y|)$ is integrable under $f(y \,|\, \boldsymbol{\theta}_0) dy$).

Take $N \in \mathbb{N}$ large enough that $\boldsymbol{\theta}_n \in \mathcal{C}$ for all $n \geq N$. By Jensen's inequality and Lemma 9.2.1 (under H_0), we deduce that

$$
\begin{aligned}
&\left| \int_0^1 \int_0^1 v \frac{1}{n} \sum_{i=1}^n U_{j,k}(Y_i, \boldsymbol{t}_{n,u,v}) du dv - \int_0^1 \int_0^1 v \frac{1}{n} \sum_{i=1}^n \overline{U}_{j,k}(\boldsymbol{\theta}_0) du dv \right| \\
&\quad \leq \int_0^1 \int_0^1 v \frac{1}{n} \sum_{i=1}^n \left| U_{j,k}(Y_i, \boldsymbol{t}_{n,u,v}) - \overline{U}_{j,k}(\boldsymbol{\theta}_0) \right| du dv \\
&\quad \leq \frac{1}{2} \sup_{\boldsymbol{t} \in B_{\|\boldsymbol{\theta}_n - \boldsymbol{\theta}_0\|_2}[\boldsymbol{\theta}_0]} \frac{1}{n} \sum_{i=1}^n \left| U_{j,k}(Y_i, \boldsymbol{t}) - \overline{U}_{j,k}(\boldsymbol{\theta}_0) \right| \stackrel{\mathbb{P}_{H_{0,n}}}{\longrightarrow} 0.
\end{aligned} \tag{9.4.36}
$$

By definition of the matrix J in (9.3.4), note that $\overline{U}_{j,k}(\boldsymbol{\theta_0}) = -J_{\theta_j \theta_k}$ (this can be seen by integrating by parts). Hence, (9.4.36) shows that the second term on the right-hand side of (9.4.34) is equal to $-\frac{1}{2} \boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}} \, \boldsymbol{\delta} + o_{\mathbb{P}_{H_{0,n}}}(1)$. We deduce that

$$
\log \left(\frac{d\mathbb{P}_{H_{1,n}}}{d\mathbb{P}_{H_{0,n}}} \right) \stackrel{\mathbb{P}_{H_{0,n}}}{\rightsquigarrow} \mathcal{N}(-\frac{1}{2} \boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}} \, \boldsymbol{\delta}, \boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}} \, \boldsymbol{\delta}). \tag{9.4.37}
$$

Define a random variable $V > 0$ such that $\log(V) \stackrel{\mathbb{P}_{H_{0,n}}}{\sim} \mathcal{N}(-\frac{1}{2} \boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}} \, \boldsymbol{\delta}, \boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}} \, \boldsymbol{\delta})$. The continuous mapping theorem implies that

$$
\frac{d\mathbb{P}_{H_{1,n}}}{d\mathbb{P}_{H_{0,n}}} \stackrel{\mathbb{P}_{H_{0,n}}}{\rightsquigarrow} V. \tag{9.4.38}
$$

By the definition of V, we have $\mathbb{E}_{H_{0,n}}[V] = 1$. This shows (iii) in Lemma 9.4.1 with $P_n = \mathbb{P}_{H_{0,n}}$ and $Q_n = \mathbb{P}_{H_{1,n}}$, which implies $\mathbb{P}_{H_{1,n}} \triangleleft \mathbb{P}_{H_{0,n}}$ by (i). Define $U \mathrel{\hat{=}} V$ and note that $\mathbb{P}_{H_{0,n}}(U > 0) = 1$ by definition of V. This shows (ii) in Lemma 9.4.1 where the roles of P_n and Q_n have been interchanged, which implies $\mathbb{P}_{H_{0,n}} \triangleleft \mathbb{P}_{H_{1,n}}$ by (i). We conclude that the sequences $(\mathbb{P}_{H_{0,n}}, n \in \mathbb{N})$ and $(\mathbb{P}_{H_{1,n}}, n \in \mathbb{N})$ are mutually contiguous, which we denote by $\mathbb{P}_{H_{0,n}} \triangleleft \triangleright \mathbb{P}_{H_{1,n}}$. The conclusion follows from (iv). $\square$

Proof of Proposition 9.3.7. From the expressions that we found for the two terms on the right-hand side of (9.4.34) in the proof of Proposition 9.3.5, we have

$$\begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) \\ \log\left(\frac{d\mathbb{P}_{H_{1,n}}}{d\mathbb{P}_{H_{0,n}}}\right) \end{pmatrix} = \begin{pmatrix} \mathbf{0}_2 \\ \mathbf{0}_2 \\ -\frac{1}{2}\boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}}\, \boldsymbol{\delta} + o_{\mathbb{P}_{H_{0,n}}}(1) \end{pmatrix} + \begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) \\ (1+o(1))\, \boldsymbol{\delta}^\top \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \end{pmatrix},$$

where $\mathbf{0}_2 \mathrel{\hat{=}} (0,0)^\top$. By the central limit theorem (see the definition of J in Proposition 9.3.1), we obtain that, under H_0,

$$\begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) \\ \log\left(\frac{d\mathbb{P}_{H_{1,n}}}{d\mathbb{P}_{H_{0,n}}}\right) \end{pmatrix} \overset{\mathbb{P}_{H_0}}{\rightsquigarrow} \mathcal{N}_5\left(\begin{pmatrix} \mathbf{0}_2 \\ \mathbf{0}_2 \\ -\frac{1}{2}\boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}}\, \boldsymbol{\delta} \end{pmatrix}; \begin{pmatrix} J_{\boldsymbol{\theta}\boldsymbol{\theta}} & J_{\boldsymbol{\theta}\boldsymbol{\kappa}} & J_{\boldsymbol{\theta}\boldsymbol{\theta}}\boldsymbol{\delta} \\ J_{\boldsymbol{\theta}\boldsymbol{\kappa}} & J_{\boldsymbol{\kappa}\boldsymbol{\kappa}} & J_{\boldsymbol{\theta}\boldsymbol{\kappa}}\boldsymbol{\delta} \\ \boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}} & \boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\kappa}} & \boldsymbol{\delta}^\top J_{\boldsymbol{\theta}\boldsymbol{\theta}}\boldsymbol{\delta} \end{pmatrix}\right).$$

Then, by Le Cam's third lemma,

$$\begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\theta}}(Y_i) \\ \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{d}_{\boldsymbol{\kappa}}(Y_i) \end{pmatrix} \overset{\mathbb{P}_{H_{1,n}}}{\rightsquigarrow} \mathcal{N}_4\left(\begin{pmatrix} J_{\boldsymbol{\theta}\boldsymbol{\theta}}\boldsymbol{\delta} \\ J_{\boldsymbol{\theta}\boldsymbol{\kappa}}\boldsymbol{\delta} \end{pmatrix}; \begin{pmatrix} J_{\boldsymbol{\theta}\boldsymbol{\theta}} & J_{\boldsymbol{\theta}\boldsymbol{\kappa}} \\ J_{\boldsymbol{\theta}\boldsymbol{\kappa}} & J_{\boldsymbol{\kappa}\boldsymbol{\kappa}} \end{pmatrix}\right).$$

This ends the proof. □

9.5. Appendix

Lemma 9.5.1. *Let $X_1, X_2, X_3, \ldots$ be a sequence of i.i.d. random variables such that $X_1 \sim APD(\boldsymbol{\theta}_0, \boldsymbol{\kappa})$, where $\lambda = 1$, $\mu \in \mathbb{R}$ and $\sigma > 0$, i.e. the density of X_1 is given by*

$$f_{X_1}(x) \mathrel{\hat{=}} \frac{1}{4\sigma} e^{-\frac{1}{2}\left|\frac{x-\mu}{\sigma}\right|}, \quad x \in \mathbb{R}. \tag{9.5.1}$$

Define $H : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ by

$$H(y) \mathrel{\hat{=}} sign(y)(\log|y| + 1). \tag{9.5.2}$$

Let $\{\hat{\mu}_n\}_{n\in\mathbb{N}}$ and $\{\hat{\sigma}_n\}_{n\in\mathbb{N}}$ be the sequences of maximum likelihood estimators found in (9.2.11) for $\lambda = 1$:

$$\hat{\mu}_n \mathrel{\hat{=}} median(X_1, X_2, \ldots, X_n) \quad and \quad \hat{\sigma}_n = \frac{1}{n}\sum_{i=1}^n \frac{1}{2}|X_i - \hat{\mu}_n|. \tag{9.5.3}$$

The median is defined in (9.2.12). For $v \in [0,1]$, let $\mu^\star_{n,v} \mathrel{\hat=} \mu + v(\hat{\mu}_n - \mu)$ and $\sigma^\star_{n,v} \mathrel{\hat=} \sigma + v(\hat{\sigma}_n - \sigma)$. Then,

$$\lim_{n\to\infty} \sup_{v\in[0,1]} \mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i \neq \mu^\star_{n,v}\}} H\left(\frac{X_i - \mu^\star_{n,v}}{\sigma^\star_{n,v}}\right) - \mathbb{E}\left[H\left(\frac{X_1 - \mu}{\sigma}\right)\right]\right| = 0. \tag{9.5.4}$$

Proof. Without loss of generality, assume that $\mu = 0$. Since $\sigma > 0$ and $\hat{\sigma}_n > 0$ a.s., we have $\sigma^\star_{n,v} > 0$ a.s. for any $v \in [0,1]$, which implies that the factors $\sigma^\star_{n,v}$ and σ in the sign function of H are irrelevant. Also, f_{X_1} is symmetric, so $\mathbb{E}[\mathrm{sign}(X_1)] = 0$. Combining these facts together, the supremum in (9.5.4) is bounded from above by

$$\begin{aligned}(c) + (d) \mathrel{\hat=} &\sup_{v\in[0,1]} \mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i \neq \mu^\star_{n,v}\}} h(X_i - \mu^\star_{n,v}) - \mathbb{E}\Big[h(X_1)\Big]\right| \\ &+ \sup_{v\in[0,1]} \mathbb{E}\left|\Big(1 - \log \sigma^\star_{n,v}\Big) \cdot \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i \neq \mu^\star_{n,v}\}} \mathrm{sign}(X_i - \mu^\star_{n,v})\right|,\end{aligned} \tag{9.5.5}$$

where $h(y) \mathrel{\hat=} \mathrm{sign}(y) \log|y|$. By Lemma 3.1 in Lafaye de Micheaux and Ouimet (2018), we have $(c) \to 0$.

It remains to prove that $(d) \to 0$ in (9.5.5). By the Cauchy-Schwarz inequality,

$$\begin{aligned}(d)^2 &\le \mathbb{E}\left[\sup_{v\in[0,1]} \Big(1 - \log \sigma^\star_{n,v}\Big)^2\right] \cdot \mathbb{E}\left[\sup_{v\in[0,1]} \left(\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i \neq \mu^\star_{n,v}\}} \mathrm{sign}(X_i - \mu^\star_{n,v})\right)^2\right] \\ &\mathrel{\hat=} (d.1) \cdot (d.2).\end{aligned} \tag{9.5.6}$$

We show that $(d.1)$ is bounded and $(d.2)$ tends to zero as $n \to \infty$. We start with $(d.2)$. For every $\omega \in \Omega$, the function

$$v \mapsto \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i(\omega) \neq \mu^\star_{n,v}(\omega)\}} \mathrm{sign}(X_i(\omega) - \mu^\star_{n,v}(\omega)) \tag{9.5.7}$$

is monotone and equal to 0 at $v = 1$ (by definition of $\hat{\mu}_n$, recall (9.2.10)). Therefore, for each $\omega \in \Omega$, the supremum of the square in $(d.2)$ is always attained at $v = 0$. We deduce that

$$(d.2) = \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i \neq 0\}} \mathrm{sign}(X_i)\right)^2\right] \longrightarrow 0, \quad n \to \infty, \tag{9.5.8}$$

by law of large numbers in L^2 ($\mathbb{E}[\mathbb{1}_{\{X_1 \neq 0\}} \mathrm{sign}(X_1)] = 0$ and the sequence of averages is uniformly bounded).

Now we show that $(d.1)$ is bounded. By successively using the inequality $(\alpha-\beta)^2 \leq 2\alpha^2+2\beta^2$, the fact that $z \mapsto (\log z)^2$ always maximizes at one of the two end points on any closed sub-interval of $(0,\infty)$, and the inequality $\max\{a,b\} \leq a+b$ for $a,b \geq 0$, we have

$$(d.1) \leq \mathbb{E}\left[\sup_{v\in[0,1]} 2+2\,(\log \sigma^{\star}_{n,v})^2\right] \leq 2+2\,(\log \sigma)^2 + 2\,\mathbb{E}\Big[(\log \hat{\sigma}_n)^2\Big]. \tag{9.5.9}$$

It remains to show that $\mathbb{E}[(\log \hat{\sigma}_n)^2] < \infty$. Since $\hat{\sigma}_n$ is a mean of integrable terms (see (9.5.3)), we expect, at least heuristically (because of large deviations), that, as $n \to \infty$, its density concentrates more and more around σ and decays exponentially faster and faster in the right tail. The specific form of the density of $\hat{\sigma}_n$ is given in Equation (32) of Karst and Polowy (1963) and confirms the intuition. For $N \in \mathbb{N}$ large enough (depending on σ), there exists $\lambda_\sigma > 0$ small enough that, for all $n \geq N$,

$$\begin{aligned}
\mathbb{E}\Big[(\log \hat{\sigma}_n)^2\Big] &= \int_{(0,\sigma/2)\cup(\sigma/2,(3\sigma/2)\vee 1)\cup((3\sigma/2)\vee 1,\infty)} (\log s)^2 \cdot f_{\hat{\sigma}_n}(s)ds \\
&\leq \underbrace{\int_0^{\sigma/2} (\log(s))^2 \cdot 1\, ds}_{<\infty} + M_\sigma \underbrace{\int_{\sigma/2}^{(3\sigma/2)\vee 1} f_{\hat{\sigma}_n}(s)ds}_{\leq 1} + \underbrace{\int_{(3\sigma/2)\vee 1}^{\infty} s\cdot e^{-\lambda_\sigma s}ds}_{<\infty} \\
&< \infty,
\end{aligned} \tag{9.5.10}$$

where $a \vee b \mathrel{\hat{=}} \max\{a,b\}$ and $M_\sigma \mathrel{\hat{=}} \max_{s\in[\sigma/2,(3\sigma/2)\vee 1]}(\log s)^2 < \infty$. This ends the proof. $\square$

Notation

$\mathrel{\hat{=}}$	A definition or an equality that holds by definition
$\mathbf{1}_d$	The d-dimensional vector $(1,1,\ldots,1)^\top$
I_d	The identity matrix of order d
$\mathcal{N}_d(\,\cdot\,,\cdot\,)$	A d-dimensional normal distribution
$\overset{P}{\rightsquigarrow}$	Convergence in law under the measure P
$\overset{P}{\longrightarrow}$	Convergence in probability under the measure P

$\mathbb{P}_{H_0}$	Measure $\mathbb{P}$ conditional on the hypothesis H_0
$\mathbb{P}_{H_1}$	Measure $\mathbb{P}$ conditional on the hypothesis H_1
χ_2^2	Chi-square distribution with 2 degrees of freedom
$\chi_2^2(\gamma)$	χ_2^2 distribution with noncentrality parameter $\gamma \in \mathbb{R}$

Bibliography

M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964. MR0167642.

A. Desgagné and P. Lafaye de Micheaux. A powerful and interpretable alternative to the Jarque-Bera test of normality based on 2nd-power skewness and kurtosis, using the Rao's score test on the APD family. *J. Appl. Stat.*, 45(13):2307–2327, 2018. MR3841677.

M. Falk, A. Guillou, and G. Toulemonde. A LAN based Neyman smooth test for Pareto distributions. *J. Statist. Plann. Inference*, 138(10):2867–2886, 2008. MR2442221.

T. S. Ferguson. *A course in large sample theory.* Texts in Statistical Science Series. Chapman & Hall, London, 1996. MR1699953.

C. Fernández, J. Osiewalski, and M. F. J. Steel. Modeling and inference with v-spherical distributions. *J. Amer. Statist. Assoc.*, 90(432):1331–1340, 1995. MR1379475.

O. J. Karst and H. Polowy. Sampling properties of the median of a Laplace distribution. *Amer. Math. Monthly*, 70:628–636, 1963. MR0153074.

I. Komunjer. Asymmetric power distribution: theory and applications to risk measurement. *J. Appl. Econometrics*, 22(5):891–921, 2007. MR2395888.

S. Kotz, T. J. Kozubowski, and K. Podgórski. *The Laplace distribution and generalizations.* Birkhäuser Boston, Inc., Boston, MA, 2001. MR1935481.

P. Lafaye de Micheaux and F. Ouimet. A uniform L^1 law of large numbers for functions of i.i.d. random variables that are translated by a consistent estimator. *Statist. Probab. Lett.*, 142:109–117, 2018. MR3842623.

S. Nadarajah. A generalized normal distribution. *J. Appl. Stat.*, 32(7):685–694, 2005. MR2119411.

H. Rubin and A. L. Rukhin. Convergence rates of large deviations probabilities for point estimators. *Statist. Probab. Lett.*, 1(4):197–202, 1983. MR709248.

A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. MR1652247.

Conclusion

10.1. Conjectures

Here is a partial list of conjectures that could be reachable with some effort. It should be possible to prove (or disprove) them by following/extending some of the strategies presented in this thesis and/or the references on which it is based.

Conjecture 10.1.1 (Limiting two-overlap distribution for the Riemann zeta function)**.** *Let $\theta > -1$ be given. For $h, h' \in I \mathrel{\hat{=}} [-\log^\theta T, \log^\theta T]$, define the overlaps by*

$$\rho(h,h') \mathrel{\hat{=}} \frac{\mathbb{E}\Big[\log|\zeta(\tfrac{1}{2}+i\tau+ih)| \, \log|\zeta(\tfrac{1}{2}+i\tau+ih')|\Big]}{\sqrt{\mathbb{E}\Big[(\log|\zeta(\tfrac{1}{2}+i\tau+ih)|)^2\Big]}\sqrt{\mathbb{E}\Big[(\log|\zeta(\tfrac{1}{2}+i\tau+ih')|)^2\Big]}}, \tag{10.1.1}$$

where $\tau \sim$ Uniform$(T, 2T)$ under $\mathbb{P}$. For $\beta > 0$, define the Gibbs measure by

$$\mathcal{G}_{\beta,T}(A) \mathrel{\hat{=}} \int_A \frac{|\zeta(\frac{1}{2}+i\tau+ih)|^\beta}{\int_I |\zeta(\frac{1}{2}+i\tau+ih')|^\beta dh'} dh, \quad A \in \mathcal{B}(I), \tag{10.1.2}$$

where $\mathcal{B}(I)$ denotes the Borel σ-algebra on I. For

$$\beta_c(\theta) \mathrel{\hat{=}} \begin{cases} 2, & \text{if } \theta \leq 0, \\ 2\sqrt{1+\theta}, & \text{if } \theta > 0, \end{cases} \tag{10.1.3}$$

and any Borel set $A \in \mathcal{B}(\mathbb{R})$, we have

$$\lim_{T\to\infty} \mathbb{E}\mathcal{G}_{\beta,T}^{\times 2}\Big[\mathbf{1}_{\{\rho(h,h')\in A\}}\Big] = \begin{cases} \delta_{|0\wedge\theta|}(A), & \text{if } \beta \leq \beta_c(\theta), \\ \frac{\beta_c(\theta)}{\beta}\delta_{|0\wedge\theta|}(A) + \Big(1 - \frac{\beta_c(\theta)}{\beta}\Big)\delta_1(A), & \text{if } \beta > \beta_c(\theta). \end{cases} \tag{10.1.4}$$

To state the second conjecture, recall the definition of a *Poisson-Dirichlet variable*. For any $\lambda \in (0,1)$, let $\eta \mathrel{\hat{=}} (\eta_k)_{k\in\mathbb{N}}$ denote the atoms of a Poisson point process on $(0,\infty)$ with intensity $t \mapsto \lambda t^{-\lambda-1}$. A Poisson-Dirichlet variable $\xi \mathrel{\hat{=}} (\xi_k)_{k\in\mathbb{N}}$ of parameter λ is a random variable on the space of decreasing weights,

$$\Big\{(x_1, x_2, \dots) \in [0,1]^{\mathbb{N}} : 1 \geq x_1 \geq x_2 \geq \dots \geq 0 \ \text{ and } \ \sum_{k=1}^\infty x_k = 1\Big\}, \tag{10.1.5}$$

which has the same law as

$$\left(\frac{\eta_k}{\sum_{j=1}^\infty \eta_j}, \ k \in \mathbb{N}\right)_{\downarrow}, \tag{10.1.6}$$

where $\downarrow$ stands for the decreasing rearrangement.

Conjecture 10.1.2 (Limiting joint distribution of the overlaps for the Riemann zeta function)**.** *Let $\theta > -1$ and $\beta > 0$ be given, and let $\xi = (\xi_k)_{k\in\mathbb{N}}$ be a Poisson-Dirichlet variable of parameter $\beta_c(\theta)/\beta$. Denote by E the expectation with respect to ξ. Then, for any $n \in \mathbb{N}$ and any continuous function $\psi : [|\theta \wedge 0|, 1]^{n(n-1)/2} \to \mathbb{R}$ of the overlaps of n points,*

$$
\begin{aligned}
&\lim_{T\to\infty} \mathbb{E}\mathcal{G}_{\beta,N}^{\times n}\Big[\psi\Big((\rho(h_\ell, h_{\ell'}))_{1\le \ell,\ell'\le n}\Big)\Big] \\
&= \begin{cases} \psi\Big(|0\wedge\theta|\mathrm{J}_n + (1 - |0\wedge\theta|)\mathrm{Id}_n\Big), & \text{if } \beta \le \beta_c(\theta), \\ E\Big[\sum_{k_1,\dots,k_n\in\mathbb{N}} \xi_{k_1}\dots\xi_{k_n}\, \psi\Big(\Big(|0\wedge\theta| + (1 - |0\wedge\theta|)\mathbf{1}_{\{k_\ell = k_{\ell'}\}}\Big)_{1\le\ell,\ell'\le n}\Big)\Big], & \text{if } \beta > \beta_c(\theta), \end{cases}
\end{aligned}
$$

where J_n denotes the $n\times n$ matrix of 1's and Id_n denotes the $n\times n$ identity matrix.

When $\theta \le 0$, the field $(\log|\zeta(\frac{1}{2} + i\tau + ih)|,\ |h| \le \log^\theta T)$ behaves approximately like a BRW with $\log^{1+\theta} T$ Gaussian r.v.s of variance $\frac{1+\theta}{2}\log\log T$, attached to each branch. When $\theta > 0$, the field behaves approximately like a collection of $\log^\theta T$ nearly independent Gaussian BRWs where each r.v. has variance $\frac{1}{2}\log\log T$, or equivalently, as an appropriately scaled 2-levels IBRW where $\sigma_1 = 0$. The following conjecture should hold.

Conjecture 10.1.3 (Second order of the maximum for the Riemann zeta function)**.** *Using the notation from Article 6 and the definition of $\beta_c(\theta)$ from (10.1.3),*

$$
\lim_{T\to\infty} \frac{\max_{|h|\le\log^\theta T} \log|\zeta(\frac{1}{2} + i\tau + ih)| - m(\theta)\log\log T}{\log\log\log T} = \begin{cases} -\frac{3}{2\beta_c(\theta)}, & \text{if } \theta \le 0, \\ -\frac{1}{2\beta_c(\theta)}, & \text{if } \theta > 0, \end{cases}
$$

where $\tau \sim$ Uniform$(T, 2T)$ under $\mathbb{P}$ and the convergence holds in $\mathbb{P}$-probability.

The following is analogous to the tightness of the recentered maximum that Chhaibi et al. (2018) proved in the context of log-characteristic polynomials of the CUE field.

Conjecture 10.1.4 (Tightness of the recentered maximum for the randomized Riemann zeta function)**.** *Using the notation from Article 5, for any $\varepsilon > 0$, there exists $K = K(\varepsilon) > 0$ large enough that*

$$
\mathbb{P}\left(\left|\max_{h\in[0,1]} X(h) - (\log\log T - \frac{3}{4}\log\log T)\right| > K\right) < \varepsilon, \tag{10.1.7}
$$

for all $T \ge 2$.

10.2. List of open problems

In this section, I give a partial list of (reasonable) open problems of interest for which I do not have a precise mathematical statement.

(1) Can we find a good model for the large values of $t \mapsto \frac{d}{dt} \log |\zeta(\frac{1}{2} + it)|$ in the sense of Proposition 1 in Harper (2013) ? If so, can the results from Arguin et al. (2017b), Arguin and Tai (2018) and Articles 4 - 6 be extended to this random model ? What about the logarithmic derivatives of higher order ?

(2) Are the large values of $\log |\zeta(\frac{1}{2} + i\tau)|$ and $\frac{d}{dt} \log |\zeta(\frac{1}{2} + it)|\Big|_{t=\tau}$ approximately independent in some sense under $\mathbb{P}$? Can we split the joint characteristic function ?

(3) Find the asymptotic properties (bias, variance, mean squared error, integrated mean squared error, asymptotic normality) of the Bernstein estimator of the density function and the c.d.f. (respectively) on the d-dimensional simplex.

Hint : Partial answers can be found in Tenbusch (1994) for the two-dimensional simplex. Also, reading Belalia (2016) may be useful since the Bernstein estimator of the c.d.f. on the general hypercube is treated. An idea that might be fruitful is to generalize the continuity correction result of Cressie (1978) to multinomial distributions (this alone is worth investigating) and generalize some of the calculations in the appendix of Leblanc (2012) to the d-dimensional simplex.

(4) If we look at the product of Gibbs measures sampling at different temperatures as in Kurkova (2003) and Pain and Zindy (2018), how does the mean overlap of the GREM compares with the mean overlap of the variable speed BBM of Bovier and Hartung (2014, 2015, 2019) or the mean overlap of the IGFF from Articles 1 - 2 ? Can we find an explicit expression for the mean overlap of the REM or the GREM ?

(5) What can we say about the complex moments of the Riemann zeta function ? Can we prove the analogue of Theorem 1.2 in Hartung and Klimovsky (2018) ?

(6) Can the results of Arguin et al. (2019c) (Article 6) be used in any way to make progress towards Karatsuba's conjectures listed in Feng (2004) ?

10.3. Errata for the published articles

- Ouimet, F. (2017). Geometry of the Gibbs measure for the discrete 2D Gaussian free field with scale-dependent variance. ALEA, Lat. Am. J. Probab. Math. Stat. 14, no. 2, 851-902.

  The corrections are (they all stem from the first point) :

  — On the second line of **(B.35)** and the second line of **(B.55)**, replace

  $$\beta(\bar{\sigma}_{u,j^\star} - \tfrac{\beta}{4}\bar{\sigma}_{j^\star}^2) \quad \text{by} \quad \beta\bar{\sigma}_{u,j^\star} - \left(1 + \tfrac{\beta^2\bar{\sigma}_{j^\star}^2}{4}\right).$$

  — The previous correction implies that on the second and third line of **(B.35)**, the second line of **(B.47)**, the third and fourth line of **(B.55)**, and the second line of **(B.64)**, we should replace $\beta\bar{\sigma}_{j^\star}(1 - \frac{\beta}{4}\bar{\sigma}_{j^\star})$ by $-(1 - \frac{\beta}{2}\bar{\sigma}_{j^\star})^2$.

  — The previous correction then implies that the left and right derivatives coincide in **(B.38), (B.49), (B.58)** and **(B.66)**, since the right derivatives are equal to

  $$\tfrac{\beta\sigma_{i^\star}(\alpha' - \alpha)}{\bar{\sigma}_{j^\star}} \quad \text{instead of} \quad +\infty.$$

  — **Lemma B.4 :** As a consequence of the corrections above, the left and right derivatives coincide for all $\beta > 0$ and thus Lemma B.4 is valid for all $\beta > 0$.

  — **Remark 6.3 :** As a consequence of Lemma B.4 being valid for all $\beta > 0$, all the results of the article are valid for all $\beta > 0$. Remark 6.3 can thus be removed.

- Ouimet, F. (2018). Poisson-Dirichlet statistics for the extremes of a randomized Riemann zeta function. Electron. Commun. Probab. 23, no. 46, 1-15.

  The corrections are :

  — **(3.1) :** $G_{\beta,T}^{\times\infty}$ instead of $G_{\beta,T}$;

  — **(3.2) :** $G_{\beta,T}^{\times 2}$ instead of $G_{\beta,T}$;

  — **Lemma 5.2 :** The condition should be
  $\max_{1\leq j\leq s}\left\{\|\partial_{z_j}^2 F\|_\infty \vee \|\partial_{\overline{z_j}}^2 F\|_\infty \vee \|\partial_{z_j}\partial_{\overline{z_j}} F\|_\infty \vee \|\partial_{\overline{z_j}}\partial_{z_j} F\|_\infty\right\} \leq M$;

  — **(5.20), (5.27) and (5.28) :** $\omega_p(h_j)$ instead of $\omega(h_j)$ in some places;

  — **(5.29) :** $(\boldsymbol{z}, \overline{\boldsymbol{z}})$ instead of $(z, \overline{z})$.

  — **p.13 line -8:** The condition $0 < \alpha < 1$ should be removed.

Appendix

11.1. Two useful lemmas

We present here two well-known lemmas that we use in the Introduction (Part 0) and in some of the articles in Part 1 and Part 2.

The first lemma bounds the probability that a Gaussian r.v. is larger than a fixed $t > 0$. The estimates are precise when t is large.

Lemma 11.1.1 (Gaussian tail estimates, see e.g. page 9 in Adler and Taylor (2007))**.** *Let $\xi \sim \mathcal{N}(0, \sigma^2)$ where $\sigma > 0$. Then, for all $t > 0$,*

$$\frac{\sigma}{t}\left(1 - \frac{\sigma^2}{t^2}\right)\varphi\left(\frac{t}{\sigma}\right) \leq \mathbb{P}(\xi \geq t) \leq \frac{\sigma}{t}\varphi\left(\frac{t}{\sigma}\right), \tag{11.1.1}$$

where φ denotes the standard Gaussian density function.

Proof. After making the change of variable $z = \frac{t}{\sigma}$, it suffices to prove that, for all $z > 0$,

$$\frac{1}{z}\left(1 - \frac{1}{z^2}\right)\varphi(z) \leq \Psi(z) \leq \frac{1}{z}\varphi(z), \tag{11.1.2}$$

where $Z \mathrel{\hat{=}} \frac{\xi}{\sigma} \sim \mathcal{N}(0,1)$ and $\Psi(z) \mathrel{\hat{=}} \mathbb{P}(Z \geq z)$. By integrating by parts,

$$\begin{aligned} 0 &\leq \int_z^\infty \frac{1}{y^2}\varphi(y)dy = \frac{1}{z}\varphi(z) + \int_z^\infty \frac{1}{y}\varphi'(y)dy \\ &= \frac{1}{z}\varphi(z) - \Psi(z), \end{aligned} \tag{11.1.3}$$

since $\frac{1}{y}\varphi'(y) = -\varphi(y)$. This shows the upper bound in (11.1.2). Similarly,

$$\begin{aligned} 0 &\leq \int_z^\infty \frac{3}{y^4}\varphi(y)dy = \frac{1}{z^3}\varphi(z) + \int_z^\infty \frac{1}{y^3}\varphi'(y)dy \\ &= \frac{1}{z^3}\varphi(z) - \frac{1}{z}\varphi(z) + \Psi(z), \end{aligned} \tag{11.1.4}$$

from (11.1.3) and the fact that $\frac{1}{y^3}\varphi'(y) = -\frac{1}{y^2}\varphi(y)$. This shows the lower bound in (11.1.2). This ends the proof. ☐

The second lemma is derived from the Cauchy-Schwarz inequality. Its purpose is to find lower bounds on the probability that a non-negative r.v. is larger than its expectation by a multiplicative factor θ between 0 and 1. The inequality is recurrent in the applications of *second-moment methods*, and also complements nicely Chebyshev's inequality.

Lemma 11.1.2 (Paley-Zygmund inequality, see Paley and Zygmund (1932))**.** *Let X be a non-negative r.v. that satisfies $\mathbb{P}(X > 0) > 0$ and $\mathbb{E}[X^2] < \infty$. Then, for all $\theta \in [0,1]$,*

$$\mathbb{P}(X \geq \theta\,\mathbb{E}[X]) \geq (1-\theta)^2 \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}. \tag{11.1.5}$$

Proof. Since $\mathbb{E}\left[X\mathbf{1}_{\{X<\theta\,\mathbb{E}[X]\}}\right] \leq \theta\,\mathbb{E}[X]$, then

$$(1-\theta)\mathbb{E}[X] \leq \mathbb{E}[X] - \mathbb{E}\left[X\mathbf{1}_{\{X<\theta\,\mathbb{E}[X]\}}\right] = \mathbb{E}\left[X\mathbf{1}_{\{X\geq\theta\,\mathbb{E}[X]\}}\right]. \tag{11.1.6}$$

By taking the square on each side of the inequality and by applying the Cauchy-Schwarz inequality on the right-hand side, we find

$$(1-\theta)^2(\mathbb{E}[X])^2 \leq \left(\mathbb{E}\left[X\mathbf{1}_{\{X\geq\theta\,\mathbb{E}[X]\}}\right]\right)^2 \leq \mathbb{E}\left[X^2\right]\mathbb{P}(X \geq \theta\,\mathbb{E}[X]). \tag{11.1.7}$$

This ends the proof. □

11.2. Codes for the simulations

In this section, the reader can find the codes that I wrote to simulate the log-correlated random fields presented in Section 0.3.3 of the Introduction.

11.2.1. Matlab

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% REM simulation                                                           %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

n = 6; N = 2^n;
M = vertcat(zeros(1,N),cumsum(normrnd(0,1,n,N)));
h = figure;
plot((0:n)',M);
saveas(h,'REM','jpg');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% GREM simulation                                                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

n1 = 3; N1 = 2^n1;
n = 6; N = 2^n;
sigma1 = 1; sigma2 = 10;
M1 = cumsum(normrnd(0,sigma1,n1,N));
```

```
M = vertcat(zeros(1,N),M1,M1+cumsum(normrnd(0,sigma2,n-n1,N)));
for k = 1:N1
    M(2:(n1+1),(1+((k-1)*N/N1)):(k*N/N1)) = ...
        repmat(M(2:(n1+1),k*N/N1),1,N/N1);
end
h = figure;
plot((0:n)',M);
saveas(h,'GREM','jpg');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% BRW simulation                                                           %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

n = 6; N = 2^n;
M = zeros(n+1,N);

for i=1:n
    Ni = 2^i;
    g = normrnd(0,1,1,Ni);
    for k=1:Ni
        M(i+1,(1+(k-1)*N/Ni):(k*N/Ni)) = ...
            M(i,(1+(k-1)*N/Ni):(k*N/Ni)) + repmat(g(k),1,N/Ni);
    end
end
h = figure;
plot((0:n)',M);
saveas(h,'BRW','jpg');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% IBRW simulation                                                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

n = 6; N = 2^n;
M = zeros(n+1,N);
sigma1 = 1; sigma2 = 10;

for i=1:n
    sigma = sigma1*(i <= n/2) + sigma2*(i > n/2);
    Ni = 2^i;
    g = normrnd(0,sigma,1,Ni);
    for k=1:Ni
        M(i+1,(1+(k-1)*N/Ni):(k*N/Ni)) = ...
            M(i,(1+(k-1)*N/Ni):(k*N/Ni)) + repmat(g(k),1,N/Ni);
    end
end
```

```
h = figure;
plot((0:n)',M);
saveas(h,'IBRW','jpg');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% BBM simulation (conditioned on having 64 leaves)                        %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

verif = 0;

while(verif == 0)
    T = 6;
    n = T*1000;
    pas = T/n;
    num = 1;
    time = 0;
    B = zeros(n+1,64);
    count = 1;

    while(count < n + 1)
        E = exprnd(1/(log(2)*num),1,1);
        time_old = time;
        while((time < time_old + E) && (count < n + 1))
            for j = 1:num
                B(count + 1,j) = B(count,j) + normrnd(0,sqrt(pas),1,1);
            end
            count = count + 1;
            time = time + pas;
        end
        r = randi([1,num],1,1);
        num = num + 1;
        if(num > 64)
            break;
        end
        B(:,num) = B(:,r);
    end
    if(num == 64)
       verif = 1;
    end
end

h = figure;
plot((0:pas:T)',B(:,1:num));
saveas(h,'BBM','jpg');
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% VSBBM simulation (conditioned on having 64 leaves)                      %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

verif = 0;

while(verif == 0)
    T = 6;
    n = T*1000;
    pas = T/n;
    num = 1;
    time = 0;
    B = zeros(n+1,64);
    count = 1;
    sigma1 = 1; sigma2 = 10;

    while(count < n + 1)
        E = exprnd(1/(log(2)*num),1,1);
        time_old = time;
        while((time < time_old + E) && (count < n + 1))
            sigma = sigma1*(count <= n/2) + sigma2*(count > n/2);
            for j = 1:num
                B(count + 1,j) = ...
                    B(count,j) + normrnd(0,sigma*sqrt(pas),1,1);
            end
            count = count + 1;
            time = time + pas;
        end
        r = randi([1,num],1,1);
        num = num + 1;
        if(num > 64)
            break;
        end
        B(:,num) = B(:,r);
    end
    if(num == 64)
        verif = 1;
    end
end

h = figure;
plot((0:pas:T)',B(:,1:num));
saveas(h,'VSBBM','jpg');
```

11.2.2. Mathematica

```
(*GFF and IGFF simulation*)

sim = 1000; n = 32;
sigma1 = 40; sigma2 = 1;
F = Re[Fourier[Table[(InverseErf[2 Random[] - 1]
        + I InverseErf[2 Random[] - 1])*
      If[j + k == 2, 0, 1/Sqrt[(Sin[(j - 1)*Pi/n]^2
           + Sin[(k - 1)*Pi/n]^2)]], {j, n}, {k, n}]]];
G = ConstantArray[0, {n, n}];
For[k = 1, k <= n, k++,
  For[l = 1, l <= n, l++,
    val = ConstantArray[0, sim];
    For[m = 1, m <= sim, m++,
     pos = {k, l};
     While[(pos[[1]] != 1) && (pos[[1]] != n)
        && (pos[[2]] != 1) && (pos[[2]] != n),
      pos = pos + Flatten[
                   RandomSample[{{1,0}, {-1,0}, {0,1}, {0,-1}}, 1]];
      ];
     val[[m]] = F[[pos[[1]], pos[[2]]]];
     ];
    G[[k, l]] = Mean[val];
    ];
  ];
GFF = F - G; ListPlot3D[GFF]

dist = 4;
GFF1 = ConstantArray[0, {n, n}];
For[k = 1, k <= n, k++,
  For[l = 1, l <= n, l++,
    val = ConstantArray[0, sim];
    For[m = 1, m <= sim, m++,
     pos = {k, l};
     While[(pos[[1]] != Max[1,k-dist]) && (pos[[1]] != Min[n,k+dist])
        && (pos[[2]] != Max[1,l-dist]) && (pos[[2]] != Min[n,l+dist]),
      pos = pos + Flatten[
                   RandomSample[{{1,0}, {-1,0}, {0,1}, {0,-1}}, 1]];
      ];
     val[[m]] = GFF[[pos[[1]], pos[[2]]]];
     ];
    GFF1[[k, l]] = Mean[val];
    ];
  ];
IGFF = sigma2*(GFF - GFF1) + sigma1*GFF1; ListPlot3D[IGFF]
```

```
(*MM simulation*)
n = 128;
F = Re[Fourier[Table[(InverseErf[2 Random[] - 1]
        + I InverseErf[2 Random[] - 1])*
            If[j + k == 2, 0, 1/(Sin[(j - 1)*Pi/n]^2
                + Sin[(k - 1)*Pi/n]^2)], {j, n}, {k, n}]]];
ListPlot3D[F]

(*RLM-RZF simulation*)
PT = 100000;
list = Prime[Range[PT]];
unif = RandomReal[{0, 2*Pi}, PT];
Plot[Sum[Re[E^(I*(unif[[k]] - h*Log[list[[k]]]))/Sqrt[list[[k]]]],
    {k,1,PT}], {h,0,2*Pi}, PlotRange -> {-6,6}, AxesOrigin -> {0,-6}]

(*LM-RZF simulation*)
T = 100000;
Plot[Re[Log[Zeta[1/2 + I*(RandomReal[{T,2T}] + h)]]], {h,0,2*Pi},
    PlotRange -> {-6,6}, AxesOrigin -> {0,-6}]

(*LCP-CUE simulation*)
n = 2^10;
CUE = RandomVariate[CircularUnitaryMatrixDistribution[n]];
II = IdentityMatrix[n];
Plot[Re[Log[Det[E^(I*h)*II - CUE]]], {h,0,2*Pi},
    PlotRange -> {-6,6}, AxesOrigin -> {0,-6}]
```

11.3. Declarations and permissions

This section contains, in order, the following documents :

(1) Declarations of the coauthors :

  (a) Louis-Pierre Arguin for Articles 1, 5 and 6,

  (b) Maksym Radziwiłł for Article 6,

  (c) Pierre Lafaye de Micheaux for Article 8.

(2) Permissions from the editors :

  (a) Victor Perez-Abreu (Editor in Chief, ALEA) for Articles 1 and 2,

  (b) Luis Renato Fontes (Theory and Methods Editor, BJPS) for Article 3,

  (c) Giambattista Giacomin (Editor in Chief, ECP) for Article 4,

  (d) Steven G. Krantz (Editor in Chief, JMAA) for Articles 5 and 7,

  (e) Yimin Xiao (Editor in Chief, SPL) for Article 8.

(3) Permission from Anatoliĭ T. Fomenko to include his drawing *From chaos to order*.

DECLARATION FROM A COAUTHOR FOR THE
INCLUSION OF ARTICLES IN A DOCTORAL THESIS

1. **Identification of the coauthor (complete name and location)**

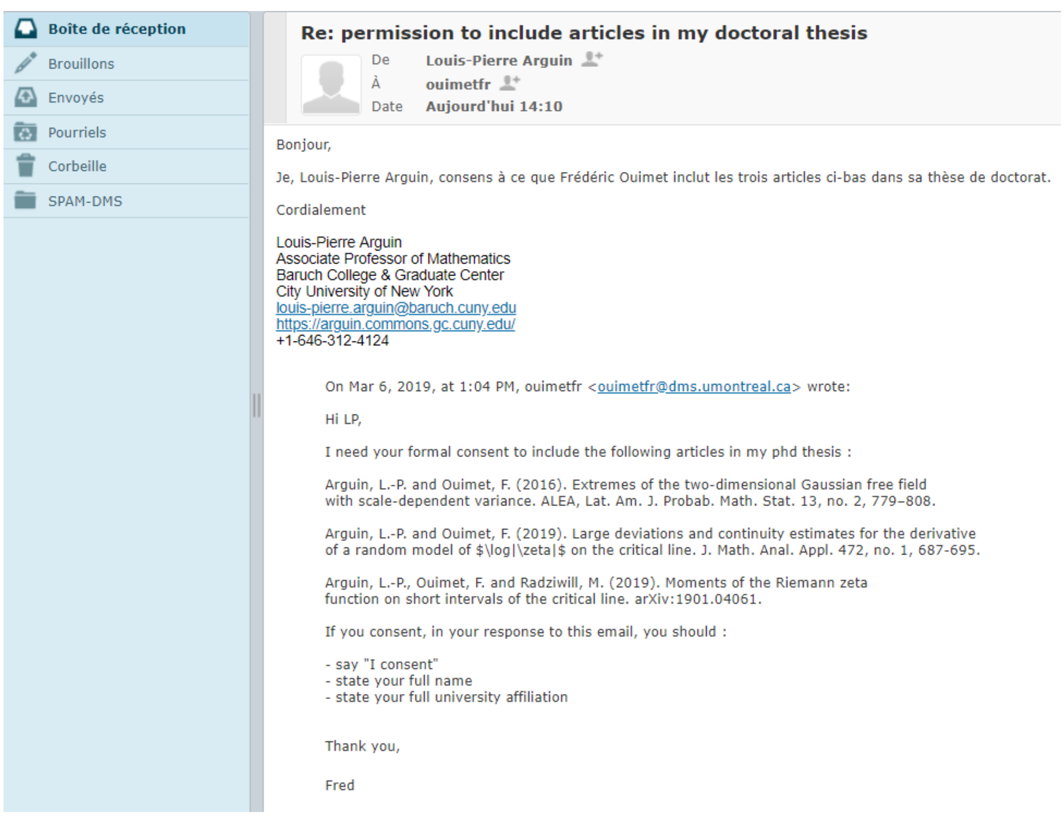
   Louis-Pierre Arguin (Associate Professor of Mathematics)
   Baruch College and Graduate Center (CUNY)
   55 Lexington Ave
   New York, NY 10010
   USA

2. **Identification of the articles**

   - Arguin, L.-P. and Ouimet, F. (2016). Extremes of the two-dimensional Gaussian free field with scale-dependent variance. ALEA, Lat. Am. J. Probab. Math. Stat. 13, no. 2, 779–808.
   - Arguin, L.-P. and Ouimet, F. (2019). Large deviations and continuity estimates for the derivative of a random model of $\log|\zeta|$ on the critical line. J. Math. Anal. Appl. 472, no. 1, 687-695.
   - Arguin, L.-P., Ouimet, F. and Radziwill, M. (2019). Moments of the Riemann zeta function on short intervals of the critical line. arXiv:1901.04061.

3. **Declaration**

   As a coauthor of the above articles, I authorize Frédéric Ouimet to include them in his doctoral thesis.



Boîte de réception
Brouillons
Envoyés
Pourriels
Corbeille
SPAM-DMS

Re: permission to include articles in my doctoral thesis

De Louis-Pierre Arguin
À ouimetfr
Date Aujourd'hui 14:10

Bonjour,

Je, Louis-Pierre Arguin, consens à ce que Frédéric Ouimet inclut les trois articles ci-bas dans sa thèse de doctorat.

Cordialement

Louis-Pierre Arguin
Associate Professor of Mathematics
Baruch College & Graduate Center
City University of New York
louis-pierre.arguin@baruch.cuny.edu
https://arguin.commons.gc.cuny.edu/
+1-646-312-4124

On Mar 6, 2019, at 1:04 PM, ouimetfr <ouimetfr@dms.umontreal.ca> wrote:

Hi LP,

I need your formal consent to include the following articles in my phd thesis :

Arguin, L.-P. and Ouimet, F. (2016). Extremes of the two-dimensional Gaussian free field
with scale-dependent variance. ALEA, Lat. Am. J. Probab. Math. Stat. 13, no. 2, 779–808.

Arguin, L.-P. and Ouimet, F. (2019). Large deviations and continuity estimates for the derivative
of a random model of $\log|\zeta|$ on the critical line. J. Math. Anal. Appl. 472, no. 1, 687-695.

Arguin, L.-P., Ouimet, F. and Radziwill, M. (2019). Moments of the Riemann zeta
function on short intervals of the critical line. arXiv:1901.04061.

If you consent, in your response to this email, you should :

- say "I consent"
- state your full name
- state your full university affiliation

Thank you,

Fred

1. **Identification of the coauthor (complete name and location)**

   Maksym Radziwill (Professor of Mathematics)
   Caltech, 358 Linde Hall (Building #37)
   1200 E California Blvd
   Pasadena, CA, 91125
   USA

2. **Identification of the article**

   - Arguin, L.-P., Ouimet, F. and Radziwill, M. (2019). Moments of the Riemann zeta function on short intervals of the critical line. arXiv:1901.04061.

3. **Declaration**

   As a coauthor of the above article, I authorize Frédéric Ouimet to include it in his doctoral thesis.

Boîte de réception
Brouillons
Envoyés
Pourriels
Corbeille
SPAM-DMS

Re: permission to include an article in my doctoral thesis

De Radziwill Maksym
À ouimetfr
Date Aujourd'hui 14:35

Hi Fred,

Sure

I consent
Maksym Radziwill
California Institute of Technology

Best,
Maks

On Wed, Mar 6, 2019 at 10:01 AM ouimetfr <ouimetfr@dms.umontreal.ca> wrote:

> Hi Maks,
>
> I need your formal consent to include the following article in my phd thesis :
>
> Arguin, L.-P., Ouimet, F. and Radziwill, M. (2019). Moments of the Riemann zeta
> function on short intervals of the critical line. arXiv:1901.04061.
>
> If you consent, in your response to this email, you should :
>
> - say "I consent"
> - state your full name
> - state your full university affiliation
>
> Thank you,
>
> Fred

DECLARATION FROM A COAUTHOR FOR THE
INCLUSION OF AN ARTICLE IN A DOCTORAL THESIS

1. **Identification of the coauthor (complete name and location)**

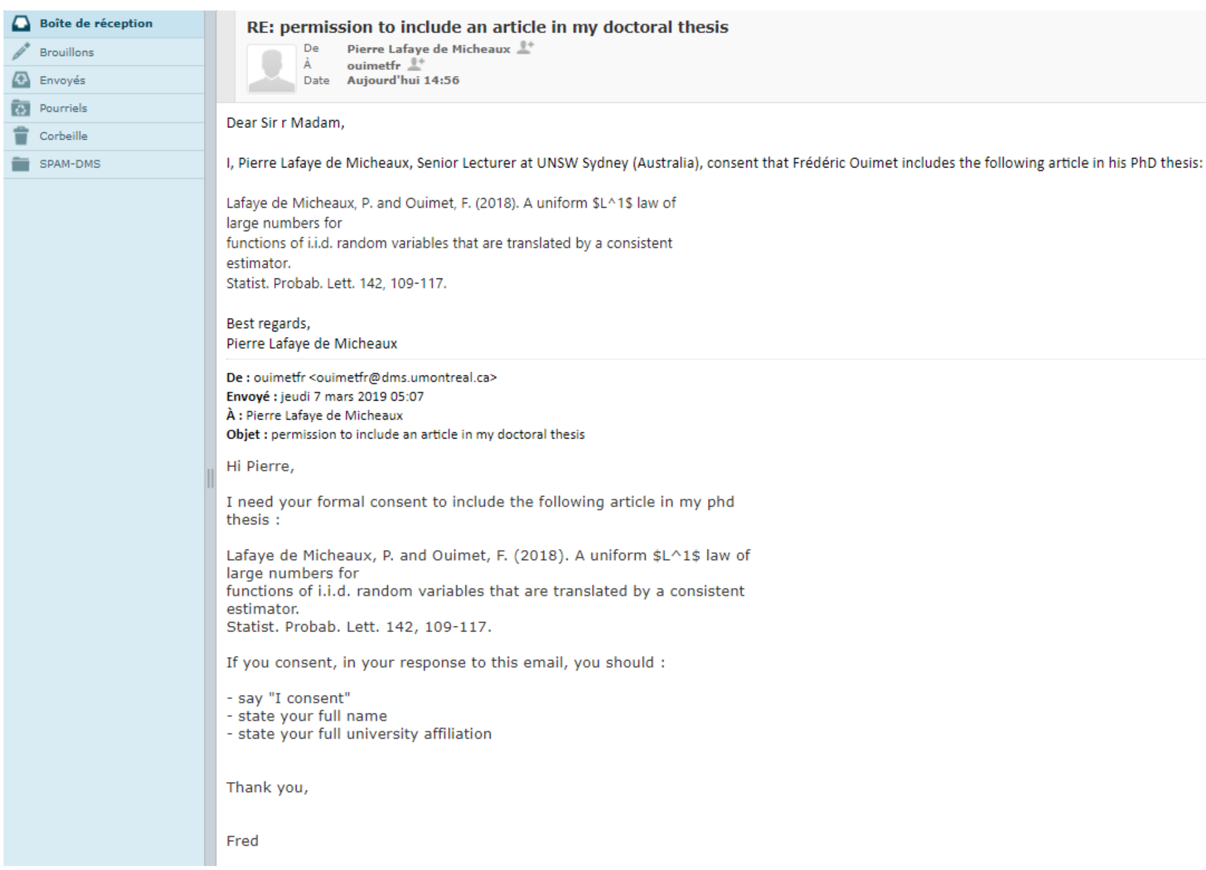
   Pierre Lafaye de Micheaux (Senior Lecturer in Statistics)
   School of Mathematics and Statistics, UNSW Sydney
   Office 2050, The Red Centre, Centre Wing, Kensington
   Sydney, NSW 2052
   Australia

2. **Identification of the article**

   - Lafaye de Micheaux, P. and Ouimet, F. (2018). A uniform L^1 law of large numbers for functions of i.i.d. random variables that are translated by a consistent estimator. Statist. Probab. Lett. 142, 109-117.

3. **Declaration**

   As a coauthor of the above article, I authorize Frédéric Ouimet to include it in his doctoral thesis.



Boîte de réception
Brouillons
Envoyés
Pourriels
Corbeille
SPAM-DMS

RE: permission to include an article in my doctoral thesis
De Pierre Lafaye de Micheaux
À ouimetfr
Date Aujourd'hui 14:56

Dear Sir r Madam,

I, Pierre Lafaye de Micheaux, Senior Lecturer at UNSW Sydney (Australia), consent that Frédéric Ouimet includes the following article in his PhD thesis:

Lafaye de Micheaux, P. and Ouimet, F. (2018). A uniform L^1 law of
large numbers for
functions of i.i.d. random variables that are translated by a consistent
estimator.
Statist. Probab. Lett. 142, 109-117.

Best regards,
Pierre Lafaye de Micheaux

De : ouimetfr <ouimetfr@dms.umontreal.ca>
Envoyé : jeudi 7 mars 2019 05:07
À : Pierre Lafaye de Micheaux
Objet : permission to include an article in my doctoral thesis

Hi Pierre,

I need your formal consent to include the following article in my phd
thesis :

Lafaye de Micheaux, P. and Ouimet, F. (2018). A uniform L^1 law of
large numbers for
functions of i.i.d. random variables that are translated by a consistent
estimator.
Statist. Probab. Lett. 142, 109-117.

If you consent, in your response to this email, you should :

- say "I consent"
- state your full name
- state your full university affiliation

Thank you,

Fred

PERMISSION FROM THE EDITOR OF A JOURNAL
TO INCLUDE ARTICLES IN A DOCTORAL THESIS

1. **Identification of the journal**

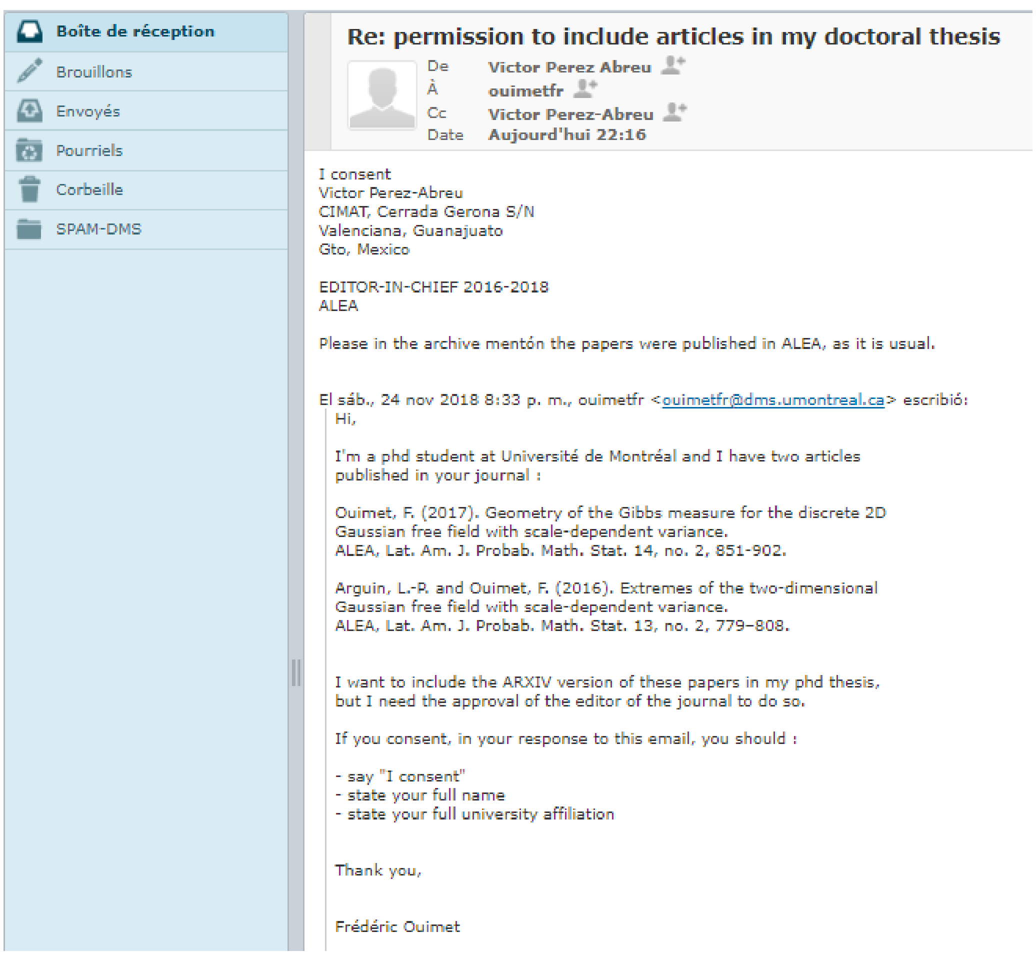
   ALEA, Latin American Journal of Probability and Mathematical Statistics

2. **Identification of the editor (complete name and location)**

   Victor Perez-Abreu (Editor in Chief, ALEA)
   CIMAT, Cerrada Gerona S/N
   Valenciana, Guanajuato
   Gto, Mexico

3. **Identification of the articles**

   - Arguin, L.-P. and Ouimet, F. (2016). Extremes of the two-dimensional Gaussian free field with scale-dependent variance. ALEA, Lat. Am. J. Probab. Math. Stat. 13, no. 2, 779–808.
   - Ouimet, F. (2017). Geometry of the Gibbs measure for the discrete 2D Gaussian free field with scale-dependent variance. ALEA, Lat. Am. J. Probab. Math. Stat. 14, no. 2, 851-902.



Boîte de réception
Brouillons
Envoyés
Pourriels
Corbeille
SPAM-DMS

Re: permission to include articles in my doctoral thesis

De Victor Perez Abreu
À ouimetfr
Cc Victor Perez-Abreu
Date Aujourd'hui 22:16

I consent
Victor Perez-Abreu
CIMAT, Cerrada Gerona S/N
Valenciana, Guanajuato
Gto, Mexico

EDITOR-IN-CHIEF 2016-2018
ALEA

Please in the archive mentón the papers were published in ALEA, as it is usual.

El sáb., 24 nov 2018 8:33 p. m., ouimetfr <ouimetfr@dms.umontreal.ca> escribió:

> Hi,
>
> I'm a phd student at Université de Montréal and I have two articles published in your journal :
>
> Ouimet, F. (2017). Geometry of the Gibbs measure for the discrete 2D Gaussian free field with scale-dependent variance.
> ALEA, Lat. Am. J. Probab. Math. Stat. 14, no. 2, 851-902.
>
> Arguin, L.-P. and Ouimet, F. (2016). Extremes of the two-dimensional Gaussian free field with scale-dependent variance.
> ALEA, Lat. Am. J. Probab. Math. Stat. 13, no. 2, 779–808.
>
> I want to include the ARXIV version of these papers in my phd thesis, but I need the approval of the editor of the journal to do so.
>
> If you consent, in your response to this email, you should :
>
> - say "I consent"
> - state your full name
> - state your full university affiliation
>
> Thank you,
>
> Frédéric Ouimet

PERMISSION FROM THE EDITOR OF A JOURNAL
TO INCLUDE AN ARTICLE IN A DOCTORAL THESIS

1. **Identification of the journal**

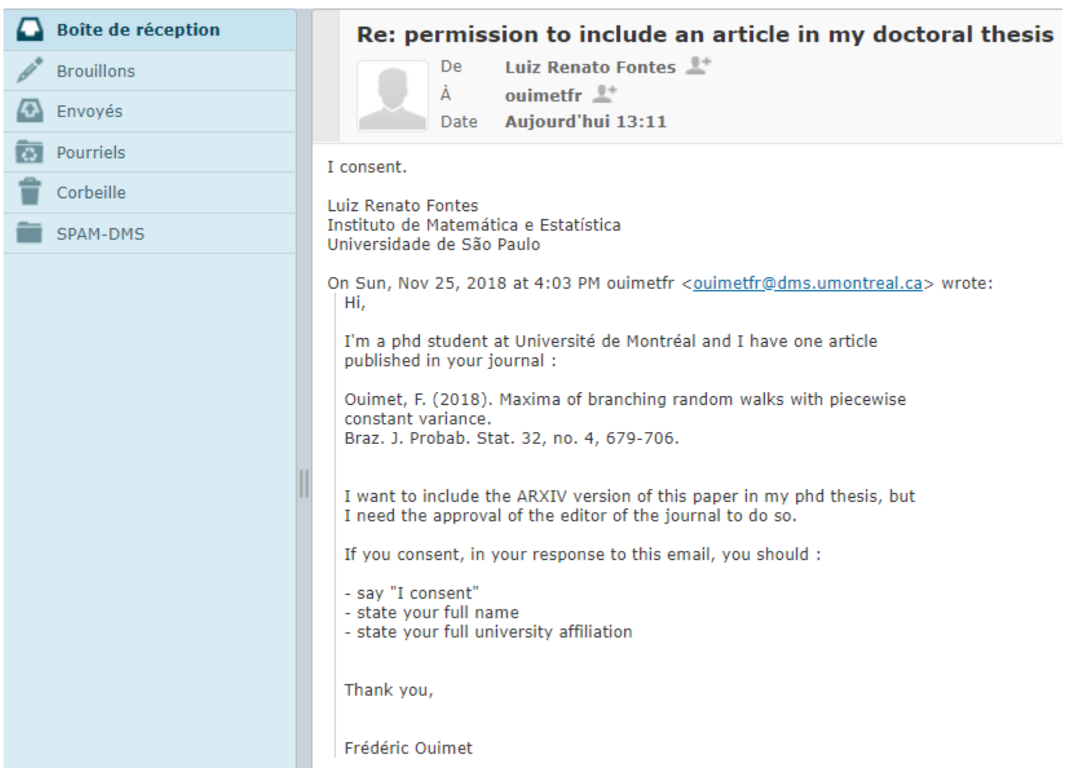
   Brazilian Journal of Probability and Statistics

2. **Identification of the editor (complete name and location)**

   Luiz Renato Fontes (Theory and Methods Editor, BJPS)
   Instituto de Matemática e Estatística
   Universidade de São Paulo
   Brazil

3. **Identification of the article**

   - Ouimet, F. (2018). Maxima of branching random walks with piecewise constant variance. Braz. J. Probab. Stat. 32, no. 4, 679-706.



Boîte de réception
Brouillons
Envoyés
Pourriels
Corbeille
SPAM-DMS

Re: permission to include an article in my doctoral thesis

De **Luiz Renato Fontes**
À **ouimetfr**
Date **Aujourd'hui 13:11**

I consent.

Luiz Renato Fontes
Instituto de Matemática e Estatística
Universidade de São Paulo

On Sun, Nov 25, 2018 at 4:03 PM ouimetfr <ouimetfr@dms.umontreal.ca> wrote:

Hi,

I'm a phd student at Université de Montréal and I have one article published in your journal :

Ouimet, F. (2018). Maxima of branching random walks with piecewise constant variance.
Braz. J. Probab. Stat. 32, no. 4, 679-706.

I want to include the ARXIV version of this paper in my phd thesis, but I need the approval of the editor of the journal to do so.

If you consent, in your response to this email, you should :

- say "I consent"
- state your full name
- state your full university affiliation

Thank you,

Frédéric Ouimet

PERMISSION FROM THE EDITOR OF A JOURNAL
TO INCLUDE AN ARTICLE IN A DOCTORAL THESIS

1. **Identification of the journal**

   Electronic Communications in Probability

2. **Identification of the editor (complete name and location)**

   Giambattista Giacomin (Editor in Chief, ECP)
   Université Paris Diderot
   UFR de Mathématiques, case 7012
   Bâtiment Sophie Germain
   75205 Paris Cedex 13
   France

3. **Identification of the article**

   - Ouimet, F. (2018). Poisson-Dirichlet statistics for the extremes of a randomized Riemann zeta function. Electron. Commun. Probab. 23, no. 46, 1-15.

Boîte de réception
Brouillons
Envoyés
Pourriels
Corbeille
SPAM-DMS

Re: permission to include an article in my doctoral thesis

De **Giambattista Giacomin**
À **ouimetfr**
Date **Aujourd'hui 04:38**

Le 25 nov. 2018 à 03:40, ouimetfr <ouimetfr@dms.umontreal.ca> a écrit :

Hi,

I'm a phd student at Université de Montréal and I have one article published in your journal :

Ouimet, F. (2018). Poisson-Dirichlet statistics for the extremes of a randomized Riemann zeta function.
Electron. Commun. Probab. 23, no. 46, 1-15.

I want to include the ARXIV version of this paper in my phd thesis, but I need the approval of the editor of the journal to do so.

If you consent, in your response to this email, you should :

- say "I consent"
- state your full name
- state your full university affiliation

Thank you,

Frédéric Ouimet

I consent

Giambattista Giacomin
Postal Address: Université Paris Diderot
UFR de Mathématiques, case 7012
Bâtiment Sophie Germain
75205 Paris Cedex 13
Office: 545 (entrance: SE end of Avenue de France)
tel: (33) - 01 57 27 93 17
URL: https://www.lpsm.paris/semoa/users/giacomin/index

PERMISSION FROM THE EDITOR OF A JOURNAL
TO INCLUDE ARTICLES IN A DOCTORAL THESIS

1. **Identification of the journal**

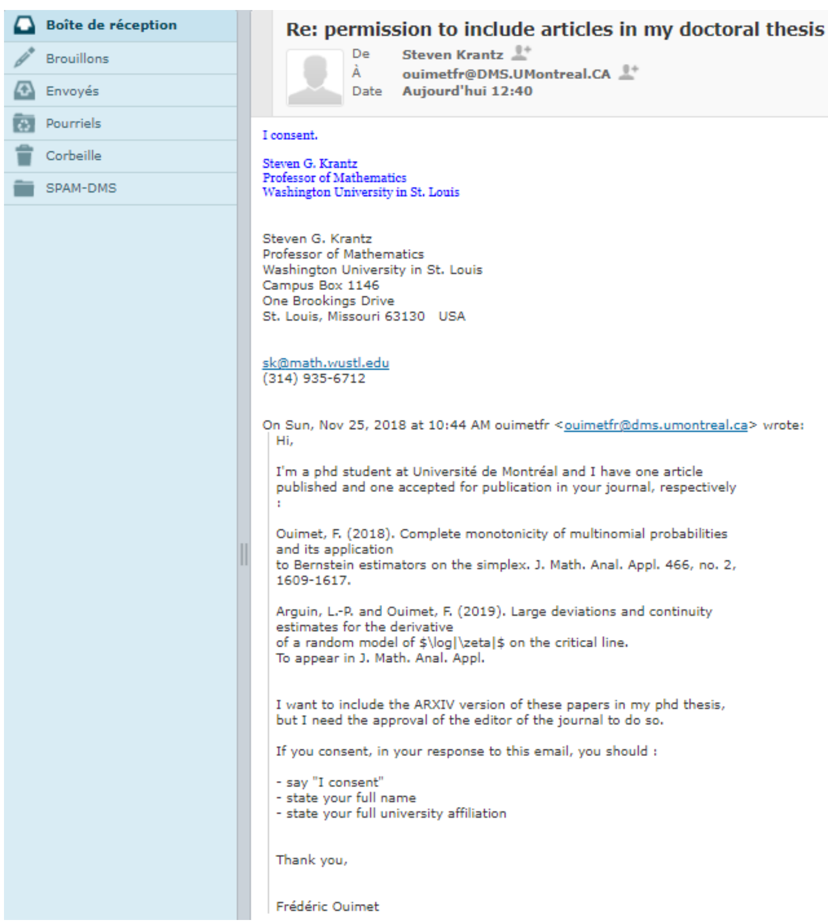
   Journal of Mathematical Analysis and Applications

2. **Identification of the editor (complete name and location)**

   Steven G. Krantz (Editor in Chief, JMAA)
   Washington University in St. Louis
   Campus Box 1146, One Brookings Drive
   St. Louis, Missouri 63130
   USA

3. **Identification of the articles**

   - Ouimet, F. (2018). Complete monotonicity of multinomial probabilities and its application to Bernstein estimators on the simplex. J. Math. Anal. Appl. 466, no. 2, 1609-1617.
   - Arguin, L.-P. and Ouimet, F. (2019). Large deviations and continuity estimates for the derivative of a random model of $\log|\zeta|$ on the critical line. J. Math. Anal. Appl. 472, no. 1, 687-695.



Boîte de réception
Brouillons
Envoyés
Pourriels
Corbeille
SPAM-DMS

Re: permission to include articles in my doctoral thesis

De Steven Krantz
À ouimetfr@DMS.UMontreal.CA
Date Aujourd'hui 12:40

I consent.

Steven G. Krantz
Professor of Mathematics
Washington University in St. Louis

Steven G. Krantz
Professor of Mathematics
Washington University in St. Louis
Campus Box 1146
One Brookings Drive
St. Louis, Missouri 63130 USA

sk@math.wustl.edu
(314) 935-6712

On Sun, Nov 25, 2018 at 10:44 AM ouimetfr <ouimetfr@dms.umontreal.ca> wrote:

Hi,

I'm a phd student at Université de Montréal and I have one article published and one accepted for publication in your journal, respectively :

Ouimet, F. (2018). Complete monotonicity of multinomial probabilities and its application to Bernstein estimators on the simplex. J. Math. Anal. Appl. 466, no. 2, 1609-1617.

Arguin, L.-P. and Ouimet, F. (2019). Large deviations and continuity estimates for the derivative of a random model of $\log|\zeta|$ on the critical line. To appear in J. Math. Anal. Appl.

I want to include the ARXIV version of these papers in my phd thesis, but I need the approval of the editor of the journal to do so.

If you consent, in your response to this email, you should :

- say "I consent"
- state your full name
- state your full university affiliation

Thank you,

Frédéric Ouimet

PERMISSION FROM THE EDITOR OF A JOURNAL
TO INCLUDE AN ARTICLE IN A DOCTORAL THESIS

1. **Identification of the journal**

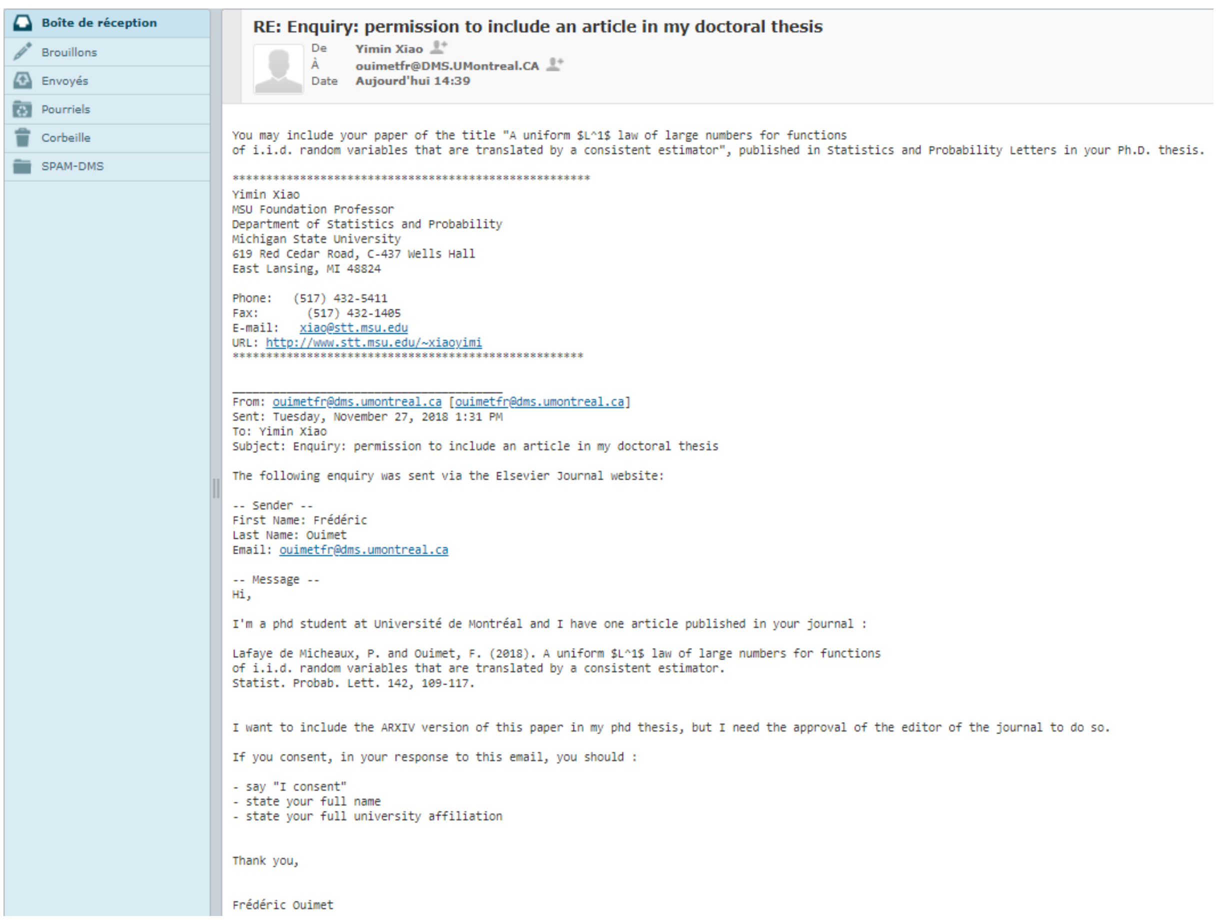
   Statistics and Probability Letters

2. **Identification of the editor (complete name and location)**

   Yimin Xiao (Editor in Chief, SPL)
   Michigan State University
   619 Red Cedar Road, C-437 Wells Hall
   East Lansing, MI 48824

3. **Identification of the article**

   - Lafaye de Micheaux, P. and Ouimet, F. (2018). A uniform L^1 law of large numbers for functions of i.i.d. random variables that are translated by a consistent estimator. Statist. Probab. Lett. 142, 109-117.



Boîte de réception
Brouillons
Envoyés
Pourriels
Corbeille
SPAM-DMS

RE: Enquiry: permission to include an article in my doctoral thesis

De Yimin Xiao
À ouimetfr@DMS.UMontreal.CA
Date Aujourd'hui 14:39

You may include your paper of the title "A uniform L^1 law of large numbers for functions
of i.i.d. random variables that are translated by a consistent estimator", published in Statistics and Probability Letters in your Ph.D. thesis.

****
Yimin Xiao
MSU Foundation Professor
Department of Statistics and Probability
Michigan State University
619 Red Cedar Road, C-437 Wells Hall
East Lansing, MI 48824

Phone: (517) 432-5411
Fax: (517) 432-1405
E-mail: xiao@stt.msu.edu
URL: http://www.stt.msu.edu/~xiaoyimi
****

____
From: ouimetfr@dms.umontreal.ca [ouimetfr@dms.umontreal.ca]
Sent: Tuesday, November 27, 2018 1:31 PM
To: Yimin Xiao
Subject: Enquiry: permission to include an article in my doctoral thesis

The following enquiry was sent via the Elsevier Journal website:

-- Sender --
First Name: Frédéric
Last Name: Ouimet
Email: ouimetfr@dms.umontreal.ca

-- Message --
Hi,

I'm a phd student at Université de Montréal and I have one article published in your journal :

Lafaye de Micheaux, P. and Ouimet, F. (2018). A uniform L^1 law of large numbers for functions
of i.i.d. random variables that are translated by a consistent estimator.
Statist. Probab. Lett. 142, 109-117.

I want to include the ARXIV version of this paper in my phd thesis, but I need the approval of the editor of the journal to do so.

If you consent, in your response to this email, you should :

- say "I consent"
- state your full name
- state your full university affiliation

Thank you,

Frédéric Ouimet

1. **Identification of the artist (complete name and location)**

   Anatoli Timofeïevitch Fomenko (Professor of Mathematics)
   Lomonosov Moscow State University
   Faculty of Mechanics and Mathematics
   GSP-1, 1 Leninskive Gory, Main Building
   119991, Moscow, Russian Federation

2. **Identification of the drawing**

   Page 105 of

   - Fomenko, A. T. Mathematical Impressions. With writing assistance and a preface by Richard Lipkin. American Mathematical Society, Providence, RI, 1990. 184 pp. ISBN: 0-8218-0162-7.

3. **Declaration**

   As the creator of the above drawing, I authorize Frédéric Ouimet to include it in his doctoral thesis.

Boîte de réception
Brouillons
Envoyés
Pourriels
Corbeille
SPAM-DMS

Re: Special request for my PhD thesis

De atfomenko@mail.ru
À ouimetfr
Date Aujourd'hui 04:05
Priorité Normale

Dear Frédéric Ouimet,
I agree and give to you the permission to use my drawing in your work. I wish to you the success in the science.
Wirh best regards
A.T.Fomenko

> Суббота, 23 марта 2019, 3:34 +03:00 от ouimetfr <ouimetfr@dms.umontreal.ca>:
>
> Dear Prof. Fomenko,
>
> I'm a PhD student at University of Montreal in the field of Probability
> and I will defend my thesis very soon.
>
> The first submission of my thesis is due in about a week and I have a
> special request for you.
> I first found out about your drawings by reading Albert Shiryaev's
> graduate book on probability and immediately bought your AMS book titled
> "Mathematical Impressions" because I wanted to see more.
>
> I have a rather substantial thesis (I had to break it down into 2 pieces
> to upload it with this email !), so I thought that instead of including
> the usual math quote at the beginning, I would like (with your
> permission of course) to include your drawing named "From chaos to
> order".
>
> I have included the drawing in question on page xxvii of my thesis for
> you to see.
>
> If you don't agree about the inclusion, I will remove the drawing from
> the thesis without issue.
>
> Thank you for your consideration,
>
> Sincerely,
>
> https://sites.google.com/site/fouimet26/research