Two Random Multiplicative Processes: Multiplicative Cascades and Eigenvectors of the Random Schrödinger Operator

by

Ben Rifkind

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Mathematics
University of Toronto

Abstract

Two Random Multiplicative Processes: Multiplicative Cascades and Eigenvectors of the Random Schrödinger Operator

Ben Rifkind
Doctor of Philosophy
Graduate Department of Mathematics
University of Toronto
2014

In this thesis we focus on two distinct random processes constructed using a sequence of independent identically distributed random variables along with a product structure on the underlying space. The first, a collection of random variables attached to vertices of the rooted binary tree, uses the multiplicative structure of paths along the tree while the second, a collection of two by two random matrices, uses the usual multiplication of matrices.

In the first chapter we construct a continuous time version of the multiplicative cascade. The multiplicative cascade can be thought of as a randomization of an initial measure on the boundary of a tree, constructed from an independent, identically distributed collection of random variables attached to the tree vertices. The new random measure is constructed from the old by weighting the measure of any vertex v in the the tree by the product of the random variables along the path from the root to v. Given an initial measure with certain regularity properties, we construct a continuous time, measure-valued process whose value at each time is a cascade of the initial one. We do this by replacing the random variables on the vertices with independent increment processes satisfying certain moment assumptions. This process has a Markov property: at any given time it is a cascade of the process at any earlier time by random variables that are independent of the past. It is also a martingale and, under certain extra conditions, it is also continuous. For Gaussian independent increments processes we develop the infinite-dimensional stochastic calculus that describes the evolution of the measure process, and use it to compute the optimal Hölder exponent in the Wasserstein distance on measures.

In the second chapter, we characterize the scaling limit of a uniformly chosen eigenvector of the one-dimensional discrete random Schrödinger operator,

$$(H_n\psi)(\ell) = \psi(\ell-1) + \psi(\ell+1) + v_\ell\psi(\ell),$$

with $\psi(0) = \psi(n+1) = 0$ and $v_\ell = \omega_\ell/\sqrt{n}$. Here ω_ℓ are independent identically distributed random

variables with mean zero and variance one. Our analysis uses the well known transfer matrix formulation of the spectral problem for H_n; the spectral information is encapsulated in a process of products of independent, identically distributed two by two random matrices. The limiting diffusion for this product process was developed in [13] to study the local eigenvalue point process. We build on that framework to show that the envelope of a uniformly chosen eigenvector converges weakly to the process on $[0, 1]$,

$$\exp\left(\frac{\mathcal{B}_{t-u}}{\sqrt{2}} - \frac{|t-u|}{4}\right),$$

with u independent and uniform on $[0, 1]$ and $\mathcal{B}$ a two-sided Brownian motion started from zero.

Dedication

To Mom and Dad

Acknowledgements

There are many people I would like to thank for helping me to finish this thesis. First in my mind are my parents. Sometimes it is important to state the obvious: I have been extremely fortunate in life. And most of that luck can be traced to having you as my parents. Your love and support has given me an advantage that is impossible to quantify. And your values, especially your emphasis on education, are ingrained in me. The idea that knowledge itself is of fundamental importance is one of the main reasons I pursued mathematics. I learned that idea from you.

To my advisor, Bálint Virág, a very big thank you. Thank you for your kindness and patience in dealing with me throughout the years. You never made me feel silly for not understanding. On the contrary, you explained so many seemingly complicated mathematical concepts in simple and elegant terms. Your depth of understanding of a problem as well as your creativity and ability to distill difficult ideas inspired me. You are the kind of mathematician I wanted to be and attempted to emulate in my thesis work. I can definitely say that without you I would not have completed this thesis. It has been a pleasure working with you.

Thank you to Tom Alberts, the co-author of my first paper which is also the first chapter of this thesis. Working on math with you never really felt like work. It was mostly just a lot of fun hanging out with a good friend. Thank you for taking a young PhD student under your wing.

Thanks to the members of my thesis committee, Kostya Khanin and Jeremy Quastel, for the helpful suggestions and encouragement throughout the years. Thank you very much to the external appraiser, Wilhelm Schlag, for taking the time to review my thesis. And thank you to Dmitry Panchenko and Michael Goldstein for attending my defense and for your insightful comments and questions.

Thank you to Ida Bulat and Jemima Merisca who handled all my many bureaucratic issue and allowed me to hand in forms much later than was reasonable. You were both always so helpful and accommodating. Thank you especially to Jemima for so handily taking care of all the final graduation details.

Finally, to my fellow students in the department, thank you for your friendship. The day to day would have been a lot less fun and interesting without your company. I was very lucky to make a lot of great friends here. So lucky that there are too many to name individually but I hope you know who you are.

Contents

Chapter 1

Diffusions of Multiplicative Cascades

1.1 Introduction

Multiplicative cascades are a particular type of random measures with many interesting statistical properties. The space on which these measures live is not always the same, but there is typically a tree structure underlying their construction and so it is convenient to consider them as living on the boundary of an infinite tree. This is the situation we consider. This has the further advantage that several different models of statistical mechanics are fully described by this framework, most notably tree polymers, branching random walk, and certain models of random walk in random environment.

For simplicity we work on a rooted, infinite binary tree $\mathcal{T}$, and the boundary $\partial\mathcal{T}$ is the set of all infinite self-avoiding paths that begin at the root. Elements of $\partial\mathcal{T}$ are called rays and we denote them by ξ. The inputs to the cascade model are a positive measure Γ on $\partial\mathcal{T}$, which can be specified arbitrarily, and an i.i.d. collection of random variables $\{W(v)\}_{v\in\mathcal{T}}$ attached to the vertices of the tree. The only a priori assumption on the distribution of the W is that it is strictly positive and has mean one. These random variables are then cascaded on to Γ to produce a random measure on $\partial\mathcal{T}$; we denote it by Γ_W or sometimes

$$\Gamma_W = \mathcal{C}(\Gamma; W).$$

The cascading procedure is simple to describe: for each $n \geq 0$ one uses the random weights W up to generation n to construct a random measure via

$$d\Gamma_W^{(n)}(\xi) = \prod_{i=1}^{n} W(\xi_i)\, d\Gamma(\xi).$$

The random cascade measure is then defined as the limit

$$\Gamma_W := \lim_{n\to\infty} \Gamma_W^{(n)}. \tag{1.1.1}$$

A martingale argument shows that the limit exists almost surely for *any* choice of the initial measure Γ, in the topology of weak convergence on the space of measures. Full details are given in Section 2. As we will see there it *may* happen that Γ_W is the zero measure, but nonetheless it is well-defined, and given this the main problem is to determine the properties of Γ_W and how they depend on the input measure Γ and the cascading distribution W. Fundamental properties of cascade measures were derived in [10], and further explorations have been made in several later papers; see for example [2, 9, 15, 17, 7].

Even in the simplest cases the relationship between Γ_W and Γ is interesting. Observe that if $W = 1$ then $\Gamma_W = \Gamma$, but if the cascading distribution is not identically one then Γ_W is necessarily distinct from Γ. There are two possible alternatives:

- Γ_W may be identically the zero measure, even though Γ is not, but
- if Γ_W is not the zero measure then it is genuinely random, meaning it depends on the specific realization of the W variables, but almost surely it is singular with respect to Γ.

The positivity of Γ_W is determined by both the regularity of Γ (roughly meaning how strongly it concentrates on some rays more than others) and moment properties of the cascading distribution. Full details are given in Section 1.2. The singularity property, however, holds even if the cascading distribution is highly concentrated near one. It is a simple consequence of the fact that along any ray the density is the product of positive, iid, mean one random variables, which almost surely goes to zero as the number of terms in the product goes to infinity.

The main purpose of this work is to study what happens when the cascading distribution is highly concentrated near one and the cascading procedure is iterated. The scheme is simple: start with a positive measure Γ on $\partial\mathcal{T}$ and cascade once to produce Γ_W. Since the cascading procedure does not depend on the choice of the initial measure, we may use Γ_W as the input measure and cascade it with vertex variables $\{W^*(v)\}_{v\in\mathcal{T}}$ that are independent of the $\{W(v)\}$ collection. This iteration can be repeated indefinitely, at each time cascading with a collection of vertex variables that are independent of all previous ones, and in doing so it produces a discrete time, measure-valued Markov process.

This discrete time process is interesting in its own right, but we prefer instead to study a continuous time version. Intuitively the idea behind the continuous time process is clear: starting from some initial measure, in each infinitesimal unit of time we cascade the previous measure with an independent collection of random variables whose distribution is an infinitesimal perturbation away from the degenerate distribution at one. Repeating this scheme indefinitely builds the process.

As is usual, however, rigorously constructing the continuous time process takes more care than constructing the discrete time one, even though the basic idea is the same. Several different construction techniques could be considered; for example, the discrete time process is well-defined, and the continuous time process could be constructed by taking a weak limit as the discrete time step goes to zero and the cascading distribution concentrates near one. Alternatively, the process is essentially defined by saying that the measure at each time is a cascade of the process at an earlier time (by an independent collection of random variables); this is akin to specifying the transition probabilities of the process, and then the existence would follow from the general theory on measure-valued diffusions (see for example [6]).

In this work we propose a simpler and more direct construction procedure. Instead of appealing to the more abstract concepts above, we simply attach to the vertices of the tree a family of dynamic weights $\{t \mapsto W_t(v)\}_{v\in\mathcal{T}}$. Using the cascading procedure defined in equation (1.1.1), this gives us a process $t \mapsto \Gamma_t := \Gamma_{W_t}$ of random cascade measures. We choose the weight process $t \mapsto W_t$ so that the Γ_t process satisfies the following important Markov property: the value at any given time is a cascade of the value at any previous time, by a noise that is independent of the past of the process. More precisely, our process is defined on an interval $[0, T]$ for some $T > 0$, and has the property that for any $s, t \geq 0$

such that $t+s \leq T$, *both* of the relations

$$\Gamma_{t+s} = \mathcal{C}(\Gamma; W_{t+s}) \quad \text{and} \quad \Gamma_{t+s} = \mathcal{C}\left(\mathcal{C}\left(\Gamma; W_t\right); \frac{W_{t+s}}{W_t}\right)$$

hold. This is a fully rigorous statement, but should be regarded as a manifestation of the non-rigorous infinitesimal cascading procedure described earlier. The main focus of this work is to show that, under suitable assumptions on the i.i.d. collection of weight processes $W_t(v)$ attached to the vertices of the tree, the following is true:

Main Results. *Assume that the process $t \mapsto \log W_t$ is an independent increments process, with $W_0 = 1$, $\mathbf{E}[W_t] = 1$, and W_t always strictly positive. Assume the process is defined on an interval $[0,T]$ for some $T > 0$. If there is a $\delta > 0$ such that W_T has a finite $(1+\delta)$ moment, and the measure Γ is W_T-regular (see Definition 1.2.1), then*

- *the process $\Gamma_t := \mathcal{C}(\Gamma, W_t)$ is well-defined on $[0,T]$, i.e. the event that $\lim_{n\to\infty} \Gamma_t^{(n)}$ exists for all $0 \leq t \leq T$ has full probability,*
- *for any $s,t \geq 0$ with $t+s \leq T$, the equality $\Gamma_{t+s} = \mathcal{C}(\Gamma_t, W_{t+s}/W_t)$ also holds almost surely,*
- *the process is a martingale with respect to the filtration $\sigma(\Gamma_s : s \leq t)$,*
- *if the process $t \mapsto W_t$ is continuous, then so is the Γ_t process in the topology of weak convergence of measures.*
- *for $t \mapsto \log W_t$ a Gaussian process the measure process Γ_t is Hölder-$(\frac{1}{2} - \epsilon)$ continuous in the Wasserstein metric on measures for any $\epsilon > 0$, but not Hölder-$(\frac{1}{2} + \epsilon)$ continuous.*

These results are intuitive, but we want to emphasize that they are not immediate. It is easily seen that all four of these properties hold trivially for the finite level $t \mapsto \Gamma_t^{(n)}$ processes, but it requires some extra work to carry them over to the limit as $n \to \infty$. For fixed t and s, the Markov property, which is essentially a result about the composition of cascades, was first proven by [21] and later reproved in [8]. The existence of a discrete time Markov process would therefore follow from their work. With somewhat different analysis, we take care of the subtle difficulties in extending this notion to a continuous time process. The main technical difficulty is that the process cannot be started from an arbitrary measure; it has to be started from those which enjoy a sufficient amount of regularity. For practical applications the regularity condition we use is not at all restrictive, but we have to ensure that once the process begins it will stay within the class of sufficiently regular measures so that it can be continued. In Section 1.2 we describe exactly what we mean by sufficiently regular, and in Section 1.3 we prove that the evolution of the regularity of the process is well-behaved. This is a part of our proof of the results above.

It is also important to note that our main technique of replacing static weights with time varying processes has already been carried out for several other models. Likely the most prominent one is Dyson's Brownian motion, which is obtained by replacing the Gaussian entries of the GUE matrices with standard Brownian motions. More recently, however, the idea has been applied to the Sherrington-Kirkpatrick model of spin glasses in [4], and then re-applied to greater effect by a series of other authors [3, 19]. The paper [16] also uses the same technique in the context of lattice polymer models, which are somewhat similar to ours through the connection between tree polymers and multiplicative cascades. However, the main purpose of these papers is to use the dynamic weights technique to derive growth exponents and fluctuation behavior for partition functions of Gibbs measures as the size of the system grows large, whereas we are more concerned with showing that the infinite volume measure-valued process has the properties listed above.

We put particular emphasis on the results derived in Sections 1.4 and 1.5, where we specialize to the case when $\log W_t$ is a Brownian motion. This allows us to extend classical stochastic calculus results to this infinite-dimensional setting and use them to describe the evolution of the measures. One of our long term goals is to use these stochastic calculus techniques to compute explicit formulas for probability densities of certain quantities related to the measure; for example the total mass at any fixed time. We believe this is possible, but ultimately it will require more refined techniques that are beyond the scope of the current result. Nonetheless, interesting results can already be derived using the stochastic calculus that we develop, and in Section 1.5 we use it to show Hölder continuity of the measure process in the Wasserstein distance. We also show that the optimal Hölder exponent is $1/2$. Both are somewhat surprising facts, since for any $t \neq s$ the measures Γ_t and Γ_s are almost surely singular; hence the process $t \mapsto \Gamma_t$ is very discontinuous in the total variation distance. Given this it is not immediately clear that continuity can be expected in any topology stronger than the one induced by weak convergence, and our result should be viewed in this context.

The outline of this chapter is as follows: in Section 1.2 we set up our notation and recall some well known properties of cascade measures. In Section 1.3 we construct the process and show that it is well-defined, and give proofs for the main results listed above. In Section 1.4 we discuss the special case when the weight process is an exponential of a Brownian motion, and use stochastic calculus to describe the infinitesimal evolution of the process. This shows one advantage of our construction over the more abstract possibilities listed earlier: it allows for a full description of the evolution of the measure-valued process in terms of the input weight process $t \mapsto W_t(v)$. In Section 1.6 we describe possible applications of our process to models of tree polymers and to the KPZ formula of one-dimensional random geometry.

1.2 Background and Notation

We begin with our notation for trees. Let $\mathcal{T}$ be a rooted infinite binary tree and denote the root by ς. Given a vertex $v \in \mathcal{T}$ we let $|v|$ be its generation, by which we mean its distance from the root. Let v_L and v_R be the left and right offspring of v, respectively, and write v_p for the parent of v. We let $\mathcal{T}(v)$ be the subtree of $\mathcal{T}$ rooted at v. Note that when working on subtrees we still use $|u|$ to denote the distance from ς, not from the root of the subtree.

We will mostly be working on the boundary of $\mathcal{T}$, which we denote by $\partial\mathcal{T}$. Recall that $\partial\mathcal{T}$ is the set of all infinite self-avoiding paths in the tree that begin at the root. Elements of $\partial\mathcal{T}$ are called rays and are usually denoted by ξ. We denote by ξ_n the vertex in the n^{th} generation of the ray ξ. Given two rays ξ and ζ we let $\xi \wedge \zeta$ be the vertex of $\mathcal{T}$ that is the last common ancestor of ξ and ζ. For a given vertex $v \in \mathcal{T}$ we let $\partial\mathcal{T}(v)$ be the set of all rays passing through v.

1.2.1 Measures on $\partial\mathcal{T}$

Even though $\partial\mathcal{T}$ is an uncountable set, a measure on $\partial\mathcal{T}$ is completely determined by the countable collection of values $\{\Gamma(\partial\mathcal{T}(v))\}_{v\in\mathcal{T}}$. Hence every positive, finite measure on $\partial\mathcal{T}$ can be identified with a function $\Gamma : \mathcal{T} \to \mathbb{R}_+$ satisfying the two conditions

- $0 < \Gamma(\varsigma) < \infty$,
- for every vertex $v \in \mathcal{T}$, $\Gamma(v) = \Gamma(v_L) + \Gamma(v_R)$.

Due to this identification, measures on $\partial\mathcal{T}$ are also called flows on $\mathcal{T}$. As long as $0 < \Gamma(\varsigma) < \infty$ it is possible to normalize Γ to be a probability measure, i.e. so that $\Gamma(\varsigma) = 1$. Sometimes we denote the normalized measure by Γ^*, but in general we do not assume that we are working with probability measures.

A special measure on $\partial\mathcal{T}$ is the "Lebesgue" measure given by $\theta(v) = 2^{-|v|}$. Observe that θ is also the measure induced on $\partial\mathcal{T}$ by constructing random paths via simple random walk; that is, starting at the root and then using independent and unbiased coin flips at each vertex to decide whether to move left or right down the tree.

The topology on measures is as follows: we say that a sequence of measures Γ_n converges to Γ if $\Gamma_n(v) \to \Gamma(v)$ for all $v \in \mathcal{T}$. This is equivalent to weak convergence of $\Gamma_n \to \Gamma$, when the topology on $\partial\mathcal{T}$ is generated by the metric

$$d(\xi, \eta) = \theta(\xi \wedge \eta) = 2^{-|\xi \wedge \eta|}.$$

For a vertex $v \in \mathcal{T}$ we will write $\Gamma_{|v}$ for the measure restricted to the subtree $\mathcal{T}(v)$.

1.2.2 Random Cascade Measures on $\partial\mathcal{T}$

In this section we describe how to take a measure Γ on $\partial\mathcal{T}$ and a collection of random variables to construct a cascade measure. Let W be a random variable that is positive almost surely and has mean one. We are mostly concerned with its distribution which we call the *cascading distribution*. Assume that W is not identically one, and therefore Jensen's inequality implies that $\mathbf{E}[\log W] < 0$.

Let $\{W(v)\}_{v\in\mathcal{T}}$ be a collection of i.i.d. random variables with common distribution W. From this collection we build a random function $X : \mathcal{T} \to \mathbb{R}_+$ defined by

$$X(\xi_n) = \prod_{i=1}^{n} W(\xi_i).$$

Then for each $n \geq 0$ we construct a random measure $\Gamma_W^{(n)}$ by specifying the Radon-Nikodym derivative with respect to Γ as

$$d\Gamma_W^{(n)}(\xi) := X(\xi_n)\, d\Gamma(\xi).$$

The random cascade measure is then defined as the limit of $\Gamma_W^{(n)}$ as $n \to \infty$. Recall that the topology is pointwise in the vertices, meaning that

$$\Gamma_W(v) = \lim_{n\to\infty} \Gamma_W^{(n)}(v) \tag{1.2.1}$$

for every $v \in \mathcal{T}$. A simple martingale argument, which we now recall, shows that the limit always exists. First consider the case $v = \varsigma$, so that

$$\Gamma_W^{(n)}(\varsigma) = \int_{\partial\mathcal{T}} X(\xi_n) d\Gamma(\xi).$$

It is easy to see that $X(\xi_n)$ is a martingale with respect to the filtration

$$\mathcal{W}_n := \sigma(W(v) : |v| \leq n),$$

and therefore so is $\Gamma_W^{(n)}(\varsigma)$ by Fubini's Theorem. Since $\Gamma_W^{(n)}(\varsigma)$ is positive it converges almost surely, and since positivity of the limit does not depend on any finite collection of the $W(v)$ variables a standard 0-1 law argument shows that the limit is almost surely zero or almost surely strictly positive. In the case that $\Gamma = \theta$ Kahane and Peyriere [10] showed that

$$\mathbf{E}\left[W \log W\right] < \log 2$$

is a necessary and sufficient condition for the limit to be positive. In the case of a general measure Γ it remains an open problem to determine sharp criterion for when $\Gamma_W(\varsigma) > 0$, but there are many known sufficient conditions involving moment of W and the regularity of Γ. We will use a condition of Fan [7], for which we need the following definitions:

Definition 1.2.1. *For Γ a measure on $\partial\mathcal{T}$, define the pressure function $\lambda_\Gamma : [0, \infty) \to \mathbb{R}$ by*

$$\lambda_\Gamma(h) := \limsup_{n\to\infty} \frac{1}{n} \log \sum_{|v|=n} \Gamma(v)^h.$$

We will say that a measure Γ on $\partial\mathcal{T}$ is W-regular if

$$\mathbf{E}\left[W \log W\right] + \lambda'_\Gamma(1+) < 0.$$

We say that it is W-irregular if

$$\mathbf{E}\left[W \log W\right] + \lambda'_\Gamma(1-) > 0.$$

Observe that $\lambda_\Gamma(1) = 0$ for any Γ. Fan [7] uses the pressure function to derive the following condition:

Proposition 1.2.1 ([7])**.** *Suppose there exists a $\delta > 0$ with $\mathbf{E}\left[W^{1+\delta}\right] < \infty$ for some $\delta > 0$. Then*

- *if Γ is W-regular then $\Gamma_W(\varsigma) > 0$ almost surely,*
- *if Γ is W-irregular then $\Gamma_W(\varsigma) = 0$ almost surely.*

Observe that if λ_Γ is differentiable at $h = 1$ then the condition of W-regularity is close to sharp. For $\Gamma = \theta$ we have $\lambda_\theta(h) = (1 - h) \log 2$, and hence the condition of Kahane and Peyriere is recovered.

Remark 1.2.2. *Let W_1 and W_2 be two distinct cascading distributions, and suppose there is an $\epsilon > 0$ such that $\mathbf{E}\left[W_1^h\right] \leq \mathbf{E}\left[W_2^h\right] < \infty$ for $h \in [1, 1+\epsilon]$. Then since*

$$\mathbf{E}\left[W \log W\right] = \lim_{h\downarrow 0} \frac{\mathbf{E}\left[W^h\right] - 1}{h}$$

it follows that $\mathbf{E}\left[W_1 \log W_1\right] \leq \mathbf{E}\left[W_2 \log W_2\right]$. Hence W_2-regularity of Γ implies W_1-regularity of Γ.

Remark 1.2.3. *The assumption of W-regularity implicitly means that λ_Γ is differentiable from the right at $h = 1$.*

Remark 1.2.4. *It is important to note that W-regularity of a measure is a property that is inherited by all of its submeasures. Indeed, since λ_Γ is computed over a larger sum than $\lambda_{\Gamma_{|v}}$, it follows that $\lambda_\Gamma(h) \geq \lambda_{\Gamma_{|v}}(h)$ for all h. But also $\lambda_\Gamma(1) = \lambda_{\Gamma_{|v}}(1) = 0$, and therefore by the Mean Value Theorem*

$$0 \leq \lambda_\Gamma(1+\epsilon) - \lambda_{\Gamma_{|v}}(1+\epsilon) = \lambda'_\Gamma(1+s) - \lambda'_{\Gamma_{|v}}(1+s)$$

for some $s \in (0, \epsilon)$. Taking ϵ to zero gives

$$\lambda'_{\Gamma_{|v}}(1+) \leq \lambda'_\Gamma(1+),$$

which implies W-regularity of $\Gamma_{|v}$.

Remark 1.2.5. *We have only shown existence of the limit (1.2.1) in the $v = \varsigma$ case, and it is important to note that this only required that Γ is a positive measure on $\partial\mathcal{T}$. The regularity of Γ determines whether the limit is positive or zero. But these facts and the self-similarity of the tree also combine to give us the existence and positivity of the limit for $v \neq \varsigma$. Indeed, assume $n > |v|$, so that*

$$\begin{aligned}\Gamma_W^{(n)}(v) &= \int_{\partial\mathcal{T}} X(\xi_n)\mathbf{1}\{\xi \in \partial\mathcal{T}(v)\}\, d\Gamma(\xi) \\ &= X(v)\int_{\partial\mathcal{T}(v)} \frac{X(\xi_{n-|v|})}{X(v)} d\Gamma_{|v}(\xi).\end{aligned} \tag{1.2.2}$$

But the integral term is just the level $n - |v|$ cascade of the $\Gamma_{|v}$ measure by the random variables $W_{|v} = \{W(u) : u \in \mathcal{T}(v)\}$, hence the martingale argument for the $v = \rho$ case also shows that its limit exists as $n \to \infty$. Its positivity can again be determined by Fan's condition, and by the last remark W-regularity is inherited by all submeasures. Thus if Γ is W-regular then $\Gamma_W(v) > 0$ for all $v \in \mathcal{T}$ with $\Gamma(v) > 0$. Taking the limit as $n \to \infty$ in equation (1.2.2) gives the relation

$$\frac{\Gamma_W(u)}{X(v)} = \mathcal{C}\left(\Gamma_{|v}; W_{|v}\right)(u) \tag{1.2.3}$$

for all $u \in \mathcal{T}(v)$.

Finally we remark that even though the limits in (1.2.1) are defined pointwise at each vertex, the limiting object Γ_W is automatically a measure on $\partial\mathcal{T}$. This follows from the definition of the level n cascade as a measure, and therefore

$$\Gamma_W^{(n)}(v) = \Gamma_W^{(n)}(v_L) + \Gamma_W^{(n)}(v_R).$$

Now take limits as $n \to \infty$.

1.2.3 Rates of Convergence for the Cascading Procedure

Our analysis in this section relies on that in [[7]]. In particular we will need to assume that the cascade variable W satisfies a moment constraint and that the measure Γ is W-regular.

Assumption 1. *We assume that*

- *There is a $\delta > 0$ such that $\mathbf{E}\left[W^{1+\delta}\right] < \infty$.*

- *The measure Γ is W-regular.*

These assumptions allow for exponential control on the decay of the cascade measure.

Definition 1.2.2. *Define*

$$\alpha(h) := \alpha(h; W, \Gamma) = \limsup_{n\to\infty} \frac{1}{n} \log \sum_{|v|=n} \Gamma(v)^h \mathbf{E}\left[X(v)^h\right] = \lambda_\Gamma(h) + \log \mathbf{E}\left[W^h\right].$$

The moment assumption on W implies that $\alpha(h) < \infty$ for h in a neighborhood of 1. Since $\alpha(1) = 0$ it is straightforward to compute that Γ being W-regular implies that $\alpha'(1+) < 0$, and therefore $\alpha(1+\epsilon) < \alpha(1) = 0$ for ϵ sufficiently small. Therefore we also define

$$h_W := \sup\{h \geq 1 : \alpha(h) < 0\},$$

and by the last remarks we have $h_W > 1$ under Assumption 1.

Much of our analysis will rely on having a rate of convergence of $\Gamma_W^{(n)}$ to Γ_W. We will heavily make use of the following lemma:

Lemma 1.2.6. *For $1 \leq h \leq 2$ there is a positive constant $C = C(h)$ such that*

$$\left|\left|\Gamma_W^{(n+1)}(\varsigma) - \Gamma_W^{(n)}(\varsigma)\right|\right|_{L^h} \leq C||W||_{L^h}^{n+1} \left(\sum_{|v|=n} \Gamma(v)^h\right)^{1/h},$$

and therefore by the triangle inequality

$$\left|\left|\Gamma_W(\varsigma) - \Gamma_W^{(n)}(\varsigma)\right|\right|_{L^h} \leq C \sum_{m>n} ||W||_{L^h}^m \left(\sum_{|v|=m} \Gamma(v)^h\right)^{1/h}.$$

The proof relies on the following inequality of von Bahr and Esseen:

Lemma 1.2.7 ([20])**.** *Let $\{U_i\}$ and $\{V_i\}$ be sequences of random variables that are independent of each other. Also assume that the $\{V_i\}$ are mutually independent, and that $\mathbf{E}[V_i] = 0$ for all i. Then for $1 \leq h \leq 2$ there is a universal constant $c = c(h)$ such that*

$$\mathbf{E}\left[\left(\sum_i U_i V_i\right)^h\right] \leq c \sum_i \mathbf{E}\left[U_i^h\right] \mathbf{E}\left[V_i^h\right].$$

Proof of Lemma 1.2.6. We have the trivial identity

$$\begin{aligned}
\Gamma^{(n+1)}(\varsigma) - \Gamma^{(n)}(\varsigma) &= \int \left(X(\xi_{m+1}) - X(\xi_m)\right) d\Gamma(\xi) \\
&= \int X(\xi_m)(W(\xi_{m+1}) - 1)\, d\Gamma(\xi) \\
&= \sum_{|v|=m+1} \Gamma(v) X(v_p)(W(v) - 1).
\end{aligned}$$

The von Bahr-Esseen inequality applies to the latter sum, and therefore

$$\begin{aligned}\mathbf{E}\left[\left|\Gamma_W^{(n+1)}(\varsigma)-\Gamma_W^{(n)}(\varsigma)\right|^h\right] &\leq c(h)\sum_{|v|=n+1}\Gamma(v)^h\mathbf{E}\left[W^h\right]^n\mathbf{E}\left[|W-1|^h\right] \\ &\leq 2c(h)\mathbf{E}\left[W^h\right]^{n+1}\sum_{|v|=n+1}\Gamma(v)^h.\end{aligned}$$

□

The next lemma implies the L^h convergence of the total mass of the cascade measure, and therefore that $\Gamma_W(\varsigma) > 0$ almost surely.

Corollary 1.2.8. *Under Assumption 1 we have that for all $h \in (1, h_W)$,*

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbf{E}\left[|\Gamma_W(\varsigma)-\Gamma_W^{(n)}(\varsigma)|^h\right] \leq \lambda_\Gamma(h)+\log\mathbf{E}[W^h] < 0.$$

Proof. Since Γ is W-regular, by Definition 1.2.2 for all $h \in (1, h_W)$, $\alpha(h) = \lambda_\Gamma(h) + \log\mathbf{E}\left[W^h\right] < 0$. Hence for each $\gamma > 0$ such that $\alpha(h) + \gamma < 0$ there is a positive constant C such that

$$\sum_{|v|=n}\Gamma(v)^h\mathbf{E}\left[X(v)^h\right] \leq Ce^{n(\alpha(h)+\gamma)}$$

for all n. Applying the second statement of Lemma 1.2.6 completes the proof. □

We now extend Corollary 1.2.8 to show that the exponential rate of convergence is uniform for all vertices on a fixed generation of the tree.

Lemma 1.2.9. *Under Assumption 1 we have that for $h \in (1, h_W)$ and for $1 \leq i \leq n$*

$$\limsup_{n\to\infty}\frac{1}{n}\log\sum_{|v|=i}\mathbf{E}\left[\left|\Gamma_W(v)-\Gamma_W^{(n)}(v)\right|^h\right] \leq \lambda_\Gamma(h)+\log\mathbf{E}\left[W^h\right] < 0.$$

And moreover,

$$\limsup_{n\to\infty}\frac{1}{n}\log\sum_{i=1}^{n}\sum_{|v|=i}\mathbf{E}\left[\left|\Gamma_W(v)-\Gamma_W^{(n)}(v)\right|^h\right] \leq \lambda_\Gamma(h)+\log\mathbf{E}[W^h] < 0.$$

Proof. From (1.2.2) we have, for $|v| = i \leq n$,

$$\Gamma_W^{(n)}(v) = X(v)\mathcal{C}\left(\Gamma_{|v}; W_{|v}\right)^{(n-i)}(v).$$

Combining this with (1.2.3) and using that $X(v)$ is independent of the cascade on $\mathcal{T}(v)$ gives

$$\mathbf{E}\left[|\Gamma_W(v)-\Gamma_W^{(n)}(v)|^h\right] = \mathbf{E}\left[X(v)^h\right]\mathbf{E}\left[\left|\mathcal{C}(\Gamma_{|v}; W_{|v})(v)-\mathcal{C}(\Gamma_{|v}; W_{|v})^{(n-i)}(v)\right|^h\right].$$

Applying Lemma 1.2.6 to the second factor and combining with the first factor gives

$$\mathbf{E}\left[|\Gamma_W(v) - \Gamma_W^{(n)}(v)|^h\right] \leq C \left(\sum_{k>n-i} \left(\sum_{\substack{w \in \mathcal{T}(v) \\ |w|=|v|+k}} \Gamma(w)^h \mathbf{E}\left[W^h\right]^{|v|+k} \right)^{1/h} \right)^h, \tag{1.2.4}$$

where C depends only on h. Now define $a_k(v)$ by

$$a_k(v) = \sum_{\substack{w \in \mathcal{T}(v) \\ |w|=|v|+k}} \Gamma(w)^h \mathbf{E}\left[W^h\right]^{|v|+k}$$

and $\mathbf{a_n}(v) = (a_{n+1}(v), a_{n+2}(v), a_{n+3}(v), \ldots)$. Then equation (1.2.4) is equivalent to

$$\mathbf{E}\left[|\Gamma_W(v) - \Gamma_W^{(n)}(v)|^h\right] \leq C||\mathbf{a_{n-i}}(v)||_{\ell^{1/h}},$$

with $\ell^{1/h}$ denoting the usual sequence space. Summing over $|v| = i$ gives

$$\sum_{|v|=i} \mathbf{E}\left[|\Gamma_W(v) - \Gamma_W^{(n)}(v)|^h\right] \leq C \sum_{|v|=i} ||\mathbf{a_{n-i}}(v)||_{\ell^{1/h}} \leq C \left|\left| \sum_{|v|=i} \mathbf{a_{n-i}}(v) \right|\right|_{\ell^{1/h}}. \tag{1.2.5}$$

The last inequality is the Minkowski inequality for $\ell^{1/h}$ (recall $h \geq 1$). By definition of $\mathbf{a}$ we have

$$\sum_{|v|=i} \mathbf{a_{n-i}}(v) = (a_{n+1}(\varsigma), a_{n+2}(\varsigma), a_{n+3}(\varsigma), \ldots) = \mathbf{a_n}(\varsigma). \tag{1.2.6}$$

By definition of $\alpha(h)$ we have, for each $\gamma > 0$,

$$a_n(\varsigma) = \sum_{|v|=n} \Gamma(v)^h \mathbf{E}\left[W^h\right]^n \leq e^{n(\alpha(h)+\gamma)}$$

for n sufficiently large. Under Assumption 1 and using Remark 1.2.2 we have $\alpha(h) < 0$ for $h \in (1, h_W)$. Choosing γ such that $\alpha(h) + \gamma < 0$, this gives

$$||\mathbf{a_n}(\varsigma)||_{\ell^{1/h}} \leq Ce^{n(\alpha(h)+\gamma)}$$

for n sufficiently large. Combining this with (1.2.5) and (1.2.6) and sending γ to zero gives the first statement of the lemma. For the second part, simply observe that by (1.2.5) we have

$$\sum_{i=1}^{n} \sum_{|v|=i} \mathbf{E}\left[|\Gamma_W(v) - \Gamma_W^{(n)}(v)|^h\right] \leq Cn||\mathbf{a_n}(\varsigma)||_{\ell^{1/h}}.$$

□

This easily implies a uniform control of some moment of the cascade measure over all the vertices v in the tree.

Corollary 1.2.10. *Under Assumption 1 we have that for all $h \in (1, h_W)$*

$$\limsup_{n\to\infty} \frac{1}{n} \log \sum_{|v|=n} \mathbf{E}\left[|\Gamma_W(v)|^h\right] \leq \lambda_\Gamma(h) + \log \mathbf{E}\left[W^h\right] < 0$$

Proof. From the inequality $|a+b|^h \leq 2^h(|a|^h + |b|^h)$ we have

$$\mathbf{E}\left[\Gamma_W(v)^h\right] \leq 2^h \left(\mathbf{E}\left[|\Gamma_W(v) - X(v)\Gamma(v)|^h\right] + \Gamma(v)^h \mathbf{E}\left[X(v)^h\right]\right).$$

Summing over $|v| = n$ and taking logarithms we get

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log \sum_{|v|=n} \mathbf{E}\left[\Gamma_W(v)^h\right] &\leq \limsup_{n\to\infty} \frac{1}{n} \log \sum_{|v|=n} \left(\mathbf{E}\left[|\Gamma_W(v) - X(v)\Gamma(v)|^h\right] + \Gamma(v)^h \mathbf{E}\left[X(v)^h\right]\right) \\ &\leq \lambda_\Gamma(h) + \log \mathbf{E}\left[W^h\right].\end{aligned}$$

The last inequality is a consequence of the fact that the two terms in the line above both have the same exponential rate of decay, which is itself a consequence of Lemma 1.2.9. ☐

1.3 A Markovian Random Cascade Process

1.3.1 Dynamic Random Weights

The main idea of this work is to replace the random weights W on the vertices of the tree with random weight processes $t \mapsto W_t$ that evolve in time. As usual we require a moment of the cascade variable W_t as well as regularity of the measure Γ. To this we also add an independence condition. Throughout we assume the following properties of the weight processes and the initial measure Γ.

Assumption 2. *The weight process $t \mapsto W_t$ is defined in an interval $[0, T]$ with $T > 0$, and*

- *there is a $\delta > 0$ such that $\mathbf{E}\left[W_T^{1+\delta}\right] < \infty$,*
- *the measure Γ is W_T-regular.*
- *$W_0 = 1$,*
- *$W_t > 0$ and $\mathbf{E}[W_t] = 1$ for each $t \geq 0$,*
- *$t \mapsto \log W_t$ has independent increments.*

Remark 1.3.1. *Observe that for $p > 1$*

$$\mathbf{E}[W_T^p] = \mathbf{E}[W_t^p]\, \mathbf{E}\left[\frac{W_T}{W_t}\right]^p \geq \mathbf{E}[W_t^p],$$

and so, by Remark 1.2.2, the moment and regularity assumptions are inherited for W_t with $t < T$.

Such processes are easy to construct, for example exponentials of Brownian motion or exponentials of Levy processes (both properly normalized so that $\mathbf{E}[W_t] = 1$). Note, however, that in both of these examples the increments of $\log W_t$ are stationary, but that our results do *not* require this. For $s, t \geq 0$ we define

$$W_{t,t+s} := \frac{W_{t+s}}{W_t}.$$

The independent increments assumption gives that $W_{t,t+s}$ is independent of W_t. Moreover the process $t \mapsto W_t$ is a martingale, that is

$$\mathbf{E}\left[W_t | \sigma(W_r : r \leq s)\right] = W_s.$$

Now to each vertex $v \in \mathcal{T}$ attach a copy $W_t(v)$ of this process such that the collection $\{W_t(v)\}_{v \in \mathcal{T}}$ is independent. The main idea of this work is to use the cascading procedure of the last section to construct a random cascade measure Γ_{W_t} at each time $t \geq 0$, and then show that the resulting process $t \mapsto \Gamma_{W_t}$ is Markovian. This is carried out in Section 1.3.3, and the rest of the work studies properties of the process. To simplify notation we write

$$\Gamma_t := \Gamma_{W_t} = \mathcal{C}(\Gamma; W_t).$$

Observe that $\Gamma_0 = \Gamma$. We also define functions $X_{t,t+s} : \mathcal{T} \to \mathbb{R}_+$ by

$$X_{t,t+s}(\xi_n) = \prod_{i=1}^{n} W_{t,t+s}(\xi_i),$$

and the filtrations

$$\mathcal{W}_t = \sigma\left(W_s(v) : v \in \mathcal{T}, s \leq t\right), \quad \mathcal{F}_t = \sigma\left(\Gamma_s(v) : v \in \mathcal{T}, s \leq t\right).$$

In general $\mathcal{F}_t \subset \mathcal{W}_t$ and the inclusion in strict, since by knowing the weights one can construct the measure, but knowing the measure does not generally give full information on the weights.

To simplify notation, we will often drop references to the weights W_t or the initial measure Γ.

Definition 1.3.1. *For $t \in [0, T]$, let*

$$\alpha_t(h) := \limsup_{n \to \infty} \frac{1}{n} \log \sum_{|v|=n} \Gamma(v)^h \mathbf{E}\left[X_t(v)^h\right] = \lambda_\Gamma(h) + \log \mathbf{E}\left[W_t^h\right].$$

Furthermore let

$$h_t := \sup\{h \geq 1 \text{ : such that } \alpha_t(h) < 0\}.$$

Remark 1.3.2. *Under Assumption 2, $h_t > 1$ for $t \in [0, T]$. Further note that h_t is decreasing with t.*

1.3.2 Construction and Basic Properties of the Process

Before proving that the $t \mapsto \Gamma_t$ process is Markov we first deal with a technical issue. Above we said that we construct the process $t \mapsto \Gamma_t$ by applying the random cascading procedure of Section 2 at each fixed time t. However the existence of the random cascade measure is only an almost sure statement, and the event that it does not exist could conceivably depend on t. Since we are now working in continuous time it is possible that there is an exceptional set of times for which the cascade does *not* exist, which would leave our cascade process ill-defined. We begin by showing that this is not the case.

To this end first note that for each $n > 0$ the finite level measure processes $t \mapsto \Gamma_t^{(n)}$ are well-defined, and in fact are martingales in t with respect to the filtration $\mathcal{W}_t$. Indeed

$$\mathbf{E}\left[d\Gamma_{t+s}^{(n)}(\xi)|\mathcal{W}_t\right] = X_t(\xi_n)d\Gamma(\xi)\mathbf{E}\left[X_{t,t+s}(\xi_n)\right] = d\Gamma_t^{(n)}(\xi),$$

by the fact that $X_{t,t+s}$ is independent of $\mathcal{W}_t$ and has mean one. We will show that this martingale property, together with the exponential L^p convergence of the finite level measures, gives that the Γ_t process is well-defined. Moreover, the martingale property of the finite level measures is inherited by the limit.

Theorem 1.3.3. *Under Assumptions 2, the event*

$$\left\{\lim_{n\to\infty} \Gamma_t^{(n)}(v) \text{ exists for all } v \in \mathcal{T}, t \le T\right\}$$

has probability one. Moreover,

1. *for each $v \in \mathcal{T}$ the process $t \mapsto \Gamma_t(v)$ is a martingale with respect to $\mathcal{W}_t$, and hence $\mathcal{F}_t$, and,*
2. *if the weight process $t \mapsto W_t$ is continuous then so is $\Gamma_t(v)$ for each $v \in \mathcal{T}$.*

Proof. We concentrate first on the case $v = \varsigma$. Fix $h \in (1, h_T)$. Since the difference $\Gamma_t^{(n+1)}(\varsigma) - \Gamma_t^{(n)}(\varsigma)$ is a martingale in t (with respect to the filtration $\mathcal{W}_t$), we may apply Doob's maximal L^h inequality to get that

$$\begin{aligned}
P\left(\sup_{0\le t\le T} |\Gamma_t^{(n+1)}(\varsigma) - \Gamma_t^{(n)}(\varsigma)| > \beta^n\right) &\le \beta^{-nh}\mathbf{E}\left[\sup_{0\le t\le T} |\Gamma_t^{(n+1)}(\varsigma) - \Gamma_t^{(n)}(\varsigma)|^h\right] \\
&\le \beta^{-nh}\left(\frac{h}{h-1}\right)^h \mathbf{E}\left[|\Gamma_T^{(n+1)}(\varsigma) - \Gamma_T^{(n)}(\varsigma)|^h\right] \\
&\le C\beta^{-nh}\mathbf{E}\left[W_T^h\right]^n \sum_{|v|=n} \Gamma(v)^h,
\end{aligned}$$

with the last inequality coming from Lemma 1.2.6. Therefore by taking logarithms we get

$$\limsup_{n\to\infty} \frac{1}{n}\ \log P\left(\sup_{0\le t\le T} |\Gamma_t^{(n+1)}(\varsigma) - \Gamma_t^{(n)}(\varsigma)| > \beta^n\right) \le -h\log\beta + \alpha_T(h).$$

As $\alpha_T(h) < 0$, we can pick $\beta < 1$ so that the right hand side is less than zero. Now Apply Borel-Cantelli to conclude that $\Gamma_t^{(n)}(\varsigma)$ is a Cauchy sequence in n, with a rate of convergence that is uniform in t. This proves the first part of the theorem.

To prove the martingale property, simply note that by Corollary 1.2.8 there is an $h > 1$ such that $\Gamma_t^{(n)}(\varsigma)$ converges to $\Gamma_t(\varsigma)$ in L^h, and hence in L^1. Thus for $A \in \mathcal{W}_s$

$$\mathbf{E}\left[(\Gamma_{t+s}(\varsigma) - \Gamma_t(\varsigma))\mathbf{1}_A\right] = \lim_{n\to\infty} \mathbf{E}\left[(\Gamma_{t+s}^{(n)}(\varsigma) - \Gamma_t^{(n)}(\varsigma))\mathbf{1}_A\right] = 0,$$

with the last equality using the martingale property of the finite level measure process. This proves that Γ_t is a martingale with respect to $\mathcal{W}_t$, but since $\mathcal{F}_t \subset \mathcal{W}_t$ and Γ_t is $\mathcal{F}_t$-measurable, it is automatically a martingale with respect to $\mathcal{F}_t$ also.

For the continuity statement observe that if W_t is continuous then so is $\Gamma_t^{(n)}(\varsigma)$, since it is a finite product and sum of continuous functions. The Borel-Cantelli argument above gives continuity of $\Gamma_t(\varsigma)$ by completeness of $C([0,T])$ under the sup norm.

Finally, if $v \neq \varsigma$ then the proofs above are easily extended by noting that W_T-regularity is inherited by the submeasures $\Gamma_{|v}$ (see Remark 1.2.4). The simple relation $\Gamma_t(v) = X_t(v)\mathcal{C}(\Gamma_{|v}, W_t)(v)$ finishes the argument, and since there are only countably many vertices on the tree the proof is completed. □

1.3.3 The Markov Property

In this section we show that the Γ_t process has the Markov property. For a given weight process W_t on $[0,T]$, let $\mathcal{M}_T$ be the space of measures Γ that satisfy Assumption 2. The Markov property can be formally stated by saying that for any bounded, measurable $F : \mathcal{M}_T \to \mathbb{R}$ we have

$$\mathbf{E}\left[F(\Gamma_{t+s})\middle|\,\mathcal{F}_t\right] = \mathbf{E}\left[F(\Gamma_{t+s})\middle|\,\Gamma_t\right],$$

for $s,t \geq 0$ such that $s+t \leq T$. By a density argument it is sufficient to consider the functions of the form $F_v(\Gamma) = \Gamma(v)$ for $v \in \mathcal{T}$. For these functions we will actually prove the stronger statement

$$\mathbf{E}\left[F_v(\Gamma_{t+s})\middle|\,\mathcal{W}_t\right] = \mathbf{E}\left[F_v(\Gamma_{t+s})\middle|\,\Gamma_t\right],$$

the difference between the two being that $\mathcal{W}_t$ is a coarser σ-algebra than $\mathcal{F}_t$. Since the weight processes $s \mapsto W_{t,t+s}$ are independent of $\mathcal{W}_t$, it is sufficient to prove the following:

Theorem 1.3.4. *Under Assumptions 2, for fixed $s,t \geq 0$ such that $t+s \leq T$, we have with probability one that*

$$\Gamma_{t+s} = \mathcal{C}\left(\Gamma_t; W_{t,t+s}\right).$$

Proof. We will show that for every v in $\mathcal{T}$,

$$\Gamma_{t+s}(v) = \mathcal{C}(\Gamma_t; W_{t,t+s})(v). \tag{1.3.1}$$

We first concentrate on the case $v = \varsigma$. Note that both sides of equation (1.3.1) are defined as limits, and it suffices to prove that

$$\lim_{n\to\infty} \mathcal{C}(\Gamma_t; X_{t,t+s})^{(n)}(\varsigma) - \Gamma_{t+s}^{(n)}(\varsigma) = 0. \tag{1.3.2}$$

We will show that the left hand side of (1.3.2) goes to zero in L^h for any $h \in (1, h_T)$, and therefore the a.s. limit must be zero as well. Fix $h \in (1, h_T)$ and recall that

$$\mathcal{C}(\Gamma_t; X_{t,t+s})^{(n)}(\varsigma) = \sum_{|v|=n} \Gamma_t(v) X_{t,t+s}(v) = \sum_{|v|=n} \frac{\Gamma_t(v)}{X_t(v)} X_{t+s}(v).$$

The last equality follows since $X_t X_{t,t+s} = X_{t+s}$ by construction. Therefore, by definition of $\Gamma^{(n)}_{t+s}$,

$$\mathcal{C}(\Gamma_t; X_{t,t+s})^{(n)}(\varsigma) - \Gamma^{(n)}_{t+s}(\varsigma) = \sum_{|v|=n} \left(\frac{\Gamma_t(v)}{X_t(v)} - \Gamma(v) \right) X_{t+s}(v).$$

Now note that the random variables $\{\Gamma_t(v)/X_t(v) - \Gamma(v) : |v| = n\}$ are mean zero, and each depends only on the W_t weights in the subtree $\mathcal{T}(v)$. Hence they are independent of each other *and* of all the weight processes $t \mapsto W_t(v)$ with $|v| \le n$. In particular each $X_{t+s}(v)$, for $|v| = n$, is independent of these random variables. Thus we can apply the von Bahr-Esseen inequality to the difference above to get

$$\begin{aligned}\mathbf{E}\left[\left|\mathcal{C}(\Gamma_t; X_{t,t+s})^{(n)}(\varsigma) - \Gamma^{(n)}_{t+s}(\varsigma)\right|^h\right] &\le \sum_{|v|=n} \mathbf{E}\left[X_{t+s}(v)^h\right] \mathbf{E}\left[\left|\frac{\Gamma_t(v)}{X_t(v)} - \Gamma(v)\right|^h\right] \\ &= \sum_{|v|=n} \mathbf{E}\left[X_{t,t+s}(v)^h\right] \mathbf{E}\left[\left|\Gamma_t(v) - X_t(v)\Gamma(v)\right|^h\right].\end{aligned}$$

Recognizing that $X_t(v)\Gamma(v) = \Gamma^{(n)}_t(v)$ and applying Lemma 1.2.9 finishes the proof, since for $h \in (1, h_T)$ we have

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log \mathbf{E}\left[\left|\mathcal{C}(\Gamma_t; X_{t,t+s})^{(n)}(\varsigma) - \Gamma^{(n)}_{t+s}(\varsigma)\right|^h\right] &\le \log \mathbf{E}\left[W^h_{t,t+s}\right] + \lambda_\Gamma(h) + \log \mathbf{E}\left[W^h_t\right] \\ &= \lambda_\Gamma(h) + \log \mathbf{E}\left[W^h_{t+s}\right] \\ &\le \alpha_T(h) \\ &< 0.\end{aligned}$$

The second inequality follows from Remark 1.3.1, and the last is by the W_T-regularity of Γ.

The proof for $v \neq \varsigma$ is similar, with all sums in the above statements being replaced with sums over the appropriate subtrees, and by making use of the fact that $\Gamma_{|v}$ is W_T-regular for all v. Finally, since there are only countably many vertices on the tree the statement holds for all vertices simultaneously. □

Note that Theorem 1.3.4 assumes nothing about the regularity of Γ_t, even though we applied the cascading procedure to it. Theorem 1.3.4 gives that $\mathcal{C}(\Gamma_t, W_{t,t+s})$ is indeed a non-trivial measure since it is equal to Γ_{t+s}, which was already known to be non-trivial by the W_T-regularity of the original measure Γ. However, the regularity condition is only a sufficient one, and so the fact that $\mathcal{C}(\Gamma_t, W_{t,t+s})$ is non-trivial does not imply that Γ_t is $W_{t,T}$ regular. This regularity statement is true, however, and we will now prove it. In some sense this gives a classification of the state space of the Markov process: each Γ_t lives in the space of $W_{t,T}$-regular measures, which is itself contained in the space of W_T-regular measures.

Lemma 1.3.5. *Under Assumptions 2, the measures Γ_t are $W_{t,T}$-regular for each $t \le T$.*

Proof. From Corollary 1.2.10, for $h \in (1, h_T)$, we get that

$$\limsup_{n\to\infty} \frac{1}{n} \log \sum_{|v|=n} \mathbf{E}\left[\Gamma_t(v)^h\right] \le \alpha_t(h) < \alpha_T(h) < 0.$$

Therefore an application of Borel-Cantelli implies that

$$\lambda_{\Gamma_t}(h) = \limsup_{n\to\infty} \frac{1}{n} \log \sum_{|v|=n} \Gamma_t(v)^h \le \lambda_\Gamma(h) + \log \mathbf{E}\left[W_t^h\right]$$

with probability one. This gives that

$$\lambda_{\Gamma_t}(h) + \log \mathbf{E}\left[W_{t,T}^h\right] \le \lambda_\Gamma(h) + \log \mathbf{E}\left[W_t^h\right] + \log \mathbf{E}\left[W_{t,T}^h\right] = \lambda_\Gamma(h) + \log \mathbf{E}\left[W_T^h\right]$$

for all $h \in (1, h_T)$. Now apply the Mean Value Theorem and take $h \downarrow 1$ to finish the proof. $\square$

1.4 Gaussian Weight Processes

The simplest case of weights satisfying Assumption 2 is an exponential of a Brownian motion, properly normalized. In this section we study some extra properties of the random cascade process with these weights; specifically we derive stochastic calculus formulas for the evolution of the measures as driven by the Brownian noise. We restrict ourselves to the simplest case when $t \mapsto \log W_t$ has stationary increments, so that

$$W_t(v) = \exp\{B_t(v) - t/2\},$$

where $\{B_t(v)\}_{v\in\mathcal{T}}$ is a collection of independent Brownian motions with $B_0(v) = 0$. Since the W_t variables have moments of all orders for all $t \ge 0$, we only need to assume that the initial measure Γ is W_T-regular for some $T > 0$. It is easy to compute that

$$\mathbf{E}\left[W_t \log W_t\right] = -\frac{t}{2},$$

so therefore the cascade process is well-defined on $[0, -2\lambda'_\Gamma(1+))$. It is straightforward to verify from the definition of λ_Γ that $-2\lambda'_\Gamma(1+)$ is maximal when $\Gamma = \theta$, and this maximum value is $2\log 2$. Moreover, for any $T < -2\lambda'_\Gamma(1+)$, Assumptions 2 are satisfied and hence the $t \mapsto \Gamma_t$ process is always defined on a finite time interval. By Theorems 1.3.3 and 1.3.4 the process is Markovian, *and* $t \mapsto \Gamma_t(v)$ is a continuous martingale for each $v \in \mathcal{T}$. Since

$$X_t(\xi_n) = \exp\left\{\sum_{i=1}^n B_t(\xi_n) - nt/2\right\},$$

it is easy to compute that

$$\frac{dX_t(\xi_n)}{X_t(\xi_n)} = \sum_{i=1}^n dB_t(\xi_i).$$

Therefore

$$d\Gamma_t^{(n)}(\varsigma) = \int_{\partial\mathcal{T}} X_t(\xi_n)\left(\sum_{i=1}^n dB_t(\xi_n)\right) d\Gamma(\xi) = \sum_{i=1}^n \int_{\partial\mathcal{T}} dB_t(\xi_i) d\Gamma_t^{(n)}(\xi). \tag{1.4.1}$$

This leads to the following result:

Proposition 1.4.1. *The total mass $\Gamma_t(\varsigma)$ evolves according to the stochastic differential equation*

$$d\Gamma_t(\varsigma) = \sum_{i=1}^{\infty} \int_{\partial \mathcal{T}} dB_t(\xi_i)\, d\Gamma_t(\xi) = \sum_{i=1}^{\infty} \mathbf{E}_{\Gamma_t}\left[dB_t(\xi_i)\right] = \sum_{\substack{v\in\mathcal{T}\\ v\neq\varsigma}} \Gamma_t(v)\, dB_t(v), \tag{1.4.2}$$

where all stochastic differentials are understood in the Itô sense. Equivalently

$$\frac{d\Gamma_t(\varsigma)}{\Gamma_t(\varsigma)} = \sum_{i=1}^{\infty} \mathbf{E}_{\Gamma_t^*}\left[dB_t(\xi_i)\right] = \sum_{\substack{v\in\mathcal{T}\\ v\neq\varsigma}} \Gamma_t^*(v)\, dB_t(v),$$

where Γ_t^* is Γ_t normalized to be a probability measure. The quadratic variation of the latter process is

$$\frac{d\left\langle \Gamma_t(v), \Gamma_t(v)\right\rangle}{\Gamma_t(v)^2} = \sum_{i=1}^{\infty} \mathbf{E}_{\Gamma_t^*\times\Gamma_t^*}\left[\mathbf{1}\left\{\xi_i = \xi_i'\right\}\right] = \sum_{\substack{v\in\mathcal{T}\\ v\neq\varsigma}} \Gamma_t^*(v)^2 = \sum_{\substack{v\in\mathcal{T}\\ v\neq\varsigma}} \left(\frac{\Gamma_t(v)}{\Gamma_t(\varsigma)}\right)^2.$$

Before proceeding with the proof we first note that all of the stochastic integrals

$$\int_0^t \Gamma_s^{(n)}(v)\, dB_s(v), \qquad \int_0^t \Gamma_s(v)\, dB_s(v)$$

are well-defined on $[0,T]$. Both integrands are clearly progressively measurable, and as they are continuous local martingales in time their supremum is almost surely finite on the compact interval $[0,T]$. Hence

$$\int_0^T \Gamma_s^{(n)}(v)^2\, ds < \infty \quad \text{and} \quad \int_0^T \Gamma_s(v)^2\, ds < \infty$$

with probability one, which is exactly what is required for the integrals to make sense. Note, however, that the expectations of the latter integrals will not necessarily be finite for all T.

Proof. By the definition of $\Gamma_t(\varsigma)$ as the limit of $\Gamma_t^{(n)}(\varsigma)$, and computing the difference between (1.4.1) and (1.4.2), it is sufficient to show that the process

$$t \mapsto \sum_{i=1}^{n} \sum_{|v|=i} \int_0^t \left(\Gamma_s(v) - \Gamma_s^{(n)}(v)\right) dB_s(v) + \sum_{i=n+1}^{\infty} \sum_{|v|=i} \int_0^t \Gamma_s(v)\, dB_s(v) \tag{1.4.3}$$

goes to zero in some sense as $n \to \infty$. We will show that the supremum over $[0,T]$ goes to zero almost surely. Our main tool will be the Burkholder-Davis-Gundy inequality, see [18, Ch. IV, Corollary 4.2] for details.

As the quadratic variation of the first summation in (1.4.3) is

$$Q_t := \sum_{i=1}^{n} \sum_{|v|=i} \int_0^t \left(\Gamma_s(v) - \Gamma_s^{(n)}(v)\right)^2 ds,$$

the BDG inequality gives us that for $h > 0$ there is a constant $C_h > 0$ such that

$$\mathbf{E}\left[\sup_{0\le t\le T}\left|\sum_{i=1}^{n}\sum_{|v|=i}\int_0^t\left(\Gamma_s(v)-\Gamma_s^{(n)}(v)\right)\,dB_s(v)\right|^h\right]\le C_h\mathbf{E}\left[Q_t^{h/2}\right].$$

Choose $h \le 2$ so that, by subadditivity and a supremum bound on the integral terms, the right hand side is bounded above by

$$C_hT^{h/2}\sum_{i=1}^{n}\sum_{|v|=i}\mathbf{E}\left[\sup_{0\le t\le T}|\Gamma_t(v)-\Gamma_t^{(n)}(v)|^h\right].$$

Now by choosing $h > 1$, Doob's maximal inequality gives that this is further bounded above by

$$C_h^*T^{h/2}\sum_{i=1}^{n}\sum_{|v|=i}\mathbf{E}\left[|\Gamma_T(v)-\Gamma_T^{(n)}(v)|^h\right].$$

By Lemma 1.2.9 the latter term goes to zero exponentially fast as $n \to \infty$, and then Borel-Cantelli completes the proof.

For the second summation of (1.4.3), the same argument with the BDG inequality yields that

$$\mathbf{E}\left[\sup_{0\le t\le T}\left|\sum_{i=n+1}^{\infty}\sum_{|v|=i}\int_0^t\Gamma_s(v)\,dB_s(v)\right|^h\right]\le C_h^*T^{h/2}\sum_{i=n+1}^{\infty}\sum_{|v|=i}\mathbf{E}\left[\Gamma_T(v)^h\right]. \tag{1.4.4}$$

From the proof of Lemma 1.3.5 we have that

$$\limsup_{n\to\infty}\frac{1}{n}\log\sum_{|v|=n}\mathbf{E}\left[\Gamma_T(v)^h\right]\le\lambda_\Gamma(h)+\log\mathbf{E}\left[W_T^h\right]<0$$

for h sufficiently close to 1, and hence by (1.4.4) and the Borel-Cantelli lemma the second summation of (1.4.3) term goes to zero almost surely. □

Using equation (1.2.3) this leads to the following formulas for the evolution at other vertices:

Corollary 1.4.2. *For $v \in \mathcal{T}$ the mass $\Gamma_t(v)$ evolves as*

$$\frac{d\Gamma_t(v)}{\Gamma_t(v)}=\sum_{i=1}^{n}dB_t(v_i)+\sum_{\substack{u\in\mathcal{T}(v)\\u\ne v}}\frac{\Gamma_t(u)}{\Gamma_t(v)}\,dB_t(u),$$

where $\varsigma = v_0, v_1, v_2, \ldots, v_n = v$ are the vertices from the root to v. In particular this gives that if u is not a descendant of v or vice-versa then

$$\frac{d\left\langle\Gamma_t(u),\Gamma_t(v)\right\rangle}{\Gamma_t(u)\Gamma_t(v)}=|u\wedge v|\,dt,$$

where $u \wedge v$ is the last common ancestor of the paths to u and v.

Proposition 1.4.1 says that the total mass evolves as a continuous time exponential martingale. Its logarithm accumulates quadratic variation at a rate given by the last expression of Proposition 1.4.1, and, as is well known, the time at which an exponential martingale hits zero is equivalent to the time at which the accumulated quadratic variation reaches infinity. This gives another interpretation of the lifetime of the Γ_t process:

Corollary 1.4.3. *The Γ_t process reaches the zero measure at exactly the time*

$$\sup\left\{t \geq 0 : \int_0^t \sum_{\substack{v\in\mathcal{T}\\ v\neq\varsigma}} \left(\frac{\Gamma_s(v)}{\Gamma_s(\varsigma)}\right)^2 ds < \infty\right\}.$$

Before this time, that the total mass process is an exponential martingale naturally suggests the Girsanov theory plays a role here. This leads to the following:

Corollary 1.4.4. *Let P be the measure under which the vertex processes $\{B_t(v)\}_{v\in\mathcal{T}}$ are independent Brownian motions. Assume that Γ is a probability measure. For any $T' < T$, let $\tilde{P}_{T'}$ be the probability measure whose Radon-Nikodym derivative with respect to P is $\Gamma_{T'}(\varsigma)$. Then under $\tilde{P}_{T'}$ the processes*

$$\left\{t \mapsto B_t(v) - \int_0^t \Gamma_s(v)\, ds, 0 \leq t \leq T'\right\}_{v\in T}$$

are independent Brownian motions on the tree vertices.

See [18] for background on the Girsanov theory. In Section 1.6 we describe an application of this result to the model of tree polymers.

1.5 Hölder Continuity

This section highlights an interesting application of the SDE results derived in the previous section. Once again we will assume that the weight processes are

$$W_t(v) = \exp\{B_t(v) - t/2\},$$

where $\{B_t(v)\}_{v\in\mathcal{T}}$ is a collection of independent Brownian motions with $B_0(v) = 0$. Recall from Section 1.4 that Γ_t is a well defined measure-valued process on the time interval $[0, -2\lambda'_\Gamma(1+))$. Using techniques from stochastic analysis we will show that this process Γ_t is α-Hölder in the Wasserstein metric for any $\alpha < 1/2$. This gives an interesting juxtaposition of discontinuity and continuity. On the one hand, the measures Γ_t and Γ_s are mutually singular for $t \neq s$, and hence are very discontinuous in the total variation distance. However at the same time, they satisfy a very strong continuity condition in the Wasserstein metric on probability measures on the tree.

Definition 1.5.1. *The Wasserstein distance between any two probability measures μ and ν on $\partial\mathcal{T}$ is defined as*

$$d_W(\mu, \nu) := \inf_{\rho\in\Lambda(\mu,\nu)} \int_{\partial\mathcal{T}\times\partial\mathcal{T}} d(\zeta, \eta)\, d\rho(\zeta, \eta),$$

where $\Lambda(\mu, \nu)$ is the collection of all couplings of the measures μ, ν. Recall that the distance function is $d(\zeta, \eta) = 2^{-|\zeta\wedge\eta|}$.

Note that this is a metric on probability measures on $\partial\mathcal{T}$. Our main result in this section applies to the normalized process $\tilde{\Gamma}_t := \Gamma_t/\Gamma_t(\varsigma)$.

Theorem 1.5.1. *Let $T < -2\lambda'_\Gamma(1+)$. Then for any $\alpha < 1/2$, the process $\tilde{\Gamma}_t$, $0 \le t \le T$, is α-Hölder continuous in the Wasserstein metric.*

The main step in the proof is to show Hölder continuity of each of the processes $\Gamma_t(v)$, for $v \in \mathcal{T}$, along with a bound on the Hölder constant.

Theorem 1.5.2. *Let $T < -2\lambda'_\Gamma(1+)$. Then for any $v \in \mathcal{T}$, the processes $\Gamma_t(v)$ are α-Hölder continuous on $[0,T]$ for any $\alpha < 1/2$. Moreover, there is a $\gamma < 1$ such that,*

$$\sup_{v\in\mathcal{T}} \sup_{0\le s<t\le T} \gamma^{-|v|} \frac{|\Gamma_t(v) - \Gamma_s(v)|}{|t-s|^\alpha} < \infty$$

almost surely.

We use the following version of the Kolmogorov-Chentsov Theorem, which gives a bound on the magnitude of the Hölder constant. For a statement of this theorem see [12, Theorem 2.2.8]. The statement on the control of the Hölder constant is implicit in their proof.

Theorem 1.5.3 (Kolmogorov-Chentsov Theorem)**.** *Let X_t be a continuous, stochastic process on $[0,T]$ such that for all $t, s \le T$,*

$$\mathbf{E}\,|X_t - X_s|^p < K_p|t-s|^{p/2}$$

for some $p > 2$ and some constant K_p. Then X_t is α-Hölder continuous for every $\alpha < 1/2 - 1/p$. Moreover,

$$\mathbf{P}\left(\sup_{0\le s<t\le T} \frac{|X_t - X_s|}{|t-s|^\alpha} > 1\right) \le K_p H_\alpha, \tag{1.5.1}$$

where H_α is a constant depending only on α and T.

To prove Theorem 1.5.2 we restrict the process to a sequence of stopping times, prove Hölder continuity of these stopped process, and then take a limit. We construct these stopped processes in the following lemma. Note that in this section, and this section only, the notation $\Gamma_t^{(N)}$ refers to the stopped version of the Γ_t process, not to the finite level cascade measure $\Gamma_t^{(n)}$ as in all other sections.

Lemma 1.5.4. *Let $T < -2\lambda'_\Gamma(1+)$. Then there is a $\gamma < 1$ and a sequence of measure-valued processes $\Gamma_t^{(N)}$ for $N \in \mathbb{N}$, such that*

1. *$\mathbf{P}(\Gamma_t^{(N)} \ne \Gamma_t$ for some $t \le T) \to 0$ as $N \to \infty$,*

2. *for every $v \in \mathcal{T}$, $\Gamma_t^{(N)}(v)$ is α-Hölder on $[0,T]$ for any $\alpha < 1/2$,*

3. *for $\alpha < 1/2$, we have with probability one that,*

$$\sup_{v\in\mathcal{T}} \sup_{0\le s<t\le T} \gamma^{-|v|} \frac{|\Gamma_t^{(N)}(v) - \Gamma_s^{(N)}(v)|}{|t-s|^\alpha} < \infty.$$

Remark 1.5.5. *Theorem 1.5.2 follows immediately from this lemma.*

The processes $\Gamma_t^{(N)}$ will be Γ_t stopped at an appropriate stopping time. We construct these stopping times in the following lemma.

Lemma 1.5.6. *For any $T < -2\lambda'_\Gamma(1+)$ there is a $\beta < 1$ and a sequence of stopping times τ_N with $\mathbf{P}(\tau_N < T) \to 0$ as $N \to \infty$ such that,*

$$\sup_{v\in\mathcal{T}} \sup_{0\le t\le T} \beta^{-|v|}\Gamma_{t\wedge\tau_N}(v) \le CN,$$

for some constant C depending on Γ and β.

Proof. Fix $T < -2\lambda'_\Gamma(1+)$. Now recalling Definition 1.3.1 and Remark 1.3.2, we take $h \in (1, h_T)$ and note that $\alpha_T(h) < 0$. We can therefore choose β such that $\alpha_T(h)/h < \log\beta < 0$; hence $\beta < 1$. Consider the continuous, increasing processes

$$A_t(v) := \beta^{-|v|} \sup_{0\le s\le t} \Gamma_s(v).$$

It follows from the continuity of $\Gamma_t(v)$ that $A_t(v)$ is bounded on $[0,T]$ for every $v \in \mathcal{T}$. Now define

$$A_t := \sup_{v\in\mathcal{T}} A_t(v). \tag{1.5.2}$$

It follows from our choice of β and the definition of $\alpha_T(h)$ that A_0 is non-random and finite. Clearly A_t is a non-decreasing process. Note that the statement of the lemma is equivalent to finding a sequence of stopping times τ_N such that $A_{t\wedge\tau_N} \le A_0 + N$ and with $\mathbf{P}(\tau_N < T) \to 0$ as $N \to \infty$. This will follow from the fact that A_t is continuous on $[0,T]$, which we now show. Using Markov's inequality as well as Doob's L^p inequality we get that

$$\mathbf{P}\left(A_T(v) \ge A_0 \text{ for some } |v| = n\right) \le A_0^{-h}\beta^{-nh} \sum_{|v|=n} \mathbf{E}\left[\Gamma_T(v)^h\right].$$

By Corollary 1.2.10 and the choice of β,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbf{P}\left(A_T(v) \ge A_0 \text{ for some } |v| = n\right) \le -h\log\beta + \alpha_T(h) < 0.$$

Therefore by Borel-Cantelli,

$$\mathbf{P}\left(\sup_{0\le t\le T} A_t(v) \ge A_0 \text{ for only finitely many } v \in \mathcal{T}\right) = 1.$$

Take $S = \{v \in \mathcal{T} : A_T(v) \ge A_0\}$ to be the (random) set of vertices from the tree at which this inequality fails; clearly S is finite. Moreover, since the processes $A_t(v)$ and A_t are non-decreasing in t, it follows that the supremum in (1.5.2) can only be achieved at one of the vertices of S, i.e.

$$A_t = \max_{v\in S} A_t(v).$$

Hence A_t is itself continuous on $[0,T]$, since it is the maximum of a finite number of continuous processes.

We then take

$$\tau_N := \inf\{t \in [0,T] \,:\, A_t > A_0 + N\}$$

to be our sequence of stopping times. The continuity of A_t implies that $A_{t\wedge\tau_n} \le A_0+N$ as well as the fact that A_t is almost surely bounded on $[0,T]$. The boundedness on $[0,T]$ also gives that $\mathbf{P}(\tau_N < T) \to 0$ as $N \to \infty$. ☐

Proving Lemma 1.5.4 now becomes an application of the Kolmogorov-Chentsov Theorem.

Proof of Lemma 1.5.4. First take $\beta < 1$ and τ_N as in Lemma 1.5.6 and define $\Gamma_t^{(N)} := \Gamma_{t\wedge\tau_N}$ to be the stopped version of the measure-valued process. From Lemma 1.5.6, part (i) of this lemma is immediate.

Next, recall that by Corollary 1.4.2

$$d\Gamma_t(v) = \sum_{i=1}^{|v|} \Gamma_t(v)\, dB_t(v_i) + \sum_{\substack{u\in\mathcal{T}(v)\\ u\neq v}} \Gamma_t(u)\, dB_t(u).$$

Now fix $v \in \mathcal{T}$, $p > 2$ and take any $0 \le s < t \le T$. We apply the Burkholder-Davis-Gundy inequality, use the bound on the process $\Gamma_t^{(N)}$ from Lemma 1.5.6, and the fact that it is a flow on $\mathcal{T}$ to get

$$\begin{aligned}
\mathbf{E}\left|\Gamma_t^{(N)}(v) - \Gamma_s^{(N)}(v)\right|^p &\le \mathbf{E}\left(\sum_{i=1}^{|v|}\int_{s\wedge\tau_n}^{t\wedge\tau_n} \Gamma_r(v)^2\, dr + \sum_{\substack{u\in\mathcal{T}(v)\\ u\neq v}} \int_{s\wedge\tau_N}^{t\wedge\tau_n} \Gamma_r(v)^2\, dr\right)^{p/2}\\
&\le \mathbf{E}\left(\sum_{i=1}^{|v|} CN\beta^{|v|}\int_{s\wedge\tau_n}^{t\wedge\tau_n} \Gamma_r(v)\, dr + \sum_{k=1}^{\infty}\sum_{\substack{u\in\mathcal{T}(v)\\ u=|v|+k}} N\beta^{|u|}\int_{s\wedge\tau_N}^{t\wedge\tau_n} \Gamma_r(u)\, dr\right)^{p/2}\\
&= \mathbf{E}\left(|v|CN\beta^{|v|}\int_{s\wedge\tau_n}^{t\wedge\tau_n} \Gamma_r(v)\, dr + \sum_{k=1}^{\infty} N\beta^{|v|+k}\int_{s\wedge\tau_N}^{t\wedge\tau_n} \Gamma_r(v)\, dr\right)^{p/2}.
\end{aligned}$$

Now again use the upper bound $\Gamma_t^{(N)}(v) \le CN\beta^{|v|}$ and the fact that $\beta < 1$ to get the desired Kolmogorov-Chentsov inequality,

$$\begin{aligned}
\mathbf{E}\left|\Gamma_t^{(N)}(v) - \Gamma_s^{(N)}(v)\right|^p &\le \mathbf{E}\left(C^2N^2\beta^{2|v|}(|v| + C_\beta)\int =_{s\wedge\tau_N}^{t\wedge\tau_N} \mathrm{d}s\right)^{p/2}\\
&\le C^pN^p\beta^{p|v|}(|v| + C_\beta)^{p/2}(t-s)^{p/2}, \qquad (1.5.3)
\end{aligned}$$

where C_β is a constant depending only on β. Since this inequality holds for every $p > 2$,we get that for every $v \in \mathcal{T}$, $\Gamma_t^{(N)}(v)$ is α-Hölder continuous for any $\alpha < 1/2$.

Finally, fix $\alpha < 1/2$ and take $\gamma \in (\beta, 1)$. The L^p bound (1.5.3), applied to the process $\gamma^{-|v|}\Gamma_t^{(N)}(v)$, along with the Kolmogorov-Chentsov bound (1.5.1) gives that

$$\mathbf{P}\left(\sup_{0\le s<t\le T} \gamma^{-|v|}\frac{|\Gamma_t^{(N)}(v) - \Gamma_s^{(N)}(v)|}{|t-s|^\alpha} > 1\right) < K_p^N(v)H_\alpha, \qquad (1.5.4)$$

where $K_p^N(v) = C^pN^p(\beta/\gamma)^{p|v|}(|v|+C_\beta)^{p/2}$. Notice that $K_p^N(v) \to 0$ as $p \to \infty$, at least for all $|v| > M$

where $M > 0$ depends on only β, γ, and N. Therefore, since (1.5.4) is true for every $p > 2$, it follows that for all $|v| > M$ we have

$$\mathbf{P}\left(\sup_{0\le s<t\le T} \gamma^{-|v|} \frac{|\Gamma_t^{(N)}(v) - \Gamma_s^{(N)}(v)|}{|t-s|^\alpha} > 1\right) = 0,$$

which implies part (iii) of the lemma. □

The last ingredient in the proof of Theorem 1.5.1 is the following upper bound on the Wasserstein distance on $M(\mathcal{T})$.

Lemma 1.5.7. *Let μ, $\nu \in M(\mathcal{T})$ be such that for every $v \in \mathcal{T}$ we have $\mu(v), \nu(v) > 0$. Then*

$$d_W(\mu,\nu) \le \sum_{k=1} 2^{-k+1} \sum_{|v|=k-1} \nu(v) \left| \frac{\nu(v_L)}{\nu(v)} - \frac{\mu(v_L)}{\mu(v)} \right|.$$

Proof. This follows from a particular, standard coupling ρ of the two measures μ and ν. Given a ray $\xi \in \partial\mathcal{T}$, we define the probability measure ν_ξ on $\partial\mathcal{T}$ via the following iterative formula:

$$\nu_\xi(\eta_k|\eta_{k-1}) := \begin{cases} \frac{\nu(\eta_k)}{\nu(\eta_{k-1})} & \text{if } \eta_{k-1} \ne \xi_{k-1} \\ p_k(\xi) & \text{if } \eta_k = \xi_k \\ 1 - p_k(\xi) & \text{if } \eta_{k-1} = \xi_{k-1} \text{ but } \eta_k \ne \xi_k \end{cases}$$

where

$$p_k(\xi) = \left(\frac{\nu(\xi_k)}{\nu(\xi_{k-1})} \frac{\mu(\xi_{k-1})}{\mu(\xi_k)} \right) \wedge 1.$$

We define the coupling $\mathrm{d}\rho(\xi,\eta) = \mathrm{d}\mu(\xi)\mathrm{d}\nu_\xi(\eta)$. In words the coupling is the following. We first sample a ray ξ from μ. Then conditioned on ξ we sample η inductively. If η agrees with ξ on the first $k-1$ steps of the path (i.e., $\eta_k = \xi_k$), then flip a $p_k(\xi)$ coin to decide if η will agree with ξ on the k step. Once η diverges from ξ, pick the rest of its path independently from ν.

It is a matter of simple calculation to show that this is a coupling. Since ν_ξ is clearly a probability measure on $\partial\mathcal{T}$, it follows that the first marginal is μ. To compute that the second marginal is ν is straightforward.

Let $A_k(\xi) = \{\eta \in \partial\mathcal{T} : \eta_{k-1} = \xi_{k-1},\, \eta_k \ne \xi_k\}$ be the event that η agrees with ξ exactly up to level k. Hence $d(\xi,\eta) = 2^{-k}$ for $\eta \in A_k(\xi)$. Furthermore

$$\begin{aligned} \nu_\xi(A_k(\xi)) &= \prod_{i=1}^{k-1} p_i(\xi) \cdot (1 - p_k(\xi)) \\ &\le \prod_{i=1}^{k-1} \frac{\nu(\xi_i)}{\nu(\xi_{i-1})} \frac{\mu(\xi_{i-1})}{\mu(\xi_i)} \cdot \left| 1 - \frac{\nu(\xi_k)}{\nu(\xi_{k-1})} \frac{\mu(\xi_{k-1})}{\mu(\xi_k)} \right| \\ &= \frac{\nu(\xi_{k-1})}{\mu(\xi_{k-1})} \left| 1 - \frac{\nu(\xi_k)}{\nu(\xi_{k-1})} \frac{\mu(\xi_{k-1})}{\mu(\xi_k)} \right| \\ &= \frac{\nu(\xi_{k-1})}{\mu(\xi_k)} \left| \frac{\nu(\xi_k)}{\nu(\xi_{k-1})} - \frac{\mu(\xi_k)}{\mu(\xi_{k-1})} \right|. \end{aligned}$$

The first equality follows from the definition of ν_ξ while the first inequality follows from the definition of $p_i(\xi)$. A calculation now gives that

$$\begin{aligned}
d_W(\mu,\nu) &\leq \int_{\partial\mathcal{T}\times\partial\mathcal{T}} d(\xi,\eta)\mathrm{d}\mu(\xi)\mathrm{d}\nu_\xi(\eta) \\
&= \sum_{k=1}^{\infty}\int_{\partial\mathcal{T}}\int_{A_k(\xi)} d(\xi,\eta)\mathrm{d}\mu(\xi)\mathrm{d}\nu_\xi(\eta) \\
&= \sum_{k=1}\int_{\partial\mathcal{T}} 2^{-k}\mathrm{d}\mu(\xi)\nu_\xi(A_k(\xi)) \\
&\leq \sum_{k=1} 2^{-k}\int_{\partial\mathcal{T}} \frac{\nu(\xi_{k-1})}{\mu(\xi_k)}\left|\frac{\nu(\xi_k)}{\nu(\xi_{k-1})}-\frac{\mu(\xi_k)}{\mu(\xi_{k-1})}\right|\mathrm{d}\mu(\xi) \\
&= \sum_{k=1} 2^{-k}\sum_{|v|=k}\nu(v_p)\left|\frac{\nu(v_k)}{\nu(v_p)}-\frac{\mu(v_k)}{\mu(v_p)}\right|.
\end{aligned}$$

Recall that v_p denotes the parent of v in $\mathcal{T}$. Finally, noting that $\left|\frac{\nu(v_L)}{\nu(v)}-\frac{\mu(v_L)}{\mu(v)}\right| = \left|\frac{\nu(v_R)}{\nu(v)}-\frac{\mu(v_R)}{\mu(v)}\right|$, gives that

$$\sum_{|v|=k}\nu(v_p)\left|\frac{\nu(v_k)}{\nu(v_p)}-\frac{\mu(v_k)}{\mu(v_p)}\right| = 2\sum_{|v|=k-1}\nu(v)\left|\frac{\nu(v_L)}{\nu(v)}-\frac{\mu(v_L)}{\mu(v)}\right|$$

which completes the proof. □

We are finally ready to prove Theorem 1.5.1. It follows from Theorem 1.5.2 and Lemma 1.5.7.

Proof of Theorem 1.5.1. Let $0 \leq s < t \leq T$ and fix $\alpha < 1/2$. Applying Lemma 1.5.7 to the measures $\tilde{\Gamma}_t$ and $\tilde{\Gamma}_s$ gives

$$\begin{aligned}
d_W(\tilde{\Gamma}_s,\tilde{\Gamma}_t) &\leq \sum_{k=1} 2^{-k+1}\sum_{|v|=k-1}\frac{\Gamma_s(v)}{\Gamma_s(\varsigma)}\left|\frac{\Gamma_s(v_L)}{\Gamma_s(v)}-\frac{\Gamma_t(v_L)}{\Gamma_t(v)}\right| \\
&= \frac{1}{\Gamma_s(\varsigma)}\sum_{k=1} 2^{-k+1}\sum_{|v|=k-1}\left|\frac{\Gamma_s(v_L)\Gamma_t(v_R)-\Gamma_s(v_R)\Gamma_t(v_L)}{\Gamma_t(v)}\right|,
\end{aligned}$$

where we have used the fact that $\Gamma_t(v) = \Gamma_t(v_L)+\Gamma_t(v_R)$ for every t. Now note that by Theorem 1.5.2, for every $v \in \mathcal{T}$, $|\Gamma_t(v)-\Gamma_s(v)| \leq C_\alpha\gamma^{|v|}|t-s|^\alpha$ for some $\gamma < 1$. Therefore, adding and subtracting $\Gamma_t(v_R)\Gamma_t(v_L)$ gives

$$|\Gamma_s(v_L)\Gamma_t(v_R)-\Gamma_s(v_R)\Gamma_t(v_L)| \leq C_\alpha\gamma^{|v|+1}\Gamma_t(v)|t-s|^\alpha.$$

This inequality along with the fact that $\Gamma_s(\varsigma)$ is bounded away from zero on $[0,T]$ leads to the α-Hölder inequality in the Wasserstein metric,

$$\begin{aligned}
d_W(\tilde{\Gamma}_s,\tilde{\Gamma}_t) &\leq \frac{1}{\Gamma_s(\varsigma)}\sum_{k=1} 2^{-k+1}\sum_{|v|=k-1} C_\alpha\gamma^{|v|+1}|t-s|^\alpha \\
&= C_\alpha^{'}|t-s|^\alpha.
\end{aligned}$$

□

This result is optimal in the sense that $\tilde{\Gamma}_t$ is not α-Hölder for any $\alpha > 1/2$ in the Wasserstein metric.

This upper bound on the Hölder exponent follows from general arguments for martingales, which we now briefly outline.

Theorem 1.5.8. *For any interval* $[a,b] \subset [0, -2\lambda'_\Gamma(1+))$ *and for any* $\alpha > 1/2$*,*

$$\limsup_{\epsilon \to 0} \sup_{\substack{a \le s \le t \le b \\ |t-s| \le \epsilon}} \frac{d_W(\tilde{\Gamma}_t, \tilde{\Gamma}_s)}{|t-s|^\alpha} = \infty.$$

Proof. Define

$$f(\xi) = \begin{cases} 1 \text{ if } \xi_1 = \varsigma_L \\ 0 \text{ if } \xi_1 = \varsigma_R. \end{cases}$$

Since $|f(\xi) - f(\eta)| \le d(\xi, \eta)$ for any two rays ξ, $\eta \in \partial\mathcal{T}$, we have, using Jensen's inequality, that for any coupling ρ of two probability measures, μ and ν,

$$\begin{aligned} \int_{\partial\mathcal{T}\times\partial\mathcal{T}} d(\xi,\eta)\mathrm{d}\rho(\xi,\eta) &\ge \int_{\partial\mathcal{T}\times\partial\mathcal{T}} |f(\xi) - f(\eta)|\, \mathrm{d}\rho(\xi,\eta) \\ &\ge \left| \int_{\partial\mathcal{T}\times\partial\mathcal{T}} f(\xi)\mathrm{d}\rho(\xi,\eta) - \int_{\partial\mathcal{T}\times\partial\mathcal{T}} f(\eta)\mathrm{d}\rho(\xi,\eta) \right| \\ &= |\mu(\varsigma_L) - \nu(\varsigma_L)|. \end{aligned}$$

In particular, this implies that

$$d_W(\tilde{\Gamma}_t, \tilde{\Gamma}_s) \ge \left| \tilde{\Gamma}_t(\varsigma_L) - \tilde{\Gamma}_s(\varsigma_L) \right|.$$

So it remains to show that $\tilde{\Gamma}_t(\varsigma_L)$ is not α-Hölder for any $\alpha > 1/2$. Both $\Gamma_t(\varsigma_L)$ and $\Gamma_t(\varsigma)$ are non-zero and continuous on $[a,b]$ and so by Ito's formula $\tilde{\Gamma}_t(\varsigma_L)$ is a continuous semi-martingale. We will use the fact that a continuous semi-martingale whose quadratic variation is strictly increasing is not α-Hölder for any $\alpha > 1/2$ (see Lemma 1.5.9). For now we only verify that the quadratic variation is strictly increasing. Since

$$\tilde{\Gamma}_t(\varsigma_L) = \frac{\Gamma_t(\varsigma_L)}{\Gamma_t(\varsigma_L) + \Gamma_t(\varsigma_R)},$$

by Ito's formula the martingale part of $d\tilde{\Gamma}_t$ is

$$\frac{1}{\Gamma_t(\varsigma)^2}\Big(\Gamma_t(\varsigma_R)d\Gamma_t(\varsigma_L) - \Gamma_t(\varsigma_L)d\Gamma_t(\varsigma_R)\Big).$$

From Proposition 1.4.1 it follows that $d\left\langle \tilde{\Gamma}_t(\varsigma_L) \right\rangle \neq 0$. ∎

For the sake of completeness we provide the following general fact from stochastic calculus that was used in the proof of Theorem 1.5.8.

Lemma 1.5.9. *Let* X_t *be a continuous semi-martingale, i.e.* $X = X_0 + M + A$ *where* M *is a continuous local martingale,* A *a finite variation process, and* $M_0 = A_0 = 0$*. If the quadratic variation* $\langle X \rangle_t$ *is strictly increasing on some interval* $[a,b]$*, then for any* $\alpha > 1/2$ *the process* X_t *is not* α*-Hölder continuous on* $[a,b]$*.*

Proof of Lemma 1.5.9. Note that for any $\beta < 1$ and $t \in (a, b)$,

$$\limsup_{s \to t} \frac{|A_t - A_s|}{|t-s|^\beta} = 0.$$

So without loss of generality we can assume that $A = 0$ and that X is a local martingale. First consider the case where there exists a non-random $\delta > 0$ such that $\langle X \rangle_b - \langle X \rangle_a > \delta$, with probability one. Fix $\alpha > 1/2$. We will show that X_t is not α-Hölder continuous. For $n \in \mathbb{N}$ and $1 \le i \le n$, define the stopping times

$$\tau_i^n = \inf_t \left\{ \langle X \rangle_t - \langle X \rangle_a > \frac{i}{n}\delta \right\}.$$

Note that $\tau_n^n < b$ and so by the Dubins-Schwarz Theorem,

$$\left(X_{\tau_{i+1}^n} - X_{\tau_i^n},\, i = 1, ..., n \right) \overset{d}{=} \left(B\left(\frac{i+1}{n}\delta \right) - B\left(\frac{i}{n}\delta \right),\, i = 1, ..., n \right),$$

where $B(t)$ is a standard Brownian motion. Since $\mathbf{E}\left|B\left(\frac{i+1}{n}\delta\right) - B\left(\frac{i}{n}\delta\right)\right|^{\frac{1}{\alpha}} = C_p \left(\frac{\delta}{2}\right)^{\frac{2}{\alpha}} n^{-2\alpha}$, the weak law of large numbers gives that

$$\sum_{i=1}^n \left| B\left(\frac{i+1}{n}\delta \right) - B\left(\frac{i}{n}\delta \right) \right|^{\frac{1}{\alpha}} \to \infty,$$

in probability. Let

$$A_n = \bigcup_{i=1}^n \left\{ |X_{\tau_{i+1}^n} - X_{\tau_i^n}| > \left(\tau_{i+1}^n - \tau_i^n \right)^\alpha \right\}$$

be the event that X_t is not α-Hölder at level n. By the pigeonhole principle,

$$\mathbf{P}\left(A_n\right) \ge \mathbf{P}\left(\sum_{i=1}^n |X_{\tau_{i+1}^n} - X_{\tau_i^n}|^{\frac{1}{\alpha}} > (b-a) \right).$$

The convergence of the right hand side to 1 gives that $\mathbf{P}\left(A_n \text{ i.o.}\right) = 1$. Finally, since $\langle X \rangle_t$ is strictly increasing we have that with probability one $\sup_{1 \le i \le n} (\tau_{i+1}^n - \tau_i^n) \to 0$ as $n \to \infty$. This finishes the proof of this case.

Now consider the general case. For every $\epsilon > 0$, we can find a $\delta > 0$ and non-random $a < T < b$ such that $\mathbf{P}(\langle X \rangle_T - \langle X \rangle_a > \delta) \ge 1 - \epsilon$. Consider the process

$$\tilde{M}_t = \begin{cases} X_t & t \le T \\ \begin{cases} X_t & t > T,\, \langle X \rangle_T - \langle X \rangle_a \ge \delta \\ X_T + B_{\frac{\delta}{b-T}(t-T)} & t > T,\, \langle X \rangle_T - \langle X \rangle_a < \delta \end{cases} \end{cases}$$

Then $\tilde{X}_t$ is a continuous martingale with $[\tilde{X}]_b - [\tilde{X}]_a > \delta$. Therefore $\tilde{X}_t$ is not α-Hölder for any $\alpha > 1/2$ and so with probability greater than $1 - \epsilon$ neither is X_t. Since this is true for any ϵ this completes the proof. □

1.6 Applications to Other Models

1.6.1 Tree Polymers

Although this work was written in the language of multiplicative cascades it was strongly motivated by the literature on tree polymers. The polymer model is virtually identical but the language is mildly different: to the vertices of the tree attach iid random variables $\{\omega(v)\}_{v\in\mathcal{T}}$, and at inverse temperature β and level n define the polymer measure on $\partial\mathcal{T}$ by

$$d\Gamma^{(n)}_{\omega,\beta}(\xi) := \frac{1}{Z^{(n)}_{\omega,\beta}} \prod_{i=1}^{n} \exp\{\beta\omega(\xi_i)\}\ d\Gamma(\xi).$$

Here $Z^{(n)}_{\omega,\beta}$ is the partition function

$$Z^{(n)}_{\omega,\beta} = \int_{\partial\mathcal{T}} d\Gamma^{(n)}_{\omega,\beta}(\xi) = \Gamma^{(n)}_{\omega,\beta}(\varsigma).$$

In the tree polymer model we usually assume that Γ is a probability measure, and hence the partition function normalizes the polymer measure to also have mass one. Typically only the Lebesgue measure θ is used as the base measure, but we will continue to describe the model in this greater generality where any Γ can be used. The only assumption on the ω is that $e^{\lambda(\beta)} := \mathbf{E}\left[e^{\beta\omega}\right] < \infty$ for all $\beta \in \mathbb{R}$. Clearly then the polymer measure can be expressed as a cascade measure with

$$d\Gamma^{(n)}_{\omega,\beta}(\xi) = \frac{e^{n\lambda(\beta)}}{Z^{(n)}_{\omega,\beta}}\, d\Gamma^{(n)}_{W_\beta}(\xi) = \frac{d\Gamma^{(n)}_{W_\beta}(\xi)}{\Gamma^{(n)}_{W_\beta}(\varsigma)},$$

with $W_\beta(v) = \exp\{\beta\omega(v) - \lambda(\beta)\}$. If Γ is W_β-regular then Section 1.2 shows that the limiting polymer measure exists and is given by

$$\lim_{n\to\infty} d\Gamma^{(n)}_{\omega,\beta}(\xi) = \frac{d\Gamma_{W_\beta}(\xi)}{\Gamma_{W_\beta}(\varsigma)}.$$

If Γ is not W_β-regular it is still an open problem as to whether or not a limit exists. Subsequential limits automatically exists because each finite level polymer measure is normalized to be a probability measure and the tree boundary $\partial\mathcal{T}$ is compact, but the structure of the set of subsequential limits is not known. See [22] for more on this problem.

Applying our cascade process to the study of polymer measures is most helpful whenever the family of cascading distributions $W_\beta = \exp\{\beta\omega - \lambda(\beta)\}$ can be represented by a process W_t satisfying Assumption 2. By this we mean that the processes W_β and W_t have the same marginal distributions at fixed times (up to a possible change of variables between β and t), but W_t has the independent increments property of Assumption 2. In this case, the cascade process of Section 1.3 gives us a coupling of the polymer measures at different temperatures that is different from the standard one obtained by simply multiplying the same variables by a different factor. The advantage of our coupling is that it has the Markov property implied by Section 1.3.3. In polymer language this Markov property has a nice interpretation: the polymer measure at a given temperature can be constructed by choosing a polymer at any higher temperature and then placing it in a new and independent environment. Most importantly, the higher temperature does *not* have to be infinite.

The simplest case of a weight process satisfying the above is the Gaussian weights of Section 1.4. The scaling properties of Brownian motion imply that in this case the t variable acts as both a time and an inverse temperature. This gives a nice interpretation to the stochastic calculus results of Proposition 1.4.1. The SDE for $\Gamma_t(\varsigma)$ tells us that the total mass at the root evolves according to a weighted measure of the Brownian noise being inputted, with the weights prescribed by the polymer measure at the time infinitesimally beforehand. The formula for the quadratic variation tells us that it evolves according to the *overlap* of the polymer measure, that is the expected amount of time that two polymers paths chosen independently under Γ_t^* will spend together before eventually splitting. The explosion time of the cascade process is exactly when the accumulated overlap reaches infinity.

The Girsanov theory is also useful in this context. The tree polymer model can be thought of as a model of random walk in a random environment, where the random variables ω act as the environment. For this part we assume that $\Gamma = \theta$, and under the measure $\theta^*_{W_\beta}$ the process $\xi_0, \xi_1, \xi_2, \ldots$, is Markov with transition probabilities given by

$$\theta^*_{W_\beta}\left(\xi_{i+1} = (\xi_i)_L | \xi_0, \xi_1, \ldots, \xi_i\right) = \frac{\theta_{W_\beta}((\xi_i)_L)}{\theta_{W_\beta}(\xi_i)}.$$

To study this type of RWRE one typically uses the “point of view of the particle”, which is the study of the environment Markov chain defined by

$$Z_n = \{\omega(u)\}_{u \in \mathcal{T}(\xi_n)}.$$

Note that Z_n takes values in the space of environments. It is straightforward to verify that if Q is a measure under which the ω are iid random variables and ξ is chosen according to the polymer measure $\theta^*_{W_\beta}$, then Z_n is a stationary Markov process with the same transition probabilities as the ξ_i Markov chain, i.e.

$$P\left(Z_{i+1} = \{\omega(u)\}_{u \in \mathcal{T}((\xi_i)_L)} | Z_0, \ldots, Z_i\right) = \theta^*_{W_\beta}\left(\xi_{i+1} = (\xi_i)_L | \xi_0, \xi_1, \ldots, \xi_i\right).$$

See [23] for more on the environment Markov chain. It begins in stationarity, with the stationary distribution being $\theta_{W_\beta}(\varsigma)\, dQ(\omega)$. The Girsanov theory of Corollary 1.4.4 gives a way to analyze this stationary distribution. Assume that under Q the ω are iid $N(0, T')$ for some $T' < 2\log 2$. Then under $\theta_\omega(\varsigma)\, dQ(\omega)$ the variables ω have the law of

$$\int_0^{T'} \frac{\theta_s(v)}{\theta_s(\varsigma)}\, ds + \tilde{B}_{T'}(v),$$

where the $\tilde{B}_t(v)$ are iid Brownian motions on the vertices of the tree. This gives an alternate description of the stationary measure for the environment Markov Chain.

1.6.2 One-Dimensional Random Geometry and KPZ

Multiplicative cascades have also been used as a toy model for studies of random geometry, most notably in [1]. There one considers the pushforward of Γ_W onto the interval $[0, 1]$ via binary expansion; left turns in ξ correspond to zeros in the binary expansion and right turns to ones. We use Γ_W to also denote the distribution function of the measure on $[0,1]$, i.e.

$$\Gamma_W(x) = \Gamma_W([0,x]).$$

If Γ_W is strictly positive, then $\Gamma_W(x)$ is a continuous, non-decreasing function on $[0,1]$. If $\Gamma_W(v) > 0$ for every $v \in \mathcal{T}$, then $x \mapsto \Gamma_W(x)$ is strictly increasing and hence a continuous bijection of $[0,1]$ onto $[0, \Gamma_W(1)]$. In the case $\Gamma = \theta$, Benjamini and Schramm used this map to establish a relation between the Hausdorff dimension of a set and its random image under θ_W. Specifically they show the following:

Theorem 1.6.1 ([1]). *Let W be a cascading distribution with $\mathbf{E}[W \log W] < \log 2$ (so that θ is W-regular), and assume that $\mathbf{E}[W^{-s}] < \infty$ for all $s \in [0,1)$. Let $K \subset [0,1]$ be some non-empty, deterministic set. Then there is the following* ***KPZ formula****:*

$$\dim_H K = \phi_W \left(\dim_H \theta_W(K)\right),$$

where $\theta_W(K)$ is the (random) image of K via the distribution function θ_W, and ϕ_W is the deterministic bijection from $[0,1]$ onto $[0,1]$ given by

$$\phi_W(h) = h - \log_2 \mathbf{E}\left[W^h\right].$$

Applying our process to this setup gives some interesting interpretations. Let θ_t and ϕ_t denote the corresponding cascade process and bijection when we replace W by dynamic weights W_t. As time evolves, the image set $\theta_t(K)$ moves about the line and its Hausdorff changes with it, yet the dimension evolves deterministically even though the set evolves randomly. Remark 1.3.1 and the formula above tell us that $\phi_t(h)$ is a decreasing function of t for each fixed h, and hence Hausdorff dimensions get smaller as time evolves. Using our process it is possible to understand the infinitesimal evolution of the dimension. Indeed write $d(t) = \dim_H \theta_t(K)$, and then the KPZ formula becomes

$$d(0) = \phi_t(d(t)).$$

Differentiating both sides with respect to t leads to an ODE for $d(t)$:

$$\dot{d}(t) = -\frac{\dot{\phi}_t(d(t))}{\phi_t'(d(t))}.$$

The particulars of this ODE depends on the type of weight process being used. For example in the case of Gaussian weights as in Section 1.4 it becomes

$$\dot{d} = -\frac{d(1-d)}{2\log 2 - t(2d-1)} =: \psi_t(d).$$

This ODE has many interesting aspects. First note that the $2\log 2$ appears because it is the lifetime of the θ_t process, that is the time at which it collapses to the zero measure. Further, by the presence of the t term in the denominator the ODE is non-autonomous, except at $d = 1/2$ where the non-autonomous term strangely disappears. It can also be shown that

$$\lim_{t \uparrow 2\log 2} = 1 - \sqrt{1 - d(0)},$$

so that even as θ_t approaches the zero measure the Hausdorff dimension of the random set stays bounded away from zero.

Although the work of Benjamini and Schramm can be used to derive the infinitesimal evolution of the Hausdorff dimension, in principle it should be possible to derive it separately and use it to give an alternate proof of their KPZ formula. All that needs to be found is a proof of the relation

$$\dim_H \theta_{t+\delta}(K) = \dim_H \theta_t(K) + \psi_t(\dim_H \theta_t(K))\delta + o(\delta)$$

that does not use the Benjamini and Schramm statement (although many of the techniques of their proof would probably be incorporated), and then the Markov property of the θ_t process turns this infinitesimal relation at a fixed time into the ODE that holds at all times. We have attempted to derive this relation but thus far been unable to, although we hope a proof will be at hand soon. In fact we believe that there is a slightly more general fact lurking in the background: namely that if Γ is an initial measure and W a cascading distribution that is a small perturbation away from the degenerate distribution at one, then

$$\dim_H \Gamma_W(K) = \dim_H \Gamma(K) + \psi_{\Gamma,W}(\dim_H \Gamma(K)).$$

Here $\psi_{\Gamma,W}$ would be a deterministic function determined by the properties of Γ and the size and type of the perturbation of W away from one. The infinitesimal relation is given by the "derivative" of ψ as the cascading distribution concentrates at one. It is not clear to us exactly how the properties of Γ enter into the picture, although we expect that they must in some form. It is also not clear if the relation above will be independent of the set K for all initial measures Γ, although we expect it will be for initial measures with some type of self-similarity.

Chapter 2

Eigenvectors of the One-Dimensional Anderson Tight Binding Model

2.1 Introduction

We consider the critical model of the one-dimensional discrete random Schrödinger operators given by the matrix

$$H_n = \begin{pmatrix} v_{1,n} & 1 & & & & \\ 1 & v_{2,n} & 1 & & & \\ & 1 & \ddots & \ddots & & \\ & & \ddots & \ddots & 1 & \\ & & & 1 & v_{n-1,n} & 1 \\ & & & & 1 & v_{n,n} \end{pmatrix} \tag{2.1.1}$$

where

$$v_{k,n} = \sigma\omega_k/\sqrt{n}. \tag{2.1.2}$$

Here ω_k are independent random variables with mean 0, variance 1 and bounded third absolute moment.

With the boundary condition $\psi(0) = 0 = \psi(n+1)$, we can write the eigenvalue equation $H_n\psi = \mu\psi$ as

$$\psi(\ell-1) + v_{\ell,n}\psi(\ell) + \psi(\ell+1) = \mu\psi(\ell), \quad \ell = 1, \dots, n.$$

Notice that this gives the recursion $\psi(\ell+1) = (\mu - v_{\ell,n})\psi(\ell) - \psi(\ell-1)$. The general idea is to write this as the multiplicative recursion,

$$\begin{pmatrix} \psi(\ell+1) \\ \psi(\ell) \end{pmatrix} = T(\mu - v_{\ell,n}) \begin{pmatrix} \psi(\ell) \\ \psi(\ell-1) \end{pmatrix} = M_n^{\mu}(\ell) \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix}, \tag{2.1.3}$$

where

$$T(x) := \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix} \text{ and } M_n(\mu,\ell) := T(\mu - v_{\ell,n})T(\mu - v_{\ell-1,n}) \cdots T(\mu - v_{1,n}).$$

This product of matrices completely characterizes the eigenvalues and eigenvectors of H_n. μ is an eigenvalue of H_n if and only if

$$M_n(\mu, n)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = c\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{2.1.4}$$

for some $c \in \mathbb{R}$ or, equivalently $(M_n(\mu, n))_{11} = 0$. Moreover, notice that the corresponding eigenvector ϕ^μ with choice of normalization $\phi^\mu(1) = 1$ is given by

$$\phi^\mu(\ell) = (M_n(\mu, \ell - 1))_{11}, \quad \ell = 1, \ldots, n. \tag{2.1.5}$$

It therefore suffices to understand the scaling limit of this multiplicative recursion. The limiting diffusion for this process was developed in [13] and used to give a characterization of the limit of the local eigenvalue point process. Building on this framework, we give a characterization of the limiting eigenvectors.

If there is no noise (i.e. $\sigma = 0$) then the eigenvalues μ_k and eigenvectors ψ_k of H_n are given by

$$\mu_k = 2\cos(\pi k/(n+1)),$$
$$\psi_k(\ell) = \sin(\pi k \ell/(n+1)).$$

The asymptotic density near $E \in (-2, 2)$ is given by the arcsin law, $\frac{\rho}{2\pi}$ with

$$\rho = \rho(E) = \frac{1}{\sqrt{1 - E^2/4}}\mathbf{1}_{|E|<2}. \tag{2.1.6}$$

The fact that the eigenvectors of H_n are delocalized was shown in [13]. The eigenvectors are highly oscillatory and so we focus on their induced L^2 measure which gives an approximation of the envelope of the eigenvector. For μ an eigenvalue of H_n and ψ^μ the corresponding normalized eigenvector $(\sum_{\ell=1}^n |\psi^\mu(\ell)|^2 = 1)$, we consider the measure on $[0, 1]$ whose density is $|\psi^\mu(\lfloor nt \rfloor)|^2\, dt$.

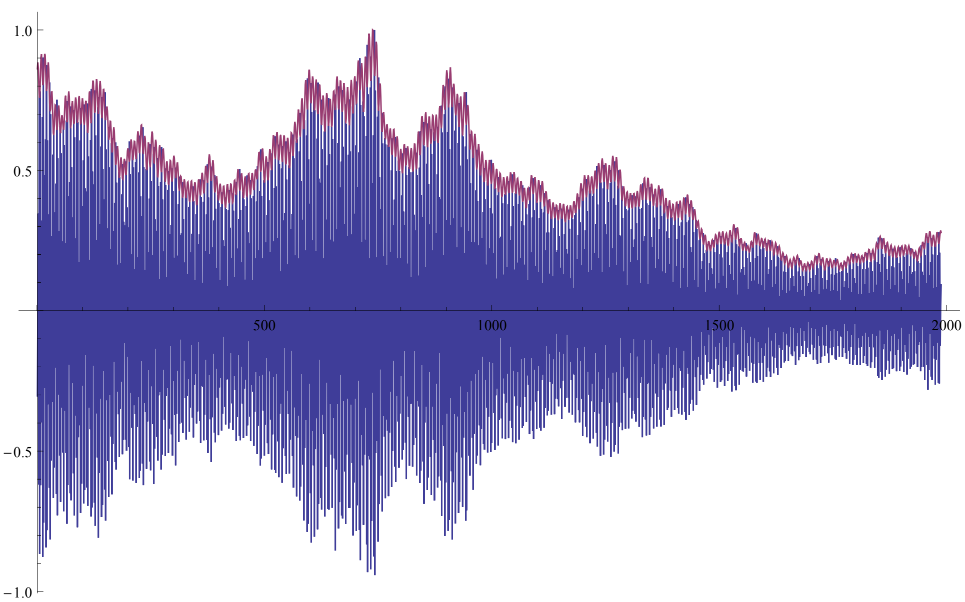

Figure 2.1: $n = 2000$, $v_\ell \sim N(0, 1)$. The graph in blue is that of a particular eigenvector, while the purple graph is the approximate density of the corresponding measure

We let $\mathcal{M}([0, 1])$ be the space of finite measures on $[0, 1]$ with the weak topology. By this we mean that $\mu_n \to \mu$ if $\int f d\mu_n \to \int f d\mu$ for every $f \in C_b([0,1], \mathbb{R})$.

Our main result is a statement about the joint convergence in law of the pairs

$$\left(\mu, |\psi^\mu(\lfloor nt \rfloor)|^2 \, dt\right) \in \mathbb{R} \times \mathcal{M}[0,1]$$

when we pick μ uniformly at random from the eigenvalues of H_n.

Theorem 2.1.1. *Let $\mathcal{B}$ be a standard two-sided Brownian motion started from 0 and take*

$$S(t) = \exp\left(\frac{\mathcal{B}_t}{\sqrt{2}} - \frac{|t|}{4}\right).$$

Pick μ uniformly from the eigenvalues of H_n and let ψ^μ be the corresponding normalized eigenvector. Then letting $\tau(E) = (\sigma\rho(E))^2$,

$$\left(\mu, n\,|\psi^\mu(\lfloor nt \rfloor)|^2 \, dt\right) \Rightarrow \left(E, \frac{S\big(\tau(t-u)\big)dt}{\int_0^1 ds\, S\left(\tau(s-u)\right)}\right).$$

Here E is distributed according to the arcsine law, u is uniform from $[0,1]$, and E, u and $\mathcal{B}$ are all independent.

The organization of this chapter is the following. In section 2.2 we explain the transfer matrix framework and give the main theorem of [13] along with our slight modification. In Section 2.3, we give a local version of Theorem 2.1.1. And finally in Section 2.4 we show how this local result gives the proof of the main theorem.

2.2 Transfer Matrix

[13] showed that the transfer matrix framework has a limiting evolution; it is this limiting object that enabled them to characterize the limiting eigenvalue process. Our main technical result is a slight strengthening of the convergence in that theorem. Our analysis will make use of this convergence and the correspondence between eigenvectors and transfer matrices. Recall the transfer matrix description of the spectral problem for H_n. We defined,

$$\begin{aligned} M_n(\mu, \ell) &:= T(\mu - v_{\ell,n})T(\mu - v_{\ell-1,n}) \cdots T(\mu - v_{1,n}), \\ T(x) &:= \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

Then μ is an eigenvalue of H_n if and only if

$$M_n(\mu, n) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = c \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{2.2.1}$$

for some $c \in \mathbb{R}$ or, equivalently $(M_n(\mu, n))_{11} = 0$. The corresponding normalized eigenvector ψ^μ is given by

$$\psi^{\mu}(\ell) = \frac{m_n^{\mu}(\ell-1)}{\sqrt{\sum_{k=1}^{n} |m_n^{\mu}(k)|^2}}, \quad \ell = 1, \ldots, n, \tag{2.2.2}$$

where we have written $m_n^{\mu}(\ell) = (M_n(\mu, \ell))_{11}$.

The main result of [13] is a limiting diffusion for M_n locally around any fixed $0 < |E| < 2$. In view of (2.1.6) we parametrize $\mu = E + \frac{\lambda}{\rho(E)n}$. We will use the notation $M_{n,E}(\lambda, \ell)$ to emphasize dependence on λ and E, and use the similar notation for other quantities. Sometimes we will drop E from our notation and when we do so we are implicity assuming that there is a fixed $0 < |E| < 2$ in the background. Setting

$$\epsilon_{\ell,n} = \frac{\lambda}{\rho n} - \frac{\sigma\omega_\ell}{\sqrt{n}}, \tag{2.2.3}$$

we have

$$M_{n,E}(\lambda, \ell) = T(E + \epsilon_{\ell,n})T(E + \epsilon_{\ell-1,n}) \cdots T(E + \epsilon_{1,n}) \text{ for } 0 \le \ell \le n. \tag{2.2.4}$$

As $T(E+\epsilon_{\ell,n})$ is a perturbation of $T(E)$, we follow the evolution in the coordinates that diagonalize $T(E)$. For $|E| < 2$, we can write $T(E) = ZDZ^{-1}$ with

$$D = \begin{pmatrix} \overline{z} & 0 \\ 0 & z \end{pmatrix}, \quad Z = \frac{i\rho(E)}{2} \begin{pmatrix} \overline{z} & z \\ 1 & 1 \end{pmatrix}, \quad z = E/2 + i\sqrt{1-(E/2)^2}. \tag{2.2.5}$$

From this we can see that for $|E| < 2$, $M_{n,E}(\lambda, \ell)$ is a perturbation of the rotation matrix D^{ℓ} and so we cannot hope for a limiting process. However, if we regularize the evolution by undoing the rotation and consider instead

$$Q_{n,E}(\lambda, \ell) = T^{-\ell}(E) M_{n,E}(\lambda, \ell), \tag{2.2.6}$$

then we have the the following scaling limit from [13].

Theorem 2.2.1. *Assume $0 < |E| < 2$. Let $\mathcal{B}(t), \mathcal{B}_2(t), \mathcal{B}_3(t)$ be independent standard Brownian motions in $\mathbb{R}$, $\mathcal{W}(t) = \frac{1}{\sqrt{2}}(\mathcal{B}_2(t) + i\mathcal{B}_3(t))$. Then the stochastic differential equation*

$$dQ(\lambda, t) = \frac{1}{2} Z \left(\begin{pmatrix} i\lambda & 0 \\ 0 & -i\lambda \end{pmatrix} dt + \begin{pmatrix} id\mathcal{B} & d\mathcal{W} \\ d\overline{\mathcal{W}} & -id\mathcal{B} \end{pmatrix} \right) Z^{-1} Q(\lambda, t), \qquad Q(\lambda, 0) = I \tag{2.2.7}$$

has a unique strong solution $Q(\lambda, t) : \lambda \in \mathbb{C}, t \ge 0$, which is analytic in λ.

Moreover, let $\tau = (\sigma\rho(E))^2$, then

$$\Big(Q_{n,E}\Big(\lambda, \lfloor nt/\tau \rfloor\Big), 0 \le t \le \tau\Big) \Rightarrow (Q(\lambda/\tau, t), 0 \le t \le \tau),$$

in the sense of finite dimensional distributions for λ and uniformly in t. Moreover, the random analytic functions $Q_{n,E}(\lambda, t)$ converge in law to $Q(\lambda/\tau, t)$ with respect to the local uniform topology on $\mathbb{C} \times [0, \tau]$.

Remark 2.2.1. *The main part of this theorem is proven in [13]. The work we have done here is to strengthen the tightness argument which allows us to get convergence in law with respect to the local* *uniform topology on* $\mathbb{C} \times [0,\tau]$*. The extra tightness argument along with how this implies the result is in Section 2.5.*

2.3 Local Limits of Eigenvalue-Eigenvector Pairs

In this section we prove a local version of Theorem 2.1.1. We will zoom in on the eigenvalue point process around a fixed $0 < |E| < 2$. Recall that the eigenvalue spacings near E are like $1/(n\rho(E))$ and so we consider the operator $n\rho(E)(H_n - E)$. Our local result is about the joint convergence of eigenvalue, eigenvectors pairs of this scaled operator. As with our global limit we consider the induced L^2 measure on $[0,\tau]$ coming from the eigenvector since it is otherwise too irregular to have a scaling limit. We think of these pairs as a point process on $X = \mathbb{R} \times \mathcal{M}[0,\tau]$,

$$\mathcal{P}_{n,E} = \Big\{ \Big(n\rho(E)(\mu - E) + \theta, \frac{n}{\tau} \left| \psi^\mu(\lfloor nt/\tau \rfloor) \right|^2 dt \Big) : \mu \text{ an eigenvalue of } H_n \Big\}.$$

With the usual product topology X is a complete, separable metric space. Let $\mathcal{M}(X)$ be the set of locally finite measures on X with the local weak topology. In other words, we say $\mu_n \in \mathcal{M}(X)$ converges to $\mu \in \mathcal{M}(X)$ if for every continuous function $f : X \to \mathbb{R}$ with compact support, $\int f d\mu_n \to \int \psi d\mu$. A random measure on $\mathcal{M}(X)$ is a measurable map $\omega \to \mu \in \mathcal{M}(X)$, with the Borel σ-algebra on $\mathcal{M}(X)$. By the point process $\mathcal{P}_{n,E}$ we mean the random measure in $\mathcal{M}(X)$ given by the sum of the delta masses corresponding to points in the set. And by convergence in law of a sequence of point processes on X we mean the usual notion of weak convergence of the corresponding random measures on $\mathcal{M}(X)$.

Theorem 2.3.1. *Fix* $0 < |E| < 2$ *and take* $\tau = \tau(E) = (\sigma\rho(E))^2$*. Let* θ *be uniform on* $[0, 2\pi]$*. Then, the point process on* $\mathbb{R} \times \mathcal{M}[0,\tau]$

$$\Big\{ \Big(n\rho(E)(\mu - E) + \theta, \frac{n}{\tau} \left| \psi^\mu(\lfloor nt/\tau \rfloor) \right|^2 dt \Big) : \mu \textit{ an eigenvalue of } H_n \Big\}$$

converges in law to a point process $\mathcal{P}_E$*.*
Moreover, for $t \in \mathbb{R}$*, let*

$$S(t) = \exp\left(\mathcal{Z}_t/\sqrt{2} - |t|/4 \right),$$

where $\mathcal{Z}$ *is a two sided Brownian motion started from* 0*.*
And define a measure μ_E *on* X *such that for every* $F \in C_b\left(\mathbb{R} \times \mathcal{M}[0,\tau]\right)$*,*

$$\int F(\lambda,\nu)\, d\mu_E(\lambda,\nu) = \frac{1}{2\pi} \int d\lambda \mathbf{E} F\left(\lambda, \frac{S(t-u)dt}{\int_0^\tau ds\, S(s-u)} \right),$$

with u *independent, uniform on* $[0,\tau]$*. Then the intensity measure of* $\mathcal{P}_E$ *is* μ_E*.*

Remark 2.3.2. *We note that [13] proved the convergence of the local eigenvalue point process and characterized the limit. Our result is an extension to the eigenvalue-eigenvector pairs.*

The proof of weak convergence proceeds in the usual steps. We first show subsequential convergence and then that the limit does not depend on the subsequence. We calculate the intensity measure in a separate lemma.

In order to characterize the limiting point process, we introduce two limiting random processes. Note that for $0 < |E| < 2$, for any $a, b \in \mathbb{R}^2$ we have

$$Z^{-1}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} i(a - bz) \\ \overline{i(a - bz)} \end{pmatrix}.$$

So Z^{-1} maps real vectors to vectors with conjugate entries. Since for $\lambda \in \mathbb{R}$ the transfer matrix $Q_{n,E}(\lambda, \ell)$ is real valued the process $Q(\lambda, t)$ will also be real valued. Therefore, we can write for $\lambda \in \mathbb{R}$,

$$\begin{pmatrix} iq^\lambda(t) \\ \overline{iq^\lambda(t)} \end{pmatrix} := Z^{-1}Q(\lambda, t)\begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.3.1}$$

for some complex numbers $q^\lambda(t)$ where $q^\lambda(0) = 1$ (the extra i in the above definition makes this and some upcoming formulas nicer). We will show that q^λ determines both the limiting eigenvalue point process and the limiting eigenvector shape. It will be useful to write $q = re^{i\theta}$ in its polar coordinates and so we make the following definition/lemma.

Lemma 2.3.3. *For $\lambda \in \mathbb{R}$, we let $\theta^\lambda(t) := 2\arg q^\lambda(t)$ uniquely defined as a continuous function and $r^\lambda(t) := \ln\left|q^\lambda(t)\right|^2$. Then r and θ uniquely satisfy the following stochastic differential equations,*

$$d\theta^\lambda(t) = \lambda dt + d\mathcal{B} + \operatorname{Im}\left[e^{-i\theta^\lambda(t)}d\mathcal{W}\right], \quad \theta^\lambda(0) = 0 \tag{2.3.2}$$

$$dr^\lambda(t) = \frac{dt}{4} + \operatorname{Re}\left[e^{-i\theta^\lambda(t)}d\mathcal{W}\right], \quad r^\lambda(0) = 0. \tag{2.3.3}$$

coupled together for all values of $\lambda \in \mathbb{R}$ where $\mathcal{B}$ and $\mathcal{W}$ are standard real and complex Brownian motions. Moreover $\theta^\lambda(t)$ is almost surely real analytic in λ and $\phi^\lambda(t) := \frac{\partial\theta^\lambda(t)}{\partial\lambda}$ satisfies the SDE

$$d\phi^\lambda(t) = dt - \operatorname{Re}(e^{-i\phi^\lambda(t)}d\mathcal{W})\phi^\lambda(t).$$

Our first step in proving Theorem 2.3.1 is to show convergence in law along subsequences.

Lemma 2.3.4. *Fix $0 < |E| < 2$. For $\lambda \in \mathbb{R}$, let $\boldsymbol{m}_n^\lambda$, $\boldsymbol{q}^\lambda$ be measures on $[0, \tau]$ with densities*

$$d\boldsymbol{m}_n^\lambda(t) = \left|\Big((2/\rho(E))M_{n,E}\left(\lambda, \lfloor nt/\tau\rfloor\right)\Big)_{11}\right|^2 dt,$$
$$d\boldsymbol{q}^\lambda(t) = \left|q^\lambda(t)\right|^2 dt.$$

Suppose that n_j is a subsequence along which $z(E)^{n_j} \to \tilde{z}$. Then the point process on X,

$$\left\{\left(\lambda, \boldsymbol{m}_n^\lambda\right) : \lambda \in \Lambda_{n_j,E}\right\},$$

converges in law to

$$\left\{\left(\lambda, 2\boldsymbol{q}^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^{2\tilde{\phi}}\right\}.$$

Here, for any $\phi \in \mathbb{R}$, we let

$$\mathsf{Sch}_\tau^\phi = \{\lambda \in \mathbb{R} : \theta(\lambda/\tau, \tau) \in 2\pi\mathbb{Z} + \phi\} \text{ and } \tilde{\phi} = \arg(z - \tilde{z}).$$

The next lemma shows that the distribution of the limit depends on the subsequence in a simple way.

Lemma 2.3.5. *Fix $\tau > 0$ and u uniform in $[0, 2\pi]$. Then for any $\phi \in \mathbb{R}$,*

$$\{\left(\lambda + u, \boldsymbol{q}^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^\phi\} =^d \left\{\left(\lambda, \boldsymbol{q}^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^u\right\}.$$

And finally we need the following lemma to help calculate the intensity measure of the limiting point process.

Lemma 2.3.6. *For every $G \in C_b\left(\mathbb{R} \times C[0,\tau]\right)$,*

$$\mathbf{E} \sum_{\lambda \in \mathsf{Sch}_\tau^*} G(\lambda, r^{\lambda/\tau}) = \frac{1}{2\pi} \int d\lambda \, \mathbf{E}\left[G\left(\lambda, \frac{\mathcal{B}}{\sqrt{2}} + \frac{f^u}{2}\right)\right],$$

with $\mathcal{B}$ a standard Brownian motion started at zero, u independent, uniform on $[0,\tau]$, and $f^u(t) = \frac{1}{2}(u - |u-t|)$.

The above three lemmas give the proof of Theorem 2.3.1.

Proof of Theorem 2.3.1. Lemma 2.3.4 gives that along a subsequence n_j such that z^{n_j} converges to $\tilde{z}$, we have that

$$\begin{aligned}\left\{(\lambda + u, \boldsymbol{m}_n^\lambda) : \lambda \in \Lambda_{n_j,E}\right\} &\Rightarrow \left\{\left(\lambda + u, 2\boldsymbol{q}^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^{\tilde{\phi}}\right\}\right\} \\ &=^d \left\{\left(\lambda, 2\boldsymbol{q}^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^*\right\}\end{aligned}$$

with the equality following by Lemma 2.3.5. Since from any subsequence we can extract a further subsequence n_j such that z^{n_j} converges, this gives that

$$\left\{(\lambda + u, \boldsymbol{m}_n^\lambda) : \lambda \in \Lambda_{n_j,E}\right\} \Rightarrow \left\{\left(\lambda, 2\boldsymbol{q}^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^*\right\}.$$

Now recall that for $\lambda \in \Lambda_{n,E}$, $\lambda = n\rho(E)(\mu - E)$ for μ an eigenvalue of H_n and the corresponding normalized eigenvector is

$$\psi^\mu(\ell) = \frac{\left(M_{n,E}(\lambda,\ell)\right)_{11}}{\sqrt{\sum_{k=1}^n \left|\left(M_{n,E}(\lambda,k)\right)_{11}\right|^2}}, \qquad \ell = 1, \ldots, n.$$

And so since $d\boldsymbol{m}_n^\lambda(t) = \left|\left((\rho/2)M_{n,E}(\lambda, (\lfloor nt/\tau \rfloor)\right)_{11}\right|^2 dt$,

$$\frac{n}{\tau}\left|\psi^\mu(\lfloor nt/\tau \rfloor)\right|^2 dt = \frac{d\boldsymbol{m}_n^\lambda(t)}{\boldsymbol{m}_n^\lambda[0,\tau]}.$$

Since the function from $\mathcal{M}[0,1]$ to itself given by $\mu \mapsto \mu/\mu[0,1]$ is continuous except at zero and the probability that $\boldsymbol{m}_n^\lambda \equiv 0$ is zero, this gives the convergence in law,

$$\left\{\left(n\rho(E)(\mu - E) + \theta, \frac{n}{\tau}\left|\psi^\mu(\lfloor nt/\tau \rfloor)\right|^2 dt\right) : \mu \text{ an eigenvalue of } H_n\right\} \Rightarrow \left\{\left(\lambda, \frac{\boldsymbol{q}^{\lambda/\tau}}{\boldsymbol{q}^{\lambda/\tau}([0,\tau])}\right) : \lambda \in \mathsf{Sch}_\tau^{\tilde{\phi}}\right\}$$

Now note that,

$$\frac{\exp(\mathcal{B}_t + \frac{1}{2}(u - |u-t|))}{\int_0^\tau ds \exp\left(\mathcal{B}_t + \frac{1}{2}(u - |u-t|)\right)} =^d \frac{\exp\left(\mathcal{Z}_{t-u} - |u-t|/2\right)}{\int \exp\left(\mathcal{Z}_{s-u} - |u-s|/2\right)},$$

as processes on $[0,\tau]$, where $\mathcal{B}$ is a standard Brownian motion while $\mathcal{Z}$ is a two sided Brownian motion started from zero. And so from Lemma 2.3.6, we have the intensity measure of the limiting point process. □

We now present the proofs of the three lemmas of this section.

Proof of Lemma 2.3.4. We are trying to show convergence in law of random point measures on $X = \mathbb{R} \times \mathcal{M}[0,\tau]$. In other words, we want to show that $\mu_{n_j} = \sum_{\lambda\in\Lambda_{n_j,E}} \delta(\lambda)\delta(\boldsymbol{m}^{\lambda}_{n_j})$ converges in law to $\mu = \sum_{\lambda\in\mathsf{Sch}^{\tilde\phi}_{\tau}} \delta(\lambda)\delta(\boldsymbol{q}^{\lambda/\tau})$ with respect to the local weak topology. By the general theory of point processes (see Proposition 11.1.VIII, [5]) it suffices to show that for any $h \in C_c(X,\mathbb{R})$, the real valued random variables $\int h d\mu_{n_j}$ converge in law to $\int h d\mu$.

First, for all $w \in \mathbb{C}$, we let

$$F_n(w,t) := \begin{pmatrix} F_n^1(w,t) \\ F_n^2(w,t) \end{pmatrix} := Z^{-1}Q_{n,E}(w,\lfloor nt/\tau \rfloor)\begin{pmatrix} 1 \\ 0 \end{pmatrix},$$
$$F(w,t) := \begin{pmatrix} F^1(w,t) \\ F^2(w,t) \end{pmatrix} := Z^{-1}Q(w,t)\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

By Lemma 2.2.1 we have that $Q_n(w,\lfloor nt/\tau \rfloor)$ converges in law with respect to the local uniform topology on $\mathbb{C}\times[0,\tau]$ (see Section 2.5) to $Q(w/\tau,t)$. Since Z is a deterministic transform, we also have that $F_n(w,t)$ converges in law to $F(w/\tau,t)$. We first show that μ_n is determined by F_n while μ is determined by F.

Recall that we defined

$$Q_{n,E}(w,\ell) = T^{-\ell}(E)M_{n,E}(w,\ell),$$

and so

$$\frac{2}{\rho(E)}\left(M_{n,E}(w,\lfloor nt/\tau\rfloor)\right)_{11} = \begin{pmatrix} 1 & 0 \end{pmatrix}\left(\frac{2}{\rho(E)}Z\right)D^{\lfloor nt/\tau\rfloor}Z^{-1}Q_{n,E}(w,\lfloor nt/\tau\rfloor)\begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.3.4}$$
$$= z^{\lfloor nt/\tau\rfloor-1}F_n^1(w,t) + \overline{z}^{\lfloor nt/\tau\rfloor-1}F_n^2(w,t),. \tag{2.3.5}$$

In other words $\boldsymbol{m}_n^{\lambda}$ is a function of F_n. Moreover, for $\lambda\in\mathbb{R}$, we have by equation (2.3.1) that $2\left|q^{\lambda}(t)\right|^2 = \left|F^1(\lambda,t)\right|^2 + \left|F^2(\lambda,t)\right|^2$ and so $\boldsymbol{q}^{\lambda}$ is a function of F.

Moreover, $\Lambda_{n,E} = \{w\in\mathbb{R} : m_n(w,\tau) = 0\}$, which again is determined by F_n. And in fact, $(2/\rho)m_{n_j}(w,\tau)$ converges in law to

$$\tilde m(w) := \lim_{n_j\to\infty} z^{n_j-1}F^1_{n_j}(w,t) + \overline{z}^{n_j-1}F_n^2(w,t)$$
$$:= \tilde z\overline{z}F^1(w/\tau,\tau) + \overline{\tilde z}zF^2(w/\tau,\tau).$$

And now notice that for $\lambda\in\mathbb{R}$, by equation (2.3.1)

$$\tilde m(\lambda,\tau) = 0 \iff \tilde z\overline{z}iq(\lambda/\tau,\tau) + \overline{\tilde z\overline{z}iq(\lambda/\tau,\tau)} = 0 \iff \arg q(\lambda/\tau,\tau) + \arg(\tilde z - z) + \frac{\pi}{2} = 0.$$

In other words $\mathsf{Sch}^{\phi}_{\tau}$ is the zero set of $\tilde m$, which is determined by F.

We have shown that $\int h d\mu_n$ is a measurable function of F_n while $\int h d\mu$ is a measurable function of F. Since F_n converges in law to F, the continuous mapping theorem (eg. [11], Theorem 3.27) allows us to remove the randomness from the problem. We may assume that F_n converges to F in the local uniform topology and simply show that this implies that $\int h d\mu_{n_j}$ converges to $\int h d\mu$. We may also assume that $h = h_1 \cdot h_2$, with $h_1 \in C_c(\mathbb{C})$ and $h_2 \in C(\mathcal{M}[0,\tau])$,

First notice that if $\lambda_n \to \lambda \in \mathbb{R}$, then as measures on $[0,\tau]$, $\boldsymbol{m}_n^{\lambda_n}$ converges weakly to $\boldsymbol{q}^{\lambda/\tau}$ (and so $h_2(\boldsymbol{m}_n^{\lambda_n})$ converges to $h_2(\boldsymbol{q}^{\lambda/\tau})$). Take $u \in C[0,\tau]$, then

$$\int u\, d\boldsymbol{m}_n^{\lambda_n} = \int_0^\tau u(t) \left| z^{\lfloor nt/\tau \rfloor} F_n^1(\lambda_n, t) + \bar{z}^{\lfloor nt/\tau \rfloor} F_n^2(\lambda_n, t) \right|^2 dt.$$

Expanding the absolute value, noting that $F_n(\lambda_n, t)$ converge uniformly on $[0,\tau]$ to $F(\lambda/\tau, t)$, and applying Lemma (2.7.1) gives that

$$\begin{aligned} \lim_n \int u\, d\boldsymbol{m}_n^{\lambda_n} &= \int_0^\tau u(t) \left(\left|F^1(\lambda/\tau, t)\right|^2 + \left|F^2(\lambda/\tau, t)\right|^2 \right) dt \\ &= \int_0^\tau u(t)\, d\boldsymbol{q}^{\lambda/\tau}(t). \end{aligned}$$

Moreover since F_n converges to F and z^{n_j} converges to $\tilde{z}$, the analytic functions on $\mathbb{C}$, $m_{n_j}(w,\tau)$ converge in the local uniform topology to $\tilde{m}(w/\tau)$. By Hurwitz's theorem this gives that the zeros of these functions converge pointwise. And the real valued zeros converge to real valued zeros. And so,

$$\lim_{n_j} \sum_{\lambda \in \mathbb{R}: m_{n_j}(\lambda,\tau)=0} h_1(\lambda) h_2(\boldsymbol{m}_{n_j}^\lambda) = \sum_{\lambda \in \mathbb{R}: \tilde{m}(\lambda/\tau)=0} h_1(\lambda) h_2(\boldsymbol{q}^{\lambda/\tau}),$$

which completes the proof. □

Proof of Lemma 2.3.5. Recall that $r^\lambda = \ln\left|q^\lambda\right|^2$. It therefore suffices to show that for u uniform on $[0, 2\pi]$,

$$\left\{ \left(\lambda + u, r^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^\phi \right\} =^d \left\{ \left(\lambda, r^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^u \right\}$$

We first show that for $u \in \mathbb{R}$ fixed,

$$\left\{ \left(\lambda + u, r^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^\phi \right\} =^d \left\{ \left(\lambda, r^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^{\phi+u} \right\}$$

Recall the SDEs from Lemma 2.3.3,

$$d\theta^\lambda = \lambda dt + d\mathcal{B} + \operatorname{Im}\left[e^{-i\theta^\lambda(t)} d\mathcal{W}\right], \quad \theta^\lambda(0) = 0 \tag{2.3.6}$$

$$dr^\lambda = \frac{dt}{4} + \operatorname{Re}\left[e^{-i\theta^\lambda(t)} d\mathcal{W}\right], \quad r^\lambda(0) = 0. \tag{2.3.7}$$

coupled together for all values of $\lambda \in \mathbb{R}$ where $\mathcal{B}$ and $\mathcal{W}$ are standard real and complex Brownian motions. We let $\tilde{\theta}^\lambda(t) := \theta^{\lambda - u/\tau}(t) + (u/\tau)t$ and $\tilde{r}^\lambda(t) := r^{\lambda - u/\tau}(t)$ and notice that $\tilde{\theta}^\lambda$ and $\tilde{r}^\lambda$ jointly solve equations (2.3.6) and (2.3.7).

And so, since $\theta^{(\lambda-u)/\tau}(\tau) = \tilde{\theta}^{\lambda/\tau}(\tau) - u$,

$$\begin{aligned}\mathsf{Sch}_\tau^\phi + u &= \{\lambda : \theta^{(\lambda-u)/\tau}(\tau) \in 2\pi\mathbb{Z} + \phi\} \\ &= \{\lambda : \tilde{\theta}^{\lambda/\tau}(\tau) - u \in 2\pi\mathbb{Z} + \phi\}.\end{aligned}$$

Therefore

$$\begin{aligned}\left\{\left(\lambda + u, r^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^\phi\right\} &= \left\{\left(\lambda, r^{(\lambda-u)/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^\phi + u\right\} \\ &= \left\{\left(\lambda, \tilde{r}^{\lambda/\tau}\right) : \tilde{\theta}^{\lambda/\tau}(\tau) \in 2\pi\mathbb{Z} + \phi + u\right\} \\ &=^d \left\{\left(\lambda, r^{\lambda/\tau}\right) : \lambda \in \mathsf{Sch}_\tau^{\phi+u}\right\}\end{aligned}$$

by the uniqueness of solutions. Now if u is uniform on $[0, 2\pi]$, then $u + \phi \mod 2\pi$ is still uniform on $[0, 2\pi]$ and so $\mathsf{Sch}_\tau^{\phi+u} =^d \mathsf{Sch}_\tau^u$ which finishes the proof. □

Proof of Lemma 2.3.6. Recall that $\mathsf{Sch}_\tau^* = \{\lambda : \theta^{\lambda/\tau}(\tau) \in 2\pi\mathbb{Z} + v\}$, where v is uniform on $[0, 2\pi]$. Integrate out v to get

$$\begin{aligned}\mathbf{E} \sum_{\lambda \in \mathsf{Sch}_\tau^*} G(\lambda, r^{\lambda/\tau}) &= \frac{1}{2\pi} \mathbf{E} \int_0^{2\pi} du \sum_{\lambda : \theta^{\lambda/\tau}(\tau) \in 2\pi\mathbb{Z} + u} G(\lambda, r^{\lambda/\tau}) \\ &= \frac{1}{2\pi} \mathbf{E} \int_{-\infty}^{\infty} du \sum_{\lambda : \theta^{\lambda/\tau}(\tau) = u} G(\lambda, r^{\lambda/\tau}).\end{aligned}$$

Now using Lemma 2.3.3 we have that $\theta^{\lambda/\tau}(\tau)$ is almost surely a real analytic function in λ while $r^{\lambda/\tau}$ is continuous in λ and so we can apply the co-area formula and then Fubini to get

$$\frac{1}{2\pi} \mathbf{E} \int_{-\infty}^{\infty} du \sum_{\lambda : \theta^{\lambda/\tau}(\tau) = u} G(\lambda, r^{\lambda/\tau}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda \mathbf{E} \left[G(\lambda, r^{\lambda/\tau}) \left| \frac{\partial \theta^{\lambda/\tau}(\tau)}{\partial \lambda} \right| \right] \tag{2.3.8}$$

From Lemma 2.3.3, we have that the evolution of $r^{\lambda/\tau}$ is given by

$$dr^{\lambda/\tau}(t) = \frac{dt}{4} + \mathrm{Re}(e^{-i\theta^{\lambda/\tau}(t)} d\mathcal{W}).$$

And moreover, $\phi^{\lambda/\tau}(t) = \frac{\partial \theta^{\lambda/\tau}(t)}{\partial \lambda}$ is well defined, with SDE

$$d\phi^{\lambda/\tau} = \frac{dt}{\tau} - \mathrm{Re}(e^{-i\theta^{\lambda/\tau}} d\mathcal{W}) \phi^{\lambda/\tau}$$

Now fix λ and notice that $e^{-i\theta/\tau^\lambda} d\mathcal{W} =^d d\mathcal{W}$ and so $r^{\lambda/\tau}$ and $\phi^{\lambda/\tau}$ do not depend on λ. We can therefore drop the superscript and jointly solve for r and ϕ to get

$$\begin{aligned}r_t &= \frac{t}{4} + \frac{\mathcal{B}_t}{\sqrt{2}} \\ \phi_t &= \frac{1}{\tau} \int_0^t du e^{(r_u - r_t)}.\end{aligned}$$

And so by Fubini,

$$\mathbf{E}\left[G(\lambda, r^{\lambda/\tau})\left|\frac{\partial\theta^{\lambda/\tau}(\tau)}{\partial\lambda}\right|\right] = \frac{1}{\tau}\int_0^\tau du\mathbf{E}\left[e^{(r_u - r_\tau)}G(\lambda, r)\right],$$

Fix $u \in [0,\tau]$ and for simplicity, consider the process $\tilde{r}_t = \mathcal{B}_t + t/2$. This is just the time change $t \to 2t$. We will calculate the distribution of the path $\tilde{r}$ on $[0,\tau]$ weighted by $\exp(\tilde{r}_u - \tilde{r}_\tau)$. In other words if we take $\mathcal{R}$ to be the law of $\tilde{r}$ on $C[0,\tau]$, we need to characterize the measure on $C[0,\tau]$ given by,

$$\exp(\omega_u - \omega_\tau)\mathrm{d}\mathcal{R}(\omega).$$

By standard Girsanov theory, if we take $\mathcal{P}$ to be the law of Brownian motion on $C[0,\tau]$, then $\mathrm{d}\mathcal{R}(\omega) = \exp\left(\frac{\omega_\tau}{2} - \frac{\tau}{8}\right) d\mathcal{P}(\omega)$ and so

$$\exp(\omega_u - \omega_\tau)d\mathcal{R}(\omega) = \exp\left(\omega_u - \frac{\omega_\tau}{2} - \frac{\tau}{8}\right)d\mathcal{P}(\omega). \tag{2.3.9}$$

Now if we let $x^u := x^u(\omega)$ be the Brownian path reflected at u, we have that the corresponding exponential martingale of $x^u/2$ at τ is

$$\exp\left(\frac{x^u_\tau}{2} - \frac{[x^u]_\tau}{8}\right) = \exp\left(\omega_u - \frac{\omega_\tau}{2} - \frac{\tau}{8}\right)$$

where $[x^u]_t$ is the quadratic variation of x^u at t. Therefore, by another application of Girsanov, if we let $f^u_t = [x^u/2, \omega]_t = \frac{1}{2}(u - |u-t|)$, then under the measure $\exp(\omega_u - \omega_\tau)d\mathcal{R}(\omega)$ on $C[0,\tau]$ a path ω is distributed like $\mathcal{B} + f^u$ where $\mathcal{B}$ is a standard Brownian motion. Undoing the time change and applying Brownian scaling gives that,

$$\mathbf{E}\left[e^{(r_u - r_\tau)}G(\lambda, r)\right] = \mathbf{E}\left[G\left(\lambda, \frac{\mathcal{B}}{\sqrt{2}} + \frac{f^u}{2}\right)\right],$$

which completes the proof.

□

Proof of Lemma 2.3.3. We let $X^\lambda(t) = Z^{-1}Q(\lambda, t)$. From equation (2.2.7) we have the following stochastic differential equation,

$$dX^\lambda = \frac{1}{2}\left(\begin{pmatrix} i\lambda & 0 \\ 0 & -i\lambda \end{pmatrix} dt + \begin{pmatrix} id\mathcal{B} & d\mathcal{W} \\ d\overline{\mathcal{W}} & -id\mathcal{B} \end{pmatrix}\right) X^\lambda, \qquad X^\lambda(0) = Z^{-1}.$$

This gives that

$$dX^\lambda_{11} = \frac{1}{2}\left(i\lambda X^\lambda_{11}\, dt + iX^\lambda_{11} d\mathcal{B} + X^\lambda_{21} d\mathcal{W}\right).$$

If $\lambda \in \mathbb{R}$, then $X^\lambda_{11} = \overline{X^\lambda_{21}}$ and moreover $iq^\lambda = X^\lambda_{11}$. We fix $\lambda \in \mathbb{R}$ and drop it from our notation to get

$$dq = \frac{i\lambda}{2} q\, dt + \frac{1}{2}\left(iqd\mathcal{B} - \overline{q}d\mathcal{W}\right) \quad q(0) = 1$$

Ito's formula then gives that

$$\begin{aligned} d\ln q &= \frac{dq}{q} - \frac{1}{2}\frac{(dq)^2}{q^2} \\ &= \frac{i\lambda}{2}dt + \frac{i}{2}d\mathcal{B} + \frac{1}{2}\frac{\overline{q}}{q}d\mathcal{W} + \frac{dt}{8} \end{aligned}$$

Since $r = 2\mathrm{Re}\ln q$ and $\theta = 2\mathrm{Im}\ln q$, this yields the following SDEs,

$$\begin{aligned} dr &= \mathrm{Re}\left(\frac{\overline{q}}{q}d\mathcal{W}\right) + \frac{dt}{4}, \\ d\theta &= \lambda dt + d\mathcal{B} + \mathrm{Im}\left(\frac{\overline{q}}{q}d\mathcal{W}\right). \end{aligned}$$

Noting that $\frac{\overline{q}}{q} = \exp(-i\theta)$ finishes the proof. □

2.4 Proof of Main Theorem

We are now in a position to prove the main theorem of this work. We will average the local result of Theorem 2.3.1 to get the more macroscopic version of the theorem. In order to do so we need to be able to control the number of eigenvalues in a microscopic interval (of size $1/(\rho n)$) around E. We need the following lemma whose proof is given in Section 2.6.

Lemma 2.4.1. *Fix $R > 0$ and let $\Delta_n(E) = \left(E - \frac{R}{n\rho(E)}, E + \frac{R}{n\rho(E)}\right)$. Furthermore, let $N_n(E) = |\Lambda_n \cap \Delta_n(E)|$ be the number of eigenvalues of H_n in $\Delta_n(E)$. Then for any $\epsilon > 0$,*

$$\sup_n \sup_{E\in(-2+\epsilon,2-\epsilon)} \mathbf{E}\left[N_n(E)\right]^{3/2} < \infty.$$

We now give the proof of Theorem 2.1.1.

Proof of Theorem 2.1.1. Take θ uniform on $[0, 2\pi]$ and let $\boldsymbol{\psi}_n^\mu \in \mathcal{M}[0,1]$ with density $\left|\psi^\mu\left(\lfloor nt\rfloor\right)\right|^2 dt$. Using Theorem 2.3.1 and the time change $t \to \tau t$, we have that for $0 < |E| < 2$, the point process

$$\mathcal{P}_{E,n} = \left\{\left(n\rho(E)(\mu - E) + \theta, n\boldsymbol{\psi}_n^\mu\right) : \mu \in \Lambda_n\right\}.$$

converges in law to a limiting point process $\mathcal{P}_\tau$.
In particular, if we fix $g_1 = (1-|x|)\mathbf{1}_{[|x|\le 1]}$, $g_2 \in C_b(\mathbb{R}\times\mathcal{M}[0,1])$ and let

$$G_n(E) := \sum_{\mu\in\Lambda_n} g_1\left(n\rho(E)(\mu - E)\right) g_2\left(\mu, \boldsymbol{\psi}_n^\mu\right).$$

Then for fixed $|E| < 2$, $G_n(E)$ converges in distribution to $G(E)$ and

$$\mathbf{E}G(E) = \frac{1}{2\pi}\mathbf{E}g_2\left(E, \frac{S(\tau(t-u))dt}{\int_0^1 ds\, S(\tau(s-u))}\right). \tag{2.4.1}$$

We now show that $\int \mathbf{E}G_n(E)d\rho(E)$ converges to $\int \mathbf{E}G(E)d\rho(E)$ from which the result will follow.

Fix $\epsilon > 0$. Since $\operatorname{supp} g_1 \subset [-1,1]$, we let

$$N_n(E) = \{\mu \in \Lambda_n : |\mu - E| \leq 1/(n\rho(E)\},$$

which gives that $G_n(E) \leq \|g_1\|_\infty \|g_2\|_\infty N_n(E)$. And so from Theorem 2.4.1,

$$\sup_n \sup_{0<|E|<2-\epsilon} \mathbf{E}\left[G_n(E)\right]^{3/2} < \infty.$$

Therefore $G_n(E)\mathbf{1}_{|E|<2-\epsilon}$ is uniformly integrable with respect to $\mathbf{P} \times d\rho$. And so since $G_n(E)$ converges in law to $G(E)$, we have that

$$\lim_{n\to\infty} \int d\rho(E)\mathbf{E}\Big[G_n(E)\mathbf{1}_{[|E|<2-\epsilon]}\Big] = \int d\rho(E)\mathbf{E}\Big[G(E)\mathbf{1}_{[|E|<2-\epsilon]}\Big]. \tag{2.4.2}$$

Now by Fubini,

$$\int d\rho(E)\mathbf{E}\Big[G_n(E)\mathbf{1}_{[|E|<2-\epsilon]}\Big] = \mathbf{E}\sum_{\mu\in\Lambda_n} g_2\left(\mu, \boldsymbol{\psi}_n^\mu\right) \int_{-2+\epsilon}^{2-\epsilon} d\rho(E) g_1\Big(n\rho(E)(\mu - E)\Big).$$

Fix $\delta > \epsilon$ and let $A_n(\delta) = \{\mu \in \Lambda_n : |\mu| < 2-\delta\}$, $B_n(\delta) = \{\mu \in \Lambda_n : |\mu| \geq 2-\delta\}$. We write

$$\int d\rho(E)\mathbf{E}\Big[G_n(E)\mathbf{1}_{[|E|<2-\epsilon]}\Big] = \mathbf{E}\left[\sum_{\mu\in A_n(\delta)} g(\mu)\right] + \mathbf{E}\left[\sum_{\mu\in B_n(\delta)} g(\mu)\right],$$

with

$$g(\mu) = g_2\left(\mu, \boldsymbol{\psi}_n^\mu\right) \int_{-2+\epsilon}^{2-\epsilon} d\rho(E) g_1\Big(n\rho(E)(\mu - E)\Big),$$

and deal with each piece separately.

First notice that for $k \in \mathbb{N}$, we can bound

$$\begin{aligned} |B_n(\delta)| &\leq \sum_{\mu\in B_n(\delta)} \left(\frac{\mu}{2-\delta}\right)^{2k} \\ &\leq (2-\delta)^{-2k} \sum_{\mu\in\Lambda_n} \mu^{2k}, \end{aligned}$$

We know that for fixed k,

$$\begin{aligned} \lim_{n\to\infty} \mathbf{E}\left[\frac{1}{n}\sum_{\mu\in\Lambda_n} \mu^{2k}\right] &= \frac{1}{2\pi} \int x^{2k}\rho(x)dx \\ &\leq C\frac{2^{2k}}{\sqrt{k}}. \end{aligned}$$

Taking $k = \lfloor 1/\delta \rfloor$, we have that $\big(1-(\delta/2)\big)^{-2k}$ is bounded independent of δ. And so,

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n}\mathbf{E}\,|B_n(\delta)| &\le C\frac{\big(1-(\delta/2)\big)^{-2k}}{\sqrt{k}} && (2.4.3)\\ &\le C\sqrt{\delta} && (2.4.4)\end{aligned}$$

Now use the second part of Lemma 2.7.2 to get that for $\mu \in B_n(\delta)$,

$$\int d\rho(E) g_1\Big(n\rho(E)(\mu - E)\Big) \le \frac{D}{n}.$$

And along with equation (2.4.4) this gives

$$\begin{aligned} \mathbf{E}\sum_{\mu\in B_n(\delta)} g(\mu) &\le \|g_2\|_\infty \frac{D}{n}\,|B_n(\delta)| \\ &= O(\sqrt{\delta}). \end{aligned}$$

Now for n large enough if $\mu \in A_n(\delta)$,

$$\begin{aligned} \int_{-2+\epsilon}^{2-\epsilon} g_1\Big(n\rho(x)(x-\mu)\Big)d\rho(x) &= \int_{-2}^{2} g_1\Big(n\rho(x)(x-\mu)\Big)d\rho(x) \\ &= \frac{1}{n}\int g_1(x)\,dx + o\,(1/n) \\ &= \frac{1}{n} + o\,(1/n) \end{aligned}$$

The first equality follows from the fact that for $x \in [-2,2]$, $\rho(x) \ge 1$. And so since $g_1 \in C_c(\mathbb{R})$, we have that $|x-\mu| \le D/n$ for some constant D. Since $\mu < 2-\delta$, we have that $|x| < 2-\epsilon$ for n large enough. The second equality follows from Lemma 2.7.2. And so

$$\begin{aligned} \mathbf{E}\sum_{\mu\in A_n(\delta)} g(\mu) &= \frac{1}{n}\sum_{\mu\in A_n(\delta)} \mathbf{E}g_2\big(\mu,\boldsymbol{\psi}^\mu\big) + o(1) \\ &= \frac{1}{n}\sum_{\mu\in\Lambda_n} \mathbf{E}g_2\big(\mu,\boldsymbol{\psi}^\mu\big) + O(\sqrt{\delta}) + o(1), \end{aligned}$$

with the last equality coming from equation (2.4.4). To sum up

$$\int d\rho(E)\mathbf{E}\Big[G_n(E)\mathbf{1}_{[|E|<2-\epsilon]}\Big] = \frac{1}{n}\sum_{\mu\in\Lambda_n} \mathbf{E}g_2(\mu,\boldsymbol{\psi}^\mu) + o(1) + O(\sqrt{\delta}). \quad (2.4.5)$$

On the other hand,

$$\int d\rho(E)\mathbf{E}\Big[G(E)\mathbf{1}_{[|E|<2-\epsilon]}\Big] = \int d\rho(E)\mathbf{E}\,[G(E)] + O(\epsilon).$$

And so by equation (2.4.1) along with equation (2.4.5) and the convergence from equation (2.4.2) we have that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{\mu\in\Lambda_n}\mathbf{E}\,g_2(\mu,\boldsymbol{\psi}^{\mu}) = \frac{1}{2\pi}\int d\rho(E)\mathbf{E}g_2\left(E,\frac{S(\tau(t-u))dt}{\int_0^1 ds\,S(\tau(s-u))}\right)+O(\epsilon)+O(\delta).$$

Since $\delta > \epsilon$ was arbitrary, this completes the proof. □

2.5 Tightness

In this section we discuss the underlying tightness bounds we need to prove the weak convergence in Lemma 2.2.1. We will use the following notions of convergence. Let $\mathcal{A}_d$ denote the space of continuous functions from $\mathbb{C}\times[0,1]$ to $\mathbb{C}^d$ that are also analytic in the first variable. In other words, if $f\in\mathcal{A}_d$, then for every $t\in[0,1]$, $f(\cdot,t)$ is an analytic function from $\mathbb{C}$ to $\mathbb{C}^d$. We equip $\mathcal{A}_d$ with the metric

$$d(f,g) := \sum_{r=1}^{\infty} 2^{-r}\frac{\|f-g\|_r}{1+\|f-g\|_r}, \qquad \|h\|_r := \max_{(x,z)\in D_r}\|h(z,x)\|,$$

where $D_r = B_r\times[0,1]$ and $B_r=\{w\in\mathbb{C}: |w|\le r\}$. Under this metric $\mathcal{A}_d\subset C\left([0,1]\times\mathbb{C},\mathbb{C}^d\right)$ is a complete, separable metric space.

A random function in $\mathcal{A}_d$ is a measurable mapping $\omega\to f\in\mathcal{A}_d$ from a probability space $(\Omega,\mathcal{F},P)$ to $(\mathcal{A}_d,\mathcal{B})$, where $\mathcal{B}$ is the Borel σ-field generated by the metric d. The law of f is the induced probability measure ρ_f on $(\mathcal{A}_d,\mathcal{B}_d)$. A sequence f_ℓ of random analytic functions is said to converge in law to a random $f\in\mathcal{A}_d$ if $\rho_{f_\ell}\to\rho_f$ in the usual sense of weak convergence.

Proposition 2.5.1. *Suppose f_ℓ is a sequence of random functions in $\mathcal{A}_d$ such that*
(1) *For every $w\in\mathbb{C}$, the processes $f_\ell(w,\cdot)\in C\left([0,1],\mathbb{C}^d\right)$ are tight,*
(2) *For every $r>0$,*

$$\lim_{M\to\infty}\sup_{\ell}\mathbf{P}\left(\|f_\ell\|_r > M\right)=0, \tag{2.5.1}$$

(3) *For each $m\ge 1$ and $(z,t)=((z_1,t_1),(z_2,t_2),\cdots,(z_m,t_m))\in(\mathbb{C}\times[0,1])^m$ there is a probability distribution $\nu_m^{(z,t)}$ on $(\mathbb{C}^d)^m$ and the random vector $(f_\ell(z_1,t_1),f_\ell(z_2,t_2),\cdots,f_\ell(z_m,t_m))\in\left(\mathbb{C}^d\right)^m$ converges in law to $\nu_m^{z,t}$.*
Then there is a random function f in $\mathcal{A}_d$ such that f_ℓ converges in law to f. Moreover for each $(z,t)=((z_1,t_1),(z_2,t_2),\cdots,(z_m,t_m))\in(\mathbb{C}\times[0,1])^m$, $(f(z_1,t_1),f(z_2,t_2),\cdots,f(z_m,t_m))\in\mathbb{C}^m$ has distribution $\nu_m^{(z,t)}$.

Proof. We first show that Assumptions (1) and (2) imply that the sequence f_ℓ is tight. We may assume that each $f_\ell\in\mathcal{A}_1$ since tightness in every coordinate function implies that the sequence is tight.

Fix $r > 0$, $|w|$, $|u| \le r$, and take $f \in \mathcal{A}_1$. Then, by Cauchy's integral formula,

$$\begin{aligned} f(w,t) - f(u,t) &= C_r \int_{|z|=2r} \left(\frac{f(z,t)}{w-z} - \frac{f(z,t)}{u-z} \right) dz \\ &= C_r \int_{|z|=2r} \frac{f(z,t)}{(w-z)(u-z)} (u-w)\, dz \end{aligned}$$

And so Jensen's inequality along with the fact that $|z-u|$, $|w-u| \ge r$ gives that, for every t,

$$|f(w,t) - f(u,t)| \le C_r \|f\|_{2r} |u-w| .$$

This inequality gives that for $|\zeta| \le r$,

$$|f(u,t) - f(w,s)| \le C_r \|f\|_{2r} (|u-\zeta| + |w-\zeta|) + |f(\zeta,t) - f(\zeta,s)| .$$

And so if we take any α-net $K_\alpha \subset B_r$ and take $\delta < \alpha/2$,

$$\sup_{\substack{\|(w,t)-(u,s)\|<\delta \\ |w|,|u|\le r}} |f(w,t) - f(u,s)| \le 2C_r \|f\|_{2r} \alpha + \max_{w \in K_\alpha} \sup_{|s-t|<\delta} |f(w,t) - f(w,s)| . \tag{2.5.2}$$

Now fix $\epsilon > 0$. Since $f_\ell(w,\cdot)$ is tight for $w \in \mathbb{C}$, for every $\gamma > 0$ we can find a $\delta_w > 0$ such that

$$\sup_{\ell \in \mathbb{N}} \mathbb{P} \left(\sup_{|s-t|<\delta} |f_\ell(w,t) - f_\ell(w,s)| > \epsilon \right) < \gamma.$$

In fact, just by adding probabilities, for any $\gamma, \alpha > 0$ we can find a finite α-net $K_\alpha \subset B_r$ and a $\delta_\alpha > 0$ such that,

$$\sup_{\ell \in \mathbb{N}} \mathbb{P} \left(\max_{w \in K_\alpha} \sup_{|s-t|<\delta_\alpha} |f_\ell(w,t) - f_\ell(w,s)| > \epsilon \right) < \gamma. \tag{2.5.3}$$

Now fix $\gamma > 0$. Assumption (2) means that we can find an M such that $\mathbb{P}(||f_\ell||_{2r} > M) < \gamma$. Take $\alpha < \epsilon(2MC_r)^{-1}$ and find a finite α-net K_α and a δ_α satisfying equation (2.5.3). Finally take $\delta = \min(\delta_\alpha, \alpha/2)$. Using equation (2.5.2), we get that,

$$\sup_{\ell \in \mathbb{N}} \mathbb{P} \left(\sup_{\substack{\|(w,t)-(u,s)\|<\delta \\ |w|,|u|\le r}} |f_\ell(w,t) - f_\ell(u,s)| \ge 2\epsilon \right) < 2\gamma. \tag{2.5.4}$$

Since ϵ and γ were arbitrary, this inequality along with Assumption (2) and Arzelà-Ascoli gives tightness of the sequence f_ℓ restricted to the discs D_r. And so by Prokohorov's theorem a subsequence of f_ℓ restricted to D_r converges in law. By a diagonal argument, there is a subsequence f_{ℓ_k} such that for each integer r, the restriction of f_{ℓ_k} to D_r converges to a random analytic function f_r on D_r. The distributions of the functions f_r are consistent with respect to restricting to smaller discs, and thus there is a random analytic function f on $\mathbb{C} \times [0,1]$ such that $f_{\ell_k} \to f$ in law with respect to the local uniform topology. Condition (2) is strong enough to ensure that f is unique and thus $f_\ell \to f$ in law.

□

Proof of Theorem 2.2.1. We intend to apply Lemma 2.5.1 to $Q_n(w,t) := Q_{n,E}(w, \lfloor nt/\tau \rfloor)$. We cannot apply this directly since for any $w \in \mathbb{C}$, the processes $Q_n(w,\cdot)$ are piecewise constant but not continuous. Instead, for all $w \in \mathbb{C}$ we let $\tilde{Q}_n(w,\cdot)$ be the linearized version of the process $Q_n(w,\cdot)$. By this we mean the function whose graph is given by the straight line between each consecutive jump discontinuity of $Q_n(w,\cdot)$. Since Q_n are analytic for any fixed t, $\tilde{Q}_n \in \mathcal{A}_4$. Theorem 1 of [13] gives the tightness bound (2) for $\tilde{Q}_n$. Theorem 2 of [13] and the continuous mapping theorem gives that for fixed $w \in \mathbb{C}$, $\tilde{Q}_n(w,\cdot)$ converge in law with respect to the uniform topology and so by by Prokhorov the tightness bound (1). This theorem also gives convergence of the finite dimensional distributions of Q_n and hence those of $\tilde{Q}_n$ which is condition (3). So by Lemma 2.5.1 Q_n converges in law to Q and since $d(Q_n, \tilde{Q}_n)$ goes to zero in probability we get that Q_n converges in law to Q with respect to the local uniform topology.

□

2.6 Local Eigenvalue Estimate

In this section we give the proof of Lemma 2.4.1. The moment bound on the number of eigenvalues in a macroscopic interval follows from an application of Theorem 2.2 of [14].

Theorem 2.6.1 ([14])**.** *Let $\mu < \mu'$ be consecutive eigenvalues of H_n. Then for any $E \in (\mu, \mu')$,*

$$\mu' - \mu \geq \left(\sum_{\ell=1}^{n} \|M_n(E,\ell)\|^2 \right)^{-1}. \tag{2.6.1}$$

Corollary 2.6.2. *For any interval $\Delta \subset \mathbb{R}$, let $N_n(\Delta) := |\Lambda_n \cap \Delta|$ be the number of eigenvalues of H_n in Δ. Then,*

$$N_n(\Delta) \leq 1 + |\Delta|^2 \int_\Delta dE \left(\sum_{\ell=1}^{n} \|M_n(E,\ell)\|^2 \right).$$

Proof. Fix $n \in \mathbb{N}$ and let $\tau(E) := \sum_{\ell=1}^{n} \|M_n(E,\ell)\|^2$. Take $\mu < \mu' \in \Delta$ consecutive eigenvalues of H_n. Integrating equation (2.6.1) gives

$$\begin{aligned}(\mu' - \mu) &\geq \frac{1}{\mu' - \mu} \int_{\mu}^{\mu'} \frac{dE}{\tau(E)} \\ &\geq \frac{1}{|\Delta|} \int_\Delta \frac{dE}{\tau(E)}.\end{aligned}$$

This gives a uniform lower bound on the distance between any two consecutive eigenvalues in Δ. And so by Jensen's inequality,

$$\begin{aligned}N_n(\Delta) &\leq 1 + \left(|\Delta| \Big/ \frac{1}{|\Delta|} \int_\Delta \frac{dE}{\tau(E)} \right) \\ &\leq 1 + |\Delta|^2 \int_\Delta dE \tau(E).\end{aligned}$$

□

To prove Theorem 2.4.1 via Corollary 2.6.2 we need a moment bound on the transfer matrices.

Lemma 2.6.3. *Let $\|\cdot\|$ be the Hilbert-Schmidt nnorm on $M_{2\times 2}(\mathbb{C})$. There is a continuous function f on $(-2,2)$ such for every $E \in (-2,2)$,*

$$\sup_n \max_{0\le \ell\le n} \mathbf{E}\,\|M_n(E,\ell)-I\|^3 < f(E).$$

Proof. Fix $E\in(-2,2)$ and $n\in\mathbb{N}$ and recall that for $0\le \ell\le n$,

$$M_n(E,\ell) = T(E-v_{\ell,n})T(E-v_{\ell-1,n})\cdots T(E-v_{1,n}),$$

with $T(x) := \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix}$ and $v_{\ell,n} = \frac{\sigma\omega_\ell}{\sqrt{n}}$.
We will prove a bound for the process $X_\ell = T^{-\ell}(E)M_n(E,\ell)$. Using the identity

$$T(y)T^{-1}(x) = I + \begin{pmatrix} 0 & y-x \\ 0 & 0 \end{pmatrix},$$

we have that

$$X_\ell = T^{-\ell}T(E-v_{\ell,n})T^{-1}T^{\ell}X_{\ell-1} \tag{2.6.2}$$

$$= (I - v_{\ell,n}\mathcal{E}_\ell)X_{\ell-1}, \tag{2.6.3}$$

where $\mathcal{E}_\ell = T^{-\ell}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}T^{\ell}(E)$.
We first show that

$$\|\mathcal{E}_\ell\| \le c_1(\rho(E))^2, \tag{2.6.4}$$

where c_1 does not depend on n or E and $\rho(E) = 1/\sqrt{1-(E/2)^2}$. Recall that we can write $T(E) = ZDZ^{-1}$ where

$$D = \begin{pmatrix} \overline{z} & 0 \\ 0 & z \end{pmatrix}, \quad Z = \frac{i\rho(E)}{2}\begin{pmatrix} \overline{z} & z \\ 1 & 1 \end{pmatrix}, \quad Z^{-1} = \begin{pmatrix} 1 & -z \\ -1 & \overline{z} \end{pmatrix}. \tag{2.6.5}$$

with $z = E/2 + i\sqrt{1-(E/2)^2}$.
Using the submultiplicativity of the Hilbert-Schmidt nnorm along with the fact that $|z|=1$ gives that for every $\ell\in\mathbb{Z}$,

$$\left\|T^{\ell}(E)\right\| \le 16\rho(E).$$

And since $\|\mathcal{E}_\ell\| \le \left\|T^{\ell}(E)\right\|\left\|T^{-\ell}(E)\right\|$, we get the bound (2.6.4).
Now notice that X_ℓ is a martingale with $X_0 = I$. We use the Burkholder-Davis-Gundy inequality along with Doob's Decomposition to get that for $0\le\ell\le n$,

$$\mathbf{E}\max_{k\le\ell}\|X_k - I\|^3 \le c_2\,\mathbf{E}\left(\sum_{k=1}^{\ell}\mathbf{E}\left[\|X_k - X_{k-1}\|^2\,|\mathcal{F}_{k-1}\right]\right)^{3/2},$$

Now use that $X_k - X_{k-1} = v_k \mathcal{E}_k X_{k-1}$, the bound on $\mathcal{E}_k$, and that $\mathbf{E} v_{\ell,n}^2 = \sigma^2/n$ to get that

$$\begin{aligned}\mathbf{E} \max_{k\le\ell} \|X_k - I\|^3 &= c_2\, \mathbf{E} \left(\frac{c_1 \sigma^2 \rho(E)^2}{n} \sum_{k=1}^{\ell} \|X_{k-1}\|^2 \right)^{3/2} \\ &\le c_3 \rho(E)^3 \frac{1}{n}\, \mathbf{E} \sum_{k=1}^{\ell} \|X_{k-1}\|^3 ,\end{aligned}$$

with the last inequality following from Jensen. Now using the inequality $\|A+B\|^p \le 2^p(\|A\|^p + \|B\|^p)$,

$$\mathbf{E} \max_{k\le\ell} \|X_k - I\|^3 \le \frac{c_3 \rho(E)^3}{n} \sum_{k=1}^{\ell} \left(\mathbf{E} \|X_{k-1} - I\|^3 + \|I\|^3 \right) \tag{2.6.6}$$

$$\le c_4 \rho(E)^3 \left(1 + \frac{S_{\ell-1}}{n} \right), \tag{2.6.7}$$

where we have set $S_\ell = \sum_{k=1}^{\ell} \mathbf{E} \|X_k - I\|^3$. This gives that

$$\begin{aligned} S_\ell - S_{\ell-1} &= \mathbf{E} \|X_\ell - I\|^3 \\ &\le c_4 \rho(E)^3 \left(1 + \frac{S_{\ell-1}}{n} \right), \end{aligned}$$

Finally, letting $R_\ell = 1 + S_\ell/n$, we have that $R_\ell \le R_{\ell-1}(1 + c_4 \rho(E)^3/n)$, and so $R_\ell \le \exp(c\rho(E)^3)$ for $1 \le \ell \le n$. Therefore, equation (2.6.7) gives that

$$\begin{aligned} \mathbf{E} \max_{0\le k\le n} \|X_k - I\|^3 &\le c_4 \rho(E)^3 R_{n-1} \\ &\le d_1 \rho(E)^3 \exp(d_2 \rho(E)^3), \end{aligned}$$

for some constants d_1 and d_2 that do not depend on E or n. Since $M_n(E,\ell) = T^{-\ell}(E) X_\ell$, this finishes the proof. ☐

Proof of Theorem 2.4.1. Using Corollary 2.6.2 we have that

$$|N_n(E) - 1|^{3/2} \le \max \left(\left[|\Delta_n(E)|^2 \int_{\Delta_n(E)} dx \sum_{\ell=1}^{n} \|M_n(E,\ell)\|^2 \right]^{3/2},\ 1 \right).$$

Since $|\Delta_n(E)| = 2R/(\rho(E)n$, we apply Jensen twice to get

$$\mathbf{E} \left[|\Delta_n(E)|^2 \int_{\Delta_n(E)} dx \sum_{\ell=1}^{n} \|M_n(x,\ell)\|^2 \right]^{3/2} \le \frac{g(E)}{n^3}\, \mathbf{E} \int_{\Delta_n(E)} dx \sum_{\ell=1}^{n} \|M_n(x,\ell)\|^3 .$$

Here g is continuous on $(-2,2)$. Now we use Fubini along with Lemma 2.6.3 to get that,

$$\mathbf{E}\int_{\Delta_n(E)} dx \sum_{\ell=1}^{n} \|M_n(x,\ell)\|^3 \leq \frac{R}{\rho(E)} sup_{x\in\Delta_n(E)} f(x).$$

Now fix $\epsilon > 0$ and $I_\epsilon = (-2+\epsilon, 2-\epsilon)$. There is an $N \in \mathbb{N}$ such that for any $n \geq N$ if $E \in I_\epsilon$, then $\Delta_n(E) \subset I_{\epsilon/2}$. Since f is continuous on $(-2,2)$ this means that for $n \geq N$,

$$\mathbf{E}\,|N_n(E)-1|^{3/2} \leq \max\left(\frac{C}{n^3}, 1\right)$$

□

2.7 Analytic Estimates

Lemma 2.7.1. *Let $D([0,1],\mathbb{C})$ be the space of cadlag functions from $[0,1]$ to $\mathbb{C}$. Suppose the sequence $f_n \in D([0,1],\mathbb{C})$ converges uniformly to $f \in C([0,1],\mathbb{C})$. Then for fixed $z \in \mathbb{C}$, $|z| = 1$ but $z \neq 1$,*

$$\lim_{n\to\infty} \int_0^1 f_n(t) z^{\lfloor nt \rfloor} dt = 0.$$

Proof. Since

$$\left| \int_0^1 f_n(t) z^{\lfloor nt \rfloor} dt - \int_0^1 f(t) z^{\lfloor nt \rfloor} dt \right| \leq \|f_n - f\|,$$

it suffices to show that for any continuous $f : [0,1] \to \mathbb{C}$,

$$\lim_{n\to\infty} \int_0^1 f(t) z^{\lfloor nt \rfloor} dt = 0.$$

We first assume that f is simple, by which we mean that $f := c\mathbf{1}_{[a,b)}$, for some constant c and subinterval $(a,b) \subset [0,1]$. We have that

$$\int_0^1 f(t) z^{\lfloor nt \rfloor} = \frac{c}{n} \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} z^k + o\left(\frac{1}{n}\right).$$

Since $z \neq 1$, $\sum_{k=0}^{N} z^k$ is bounded for all $N \in \mathbb{N}$, which finishes this case. Additivity then gives the result for any finite sum of piecewise, simple functions. And for a general $f \in C([0,1],\mathbb{C})$, we can find functions g_m which are finite sums of simple functions so that

$$\sup_n \left| \int_0^1 g_m(t) z^{\lfloor nt \rfloor} dt - \int_0^1 f(t) z^{\lfloor nt \rfloor} dt \right| \leq \int_0^1 |g_m(t) - f(t)|\, dt < \epsilon_m,$$

with $\epsilon_m \to 0$. This completes the proof. □

Lemma 2.7.2. *Let* $\rho(x) = 1/\sqrt{1-(x/2)^2}$. *Fix* $\epsilon > 0$ *and* $F \in C_c(\mathbb{R})$. *Then*

$$\sup_{|\mu|<2-\epsilon} \left| \int F\left(n\rho(x)(\mu - x)\right)\rho(x)dx - \int F(x)dx \right| = o\left(\frac{1}{n}\right)$$

Proof. Suppose that supp$F \subset [-R, R]$ for some $R > 0$. Then we can suppose $|\mu - x| \leq R/n$ because otherwise since $\rho \geq 1$ we have that $F(n\rho(x)(\mu-x)) = F(n\rho(\mu)(\mu-x)) = 0$. ρ is Lipschitz on any closed subset of $(-2, 2)$ and so for n large enough (depending only on ϵ) we have that

- $|\rho(\mu) - \rho(x)| \leq C/n$,
- $|n\rho(\mu)(\mu - x) - n\rho(x)(\mu - x)| \leq \frac{RC}{n}$.

This implies that

$$\begin{aligned} \int \left|F\left(n\rho(x)(\mu-x)\right)\rho(x)dx - F\left(n\rho(x)(\mu-x)\right)\rho(\mu)dx\right| &\leq \frac{C}{n}\int F\left(n\rho(x)(\mu-x)\right)dx \\ &\leq \frac{CR\left\|F\right\|}{n^2}. \end{aligned}$$

And also that,

$$\begin{aligned} \int \left|F\left(n\rho(x)(\mu-x)\right)\rho(\mu)dx - F\left(n\rho(\mu)(\mu-x)\right)\rho(\mu)dx\right| &\leq \rho(\mu) \sup_{|x-y|\leq CR/n} |F(x)-F(y)| \int \mathbf{1}[|\mu - x| < R/n]dx \\ &\leq \frac{D}{n} \sup_{|x-y|\leq CR/n} |F(x)-F(y)| \\ &= o\left(1/n\right) \end{aligned}$$

since F is uniformly continuous. These two inequalities imply

$$\sup_{|\mu|<2-\epsilon} \left| \int F\left(n\rho(x)(\mu-x)\right)\rho(x)dx - \int F\left(n\rho(\mu)(\mu-x)\right)\rho(\mu)dx \right| = o\left(1/n\right).$$

And we are done since $\int F\left(n\rho(\mu)(\mu-x)\right)\rho(\mu)dx = \int F(x)dx$. □

Lemma 2.7.3. *Let* $\rho(x) = 1/\sqrt{1-(x/2)^2}$ *and take* $F \in C_c(\mathbb{R})$ *with* $F \geq 0$ *and* $F(x) < F(y)$ *for* $|x| > |y|$. *Then,*

$$\sup_{|\mu|<2} \int_{-2}^{2} F\left(n\rho(x)(x-\mu)\right)\rho(x)dx \leq O\left(\frac{1}{n}\right).$$

Proof. By symmetry of $\rho(x)$, we can assume $\mu \geq 0$. Since $\rho(x) \geq 1$, we have that $|x - \mu| \leq R/n$, where supp$F \subset [-R, R]$. In particular, since $\mu \geq 0$, for n large enough, we have that x is bounded away from -2 independently of μ. And so we can write

$$\frac{c_1}{\sqrt{2-x}} \leq \rho(x) \leq \frac{c_2}{\sqrt{2-x}}.$$

The decreasing property of F gives that

$$\int_{-2}^{2} F\left(n\rho(x)(x-\mu)\right)\rho(x)dx \leq c_2 \int_{-2}^{2} F\left(c_1 n \frac{x-\mu}{\sqrt{2-x}}\right)\frac{dx}{\sqrt{2-x}}.$$

Writing $\gamma = 2 - \mu$ and changing variables $y = \sqrt{2-x}/\sqrt{\gamma}$,

$$\begin{aligned}\int_{-2}^{2} F\left(c_1 n \frac{x-\mu}{\sqrt{2-x}}\right) \frac{dx}{\sqrt{2-x}}. &= \sqrt{\gamma} \int_0^{2/\sqrt{\gamma}} F\left(c_1 n \sqrt{\gamma}\left(\frac{1-y^2}{y}\right)\right) dy \\ &\leq C \left\|F\right\| \sqrt{\gamma} \int_0^{\infty} \mathbf{1}\left[|y - 1/y| \leq R/(n\sqrt{\gamma})\right] dy.\end{aligned}$$

Now fix $\alpha > 0$. Notice that if $0 \leq x \leq 1$,

$$|x - 1/x| \leq 2\alpha \implies x \geq \sqrt{\alpha^2+1} - \alpha.$$

And so

$$\begin{aligned}\int_0^1 \mathbf{1}\left[|x - 1/x| \leq 2\alpha\right] &\leq 1 + \alpha - \sqrt{\alpha^2+1} \\ &\leq C\alpha.\end{aligned}$$

Similarly if $x \geq 1$, then

$$|x - 1/x| \leq 2\alpha \implies x \leq \alpha + \sqrt{\alpha^2+1}.$$

And so

$$\begin{aligned}\int_1^{\infty} \mathbf{1}\left[|x - 1/x| \leq 2\alpha\right] &\leq \alpha - 1 + \sqrt{\alpha^2+1} \\ &\leq C\alpha.\end{aligned}$$

Therefore

$$\begin{aligned}\sqrt{\gamma} \int_0^{\infty} \mathbf{1}\left[|x - 1/x| \leq R/(n\sqrt{\gamma})\right] dx &\leq C\sqrt{\gamma} \frac{R}{n\sqrt{\gamma}} \\ &= C/n.\end{aligned}$$

$\square$

Bibliography

[1] Itai Benjamini and Oded Schramm. KPZ in one dimensional random geometry of multiplicative cascades. *Comm. Math. Phys.*, 289(2):653–662, 2009.

[2] J. D. Biggins. Martingale convergence in the branching random walk. *J. Appl. Probability*, 14(1):25–37, 1977.

[3] Anton Bovier, Irina Kurkova, and Matthias Löwe. Fluctuations of the free energy in the REM and the p-spin SK models. *Ann. Probab.*, 30(2):605–651, 2002.

[4] F. Comets and J. Neveu. The Sherrington-Kirkpatrick model of spin glasses and stochastic calculus: the high temperature case. *Comm. Math. Phys.*, 166(3):549–564, 1995.

[5] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes Volume II: General Theory and Structure*. Springer, second edition, 2003.

[6] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[7] Ai Hua Fan. On Markov-Mandelbrot martingales. *J. Math. Pures Appl. (9)*, 81(10):967–982, 2002.

[8] Ai Hua Fan and J.P. Kahane. Decomposition principle in multiplicative chaos. *Preprint*, available at http://www.mathinfo.u-picardie.fr/fan/papers.html.

[9] Richard Holley and Edward C. Waymire. Multifractal dimensions and scaling exponents for strongly bounded random cascades. *Ann. Appl. Probab.*, 2(4):819–845, 1992.

[10] J.-P. Kahane and J. Peyrière. Sur certaines martingales de Benoit Mandelbrot. *Advances in Math.*, 22(2):131–145, 1976.

[11] Olav Kallenberg. *Foundations of modern probability*. Probability and its applications. Springer, New York, Berlin,, Paris, 2002. Sur la 4e de couv. : This new edition contains four new chapters as well as numerous improvements throughout the text.

[12] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[13] Eugene Kritchevski, Benedek Valkó, and Bálint Virág. The scaling limit of the critical one-dimensional random Schrödinger operator. *Communications in Mathematical Physics*, 314(3):775–806, 2012.

[14] Y. Last and B. Simon. Fine Structure of the Zeros of Orthogonal Polynomials, IV. A Priori Bounds and Clock Behavior. *ArXiv Mathematics e-prints*, June 2006.

[15] Quansheng Liu and Alain Rouault. Limit theorems for Mandelbrot's multiplicative cascades. *Ann. Appl. Probab.*, 10(1):218–239, 2000.

[16] David Márquez-Carreras, Carles Rovira, and Samy Tindel. A model of continuous time polymer on the lattice. *Commun. Stoch. Anal.*, 5(1):103–120, 2011.

[17] Mina Ossiander and Edward C. Waymire. Statistical estimation for multiplicative cascades. *Ann. Statist.*, 28(6):1533–1560, 2000.

[18] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[19] Samy Tindel. On the stochastic calculus method for spins systems. *Ann. Probab.*, 33(2):561–581, 2005.

[20] Bengt von Bahr and Carl-Gustav Esseen. Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist*, 36:299–303, 1965.

[21] E. C Waymire and S. C. Williams. Multiplicative cascades: dimension spectra and dependence. *J. Fourier Anal. Appl. Special Issue*, pages 589–609, 1995.

[22] Edward C. Waymire and Stanley C. Williams. T-martingales, size biasing, and tree polymer cascades. In *Recent developments in fractals and related fields*, Appl. Numer. Harmon. Anal., pages 353–380. Birkhäuser Boston Inc., Boston, MA, 2010.

[23] Ofer Zeitouni. Random walks in random environment. In *Lectures on probability theory and statistics*, volume 1837 of *Lecture Notes in Math.*, pages 189–312. Springer, Berlin, 2004.