Electron. J. Probab. **23** (2018), no. 70, 1–24.
ISSN: 1083-6489 https://doi.org/10.1214/18-EJP197

The phase transition in the ultrametric ensemble and local stability of Dyson Brownian motion

Per von Soosten* Simone Warzel†

Abstract

We study the ultrametric random matrix ensemble, whose independent entries have variances decaying exponentially in the metric induced by the tree topology on $\mathbb{N}$, and map out the entire localization regime in terms of eigenfunction localization and Poisson statistics. Our results complement existing works on complete delocalization and random matrix universality, thereby proving the existence of a phase transition in this model. In the simpler case of the Rosenzweig-Porter model, the analysis yields a complete characterization of the transition in the local statistics. The proofs are based on the flow of the resolvents of matrices with a random diagonal component under Dyson Brownian motion, for which we establish submicroscopic stability results for short times. These results go beyond norm-based continuity arguments for Dyson Brownian motion and complement the existing analysis after the local equilibration time.

Keywords: Dyson Brownian motion; localization transition; local statistics; ultrametric ensemble.
AMS MSC 2010: 15A52; 47B80.
Submitted to EJP on January 18, 2018, final version accepted on July 10, 2018.
Supersedes arXiv:1705.00923.

1 Introduction

One-dimensional lattice Hamiltonians with random long-range hopping provide a useful and simplified testing ground for the Anderson metal-insulator transition in more complicated systems. Two prominent examples of such models are the random band matrices [7, 19], whose entries H_{ij} are zero outside some band $|i-j| \leq W$, and the power-law random band matrices (PRBM) [17, 32], whose entries H_{ij} have variances decaying according to some power of the Euclidean distance $|i-j|$. Even for these models, the mathematically rigorous understanding is far from complete, although there has been some progress; see [36, 13, 5, 9, 15, 35] and references therein. This article is concerned with a further simplification, the ultrametric ensemble of Fyodorov, Ossipov

*Technische Universität München - Zentrum Mathematik. E-mail: vonsoost@ma.tum.edu
†Technische Universität München - Zentrum Mathematik. E-mail: warzel@ma.tum.edu

and Rodriguez [20], which is essentially obtained by replacing the Euclidean distance in the definition of the PRBM with the metric induced by the tree topology.

The index space of the ultrametric ensemble is $B_n = \{1, 2, \dots, 2^n\}$ endowed with the metric

$$d(x, y) = \min \{r \geq 0 \,|\, x \text{ and } y \text{ lie in a common member of } \mathcal{P}_r\},$$

where $\{\mathcal{P}_r\}$ is the nested sequence of partitions defined by

$$B_n = \{1, \dots, 2^r\} \cup \{2^r + 1, \dots, 2 \cdot 2^r\} \cup \cdots \cup \{2^{n-r-1}2^r + 1, \dots, 2^n\}.$$

The basic building blocks of the ultrametric ensemble are the matrices $\Phi_{n,r} : \ell^2(B_n) \to \ell^2(B_n)$ whose entries are independent (up to the symmetry constraint) centered real Gaussian random variables with variance

$$\mathbb{E}\, |\langle \delta_y, \Phi_{n,r} \delta_x \rangle|^2 = 2^{-r} \begin{cases} 2 & \text{if } d(x,y) = 0 \\ 1 & \text{if } 1 \leq d(x,y) \leq r \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

Here and throughout this paper, $\delta_x \in \ell^2$ denotes the canonical basis element defined by

$$\delta_x(u) = \begin{cases} 1 & \text{if } u = x \\ 0 & \text{if } u \neq x \end{cases}$$

and $\delta_{xy} = \langle \delta_y, \delta_x \rangle$. Thus $\Phi_{n,r}$ is a direct sum of 2^{n-r} random matrices drawn independently from the Gaussian Orthogonal Ensemble (GOE) of size 2^r. The ultrametric ensemble with parameter $c \in \mathbb{R}$ refers to the random matrix

$$H_n = \frac{1}{Z_{n,c}} \sum_{r=0}^{n} 2^{-\frac{(1+c)}{2} r} \Phi_{n,r} \tag{1.2}$$

where $\Phi_{n,r}$ and $\Phi_{n,s}$ are independent for $r \neq s$. We choose the normalizing constant $Z_{n,c}$ such that

$$\sum_{y \in B_n} \mathbb{E}\, |\langle \delta_y, H_n \delta_x \rangle|^2 = 1,$$

making the variance matrix Σ_n of H_n doubly stochastic. The original definition in [20] contains an additional parameter governing the relative strengths of the diagonal and off-diagonal disorder, but this parameter does not significantly alter our analysis and so we omit it altogether. Moreover, the authors of [20] constructed the block matrices $\Phi_{n,r}$ from the Gaussian Unitary Ensemble (GUE) and our results apply to both GOE and GUE blocks with only slight changes.

The ultrametric ensemble is a hierarchical analogue of the PRBM in a sense which was first introduced for the Ising model by Dyson [12] and studied rigorously in the context of quantum hopping systems with only diagonal disorder in [6, 24, 26, 27, 33, 38]. In particular, the definition (1.2) shows that the variance matrix Σ_n is a rescaled and shifted version of the hierarchical Laplacian, for which one can easily show that $\Sigma_n \geq 0$ by explicit diagonalization as in [25]. The normalizing factor $Z_{n,c}$ can also be calculated

$$\begin{aligned} Z_{n,c}^2 &= \sum_{y \in B_n} \mathbb{E} \left| \langle \delta_y, \left(\sum_{r=0}^{n} 2^{-\frac{(1+c)}{2} r} \Phi_{n,r} \right) \delta_x \rangle \right|^2 = \left(1 - 2^{-(1+c)(n+1)}\right) \frac{1 + \mathcal{O}(1)}{1 - 2^{-(1+c)}} \\ &= \begin{cases} \mathcal{O}(1) & \text{if } c > -1 \\ \mathcal{O}\left(2^{-(1+c)n}\right) & \text{if } c < -1 \end{cases} \end{aligned}$$

so that $Z_{n,c}$ grows exponentially in n in case $c < -1$ and $Z_{n,c}$ is asymptotically constant in case $c > -1$. Finally, the spread

$$M_n := \left(\max_{x,y \in B_n} \mathbb{E} \, |\langle \delta_y, H_n \delta_x \rangle|^2 \right)^{-1} = \begin{cases} Z_{n,c}^2 \, 2^{-o(n)} & \text{if } c \geq -2 \\ 2^{(1+o(1))n} & \text{if } c < -2 \end{cases} \tag{1.3}$$

also grows like a positive power of the system size 2^n when $c < -1$.

Because of the successful track record of hierarchical approximations in statistical physics, one might expect the core features of the PRBM phase transition to be present in the ultrametric ensemble as well. Indeed, the authors of [20] present arguments at a theoretical physics level of rigor as well as numerical evidence for a localization-delocalization transition in the eigenfunctions of H_n as the parameter changes from $c > 0$ to $c < 0$. In this paper, we pursue the point of view that the effect of the Gaussian perturbations $\Phi_{n,r}$ on the spectrum of H_n can be described dynamically by Dyson Brownian motion [11] and, in this sense, the critical point $c = 0$ is natural because it governs whether the evolution passes the local equilibration time of this system or not. In particular, we establish the localized phase by proving that the eigenfunctions remain localized and the level statistics converge to a Poisson point process if $c > 0$. In comparison to both the random band matrices and the PRBM, the hierarchical structure therefore significantly facilitates the analysis. We are not aware of any rigorous results on the localized phase of the PRBM, whereas for random band matrices the only localization result is due to Schenker [35] and no proof of Poisson statistics is known.

For the first result we recall the Wegner estimate [39], which asserts that the infinite-volume density of states measure defined by

$$\nu(f) = \lim_{n \to \infty} 2^{-n} \sum_{\lambda_j \in \sigma(H_n)} f(\lambda_j) \tag{1.4}$$

has a bounded Radon-Nikodym derivative whose values we denote by $\nu(E)$.

Theorem 1.1 (Poisson statistics)**.** *Suppose $c > 0$ and let $E \in \mathbb{R}$ be a Lebesgue point of ν. Then, the random measure*

$$\mu_n(f) = \sum_{\lambda \in \sigma(H_n)} f(2^n(\lambda - E))$$

converges in distribution to a Poisson point process with intensity $\nu(E)$ as $n \to \infty$.

The proof of Theorem 1.1 is contained in Section 5.

The second main result says that if an eigenfunction of H_n in some mesoscopic energy interval has any mass at some $x \in B_n$, then actually all but an exponentially small amount of the total mass is carried in an exponentially vanishing fraction of the volume near x with high probability. We make this precise in terms of the eigenfunction correlator

$$Q_n(x, y; W) = \sum_{\lambda \in \sigma(H_n) \cap W} |\psi_\lambda(x) \psi_\lambda(y)|,$$

which in completely delocalized regimes is typically given by $\sum_{y \neq x} Q_n(x, y; W) \approx 2^n |W|$. Thus, since $2^n |W|$ grows very large for mesoscopic spectral windows, it is a signature of localization if the eigenfunction correlator asymptotically vanishes for small enough mesoscopic intervals W, as is proved in the following theorem.

Theorem 1.2 (Eigenfunction localization)**.** *Suppose $c > 0$ and let $E \in \mathbb{R}$. Then, there exist $w, \mu, \kappa > 0$, $C < \infty$, and a sequence m_n with $n - m_n \to \infty$ such that for every $x \in B_n$ the ℓ^2-normalized eigenfunctions satisfy*

$$\mathbb{P} \left(\sum_{y \in B_n \setminus B_{m_n}(x)} Q_n(x, y; W) > 2^{-\mu n} \right) \leq C \, 2^{-\kappa n}$$

with

$$W = \left[E - 2^{-(1-w)n}, E + 2^{-(1-w)n}\right].$$

The proof of Theorem 1.2 in Section 6 makes use of a new relation (Theorem 6.1) between the eigenfunction correlator and the imaginary part of the Green function at complex energies.

The above results gain additional interest upon noting that, when $c < -1$, the ultrametric ensemble has an essential mean field character and techniques originally developed for Wigner matrices show that the energy levels agree asymptotically with those of the GOE and that the eigenfunctions are completely delocalized. We will now roughly sketch how to apply these results in the present situation and state the corresponding theorems. The results of [16] show that the semicircle law $\rho_{sc}(E) = \sqrt{(4-E^2)_+}/(2\pi)$ is valid on scales of order M_n^{-1} (c.f. (1.3)) also for the matrices

$$\widetilde{H}_n = \frac{1}{Z_{n,c}} \sum_{r=0}^{n-1} 2^{-\frac{(1+c)}{2} r} \Phi_{n,r} + \frac{1-\sqrt{T_n}}{Z_{n,c}} 2^{-\frac{(1+c)}{2} n} \Phi_{n,n}$$

with a small part of the final $\mathcal{O}(1)$ Gaussian component removed. We set $T_n = M_n^{-1+\delta}$ with $\delta \in (0,1)$. The validity of the local semicircle law already implies the complete delocalization of the eigenfunctions in mesoscopic windows in the bulk of the spectrum using the results of [13, Thm. 2.21]. The assumption of [13] that the spectrum of the variance matrix is well-separated from -1 follows simply from $\Sigma_n \geq 0$.

Theorem 1.3 (Eigenfunction delocalization; cf. [13, 16])**.** *Let $c < -1$. For any compact interval $I \subset (-2,2)$ there exist $\kappa, \epsilon > 0$ such that for all $E \in I$ the ℓ^2-normalized eigenfunctions of H_n in $[E - M_n^{-1}, E + M_n^{-1}]$ satisfy*

$$\|\psi_\lambda\|_\infty = \mathcal{O}(M_n^{-1/2+\epsilon})$$

with probability $1 - \mathcal{O}(N^{-\kappa})$.

Random matrix universality of the local statistics may be expressed by saying that the k-point correlation functions

$$\rho_{H_n}^{(k)}(\lambda_1, ..\lambda_k) = \int_{\mathbb{R}^{2^n-k}} \rho_{H_n}(\lambda_1, \ldots, \lambda_{2^n})\, d\lambda_{k+1} \ldots\, d\lambda_{2^n},$$

the k-th marginals of the symmetrized eigenvalue density ρ_{H_n}, locally agree with the corresponding objects for the GOE asymptotically. For this, we employ the work of Landon, Sosoe and Yau [28, Thm. 2.2] concerning the universality of Gaussian perturbations for

$$H_n = \widetilde{H}_n + \frac{\sqrt{T_n}}{Z_{n,c}} 2^{-\frac{(1+c)}{2} n} \Phi_{n,n}.$$

For the statement of the theorem, let

$$\begin{aligned}\Psi_{n,E}^{(k)}(\alpha_1, \ldots, \alpha_k) = \rho_{H_n}^{(k)} &\left(E + 2^{-n}\frac{\alpha_1}{\rho_{sc}(E)}, \ldots, E + 2^{-n}\frac{\alpha_k}{\rho_{sc}(E)}\right) \\ &- \rho_{GOE}^{(k)}\left(E + 2^{-n}\frac{\alpha_1}{\rho_{sc}(E)}, \ldots, E + 2^{-n}\frac{\alpha_k}{\rho_{sc}(E)}\right),\end{aligned}$$

where $\rho_{GOE}^{(k)}$ is the k-point correlation function of the $2^n \times 2^n$ GOE and ρ_{sc} is the density of the semicircle law.

Theorem 1.4 (WDM statistics; cf. [28, 16])**.** *Suppose $c < -1$, $E \in (-2,2)$ and $k \geq 1$. Then,*

$$\lim_{n\to\infty} \int_{\mathbb{R}^k} O(\alpha) \Psi_{n,E}^{(k)}(\alpha)\, d\alpha = 0$$

for every $O \in C_c^\infty(\mathbb{R}^k)$.

Summing up, these results rigorously prove the existence of a metal-insulator transition in the ensemble of ultrametric random matrices. In particular, our results allow an approach all the way to the critical point from the localized side $c > 0$. The previous arguments do not cover the regime $c \in [-1, 0)$. In this case, the eigenfunctions are still conjectured to be completely delocalized and the local eigenvalue statistics are conjectured to remain in the Wigner-Dyson-Mehta universality class as in the case $c < -1$ [20].

The Rosenzweig-Porter model is the $N \times N$ random matrix

$$H_t = V + \sqrt{t}\,\Phi_{n,n}$$

with $N = 2^n$, interpolation parameter $t = N^{-(1+c)}$, and independent random potential

$$V = \sum_x V(x)|\delta_x\rangle\langle\delta_x| \tag{1.5}$$

whose entries are drawn independently from some density $\varrho \in L^\infty$. It provides a standard interpolation between integrability and chaos and more recently has also been suggested as a toy model for many-body localization with three distinct phases [18, 23]. As a by-product of our analysis, we completely characterize the localized phase $c > 0$ of this model, which may be thought of as a hierarchical model with the "intermediate layers" removed.

Theorem 1.5 (Rosenzweig-Porter model)**.** *Suppose $t \leq N^{-(1+c)}$ with $c > 0$ and let $E \in \mathbb{R}$. Then:*

1. *As $N \to \infty$, the random measure defined by*

$$\mu_N(f) = \sum_{\lambda \in \sigma(H_t)} f(N(\lambda - E))$$

   converges in distribution to a Poisson point process with intensity $\varrho(E)$ provided E is a Lebesgue point of ϱ.
2. *There exist $w, \mu, \kappa > 0$ and $C < \infty$ such that for every $x \in \{1, \dots, N\}$ the ℓ^2-normalized eigenfunctions satisfy*

$$\mathbb{P}\left(\sum_{\lambda \in \sigma(H_t) \cap W} \sum_{y \neq x} |\psi_\lambda(x)\psi_\lambda(y)| > N^{-\mu} \right) \leq C N^{-\kappa}$$

   with

$$W = \left[E - N^{-(1-w)}, E + N^{-(1-w)} \right].$$

A proof of this result, which is similar to but simpler than those of Theorems 1.1 and 1.2, is included in Appendix A.

If $t = N^{-(1+c)}$ with $c < 0$ the prerequisites of [28, Thm. 2.2] may be verified for the Rosenzweig-Porter model by an exponential moment calculation similar to Cramér's theorem. This proves the emergence of Wigner-Dyson-Mehta statistics in the fixed-energy sense. The first point of Theorem 1.5 thus optimally complements these results and completes the mathematical understanding of the phase transition in the Rosenzweig-Porter model in terms of the local statistical behavior of the energy levels. The second point proves localization in the same sense as Theorem 1.2, and also yields an explicit relation between w, μ, κ and c, which shows that w, μ, κ may increase if c increases as well. If $c \leq -1$, the complete delocalization of the eigenfunctions was proved by Lee and Schnelli [30] as a corollary to a local law. In terms of the eigenvalue statistics, the non-ergodic delocalized phase established in [37, 3] can presumably only be detected by mesoscopic Poissonian fluctuations around the microscopic Wigner-Dyson-Mehta statistics (see also [10, 21]).

2 Local stability of Dyson Brownian motion

In the technical core of this paper, we consider a general $N \times N$ random matrix flow

$$H_t = T + V + \Phi_t, \tag{2.1}$$

where Φ_t is a Brownian motion in the space of symmetric matrices whose entries are given by

$$\langle \delta_y, \Phi_t \delta_x \rangle = \sqrt{\frac{1+\delta_{xy}}{N}} B_{xy}(t) \tag{2.2}$$

with independent standard Brownian motions $B_{xy}(t) = B_{yx}(t)$. The random potential V is given by (1.5) in terms of random variables $\{V(x)\}$ which are independent of the Brownian motions and whose conditional distributions are uniformly Lipschitz continuous, i.e.,

$$\mathbb{P}\left(V(x) \in I \,|\, \{V(y)\}_{y \neq x}\right) \leq C_V |I|, \tag{2.3}$$

for all Borel sets $I \subset \mathbb{R}$ and $x \in \{1, \ldots, N\}$ with a constant $C_V < \infty$ independent of N. Finally, T is some real symmetric $N \times N$ matrix, which may also be random provided T, V, and Φ_t remain independent. Dyson derived the equations

$$d\lambda_j(t) = \sqrt{\frac{2}{N}} dB_j(t) + \frac{1}{N} \sum_{i \neq j} \frac{dt}{\lambda_j(t) - \lambda_i(t)}$$

for the evolution of the eigenvalues $\lambda_1(t) \leq \cdots \leq \lambda_N(t)$ and conjectured $t = N^{-1}$ as the local equilibration time of this system. Thus, one expects that the local statistics of the eigenvalues agree asymptotically as $N \to \infty$ with those of the GOE if $t \gg N^{-1}$. This was first proved by Erdős, Schlein, and Yau [14] when H_0 is an independent Wigner matrix and recently for some very general, even deterministic, initial conditions by Landon, Sosoe, and Yau [28, 29]. In particular, the result [28], which establishes fixed-energy universality in the bulk of the spectrum for $t \gg N^{-1}$, covers the case $T = 0$ with asymptotically full probability. The eigenfunctions of H_t also follow a highly singular stochastic differential equation, which has been studied for Wigner initial conditions by Bourgade and Yau [4].

All of the aforementioned works rely on powerful rigidity estimates for the eigenvalues in the regime $t \gg N^{-1}$, which are not available for $t \ll N^{-1}$, where the spectral characteristics of $H_0 = T + V$ are expected to remain dominant. In this article, we study the stability of the spectral measures in the regime $t \ll N^{-1}$ by deriving a more tractable stochastic differential equation for the resolvent

$$R_t(z) = (H_t - z)^{-1},$$

which is significantly more amenable to analysis but still carries the relevant spectral information in its entries. In place of the rigidity estimates, the enabling tool for the analysis of the resolvent flow is the smoothing of spectral quantities with the external potential via spectral averaging (see Section 4 for details).

Our first result for H_t is about the normalized trace

$$S_t(z) = \frac{1}{N} \operatorname{Tr} R_t(z) = \int \frac{1}{\lambda - z} \, \nu_t(d\lambda),$$

which we have written as the Stieltjes transform of the empirical eigenvalue measure

$$\nu_t = \frac{1}{N} \sum_{\lambda \in \sigma(H_t)} \delta_\lambda.$$

Thus $S_t(z)$ contains detailed local information about the eigenvalues of H_t as z approaches the real axis. In particular, since the mean eigenvalue spacing of H_t is typically of order N^{-1}, knowledge of $S_t(z)$ with $\operatorname{Im} z \approx N^{-1}$ makes it possible to track individual eigenvalues near $\operatorname{Re} z$ along the flow (2.1). The following theorem shows that $S_t(z)$ remains stable even when $\operatorname{Im} z \ll N^{-1}$ provided that also $t \ll N^{-1}$.

Theorem 2.1. *For every $c > 0$, there exists $C < \infty$, depending only on c and C_V, such that*

$$\mathbb{E}\,|S_t\,(E+i\eta) - S_0\,(E+i\eta)| \le CN^{-c/2}\left(1 + \frac{1}{N\eta} + \frac{1}{(N\eta)^3}\right)$$

for all $t \le N^{-(1+c)}$ and $E \in \mathbb{R}$.

In essence, Theorem 2.1 asserts that for any spectral scale much larger than $t \ll N^{-1}$ the empirical eigenvalue measure is unaffected by the flow (2.1). This bound is much stronger than the one obtained from the crude norm estimate

$$\|R_t(z) - R_0(z)\| \le \|R_t(z)\Phi_t R_0(z)\| \approx \frac{\sqrt{t}}{(\operatorname{Im} z)^2} \tag{2.4}$$

and potential improvements on this theme obtained by running the Brownian motions $\langle \delta_y, \Phi_t \delta_x\rangle$ one after the other. The example $H_0 = 0$ shows that at least some regularity of the initial condition is needed for Theorem 2.1 to remain true. However, Theorem 2.1 can be proved for slightly more general H_0 under the weaker assumption that H_t satisfies the Wegner and Minami [31] estimates

$$\mathbb{E}\,\nu_t(I) \le C|I|, \qquad \mathbb{E}\,\nu_t(I)(\nu_t(I) - N^{-1}) \le C|I|^2$$

with a constant $C < \infty$ uniform in N and t. This is easily seen from the proof below. It is also possible to present Theorem 2.1 (and Theorem 2.2) as explicit bounds for arbitrary $t > 0$, but we artificially restrict to $t \le N^{-(1+c)}$ in order to keep the right hand side simple.

The properties of the eigenfunctions of H_t are encoded in the spectral measures

$$\mu_{xy} = \sum_{\lambda \in \sigma(H_t)} \psi_\lambda(x)\overline{\psi}_\lambda(y)\delta_\lambda$$

where $\{\psi_\lambda\}$ is an orthonormal basis of eigenfunctions of H_t and we have eased the notational burden by keeping the dependence of ψ_λ and μ_{xy} on t implicit. Hence, the Green functions

$$G_t(x,y;z) = \langle \delta_y, R_t(z)\delta_x\rangle = \int \frac{1}{\lambda - z}\,\mu_{xy}(d\lambda)$$

at scales $\operatorname{Im} z \approx N^{-1}$ describe the eigenfunctions of H_t locally near $\operatorname{Re} z$. The stability result analogous to Theorem 2.1 for $G_t(x,y;z)$ is contained in the following theorem.

Theorem 2.2. *For every $c > 0$, there exists $C < \infty$, depending only on c and C_V, such that*

$$\frac{1}{N}\sum_y \mathbb{E}\,|G_t\,(x,y;E+i\eta) - G_0\,(x,y;E+i\eta)| \le CN^{-c/2}\left(1 + \frac{1}{N\eta} + \frac{1}{(N\eta)^3}\right)$$

for all $t \le N^{-(1+c)}$, $E \in \mathbb{R}$, and $x \in \{1,\dots,N\}$.

The proofs of Theorems 2.1 and 2.2, which may be found in Section 4, are based on the fact that $R_t(z)$ satisfies the stochastic differential equation

$$dR_t(z) = \left(S_t(z)\frac{\partial}{\partial z}R_t(z) + \frac{1}{2N}\frac{\partial^2}{\partial z^2}R_t(z)\right)\,dt + d\tilde{M}_t, \tag{2.5}$$

where $\tilde{M}_t$ is a matrix-valued martingale whose entries can be given explicitly in terms of $R_t(z)$. The details of this equation, as well as its derivation, are contained in Section 4. The proofs rely on bounds controlling the regularized singularities of the terms in the right-hand side of (2.5) down to the scale $\eta = \mathcal{O}(N^{-1})$. In the same vein as remarked below (2.4), we stress that Theorems 2.1 and 2.2 cannot be obtained from the trivial estimate $\|\frac{\partial^n}{\partial z^n} R_t(z)\| \leq (\operatorname{Im} z)^{-(n+1)}$. Instead, we use the smoothing of singularities through the external random potential V as detailed in Section 3. Since we are interested in times scales $t \ll N^{-1}$, the potential is in fact the only possible source for this smoothing.

Let us conclude by noting that the analogue of (2.1) for perturbations drawn from the GUE,

$$\langle \delta_y, \tilde{\Phi}_t \delta_x \rangle = \sqrt{\frac{1}{N}} \begin{cases} \frac{1}{\sqrt{2}}(B_{xy}(t) + i\tilde{B}_{xy}(t)) & \text{if } x < y \\ B_{xx}(t) & \text{if } x = y \end{cases}$$

with $\tilde{B}_{xy}$ independent of B_{xy}, has also been widely studied. The analysis of this model is usually simpler because the additional symmetry enables explicit integration formulas (see [13] and references therein for a summary) and all the results and methods of this paper require only minor modifications to treat also the GUE flow.

3 Smoothing effects of the potential

Throughout this section we will let H be a general $N \times N$ random matrix of the form

$$H = T + V,$$

where V is a potential satisfying the assumption (2.3) and T is some Hermitian random matrix independent of V, which should be thought of as $T + \Phi_t$ from (2.1). Our goal is to use the smoothing effects of V on the spectral measures μ_{xy} of δ_x and δ_y for H and the empirical eigenvalue measure

$$\nu(f) = \frac{1}{N} \sum_{\lambda \in \sigma(H)} f(\lambda)$$

to control the resolvent flow (2.5). We start by recalling two staples of the theory of random Schrödinger operators, the spectral averaging principle [22] and the closely related Wegner estimate, whose proofs may be found in [1]. The former asserts that for any Borel set $I \subset \mathbb{R}$ and any $x \in \{1, \ldots, N\}$ we have

$$\mathbb{E}_x \left[\mu_x(I)\right] \leq C_V |I|, \tag{3.1}$$

where $\mathbb{E}_x$ denotes the conditional expectation with respect to the random variables $\{V(k) : k \neq x\}$. By averaging this bound over all $x \in \{1, \ldots, N\}$, we immediately obtain the latter result, namely that

$$\mathbb{E}\, \nu(I) \leq C_V |I| \tag{3.2}$$

for all Borel sets $I \subset \mathbb{R}$. The following lemma is a simple extension of these results based on the proof of Minami's estimate by Combes, Germinet, and Klein [8]. We write $|\mu|$ for the total variation measure of μ.

Lemma 3.1. *There exists $C < \infty$, depending only on C_V, such that*

1. $\mathbb{E}\, |\mu_{xy}|(I) \leq C|I|$ *and*
2. $\mathbb{E}\, [\nu(I)|\mu_{xy}|(J)] \leq C\left(|I| + \frac{2}{N}\right)|J|$

for all Borel sets $I, J \subset \mathbb{R}$ and $x, y \in \{1, \ldots, N\}$.

Proof. Notice that

$$|\mu_{xy}|(I) = \sum_{\lambda \in \sigma(H) \cap I} |\psi_\lambda(x)\psi_\lambda(y)|\,,$$

so the Cauchy-Schwarz inequality implies

$$|\mu_{xy}|(I) \le \sqrt{\mu_x(I)\mu_y(I)}.$$

Applying the Cauchy-Schwarz inequality to the expectation $\mathbb{E}_{xy}$ conditioned on $\{V(k) : k \neq x, y\}$ and using (3.1) then yield

$$\mathbb{E}_{xy}\,|\mu_{xy}|(I) \le \mathbb{E}_{xy}\,\sqrt{\mu_x(I)\mu_y(I)} \le \sqrt{\mathbb{E}_{xy}\,\mu_x(I)\mathbb{E}_{xy}\,\mu_y(I)} \le C|I|, \tag{3.3}$$

which implies the first assertion of the Lemma.

For the second claim, notice that for fixed values $\{V(k) : k \neq x, y\}$ of the potential away from x and y, the number of eigenvalues in I can change by at most two as $V(x)$ and $V(y)$ vary in $\mathbb{R}$. Hence

$$\begin{aligned}\mathbb{E}\,[\nu(I)|\mu_{xy}|(J)] &\le \mathbb{E}\,\left[\left(\nu(I) + \frac{2}{N}\right)\mathbb{E}_{xy}\,|\mu_{xy}|(J)\right] \\ &\le C|J|\mathbb{E}\,\left[\nu(I) + \frac{2}{N}\right] \le C\left(|I| + \frac{2}{N}\right)|J|,\end{aligned}$$

by (3.2) and (3.3). □

Intuitively, Lemma 3.1 asserts that the joint measure $\mathbb{E}\,[\nu \times |\mu_{xy}|]$ is continuous down to scales of order N^{-1}, which clearly has consequences for the integrals of test functions in terms of their variations on scales of order N^{-1}. The next results are a quantitative manifestation of this idea for the Stieltjes transforms

$$G(x, y; z) = \int \frac{1}{\lambda - z}\,\mu_{xy}(d\lambda)$$

and

$$S(z) = \int \frac{1}{\lambda - z}\,\nu(d\lambda),$$

which occur naturally in our study of the resolvent flow. In particular, the following theorem gives bounds for the drift.

Theorem 3.2. *There exists $C < \infty$, depending only on C_V, such that*

$$\mathbb{E}\left|\frac{1}{2N}\frac{\partial^2}{\partial z^2}G(x, y; z)\right| \le \frac{C}{N(\operatorname{Im} z)^2}$$

and

$$\mathbb{E}\left|S(z)\frac{\partial}{\partial z}G(x, y; z)\right| \le CN\left(\log N + \frac{1}{N\operatorname{Im} z}\right)\left(1 + \frac{1}{(N\operatorname{Im} z)^2}\right) + \frac{C}{\operatorname{Im} z}$$

for all $x, y \in \{1, \dots, N\}$ and $z \in \mathbb{C}_+$.

Proof. The first point of Lemma 3.1 implies that

$$\begin{aligned}\mathbb{E}\left|\frac{1}{2N}\frac{\partial^2}{\partial z^2}G(x, y; z)\right| &\le \frac{1}{N\operatorname{Im} z}\mathbb{E}\int \frac{1}{|\lambda - z|^2}\,|\mu_{x,y}|(d\lambda) \\ &\le \frac{C}{N\operatorname{Im} z}\int \frac{1}{|\lambda - z|^2}\,d\lambda \\ &\le \frac{C}{N(\operatorname{Im} z)^2},\end{aligned} \tag{3.4}$$

which is the first assertion of the theorem.

Next, let us introduce

$$f(\lambda) = \frac{1_{|\lambda - \operatorname{Re} z| \le 1}}{|\lambda - z|}, \quad \tilde{f}(\lambda) = \frac{1_{|\lambda - \operatorname{Re} z| > 1}}{|\lambda - z|}, \quad g(\lambda) = \frac{1}{|\lambda - z|^2}$$

so that

$$\mathbb{E}\left| S(z) \frac{\partial}{\partial z} G(x, y; z) \right| \le \mathbb{E} \iint \Big(f(\lambda_1) + \tilde{f}(\lambda_1) \Big)\, g(\lambda_2)\, \nu(d\lambda_1) |\mu_{xy}|(d\lambda_2).$$

Setting $I_\alpha = \operatorname{Re} z + [\alpha/N, (\alpha+1)/N)$,

$$\begin{aligned}
&\mathbb{E} \iint f(\lambda_1) g(\lambda_2)\, \nu(d\lambda_1) |\mu_{xy}|(d\lambda_2) \\
&\le \sum_{\alpha,\beta \in \mathbb{Z}} \left(\sup_{\lambda \in I_\alpha} f(\lambda) \right) \left(\sup_{\lambda \in I_\beta} g(\lambda) \right) \mathbb{E}\left[\nu(I_\alpha) |\mu_{xy}|(I_\beta) \right] \\
&\le \frac{C}{N^2} \sum_{\alpha,\beta \in \mathbb{Z}} \left(\sup_{\lambda \in I_\alpha} f(\lambda) \right) \left(\sup_{\lambda \in I_\beta} g(\lambda) \right),
\end{aligned}$$

where we used the second part of Lemma 3.1 to bound the expectations. Since f and g are symmetric about $\operatorname{Re} z$ and monotone decreasing in $|\lambda - \operatorname{Re} z|$, the previous chain of inequalities continues

$$\begin{aligned}
&\le \frac{4C}{N^2} \sum_{\alpha,\beta \in \mathbb{N}_0} f\left(\operatorname{Re} z + \frac{\alpha}{N} \right) g\left(\operatorname{Re} z + \frac{\beta}{N} \right) \\
&= CN \sum_{\alpha=0}^{N} \frac{1}{\sqrt{\alpha^2 + (N \operatorname{Im} z)^2}} \sum_{\beta \in \mathbb{N}_0} \frac{1}{\beta^2 + (N \operatorname{Im} z)^2} \\
&\le CN \left(\log N + \frac{1}{N \operatorname{Im} z} \right) \left(1 + \frac{1}{(N \operatorname{Im} z)^2} \right).
\end{aligned}$$

Finally, because $|\tilde{f}| \le 1$, the remaining summands satisfy

$$\mathbb{E} \iint \tilde{f}(\lambda_1) g(\lambda_2)\, \nu(d\lambda_1) |\mu_{xy}|(d\lambda_2) \le \mathbb{E} \int \frac{1}{|\lambda - z|^2}\, |\mu_{xy}|(d\lambda) \le \frac{C}{\operatorname{Im} z},$$

arguing as in (3.4). □

Evaluating the trace defining $S(z)$ in the site basis,

$$S(z) = \frac{1}{N} \sum_y G(y, y; z),$$

we may average the bounds furnished by Theorem 3.2 to obtain the following corollary, which gives the corresponding bounds for the drift of diffusion of the trace of the resolvent.

Corollary 3.3. *There exists $C < \infty$, depending only on C_V, such that*

$$\mathbb{E}\left| \frac{1}{2N} \frac{\partial^2}{\partial z^2} S(z) \right| \le \frac{C}{N (\operatorname{Im} z)^2}$$

and

$$\mathbb{E}\left| S(z) \frac{\partial}{\partial z} S(z) \right| \le CN \left(\log N + \frac{1}{N \operatorname{Im} z} \right) \left(1 + \frac{1}{(N \operatorname{Im} z)^2} \right) + \frac{C}{\operatorname{Im} z}$$

for all $z \in \mathbb{C}_+$.

We conclude this section with a bound in the same spirit as the previous results for a term which does not explicitly occur in the resolvent flow, but which will nevertheless prove useful in controlling the diffusion of (2.5).

Theorem 3.4. *There exists $C < \infty$, depending only on C_V, such that*

$$\mathbb{E}\left[\operatorname{Im} G(x,x;z)\operatorname{Im} S(z)\right] \le C\left(1+\frac{1}{N\operatorname{Im} z}\right)^2$$

for all $x \in \{1,\dots,N\}$ and $z \in \mathbb{C}_+$.

Proof. The proof follows along the same lines as that of Theorem 3.2. Setting $I_\alpha = \operatorname{Re} z + [\alpha/N, (\alpha+1)/N)$, letting

$$P_z(\lambda) = \operatorname{Im}\frac{1}{\lambda - z} = \frac{\operatorname{Im} z}{(\lambda - \operatorname{Re} z)^2 + (\operatorname{Im} z)^2} \tag{3.5}$$

denote the rescaled Poisson kernel, and using Lemma 3.1, we see that

$$\begin{aligned}\mathbb{E}\left[\operatorname{Im} G(x,x;z)\operatorname{Im} S(z)\right] &= \mathbb{E}\iint P_z(\lambda_1)P_z(\lambda_2)\,\nu(d\lambda_1)\mu_x(d\lambda_2)\\ &\le \sum_{\alpha,\beta\in\mathbb{Z}}\left(\sup_{\lambda\in I_\alpha} P_z(\lambda)\right)\left(\sup_{\lambda\in I_\beta} P_z(\lambda)\right)\mathbb{E}\left[\nu(I_\alpha)\mu_x(I_\beta)\right]\\ &\le \frac{C}{N^2}\sum_{\alpha,\beta\in\mathbb{Z}}\left(\sup_{\lambda\in I_\alpha} P_z(\lambda)\right)\left(\sup_{\lambda\in I_\beta} P_z(\lambda)\right).\end{aligned}$$

Since P_z is symmetric about $\operatorname{Re} z$ and monotone decreasing in $|\lambda - \operatorname{Re} z|$, the last term is in turn bounded by

$$\begin{aligned}&\le \frac{4C}{N^2}\sum_{\alpha,\beta\in\mathbb{N}_0} P_z\left(\operatorname{Re} z + \frac{\alpha}{N}\right)P_z\left(\operatorname{Re} z + \frac{\beta}{N}\right)\\ &= C\sum_{\alpha\in\mathbb{N}_0}\frac{N\operatorname{Im} z}{\alpha^2 + (N\operatorname{Im} z)^2}\sum_{\beta\in\mathbb{N}_0}\frac{N\operatorname{Im} z}{\beta^2 + (N\operatorname{Im} z)^2}\\ &\le C\left(1+\frac{1}{N\operatorname{Im} z}\right)^2.\end{aligned}$$

□

4 Proof of stability

In this section, we turn to the proofs of Theorems 2.1 and 2.2. We start by deriving the stochastic differential equations (2.5) for the resolvent $R_t(z)$ in terms of the Green functions and the normalized trace. For this, we define the martingales

$$dM_t(x,y;z) = -\frac{1}{\sqrt{N}}\sum_{u\le v}\langle\delta_y, R_t(z)P_{uv}R_t(z)\delta_x\rangle\,dB_{uv}(t),$$

where

$$P_{uv} = \frac{1}{\sqrt{1+\delta_{uv}}}\left(|\delta_u\rangle\langle\delta_v| + |\delta_v\rangle\langle\delta_u|\right) = \sqrt{N}\frac{\partial}{\partial B_{uv}}H_t$$

denotes the symmetric matrix element corresponding to $\{\delta_u, \delta_v\}$.

Theorem 4.1. *The Green function satisfies*

$$dG_t(x,y;z) = \left(S_t(z)\frac{\partial}{\partial z}G_t(x,y;z) + \frac{1}{2N}\frac{\partial^2}{\partial z^2}G_t(x,y;z)\right)dt + dM_t(x,y;z)$$

for all $x, y \in \{1,\dots,N\}$ and $z \in \mathbb{C}_+$.

Proof. By the resolvent equation,

$$\frac{\partial}{\partial B_{uv}} R_t(z) = -\frac{1}{\sqrt{N}} R_t(z) P_{uv} R_t(z),$$

so using Itô's Lemma shows that

$$\begin{aligned} dG_t(x,y;z) &= \frac{1}{N}\sum_{u\le v} \langle \delta_y, R_t(z)P_{uv}R_t(z)P_{uv}R_t(z)\delta_x\rangle \, dt \\ &\quad - \frac{1}{\sqrt{N}}\sum_{u\le v} \langle \delta_y, R_t(z)P_{uv}R_t(z)\delta_x\rangle \, dB_{uv}(t) \\ &= \frac{1}{N}\sum_{u\le v} \langle \delta_y, R_t(z)P_{uv}R_t(z)P_{uv}R_t(z)\delta_x\rangle \, dt + dM_t(x,y;z). \end{aligned}$$

We expand the drift term as

$$\begin{aligned} &\frac{1}{N}\sum_{u<v} \langle \delta_y, R_t(z)\delta_v\rangle\langle \delta_u, R_t(z)\delta_u\rangle\langle \delta_v, R_t(z)\delta_x\rangle + \langle \delta_y, R_t(z)\delta_u\rangle\langle \delta_v, R_t(z)\delta_v\rangle\langle \delta_u, R_t(z)\delta_x\rangle \\ &+ \frac{1}{N}\sum_{u<v} \langle \delta_y, R_t(z)\delta_v\rangle\langle \delta_u, R_t(z)\delta_v\rangle\langle \delta_u, R_t(z)\delta_x\rangle + \langle \delta_y, R_t(z)\delta_u\rangle\langle \delta_v, R_t(z)\delta_u\rangle\langle \delta_v, R_t(z)\delta_x\rangle \\ &+ \frac{2}{N}\sum_{u} \langle \delta_y, R_t(z)\delta_u\rangle\langle \delta_u, R_t(z)\delta_u\rangle\langle \delta_u, R_t(z)\delta_x\rangle \end{aligned}$$

and exploit that the second term in each sum is the same as the first term with u and v interchanged to rewrite these sums as

$$\begin{aligned} &= \frac{1}{N}\sum_{u,v} \langle \delta_y, R_t(z)\delta_v\rangle\langle \delta_u, R_t(z)\delta_u\rangle\langle \delta_v, R_t(z)\delta_x\rangle \\ &+ \frac{1}{N}\sum_{u,v} \langle \delta_y, R_t(z)\delta_v\rangle\langle \delta_u, R_t(z)\delta_v\rangle\langle \delta_u, R_t(z)\delta_x\rangle. \end{aligned}$$

In the second sum, we use that the spectral measures μ_{vu} are real to replace $\langle \delta_u, R_t(z)\delta_v\rangle$ with $\langle \delta_v, R_t(z)\delta_u\rangle$, which yields

$$\begin{aligned} &= \langle \delta_y, R_t(z)^2\delta_x\rangle \frac{1}{N}\operatorname{Tr} R_t(z) + \frac{1}{N}\langle \delta_y, R_t(z)^3\delta_x\rangle \\ &= S_t(z)\frac{\partial}{\partial z} G_t(x,y;z) + \frac{1}{2N}\frac{\partial^2}{\partial z^2} G_t(x,y;z). \end{aligned}$$

□

We remark that the applicability of the arguments in this paper to GUE perturbations in place of GOE perturbations is not affected by the last part of the proof, which made use of the fact that the spectral measures are real in the GOE case. This is because the additional unitary symmetry ensures that the third order term involving $\langle \delta_y, R_t(z)^3\delta_x\rangle$ vanishes completely for the GUE flow.

By averaging the evolution of $G_t(x,x;z)$ over $x \in \{1,\dots,N\}$, we obtain an equation with a diffusion given by

$$M_t(z) = \frac{1}{N}\sum_x M_t(x,x;z),$$

which is the familiar complex Burgers equation for $S_t(z)$ [2].

Corollary 4.2. *The normalized trace satisfies*

$$dS_t(z) = \left(S_t(z)\frac{\partial}{\partial z} S_t(z) + \frac{1}{2N}\frac{\partial^2}{\partial z^2} S_t(z) \right) dt + dM_t(z)$$

for all $z \in \mathbb{C}_+$.

We will now employ the results of Section 3 to smooth the resolvent flow of Theorem 4.1. Theorem 3.2 and Corollary 3.3 already accomplish this for the drift, but some further analysis based on spatial averaging is required to control the diffusion and this is the content of the next two theorems.

Theorem 4.3. *There exists a constant $C < \infty$, depending only on C_V, such that*

$$\frac{1}{N}\sum_y \mathbb{E}\,|M_t(x,y;z)| \le C\sqrt{\frac{t}{N(\operatorname{Im} z)^2}}\left(1+\frac{1}{N\operatorname{Im} z}\right)$$

for all $x \in \{1,\dots,N\}$, $z \in \mathbb{C}_+$ and $t \ge 0$.

Proof. The quadratic variation of $M_t(x,y;z)$ satisfies

$$\begin{aligned}
\langle M_t(x,y;z)\rangle &= \frac{1}{N}\int_0^t \sum_{u\le v} |\langle \delta_y, R_s(z)P_{uv}R_s(z)\delta_x\rangle|^2\, ds \\
&\le \frac{2}{N}\int_0^t \sum_{u,v} |\langle \delta_y, R_s(z)\delta_u\rangle\langle\delta_v, R_s(z)\delta_x\rangle|^2\, ds \\
&= \frac{2}{N}\int_0^t \left(\sum_u |\langle \delta_y, R_s(z)\delta_u\rangle|^2\right)\left(\sum_v |\langle\delta_v, R_s(z)\delta_x\rangle|^2\right) ds \\
&= \frac{2}{N(\operatorname{Im} z)^2}\int_0^t \operatorname{Im} G_s(x,x;z)\operatorname{Im} G_s(y,y;z)\, ds,
\end{aligned}$$

where we combined the symmetrization argument of Theorem 4.1 with the inequality $(a+b)^2 \le 2(a^2+b^2)$. Hence

$$\begin{aligned}
\frac{1}{N}\sum_y \mathbb{E}\,|\langle M_t(x,y;z)\rangle| &\le \frac{2}{N(\operatorname{Im} z)^2}\int_0^t \mathbb{E}\,[\operatorname{Im} G_s(x,x;z)\operatorname{Im} S_s(z)]\, ds \\
&\le \frac{Ct}{N(\operatorname{Im} z)^2}\left(1+\frac{1}{N\operatorname{Im} z}\right)^2
\end{aligned}$$

by Theorem 3.4. Combining the Burkholder-Davis-Gundy inequality with Jensen's inequality for $\frac{1}{N}\sum_y \mathbb{E}$ shows that

$$\begin{aligned}
\frac{1}{N}\sum_y \mathbb{E}\,|M_t(x,y;z)| &\le C\left(\frac{1}{N}\sum_y \mathbb{E}\,\langle M_t(x,y;z)\rangle\right)^{1/2} \\
&\le C\sqrt{\frac{t}{N(\operatorname{Im} z)^2}}\left(1+\frac{1}{N\operatorname{Im} z}\right).
\end{aligned}$$

□

Next, we state the corresponding result for the averaged martingale

$$M_t(z) = \frac{1}{N}\sum_x M_t(x,x;z)$$

occuring in Corollary 4.2.

Theorem 4.4. *There exists a constant $C < \infty$, depending only on C_V, such that*

$$\mathbb{E}\,|M_t(z)| \le \sqrt{\frac{Ct}{N^2(\operatorname{Im} z)^3}}$$

for all $z \in \mathbb{C}_+$ and $t \ge 0$.

Proof. By symmetrization,

$$\begin{aligned} M_t(z) &= \frac{1}{N}\sum_x M_t(x,x;z) \\ &= -\frac{1}{N^{3/2}}\sum_{u,v}\frac{1}{\sqrt{1+\delta_{uv}}}\int_0^t \sum_x \langle \delta_v, R_s(z)\delta_x\rangle\langle \delta_x, R_s(z)\delta_u\rangle \, dB_{uv}(s) \\ &= -\frac{1}{N^{3/2}}\sum_{u,v}\frac{1}{\sqrt{1+\delta_{uv}}}\int_0^t \frac{\partial}{\partial z}\langle \delta_v, R_s(z)\delta_u\rangle \, dB_{uv}(s), \end{aligned}$$

so the quadratic variation may be expressed as

$$\begin{aligned} \langle M_t(z)\rangle &= \frac{1}{N^3}\int_0^t \sum_{u,v}\frac{1}{1+\delta_{uv}}\left|\frac{\partial}{\partial z}\langle \delta_v, R_s(z)\delta_u\rangle\right|^2 ds \\ &\le \frac{1}{N^3(\operatorname{Im} z)^2}\int_0^t \sum_{u,v}\left|\langle \delta_v, R_s(z)\delta_u\rangle\right|^2 ds \\ &= \frac{1}{N^2(\operatorname{Im} z)^3}\int_0^t \operatorname{Im} S_s(z)\, ds. \end{aligned}$$

Using, in order, the Burkholder-Davis-Gundy inequality, Jensen's inequality, and the Wegner estimate (3.2) yields

$$\begin{aligned} \mathbb{E}\,|M_t(z)| &\le C\left(\mathbb{E}\,\langle M_t(z)\rangle\right)^{1/2} \\ &\le C\left(\frac{1}{N^2(\operatorname{Im} z)^3}\int_0^t \mathbb{E}\operatorname{Im} S_s(z)\, ds\right)^{1/2} \\ &\le \sqrt{\frac{Ct}{N^2(\operatorname{Im} z)^3}}. \end{aligned}$$

□

The proofs of Theorems 2.1 and 2.2 now reduce to plugging the various previous estimates into the integrated forms of Theorem 4.1 and Corollary 4.2. For the sake of completeness, we illustrate this with the proof of Theorem 2.2, but omit the very similar proof of Theorem 2.1.

Proof of Theorem 2.2. By Theorem 4.1,

$$\begin{aligned} &\frac{1}{N}\sum_y \mathbb{E}\,\left|G_t\left(x,y;E+i\eta\right) - G_0\left(x,y;E+i\eta\right)\right| \\ &\le \frac{1}{N}\sum_y \int_0^t \mathbb{E}\left|S_s(z)\frac{\partial}{\partial z}G_s(x,y;z) + \frac{1}{2N}\frac{\partial^2}{\partial z^2}G_s(x,y;z)\right| ds \\ &+ \frac{1}{N}\sum_y \mathbb{E}\,|M_t(x,y;z)|, \end{aligned}$$

which by Theorems 3.2 and 4.3 is bounded by

$$\begin{aligned} &\le CtN\left(\log N + \frac{1}{N\eta}\right)\left(1+\frac{1}{(N\eta)^2}\right) + \frac{Ct}{\eta} + \frac{Ct}{N\eta^2} \\ &+ C\sqrt{\frac{t}{N\eta^2}}\left(1+\frac{1}{N\eta}\right). \end{aligned}$$

After taking a factor $N^{-c/2}$ from $t \le N^{-(1+c)}$ to control the $\log N$ term, each term is dominated by either $1+(N\eta)^{-1}$ or $1+(N\eta)^{-3}$, which proves the theorem. □

5 Proof of Poisson statistics

In the remainder of this paper, we will show how to apply Theorems 2.1 and 2.2 to the ultrametric ensemble H_n defined in (1.2), thereby obtaining Theorems 1.1 and 1.2. When $c > 0$, the limit $\lim_{n\to\infty} Z_{n,c} \in (0,\infty)$ exists, and thus we may drop the normalizing constant $Z_{n,c}$ from the definition of H_n without any loss of generality. Similarly to our approach in [38], we will prove Theorem 1.1 by approximating $H_n \equiv \sum_{r=0}^{n} 2^{-\frac{1+c}{2}r}\Phi_{n,r}$ with the truncated Hamiltonian

$$H_{n,m} = \sum_{r=0}^{m} 2^{-\frac{1+c}{2}r}\Phi_{n,r}, \tag{5.1}$$

which has the property that, for any $m \le k \le n$,

$$H_{n,m} = \bigoplus_{j=1}^{2^{n-k}} H_{k,m}^{(j)}, \tag{5.2}$$

where each $H_{k,m}^{(j)}$ is an independent copy of H_k. Therefore

$$\mu_{n,m}(f) = \sum_{\lambda\in\sigma(H_{n,m})} f(2^n(\lambda - E))$$

consists of 2^{n-m} independent components, a fact whose relevance to Theorem 1.1 is contained in the following characterization of Poisson point processes.

Proposition 5.1. *Let $\{\mu_{n,j} \,|\, j = 1,\dots,N_n\}$ be a collection of point processes such that:*

1. *The point processes $\{\mu_{n,1},\dots,\mu_{n,N_n}\}$ are independent for all $n \ge 1$.*
2. *If $B \subset \mathbb{R}$ is a bounded Borel set, then*

$$\lim_{n\to\infty} \sup_{j\le N_n} \mathbb{P}(\mu_{n,j}(B) \ge 1) = 0.$$

3. *There exists some $c \ge 0$ such that if $B \subset \mathbb{R}$ is a bounded Borel set with $|\partial B| = 0$, then*

$$\lim_{n\to\infty} \sum_{j=1}^{N_n} \mathbb{P}(\mu_{n,j}(B) \ge 1) = c|B|$$

and

$$\lim_{n\to\infty} \sum_{j=1}^{N_n} \mathbb{P}(\mu_{n,j}(B) \ge 2) = 0.$$

Then, $\mu_n = \sum_j \mu_{n,j}$ converges in distribution to a Poisson point process with intensity c.

We recall [1] that a sequence of point processes μ_n converges in distribution to μ whenever

$$\lim_{n\to\infty} \mathbb{E}\, e^{-\mu_n(P_z)} = \mathbb{E}\, e^{-\mu(P_z)}$$

for all $z \in \mathbb{C}_+$, where P_z is the rescaled Poisson kernel

$$P_z(\lambda) = \operatorname{Im}\frac{1}{\lambda - z} = \frac{\operatorname{Im} z}{(\lambda - \operatorname{Re} z)^2 + (\operatorname{Im} z)^2}. \tag{5.3}$$

Hence, Theorem 1.1 follows by furnishing a sequence m_n such that Proposition 5.1 applies to μ_{n,m_n} and

$$\lim_{n\to\infty} \mathbb{E}\, e^{-\mu_{n,m_n}(P_z)} = \lim_{n\to\infty} \mathbb{E}\, e^{-\mu_n(P_z)} \tag{5.4}$$

for all $z \in \mathbb{C}_+$.

The difference $H_n - H_{n,n-1} = \sqrt{t}\,\Phi_{n,n}$ is a Gaussian perturbation with time parameter $t = 2^{-(1+c)n}$. Therefore, Theorem 2.1 shows that there exists $C_z < \infty$ such that for all $\ell \geq n$ we have

$$\begin{aligned}\frac{1}{2^n}\,\mathbb{E}\left|\mathrm{Tr}\,(H_n - z_\ell)^{-1} - \mathrm{Tr}\,(H_{n,n-1} - z_\ell)^{-1}\right| &\leq C_z\, 2^{-\frac{c}{2}n-1}\,\left(1 + 2^{3(\ell-n)}\right)\\ &\leq C_z\, 2^{-\frac{c}{2}n}\, 2^{3(\ell-n)} \end{aligned} \tag{5.5}$$

with $z_\ell = E + 2^{-\ell}z$. Our strategy in achieving (5.4) thus consists of applying (5.5) to the finite-volume density of states measures

$$\nu_n(f) = 2^{-n}\mathrm{Tr}\, f\,(H_n)\,, \quad \nu_{n,m}(f) = 2^{-n}\mathrm{Tr}\, f\,(H_{n,m})$$

in an iterative fashion.

Theorem 5.2. *There exist $C_z < \infty$ and $\delta > 0$ such that*

$$\mathbb{E}\,|\nu_n\,(P_{z_\ell}) - \nu_{n,m}\,(P_{z_\ell})| \leq C_z\, 2^{3(\ell-(1+\delta)m)}$$

for all $\ell \geq n$.

Proof. The estimate (5.5) proves that

$$\mathbb{E}\,|\nu_k\,(P_{z_\ell}) - \nu_{k,k-1}\,(P_{z_\ell})| \leq C_z\, 2^{3(\ell-(1+\delta)k)} \tag{5.6}$$

with $\delta = c/6$ when $\ell \geq k$. Since $\nu_n - \nu_{n,m}$ is given by a telescopic sum,

$$\nu_n(P_{z_\ell}) - \nu_{n,m}(P_{z_\ell}) = \sum_{k=m+1}^{n} \left(\nu_{n,k}(P_{z_\ell}) - \nu_{n,k-1}(P_{z_\ell})\right),$$

the decomposition (5.2) implies that

$$\nu_{n,k}(P_{z_\ell}) - \nu_{n,k-1}(P_{z_\ell}) = 2^{-(n-k)} \sum_{j=1}^{2^{n-k}} \left(\nu_k(P_{z_\ell}) - \nu_{k,k-1}(P_{z_\ell})\right). \tag{5.7}$$

Applying (5.6) to each term in (5.7) yields

$$\mathbb{E}\,|\nu_n\,(P_{z_\ell}) - \nu_{n,m}\,(P_{z_\ell})| \leq \sum_{k=m+1}^{n} C_z\, 2^{3(\ell-(1+\delta)k)} \leq C_z\, 2^{3(\ell-(1+\delta)m)}. \qquad \square$$

Theorem 5.2 has two important implications for the measures μ_n and $\mu_{n,m}$ which are based on the identities $\mu_n(P_z) = \nu_n\,(P_{z_n})$ and $\mu_{n,m}(P_z) = \nu_{n,m}\,(P_{z_n})$. The first of these enables us to find a suitable sequence μ_{n,m_n} satisfying (5.4).

Corollary 5.3. *There exists a sequence m_n with $m_n \to \infty$ and $n - m_n \to \infty$ such that*

$$\lim_{n\to\infty} \mathbb{E}\,|\mu_n(P_z) - \mu_{n,m_n}(P_z)| = 0$$

for all $z \in \mathbb{C}_+$.

Proof. Since $\delta > 0$, there exists a sequence m_n with $m_n \to \infty$, $n - m_n \to \infty$ and $n - (1+\delta)m_n \to -\infty$. By applying Theorem 5.2 with $\ell = n$, we obtain

$$\mathbb{E}\,|\mu_n(P_z) - \mu_{n,m_n}(P_z)| \leq C_z 2^{3(n-(1+\delta)m_n)} \to 0. \qquad \square$$

To finish the proof of Theorem 1.1, we need to show that μ_{n,m_n} satisfies the hypothesis of Proposition 5.1. By (5.2), μ_{n,m_n} is a sum of point processes

$$\mu_{n,m_n} = \sum_{j=1}^{2^{n-m_n}} \mu_{m_n,j}$$

with independent $\mu_{m_n,j}$. If $B \subset \mathbb{R}$ is a bounded Borel set, the theorem of Combes-Germinet-Klein [8] asserts that $\mathbb{P}(\operatorname{Tr} 1_B(H_m) \geq \ell) \leq (C\, 2^m |B|)^\ell / \ell!$ and hence for any $\ell \geq 0$:

$$\mathbb{P}(\mu_{m_n,j}(B) \geq \ell) \leq \frac{(C|B|\, 2^{m_n-n})^\ell}{\ell!}. \tag{5.8}$$

Since $n - m_n \to \infty$, the first hypothesis of Proposition 5.1 follows. Writing

$$X(n,\ell) = \sum_{j=1}^{2^{n-m}} \mathbb{P}(\mu_{m_n,j}(B) \geq \ell),$$

(5.8) implies

$$X(n,\ell) \leq 2^{n-m_n} \frac{(C|B|\, 2^{m_n-n})^\ell}{\ell!} \to 0$$

when $\ell \geq 2$. In particular, $X(n,2) \to 0$ and the last hypothesis of Proposition 5.1 is satisfied. It remains to prove the remaining hypothesis of Proposition 5.1, which is the second important consequence of Theorem 5.2 and is contained in the following theorem.

Theorem 5.4. *Let $B \subset \mathbb{R}$ be a bounded Borel set. Then,*

$$\lim_{n\to\infty} X(n,1) = \nu(E)|B|.$$

Proof. By (5.2) we have $\mathbb{E}\,\nu_{p,n} = \mathbb{E}\,\nu_n$ for any $p \geq n$, and so we conclude from Theorem 5.2 with $\ell = n$ that

$$\begin{aligned}
\lim_{n\to\infty} |\mathbb{E}\,[\nu_n(P_{z_n}) - \nu(P_{z_n})]| &= \lim_{n\to\infty} \lim_{p\to\infty} |\mathbb{E}\,[\nu_n(P_{z_n}) - \nu_p(P_{z_n})]| \\
&= \lim_{n\to\infty} \lim_{p\to\infty} |\mathbb{E}\,[\nu_{p,n}(P_{z_n}) - \nu_p(P_{z_n})]| \\
&\leq \lim_{n\to\infty} C_z\, 2^{-3\delta n} = 0.
\end{aligned}$$

This shows that the measures $\lambda_n(B) = 2^n \nu(2^{-n}B + E)$ satisfy

$$\lim_{n\to\infty} (\mathbb{E}\,\mu_n(P_z) - \lambda_n(P_z)) = 0$$

and Corollary 5.3 implies that also

$$\lim_{n\to\infty} (\mathbb{E}\,\mu_{m_n}(P_z) - \lambda_n(P_z)) = 0. \tag{5.9}$$

For any bounded Borel set $B \subset \mathbb{R}$, the indicator 1_B is in the L^1-closure of the finite linear combinations from the set $\{P_z \,|\, z \in \mathbb{C}_+\}$ and the measures $\mathbb{E}\,\mu_n$ are absolutely continuous with uniformly bounded densities by the Wegner estimate. Together, these two observations yield that (5.9) is valid for any bounded Borel set $B \subset \mathbb{R}$. Moreover, since E is a Lebesgue point of ν,

$$\lim_{n\to\infty} \lambda_n(B) = \lim_{n\to\infty} 2^n \nu(2^{-n}B + E) = \nu(E)|B|,$$

and hence we have shown that

$$\lim_{n\to\infty} \mathbb{E}\,\mu_{m_n}(B) = \nu(E)|B|. \tag{5.10}$$

Since $\mu_{n_m,j}(B)$ takes values in the non-negative integers

$$\lim_{n\to\infty} X(n,1) = \lim_{n\to\infty} \sum_{j=1}^{2^{n-m}} \mathbb{E}\,\mu_{n_m,j}(B) - \lim_{n\to\infty} \sum_{\ell\geq 2} X(n,\ell)$$

so (5.8), (5.10) and the dominated convergence theorem give

$$\lim_{n\to\infty} X(n,1) = \lim_{n\to\infty} \sum_{j=1}^{2^{n-m}} \mathbb{E}\,\mu_{n_m,j}(B) = \nu(E)|B|. \qquad \square$$

6 Proof of eigenfunction localization

In this section, we prove Theorem 1.2 by comparing the eigenfunctions of H_n with the obviously localized eigenfunctions of $H_{n,m}$. Nevertheless, we again start by considering a more general $N \times N$ random matrix $H = T + V$ with a potential satisfying (2.3) and proving an implication of local resolvent bounds for the eigenfunction correlator

$$Q(x,y;W) = \sum_{\lambda\in\sigma(H)\cap W} |\psi_\lambda(x)\psi_\lambda(y)|$$

in some mesoscopic spectral window

$$W = \left[E_0 - N^{-(1-w)}, E_0 + N^{-(1-w)}\right]$$

with $w > 0$.

Theorem 6.1. *Let $\eta = N^{-(1+\ell)}$ with $\ell > w > 0$ and let $Y \subset \{1,\dots,N\}$. Then, there exists a constant $C < \infty$, depending only on C_V, such that*

$$\mathbb{P}\left(\sum_{y\in Y} Q(x,y;W) > \frac{2}{\pi}\sum_{y\in Y}\int_W |\operatorname{Im} G(x,y;E+i\eta)|\,dE + \frac{\log N}{N^w}\right) \leq CN^{w-\ell}$$

for all $x \in \{1,\dots,N\}$.

The proof of Theorem 6.1 is based on the following two lemmas, the first of which is formulated in terms of the the Poisson kernel P_z defined in (3.5).

Lemma 6.2. *There exists a constant $C < \infty$, depending only on C_V, such that*

$$\mathbb{E}\sum_y |\mu_{xy}|\,(1_{W^c}(1_W * P_{i\eta})) \leq CN\eta\left(1 + \log\sqrt{1+\eta^{-2}|W|^2}\right)$$

for all intervals $W \subset \mathbb{R}$ and $\eta > 0$.

Proof. By spectral averaging (Lemma 3.1),

$$\begin{aligned}\sum_y \mathbb{E}\,|\mu_{xy}|(1_{W^c}(1_W * P_{i\eta})) &\leq CN\int_{W^c}(1_W * P_{i\eta})(\lambda)\,d\lambda \\ &= CN\int_{W^c}\int_W \frac{\eta}{(u-v)^2+\eta^2}\,du\,dv \\ &= CN\eta\int_{\eta^{-1}W^c}\int_{\eta^{-1}W}\frac{1}{1+(u-v)^2}\,du\,dv.\end{aligned}$$

Without loss of generality, we may assume that $\eta^{-1}W = [-a, a]$, so

$$\int_{\eta^{-1}W^c}\int_{\eta^{-1}W} \frac{1}{1+(u-v)^2}\,du\,dv = \int_{\eta^{-1}W^c} \arctan(a-v) - \arctan(v+a)\,dv$$
$$= 2\int_a^\infty \arctan(v+a) - \arctan(v-a)\,dv$$

since $\arctan v$ is an odd function of v. After the appropriate translations, this last integral is

$$= 2\lim_{R\to\infty}\int_{R-a}^{R+a} \arctan v\,dv - 2\int_0^{2a} \arctan v\,dv$$
$$= 2\left(\frac{2\pi a}{2} - \int_0^{2a} \arctan v\,dv\right)$$
$$= 2a\left(\frac{\pi}{2} - \arctan(2a)\right) + \log\sqrt{1+4a^2}.$$

The proof is completed by noting $|\arctan(x) - \pi/2| \le 1/x$ and inserting $a = \eta^{-1}|W|/2$. □

The second lemma needed for the proof of Theorem 6.1 controls the generic spacing between the eigenvalues of H in the interval W.

Lemma 6.3. *Let $W \subset \mathbb{R}$ be an interval and $|W| \ge S > 0$. Then, there exists a constant $C < \infty$, depending only on C_V, such that the event*

$$\mathcal{E} = \left\{\min_{\lambda\in\sigma(H)\cap W} d\left(\lambda, \partial W \cup \sigma(H)\setminus\{\lambda\}\right) > 2S\right\}$$

satisfies

$$\mathbb{P}(\mathcal{E}^c) \le CSN(1+|W|N).$$

Proof. We split W into a disjoint union of adjacent intervals

$$W = I_1 \cup \cdots \cup I_p$$

with $|I_k| = 2S$ for $1 \le k \le p-1$ and $|I_p| \le 2S$, and let $\tilde{I}_k$ denote the fattened interval $I_k + [-2S, 2S]$. Then $\mathcal{E}^c$ can only occur if

1. $\tilde{I}_k$ contains at least two eigenvalues of H for some $1 \le k \le p$, or
2. $\partial W + [-2S, 2S]$ contains an eigenvalue of H.

Therefore, the Wegner and Minami estimates show that

$$\mathbb{P}(\mathcal{E}^c) \le \mathbb{P}\left(|(\partial W + [-2S,2S]) \cap \sigma(H)| \ge 1\right) + \sum_{k=1}^{p} \mathbb{P}\left(\left|\tilde{I}_k \cap \sigma(H)\right| \ge 2\right)$$
$$\le CSN + Cp\left(SN\right)^2,$$

and since $p \le 2|W|/S$, this proves the lemma. □

Proof of Theorem 6.1. Let $S = \frac{8}{\pi}\eta$ so that the event $\mathcal{E}$ defined in Lemma 6.3 satisfies

$$\mathbb{P}(\mathcal{E}^c) \le CN^{w-\ell}.$$

Since the spectral measures μ_{xy} are real, we can construct the function

$$f(E) = \sum_{\lambda\in\sigma(H)\cap W} \operatorname{sgn}\left[\psi_\lambda(x)\psi_\lambda(y)\right] I_\lambda(E),$$

where I_λ denotes the indicator function of the interval $[\lambda - S, \lambda + S]$. We will prove that on the event $\mathcal{E}$ we have $\|f\|_\infty \leq 1$ and

$$\sum_{y \in Y} |\mu_{xy}|(W) \leq \frac{2}{\pi} \sum_{y \in Y} \mu_{xy}(f * P_{i\eta}) + \sum_y |\mu_{xy}| \left(1_{W^c}(1_W * P_{i\eta})\right), \tag{6.1}$$

so, since

$$\begin{aligned} \mu_{xy}(f * P_{i\eta}) &= \iint f(E) P_{\lambda+i\eta}(E)\, dE\, \mu_{xy}(d\lambda) \\ &= \int f(E) \int P_{E+i\eta}(\lambda)\, \mu_{xy}(d\lambda)\, dE \\ &\leq \|f\|_\infty \int_W |\operatorname{Im} G(x, y; E + i\eta)|\, dE, \end{aligned}$$

the theorem follows from Lemma 6.2 and Markov's inequality.

On $\mathcal{E}$, the intervals I_λ are disjoint and contained in W, so $|f| \leq 1_W$ and, in particular, $\|f\|_\infty \leq 1$. To verify (6.1), we note that

$$\begin{aligned} \mu_{xy}(f * P_{i\eta}) &= \mu_{xy}\left(1_W (f * P_{i\eta})\right) + \mu_{xy}\left(1_{W^c}(f * P_{i\eta})\right) \\ &\geq \sum_{\lambda \in \sigma(H) \cap W} \psi_\lambda(x)\psi_\lambda(y)(f * P_{i\eta})(\lambda) - |\mu_{xy}| \left(1_{W^c}(1_W * P_{i\eta})\right) \end{aligned}$$

on $\mathcal{E}$ and hence it remains only to prove that

$$\operatorname{sgn}\left[\psi_\lambda(x)\psi_\lambda(y)\right](f * P_{i\eta})(\lambda) \geq \frac{\pi}{2}$$

for all $\lambda \in \sigma(H) \cap W$. This is based on the fact that

$$\int (1 - I_\lambda(E)) P_{\lambda+i\eta}(E)\, dE \leq \frac{2\eta}{S} = \frac{\pi}{4}$$

and hence

$$\int I_\lambda(E) P_{\lambda+i\eta}(E)\, dE \geq \pi - \int (1 - I_\lambda(E)) P_{\lambda+i\eta}(E)\, dE \geq \frac{3\pi}{4}.$$

If $\lambda \in \sigma(H) \cap W$ with $\operatorname{sgn}\left[\psi_\lambda(x)\psi_\lambda(y)\right] = 1$, it follows that

$$\begin{aligned} (f * P_{i\eta})(\lambda) &= \int f(E) P_{\lambda+i\eta}(E)\, dE \\ &\geq \int I_\lambda(E) P_{\lambda+i\eta}(E)\, dE - \int (1 - I_\lambda(E)) P_{\lambda+i\eta}(E)\, dE \\ &\geq \frac{\pi}{2}, \end{aligned}$$

and similarly

$$(f * P_{i\eta})(\lambda) \leq -\frac{\pi}{2}$$

if $\operatorname{sgn}\left[\psi_\lambda(x)\psi_\lambda(y)\right] = -1$. □

The proof of the last theorem made use of the fact that the spectral measures μ_{xy} are always real for the GOE flow. It is possible to extend this result to models with complex off-diagonal spectral measures, such as the GUE flow, by using the fact that

$$\langle \delta_y, \operatorname{Im}(H - z)^{-1} \delta_x \rangle + \langle \delta_x, \operatorname{Im}(H - z)^{-1} \delta_y \rangle = \operatorname{Im} G(x, y; z) + \operatorname{Im} G(y, x; z),$$

but we omit these complications here.

With Theorem 6.1 in hand, we now turn to the proof of Theorem 1.2. As in Section 5, we drop the normalizing constant $Z_{n,c}$ from the definition of H_n. The core of this argument again consists of resolvent bounds for Gaussian perturbations, and thus we consider the Green functions

$$G_n(x,y;z) = \langle \delta_y, (H_n - z)^{-1}\delta_x\rangle, \quad G_{n,m}(x,y;z) = \langle \delta_y, (H_{n,m} - z)^{-1}\delta_x\rangle.$$

If $\eta = 2^{-(1+\ell)n}$ for some $\ell > 0$, Theorem 2.2 proves that there exists $C < \infty$ such that

$$\begin{aligned} &2^{-k} \sum_{y\in B_k(x)} \mathbb{E}\left|G_k\left(x,y;E+i\eta\right) - G_{k,k-1}\left(x,y;E+i\eta\right)\right| \\ &\leq C\, 2^{-\frac{c}{2}k}\left(1 + 2^{3((1+\ell)n-k)}\right) \\ &= C\, 2^{3(1+\ell)n - 3(1+\delta)k} \end{aligned}$$

with $\delta = c/6$ whenever $k \leq n$. Iterating this result, we see that

$$\begin{aligned} &2^{-n} \sum_{y\in B_n} \mathbb{E}\left|G_n\left(x,y;E+i\eta\right) - G_{n,m}\left(x,y;E+i\eta\right)\right| \\ &\leq 2^{-n} \sum_{k=m+1}^{n} \sum_{y\in B_n} \mathbb{E}\left|G_{n,k}\left(x,y;E+i\eta\right) - G_{n,k-1}\left(x,y;E+i\eta\right)\right| \\ &= 2^{-n} \sum_{k=m+1}^{n} \sum_{y\in B_k(x)} \mathbb{E}\left|G_k\left(x,y;E+i\eta\right) - G_{k,k-1}\left(x,y;E+i\eta\right)\right| \\ &\leq 2^{-n} \sum_{k=m+1}^{n} 2^k C\, 2^{3(1+\ell)n - 3(1+\delta)k} \leq C\, 2^{(3(1+\ell)-1)n}\, 2^{-(3(1+\delta)-1)m}. \end{aligned}$$

Since $\delta > 0$, we can choose $\ell > 0$, $\epsilon \in (0,1)$, and $w \in (0,\ell)$ such that

$$2\mu := (1-\epsilon)(3(1+\delta)-1) - (3(1+\ell)-1) - w > 0.$$

Thus, setting $m_n = (1-\epsilon)n$ and

$$W = \left[E - 2^{-(1-w)n}, E + 2^{-(1-w)n}\right],$$

and using that $G_{n,m}(x,y;z) = 0$ if $y \notin B_m(x)$ show that

$$\sum_{y\in B_n\setminus B_{m_n}(x)} \mathbb{E}\int_W \left|\mathrm{Im}\, G_n\left(x,y;E+i\eta\right)\right| \, dE \leq C\, 2^{-2\mu n}.$$

Applying Markov's inequality, we arrive at

$$\mathbb{P}\left(\sum_{y\in B_n\setminus B_{m_n}(x)} \int_W \left|\mathrm{Im}\, G_n\left(x,y;E+i\eta\right)\right| \, dE > 2^{-\mu n}\right) \leq C\, 2^{-\mu n},$$

so Theorem 1.2 follows from Theorem 6.1, which says that

$$\sum_{y\in B_n\setminus B_{m_n}(x)} Q_n(x,y;W) \leq \sum_{y\in B_n\setminus B_{m_n}(x)} \int_W \left|\mathrm{Im}\, G_n(x,y;E+i\eta)\right| \, dE + \frac{\log 2^n}{2^{wn}}$$

with probability $1 - \mathcal{O}\left(2^{(w-\ell)n}\right)$.

A Proofs for the Rosenzweig-Porter model

Proof of Theorem 1.5. As $N \to \infty$, the random measure defined by

$$\mu_{N,0}(f) = \sum_{\lambda \in \sigma(H_0)} f(N(\lambda - E))$$

converges in distribution to a Poisson point process with intensity $\varrho(E)$. Setting $z_N = E + z/N$, a simple calculation yields

$$\mu_N(P_z) = \operatorname{Im} S_t(z_N).$$

Thus,

$$\left| \mathbb{E}\, e^{-\mu_N(P_z)} - \mathbb{E}\, e^{-\mu_{N,0}(P_z)} \right| \le \mathbb{E}\, |S_t(z_N) - S_0(z_N)| \le C N^{-c/2},$$

which shows that the characteristic functionals of μ_N and $\mu_{N,0}$ asymptotically agree on the set $\{P_z : z \in \mathbb{C}_+\}$ whose linear span is dense in C_0. This proves the first point.

For the second assertion, choose $\ell > w > 0$ and $\mu_0 > 0$ such that

$$3\ell + w + 2\mu_0 \le c/2.$$

Since $G_0(x, y; z) = 0$ for $x \ne y$, Theorem 2.2 shows that with $\eta = N^{-(1+\ell)}$ we have

$$\begin{aligned} \mathbb{E} \sum_{y \ne x} \int_W |\operatorname{Im} G_t(x, y; E + i\eta)|\, dE &\le C|W| N N^{-c/2} (\eta N)^{-3} \\ &\le C N^{w + 3\ell - c/2} \le N^{-2\mu_0}. \end{aligned}$$

By Markov's inequality,

$$\mathbb{P}\left(\sum_{y \ne x} \int_W |\operatorname{Im} G_t(x, y; E + i\eta)|\, dE \ge N^{-\mu_0} \right) \le C N^{-\mu_0}$$

so choosing $0 < \mu < \min\{w, \mu_0\}$ and $\kappa = \min\{w - \ell, \mu_0\}$, Theorem 6.1 shows that

$$\mathbb{P}\left(\sum_{y \ne x} Q_N(x, y; W) > N^{-\mu} \right) \le C N^{-\kappa}. \qquad \square$$

References

[1] M. Aizenman and S. Warzel, *Random operators: Disorder effects on quantum spectra and dynamics*, Graduate Studies in Mathematics, vol. 168, American Mathematical Society, Providence, RI, 2015. MR-3364516

[2] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*, Cambridge Studies in Advanced Mathematics, vol. 118, Cambridge University Press, Cambridge, 2010. MR-2760897

[3] L. Benigni, *Eigenvectors distribution and quantum unique ergodicity for deformed Wigner matrices*, Preprint available at arXiv:1711.07103, 2017.

[4] P. Bourgade and H.-T. Yau, *The eigenvector moment flow and local quantum unique ergodicity*, Comm. Math. Phys. **350** (2017), no. 1, 231–278. MR-3606475

[5] P. Bourgade, L. Erdős, H.-T. Yau, and J. Yin, *Universality for a class of random band matrices*, Adv. Theor. Math. Phys. **21** (2017), no. 3, 739–800. MR-3695802

[6] A. Bovier, *The density of states in the Anderson model at weak disorder: a renormalization group analysis of the hierarchical model*, J. Statist. Phys. **59** (1990), no. 3-4, 745–779. MR-1063180

[7] G. Casati, L. Molinari, and F. Izrailev, *Scaling properties of band random matrices*, Phys. Rev. Lett. **64** (1990), 1851–1854. MR-1046365

[8] J.-M. Combes, F. Germinet, and A. Klein, *Generalized eigenvalue-counting estimates for the Anderson model*, J. Stat. Phys. **135** (2009), no. 2, 201–216. MR-2505733

[9] M. Disertori, H. Pinson, and T. Spencer, *Density of states for random band matrices*, Comm. Math. Phys. **232** (2002), no. 1, 83–124. MR-1942858

[10] M. Duits and K. Johansson, *On mesoscopic equilibrium for linear statistics in Dyson's Brownian motion*, To appear in Memoirs of the American Mathematical Society (2017).

[11] F. J. Dyson, *A Brownian-motion model for the eigenvalues of a random matrix*, J. Mathematical Phys. **3** (1962), 1191–1198. MR-0148397

[12] F. J. Dyson, *Existence of a phase-transition in a one-dimensional Ising ferromagnet*, Comm. Math. Phys. **12** (1969), no. 2, 91–107. MR-0436850

[13] L. Erdős, *Universality of Wigner random matrices: a survey of recent results*, Russian Mathematical Surveys **66** (2011), no. 3, 507. MR-2859190

[14] L. Erdős, B. Schlein, and H.-T. Yau, *Universality of random matrices and local relaxation flow*, Invent. Math. **185** (2011), no. 1, 75–119. MR-2810797

[15] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin, *Delocalization and diffusion profile for random band matrices*, Comm. Math. Phys. **323** (2013), no. 1, 367–416. MR-3085669

[16] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin, *The local semicircle law for a general class of random matrices*, Electron. J. Probab. **18** (2013), no. 59, 58. MR-3068390

[17] F. Evers and A. D. Mirlin, *Anderson transitions*, Rev. Mod. Phys. **80** (2008), 1355–1417.

[18] D. Facoetti, P. Vivo, and G. Biroli, *From non-ergodic eigenvectors to local resolvent statistics and back: A random matrix perspective*, EPL (Europhysics Letters) **115** (2016), no. 4, 47003.

[19] Y. V. Fyodorov and A. D. Mirlin, *Scaling properties of localization in random band matrices: A σ-model approach*, Phys. Rev. Lett. **67** (1991), 2405–2409. MR-1130103

[20] Y. V. Fyodorov, A. Ossipov, and A. Rodriguez, *The Anderson localization transition and eigenfunction multifractality in an ensemble of ultrametric random matrices*, Journal of Statistical Mechanics: Theory and Experiment **2009** (2009), no. 12, L12001.

[21] J. Huang and B. Landon, *Local law and mesoscopic fluctuations of Dyson Brownian motion for general β and potential*, arXiv:1612.06306 (2017).

[22] S. Kotani, *Lyapunov exponents and spectra for one-dimensional random Schrödinger operators*, Random matrices and their applications (Brunswick, Maine, 1984), Contemp. Math., vol. 50, Amer. Math. Soc., Providence, RI, 1986, pp. 277–286. MR-0841099

[23] V. E. Kravtsov, I. M. Khaymovich, E. Cuevas, and M. Amini, *A random matrix model with localization and ergodic transitions*, New Journal of Physics **17** (2015), no. 12, 122002.

[24] E. Kritchevski, *Hierarchical Anderson model*, Probability and mathematical physics, CRM Proc. Lecture Notes, vol. 42, Amer. Math. Soc., Providence, RI, 2007, pp. 309–322. MR-2352276

[25] E. Kritchevski, *Spectral localization in the hierarchical Anderson model*, Proc. Amer. Math. Soc. **135** (2007), no. 5, 1431–1440 (electronic). MR-2276652

[26] E. Kritchevski, *Poisson statistics of eigenvalues in the hierarchical Anderson model*, Ann. Henri Poincaré **9** (2008), no. 4, 685–709. MR-2413200

[27] S. Kuttruf and P. Müller, *Lifshits tails in the hierarchical Anderson model*, Ann. Henri Poincaré **13** (2012), no. 3, 525–541. MR-2909106

[28] B. Landon, P. Sosoe, and H.-T. Yau, *Fixed energy universality for Dyson Brownian motion*, Preprint available at arXiv:1609.09011, 2016.

[29] B. Landon and H.-T. Yau, *Convergence of Local Statistics of Dyson Brownian Motion*, Comm. Math. Phys. **355** (2017), no. 3, 949–1000. MR-3687212

[30] J. O. Lee and K. Schnelli, *Local deformed semicircle law and complete delocalization for Wigner matrices with random potential*, J. Math. Phys. **54** (2013), no. 10, 103504, 62. MR-3134604

[31] N. Minami, *Local fluctuation of the spectrum of a multidimensional Anderson tight binding model*, Comm. Math. Phys. **177** (1996), no. 3, 709–725. MR-1385082

[32] A. D. Mirlin, Y. V. Fyodorov, F.-M. Dittes, J. Quezada, and T. H. Seligman, *Transition from localized to extended eigenstates in the ensemble of power-law random banded matrices*, Phys. Rev. E **54** (1996), 3221–3230.

[33] S. Molchanov, *Hierarchical random matrices and operators. Application to Anderson model*, Multidimensional statistical analysis and theory of random matrices (Bowling Green, OH, 1996), VSP, Utrecht, 1996, pp. 179–194. MR-1463464

[34] N. Rosenzweig and C. E. Porter, *"Repulsion of energy levels" in complex atomic spectra*, Phys. Rev. **120** (1960), 1698–1714.

[35] J. Schenker, *Eigenvector localization for random band matrices with power law band width*, Comm. Math. Phys. **290** (2009), no. 3, 1065–1097. MR-2525652

[36] S. Sodin, *The spectral edge of some random band matrices*, Ann. of Math. (2) **172** (2010), no. 3, 2223–2251. MR-2726110

[37] P. von Soosten and S. Warzel, *Non-ergodic delocalization in the Rosenzweig-Porter model*, Preprint available at arXiv:1709.10313, 2017.

[38] P. von Soosten and S. Warzel, *Renormalization Group Analysis of the Hierarchical Anderson Model*, Ann. Henri Poincaré **18** (2017), no. 6, 1919–1947. MR-3649447

[39] F. Wegner, *Bounds on the density of states in disordered systems*, Z. Phys. B **44** (1981), no. 1-2, 9–15. MR-0639135

Acknowledgments. We thank Y. V. Fyodorov for introducing us to the ultrametric ensemble. This work was supported by the DFG (WA 1699/2-1).