TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 370, Number 1, January 2018, Pages 641–668
http://dx.doi.org/10.1090/tran/6977
Article electronically published on July 7, 2017

LÉVY-KHINTCHINE RANDOM MATRICES AND THE POISSON WEIGHTED INFINITE SKELETON TREE

PAUL JUNG

ABSTRACT. We study a class of Hermitian random matrices which includes Wigner matrices, heavy-tailed random matrices, and sparse random matrices such as adjacency matrices of Erdős-Rényi random graphs with $p_n \sim \frac{1}{n}$. Our $n \times n$ random matrices have real entries which are i.i.d. up to symmetry. The distribution of entries depends on n, and we require row sums to converge in distribution. It is then well-known that the limit distribution must be infinitely divisible.

We show that a limiting empirical spectral distribution (LSD) exists and, via local weak convergence of associated graphs, that the LSD corresponds to the spectral measure associated to the root of a graph which is formed by connecting infinitely many Poisson weighted infinite trees using a backbone structure of special edges called "cords to infinity". One example covered by the results are matrices with i.i.d. entries having infinite second moments but normalized to be in the Gaussian domain of attraction. In this case, the limiting graph is $\mathbb{N}$ rooted at 1, so the LSD is the semicircle law.

CONTENTS

1. INTRODUCTION

This paper jointly studies the limiting spectral distributions (LSD) for three classes of Hermitian random matrices that have appeared in the literature. The first class of random matrices are classic Wigner matrices introduced in the seminal

Received by the editors August 4, 2014 and, in revised form, March 12, 2015, February 13, 2016, and April 18, 2016.

2010 *Mathematics Subject Classification.* Primary 15B52, 60B20, 60G51.

Key words and phrases. Empirical spectral distribution, Wigner matrices, Lévy matrices, heavy-tailed random matrices, sparse random matrices, Erdős-Rényi graph, local weak convergence.

The author's research was partially supported by NSA grant H98230-14-1-0144 while at the University of Alabama Birmingham.

work of their namesake, [Wig55]. The literature on this class of random matrices is overwhelmingly abundant (see [AGZ10, BS10, Tao12]).

The second class of matrices are adjacency matrices of Erdős-Rényi random graphs on n vertices whose edges are present with probability proportional to $1/n$. The analysis of the LSD in the context of random matrices seems to have started in [RB88]. These matrices are called *sparse*[1] *random matrices*, and they can be considered a Poissonian variation of Wigner matrices. The LSD of sparse random matrices was analyzed using the "moment method" in [Rya98, BG01, KSV04, Zak06], and using the "resolvent method" in [KSV04]. An insightful modification of the latter approach led to improved results in [BL10] (see also [Küh08] for references in the physics literature).

Finally, the third class of random matrices is formed from properly normalized heavy-tailed entries, and, following [BAG08], we call them *heavy-tailed random matrices*. These are also known in the physics literature as Lévy matrices or Wigner-Lévy matrices, and they were introduced by Cizeau and Bouchaud in [CB94]. Later, they were studied more rigorously in [Sos04, BAG08, BCC11a]. These matrices are *not* to be confused with *free* Lévy matrices [BG05, BJN$^+$07].

In each of the three classes of matrices above, the entries are i.i.d. up to self-adjointness, although the distributions may differ for different n. In order to obtain non-trivial LSDs, a proper rescaling or change in distribution is needed as $n \to \infty$ (such rescaling is often implicit in the formulation). After respectively rescaling, if one sums all the entries in a single row or column and takes $n \to \infty$, then one obtains a Gaussian, Poisson, or stable distribution in each of the respective classes. These are all examples of infinitely divisible distributions, which suggests that all three classes of matrices can many times be thought of under this umbrella, and various papers (for example [Rya98, Sec. 3.1]) have done exactly that. More recently, [BGGM13] establishes a functional central limit theorem for the Cauchy-Stieltjes transforms of the LSDs of all three classes, and [Mal12] studies the joint LSDs of a pair of independent ensembles in these three classes using algebraic techniques inspired by free probability.

Here, we also view these three classes as examples from this larger class of matrix ensembles characterized by the Lévy-Khintchine formula, and, in particular, the matrices are viewed as (weighted) adjacency matrices. As was done in the heavy-tailed setting in [BCC11a], our main objective is to equate the LSD of the limiting adjacency operator with the spectral measure associated to the root (or vacuum state) vector in $L^2(V)$ where V is the vertex set of the limiting graph in the sense of local weak convergence (see below). This allows for further analysis of the LSD using the recursive structure of the limiting graph.

The ensembles we consider have i.i.d. complex entries for each n, up to self-adjointness, with zeros on the diagonal. It is well-known that any weak limit of row sums must be infinitely divisible in $\mathbb{C}$ (viewed as $\mathbb{R}^2$). Actually, the "identically distributed" condition may be weakened to require only that the moduli of the entries are identically distributed. In this weakened form one still has that the sum of the square-moduli of entries in a row, i.e., the Euclidean norm-squared of a row as a vector in $\mathbb{R}^{2n}$, converges in distribution to a positive law which is the marginal distribution of a Lévy subordinator.

[1] Here, the random number of non-zero entries in each row remains bounded in distribution as $n \to \infty$. The term "sparse" sometimes refers to what others call dilute random matrices for which the order of non-zero entries in each row is $o(n)$.

In particular, recall (see [Kyp06] or [Kal02]) that a probability measure μ on $\mathbb{R}$ is infinitely divisible with distribution $ID(\sigma^2, b, \Pi)$ and Lévy exponent Ψ,

$$e^{\Psi(\theta)} := \int_{\mathbb{R}} e^{i\theta x}\mu(dx) \quad \text{for } \theta \in \mathbb{R},$$

if and only if there exists a triplet of characteristics (σ^2, b, Π) such that

$$\Psi(\theta) := -\frac{1}{2}\theta^2\sigma^2 + i\theta b + \int_{\mathbb{R}} \left(e^{i\theta x - 1} - \frac{i\theta x}{1+x^2}\right)\Pi(dx), \tag{1.1}$$

where $\sigma^2 \geq 0, b \in \mathbb{R}$, and $\Pi(dx)$ concentrates on $\mathbb{R}\backslash\{0\}$ and satisfies

$$\int_{\mathbb{R}} (1 \wedge |x|^2)\,\Pi(dx) < \infty. \tag{1.2}$$

If μ concentrates on $(0, \infty)$, then the exponent corresponds to the subordinator characteristics (b_s, Π_s) and takes the simplified form

$$\Psi_s(\theta) := i\theta b_s + \int_{(0,\infty)} (e^{i\theta x - 1})\,\Pi_s(dx), \tag{1.3}$$

where $\Pi_s(dx)$ also concentrates on $(0, \infty)$, but instead of (1.2), it satisfies

$$\int_{(0,\infty)} (1 \wedge x)\,\Pi_s(dx) < \infty.$$

Here, the s subscript indicates the *subordinator* form of the Lévy exponent.

We say a sequence of $n \times n$ random matrices $(\mathcal{C}_n)_{n\in\mathbb{N}}$ is a **Lévy-Khintchine random matrix ensemble** with characteristics $(\sigma^2, 0, \Pi)$ if for each n, the moduli of entries $\mathcal{C}_n(j,k) = \bar{\mathcal{C}}_n(k,j), j \neq k$ are i.i.d. (up to self-adjointness, with zeros on the diagonal) and the

$$\text{weak limit} \quad \lim_{n\to\infty} \sum_{k=1}^{n} \pm|\mathcal{C}_n(1,k)| \quad \text{is infinitely divisible} \tag{1.4}$$
$$\text{with characteristics} \quad (\sigma^2, 0, \Pi),$$

where the signs $\pm$ are independent Rademacher random variables (independent also from $\mathcal{C}_n$). This implies that Π is a symmetric measure. It is not hard to see that (1.4) is true if and only if

$$\lim_{n\to\infty} \sum_{k=1}^{n} \mathcal{C}_n(1,k)$$

is infinitely divisible with some other characteristics $(\sigma^2, \tilde{b}, \tilde{\Pi})$ where σ^2 remains unchanged, but $\tilde{b}$ may be non-zero and $\tilde{\Pi}$ is not in general symmetric. An equivalent form of the above is that the

$$\text{weak limit} \quad \lim_{n\to\infty} \sum_{k=1}^{n} |\mathcal{C}_n(1,k)|^2 \quad \text{is infinitely divisible} \tag{1.5}$$
$$\text{with subordinator characteristics} \quad (\sigma^2, \hat{\Pi}_s),$$

where $\hat{\Pi}_s$ can be easily found in terms of Π (see [Kal02, Ch. 15]). Note that in this form σ^2 plays the role of the drift coefficient b_s. We note that by standard arguments, one could set the diagonal elements to any real number which converges to 0 fast enough, and this would not affect the LSD (see the remarks following

Theorem 1.1 below). For the sake of simplicity, we will always set diagonal entries to zero.

In the context of Lévy processes, the three components of the triplet (σ^2, b, Π) correspond to a Brownian component, a drift component, and a jump component (with possibly additional "compensating drift"), respectively. We will see in our context that σ^2 corresponds to a Wigner component, the drift component is inconsequential since by using the random signs it becomes 0 (cf. [BAG08, Remark 1.9]), and the Lévy measure Π generalizes both heavy-tailed and sparse random matrices.

1.1. **Main results.** For a given Lévy-Khintchine ensemble, let $\{\lambda_j\}_{j=1}^n$ denote the eigenvalues of the nth matrix in the sequence. The empirical spectral measure is defined as

$$\mu_{\mathcal{C}_n} := \frac{1}{n}\sum_{j=1}^{n} \delta_{\lambda_j}, \tag{1.6}$$

and their cumulative distribution functions are the empirical spectral distributions (ESDs).

Theorem 1.1 (Existence of the LSD)**.** *For any Lévy-Khintchine random matrix ensemble $(\mathcal{C}_n)_{n\in\mathbb{N}}$ with characteristics $(\sigma^2, 0, \Pi)$ (or alternatively with subordinator characteristics $(\sigma^2, \hat{\Pi}_s)$), there exists a symmetric non-random probability measure $\mu_{\mathcal{C}_\infty}$ depending only on the characteristics to which $(\mu_{\mathcal{C}_n})_{n\in\mathbb{N}}$ weakly converges, almost surely, as $n \to \infty$. In other words,*

$$\mathbf{P}\left(\lim_{n\to\infty}\langle \mu_{\mathcal{C}_n}, f\rangle = \langle \mu_{\mathcal{C}_\infty}, f\rangle \textit{ for all bounded continuous } f\right) = 1. \tag{1.7}$$

Moreover, the limiting measure $\mu_{\mathcal{C}_\infty}$ has bounded support if and only if Π is trivial.

Remarks.

(1) One can of course also consider a situation where the matrices of the ensemble $(\mathcal{C}_n)_{n\in\mathbb{N}}$ are not defined on the same probability space. In that case, $(\mu_{\mathcal{C}_n})_{n\in\mathbb{N}}$ weakly converges, in probability, as $n \to \infty$.

(2) As mentioned above, setting the diagonal entries to zero was a matter of convenience for the proofs below, and the result still holds true if the diagonal entries are real i.i.d. satisfying (1.4) with $\mathcal{C}_n(1,k)$ replaced by $\mathcal{C}_n(k,k)$. The argument is, briefly, that one may appeal to Lemma 5.3 below to assume that the diagonal entries are bounded and then use an argument similar to [AGZ10, Theorem 2.1.21].

(3) If $(\mathcal{A}_n)_{n\in\mathbb{N}}$ is a non-Hermitian sequence, an extension of the above proof technique to the singular values of $(\mathcal{A}_n - zI_n)_{n\in\mathbb{N}}$ in the spirit of [DS07] follows by way of Section 2 in [BCC11b] (see also [FZ97]). Briefly, the idea is to associate a $2n \times 2n$ matrix $\mathcal{B}_n \in \mathcal{M}_{2n}(\mathbb{C})$ to each $\mathcal{A}_n$ by thinking of $\mathcal{B}_n$ as an $n \times n$ quaternionic-type matrix in $\mathcal{M}_n(\mathcal{M}_2(\mathbb{C}))$ with entries given by the 2×2 matrices

$$\mathcal{B}_n(j,k) := \begin{bmatrix} 0 & \mathcal{A}_n(j,k) \\ \bar{\mathcal{A}}_n(j,k) & 0 \end{bmatrix}.$$

In place of the usual resolvent matrix, used in the proof of Theorem 1.1, one then uses

$$R_n(U) = (\mathcal{B}_n - U \otimes I_n)^{-1}$$

where $I_n \in \mathcal{M}_n(\mathbb{C})$ and $U = U(z,w) = \left[\begin{smallmatrix} w & z \\ \bar{z} & w \end{smallmatrix}\right]$. We refer the reader to [BCC11b, Sec. 2] for more details.

As indicated in the introduction, the three main examples of Wigner matrices, heavy-tailed matrices, and sparse random matrices have all been previously studied (see above for references). Let us briefly present a few examples handled by the above theorem where we know of no previous proof in the literature.

Examples.

(1) *Sparse random matrices plus noise* with characteristics $(\sigma^2, 0, \delta_1)$: This is just the case
$$\mathcal{C}_n = \mathcal{C}_n^{(1)} + \mathcal{C}_n^{(2)}$$
where $(\mathcal{C}_n^{(1)})_{n\in\mathbb{N}}$ are sparse random matrices and $(\mathcal{C}_n^{(2)})_{n\in\mathbb{N}}$ are Wigner matrices. Similar "information plus noise" random matrices have recently been considered in the literature [DS07, EK10].

(2) *Gamma(α,β) random matrices* with characteristics $(0, 0, \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x} dx)$: here the entries are such that $\mathcal{C}_n(j,k)$ has a Gamma$(\alpha/n, \beta)$ distribution and the matrix entries are i.i.d. up to the symmetry condition. In particular, one can consider the ensemble for which the first row corresponds to a Gamma process where $\mathcal{C}_n(1,k)$ denotes the increment of the process from time $\frac{k-1}{n}$ to time $\frac{k}{n}$.

(3) *Infinite-variance Wigner matrices* $(\sigma^2, 0, 0)$: One can consider Wigner-type matrices where the entries have a symmetric distribution such that
$$\mathbf{P}\left(|\mathcal{C}_n(j,k)| > x\sqrt{n \log n}\right) = x^{-2}$$
for $x \geq 1$. The entries have infinite variance, but are in the domain of attraction of a Gaussian distribution. The above result implies that the limiting spectral measure is Wigner's semicircle law (see Remark 1.10 in [BAG08]).

In the case where Π has exponential moments, an extension of the standard moment method is enough to handle the proof of Theorem 1.1, and in Section 3 we do just that under the slightly stronger assumption that Π has bounded support. When Π has some moments which are infinite and $\sigma = 0$, the proof follows by generalizing insightful *local weak convergence* arguments of [BL10, BCC11a] (see Section 4). To extend this to the general case, we combine the local weak convergence arguments with a generalized moment method and tail truncation arguments.

As a by-product of local weak convergence, one can view the limiting empirical spectral measure of the random matrix ensembles as the spectral measure of a weighted adjacency operator, at the root vector, of some new infinite graph. For ensembles with characteristics $(0, 0, \Pi)$, this idea is again a generalization of arguments in [BCC11a]. However, when $\sigma > 0$ a non-trivial generalization of Aldous' Poisson weighted infinite tree, which we call a **Poisson weighted infinite skeleton tree (PWIST)**, is required.

The idea of local weak convergence was introduced by Benjamini and Schramm [BS01] and further developed by Aldous and Steele [AS04]. Aldous and Steele describe the technique as finding "a new, infinite, probabilistic object whose local properties inform us about the limiting properties of a sequence of finite problems." When the limiting object has a tree structure, local weak convergence provides a

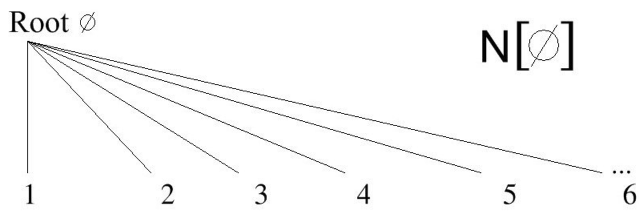

FIGURE 1. The root and its children (first generation offspring).

general framework to make the *cavity method* in physics rigorous. In our context, the cavity method was used in [CB94], and our new infinite object (with a tree structure) generalizes Aldous' Poisson infinite weighted trees (PWIT) by adding to it "cords" of infinite length which connect to independent copies of other PWITs. These cords form a *backbone structure* for a collective object which we refer to as a PWIST.

Let us first recall the definition of the PWIT(λ_Π). Start with a single root vertex $\varnothing$ with an infinite number of (first generation) children indexed by $\mathbb{N}$. The weight on the edge to the kth child is the kth arrival (ordered by absolute value) of a Poisson process on $\mathbb{R}\backslash\{0\}$ with some intensity λ. In our situation the intensity λ_Π is derived from the measure Π on $\mathbb{R}\backslash\{0\}$ by inverting:

$$\lambda_\Pi\{x : 1/x \in B\} := \Pi(B). \tag{1.8}$$

For example, if $\Pi(dx)$ is equivalent to Lebesgue measure with density $f_\Pi(x)dx$, then $\lambda_\Pi(dx)$ is also equivalent to Lebesgue measure with density $x^{-2}f_\Pi(1/x)dx$ where x^{-2} is the change-of-measure factor.

If G has a root at $\varnothing$ we write $G[\varnothing]$ for the rooted graph with (random) weights assigned to each edge. Slightly abusing notation, we denote the subgraph of a PWIT(λ_Π) formed by the root $\varnothing$, its children, and the weighted edges in between, by $\mathbb{N}[\varnothing]$; see Figure 1.

We continue now with other generations. Every vertex v in generation (or depth) $g \geq 1$ has edges to an infinite number of children indexed by $\mathbb{N}$ forming the subgraph $\mathbb{N}[v]$, with weights assigned by repeating the procedure for the weights in the first generation (for $\mathbb{N}[\varnothing]$), namely, according to the points of an *independent* Poisson random measure with intensity $\lambda_\Pi(dx)$. Therefore each $\mathbb{N}[v]$ is an i.i.d. copy of $\mathbb{N}[\varnothing]$. The union of the children vertices of $\mathbb{N}[v]$ (in other words, not including v itself) over all v in some generation $g-1$ is denoted $\mathbb{N}^g$. Thus the total vertex set is

$$\mathbb{N}^F := \bigcup_{g\geq 0} \mathbb{N}^g \tag{1.9}$$

where $\mathbb{N}^0 = \varnothing$; see Figure 2.

The PWIST will depend on both characteristics σ^2 and Π (via λ_Π). To construct a PWIST(σ, λ_Π), we start with a single PWIT(λ_Π)[$\varnothing$] rooted at $\varnothing$ and, for each vertex v of PWIT(λ_Π)[$\varnothing$], we create a new vertex ∞_v which is the root of a new independent PWIT(λ_Π)[∞_v]. We draw an edge from v to ∞_v for each v and assign this edge a non-random weight of

$$1/\sigma \in (0, \infty]. \tag{1.10}$$

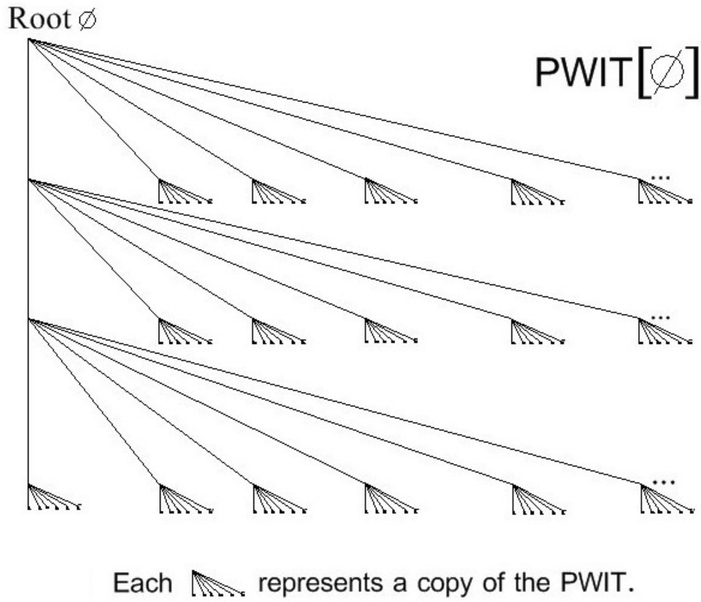

FIGURE 2. A Poisson Weighted Infinite Tree rooted at $\varnothing$. Weights on offspring edges from different vertices are determined by independent Poisson processes of intensity λ_Π.

Next, we create a new independent $\mathrm{PWIT}(\lambda_\Pi)[\infty_u]$ for each vertex u of each $\mathrm{PWIT}(\lambda_\Pi)[\infty_v]$ and draw an edge with weight $1/\sigma$ between u and ∞_u. We continue this procedure ad infinitum. If we also identify ∞_v with the integer 0 so that by concatenation, ∞_v is written $v0$, then we may write the vertex set of a $\mathrm{PWIST}(\sigma, \lambda_\Pi)$ as

$$\mathbb{N}_0^F := \bigcup_{g \geq 0} \mathbb{N}_0^g \tag{1.11}$$

where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and by concatenation we write $v = v_1 v_2 \cdots v_g \in \mathbb{N}_0^g$. As can be seen in Figure 3, edges with the weight $1/\sigma$ connect infinitely many PWITs with a backbone structure in order to form a PWIST.

Our next theorem justifies the choice (1.10) for the weight on the edge between v and ∞_v. Let us however attempt a brief heuristic explanation as to why this is the correct weight to assign to this edge. First of all, identify each weight with its absolute value so that all weights are thought of as *non-negative conductances*. Now, if $\sigma = 0$, then the connected graph containing the root $\varnothing$ is simply a $\mathrm{PWIT}(\lambda_\Pi)$ with the weights on edges representing non-negative conductances. If $\sigma > 0$, we use the interpretation that v and ∞_v are infinitely far apart, but also that there are infinitely many parallel edges (or a multi-edge) between v and ∞_v. Since distance is equivalent to resistance on electrical networks and resistance is the reciprocal of conductance, the conductance of each parallel edge is zero; however, their collective *effective* conductance is greater than 0, and in particular is of order σ. We can thus identify the multiple parallel edges with a *single* edge between v and ∞_v called a **cord to infinity** with effective resistance $1/\sigma$.

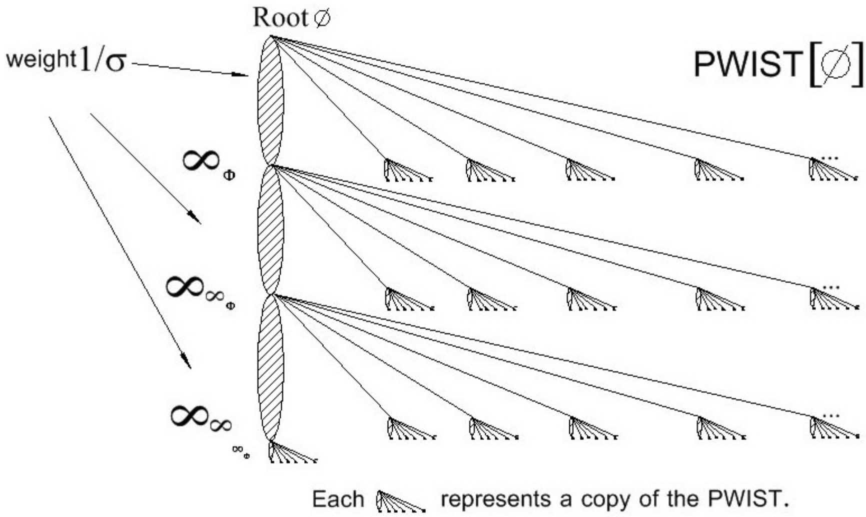

FIGURE 3. A Poisson Weighted Infinite Skeleton Tree rooted at $\varnothing$. The thick, shaded edges have a deterministic weight of $1/\sigma$. All other weights are determined randomly as before.

Let us now consider a random weighted adjacency matrix $\mathcal{C}_{G_n}$ associated to a complete **rooted geometric graph** (see Section 4 for definitions) $G_n = G_n[\varnothing] = (V_n, E_n, \mathcal{R}_n)$ where $V_n = \{1, \dots, n\}$ and $\mathcal{R}_n$ are the (possibly signed) random weights/lengths/resistances of the edges E_n. We refer to such a real-valued matrix as a *random conductance matrix* with entries given simply by the reciprocals of the signed resistances:

$$\mathcal{C}_{G_n}(j,k) := \frac{1}{\mathcal{R}_n(j,k)}. \tag{1.12}$$

When a sequence of random conductance matrices satisfies (1.4) or (1.5), it forms a Lévy-Khintchine random matrix ensemble.

This notion generalizes to a **random conductance operator** on $L^2(G_\infty) \equiv L^2(V_\infty)$ for an infinite weighted graph $G_\infty = (V_\infty, E_\infty, \mathcal{R}_\infty)$. Let the core $\mathcal{D}_{\mathrm{fs}} \subset L^2(V_\infty)$ be the set of vectors with finite support, i.e., all finite linear combinations of the basis vectors e_v which are 1 at v and 0 elsewhere. We consider the operator on $\mathcal{D}_{\mathrm{fs}}$ which is defined by

$$\mathcal{C}_{G_\infty}(u,v) = \langle e_u, \mathcal{C}_{G_\infty} e_v \rangle := \begin{cases} 1/\mathcal{R}_\infty(u,v) & \text{if } u \sim v, \\ 0 & \text{otherwise.} \end{cases} \tag{1.13}$$

This operator is closable as a graph in $L^2(V_\infty) \times L^2(V_\infty)$ since it is symmetric, i.e., Hermitian and densely defined [WS80, Theorem 5.4]. Abusing notation we also denote its unique closure by $\mathcal{C}_{G_\infty}$. In particular, we will see that the closure is self-adjoint. In the case where G_∞ is a PWIST(σ, λ_Π), by (1.8), the conductances are given by the points of a Poisson random measure with symmetric intensity $\Pi(dx)$ on $\mathbb{R}\backslash\{0\}$.

FIGURE 4. When there is no Lévy measure, the graph is just $\mathbb{N}$ with constant weights. It is well-known that the spectral measure at the root is a semicircle.

Now, recall [RS80, Secs. VII.2 and VIII.3] that the spectral measure μ_φ of a self-adjoint operator $\mathcal{C}$ associated to the vector φ is defined by the relation

$$\langle \varphi, f(\mathcal{C})\varphi \rangle =: \int_{\mathbb{R}} f(x)\mu_\varphi(dx), \quad \text{for bounded continuous } f. \tag{1.14}$$

Theorem 1.2 (LSD as the root spectral measure of a limiting operator)**.** *For any Lévy-Khintchine ensemble $(\mathcal{C}_n)_{n\in\mathbb{N}}$ with characteristics $(\sigma^2, 0, \Pi)$, the limiting empirical spectral measure $\mu_{\mathcal{C}_\infty}$ of Theorem* 1.1 *is the expected spectral measure, at the root vector $e_\varnothing$, of a self-adjoint random conductance operator $\mathcal{C}_{G_\infty}$ on $L^2(\mathbb{N}_0^F)$ where G_∞ is a PWIST(σ, λ_Π).*

Remark. The above matrix ensembles can be decomposed by the Lévy-Itō decomposition into $(\mathcal{C}_n)_{n\in\mathbb{N}}$ and $(\mathcal{C}'_n)_{n\in\mathbb{N}}$, which are independent with characteristics $(0, 0, \Pi)$ and $(\sigma^2, 0, 0)$. The sequence $(\mathcal{C}_n + \mathcal{C}'_n)_{n\in\mathbb{N}}$ then has characteristics $(\sigma^2, 0, \Pi)$. One approach is to try to generalize Voiculescu's asymptotic freeness theorem to establish the above result. However, we have been unable to do so due to the randomness of the PWIT associated to the Lévy measure Π (if the graphs were deterministic, one could use the approach of [ALS07]).

The following result is an application of the resolvent identity, and it may be used in conjunction with Theorem 1.2 to further analyze $\mu_{\mathcal{C}_\infty}$. It can be viewed as an operator version of the Schur complement formula.

Proposition 1.3 (Recursive distributional equation)**.** *Suppose that G_∞ is a PWIST(σ, λ_Π). For all $z \in \mathbb{C}_+$ the random variable*

$$R_{\varnothing\varnothing}(z) := \langle e_\varnothing, (\mathcal{C}_{G_\infty} - zI)^{-1} e_\varnothing \rangle$$

satisfies $R_{\varnothing\varnothing}(-\bar{z}) = -\bar{R}_{\varnothing\varnothing}(z)$ and the recursive distributional equation (RDE)

$$R_{\varnothing\varnothing}(z) \stackrel{d}{=} -\left(z + \sigma^2 R_{00}(z) + \sum_{k\in\mathbb{N}} |\mathcal{C}(k)|^2 R_{kk}(z) \right)^{-1} \tag{1.15}$$

where for all $k \geq 0$, R_{kk} has the same distribution as $R_{\varnothing\varnothing}$ and $\{\mathcal{C}(k)\}_{k\in\mathbb{N}}$ are the points of an independent Poisson random measure with intensity $\Pi(dx)$ *on* $\mathbb{R}\backslash\{0\}$.

Remark. For an example of how the above proposition may be used, consider Wigner matrices with i.i.d. entries with possibly infinite second moments, but normalized to be in the Gaussian domain of attraction. In this case, the Lévy measure Π is trivial and the PWIST$(\sigma, 0)$ is just $\mathbb{N}$ rooted at 1; see Figure 4.

Since the edge-weights of the limiting graph are non-random, a simple argument shows (see (4.6) below) that the resulting recursive equation is the Cauchy-Stieltjes transform (see (4.5)) of Wigner's semicircle law:

$$R_{\varnothing\varnothing}(z) = S_{\mu_{sc}}(z) = -\left(z + \sigma^2 S_{\mu_{sc}}(z)\right)^{-1}.$$

The rest of the paper is organized as follows. In the next section, we introduce a replacement procedure which creates a new sequence of matrices by modifying a given Lévy-Khintchine ensemble. This modification replaces complex values with real values and also embodies our notion of "cords to infinity". It is the key procedure which allows us to generalize PWITs to PWISTs. In Section 3, the moment method is used to prove a weak version of Theorem 1.1 in the case that the Lévy measure Π has bounded support. The main point of Section 3, however, is to show that the limiting root spectral measure of a Lévy-Khintchine ensemble is invariant under the replacement procedure of Section 2 (in preparation for proofs of the main results). In Section 4, we precisely define local weak convergence and present an adaptation of the arguments of [BCC11a]. In particular, we show that the local weak convergence argument proves Theorem 1.2 for real Lévy-Khintchine ensembles with $\sigma = 0$. Finally, in Section 5, we combine the arguments of Sections 3 and 4 to prove the main results in the general case. In the appendix we gather some known results which are needed along the way.

2. A replacement procedure for cords to infinity

In this section, we define an important sequence of *modified* matrices $(\mathcal{C}_n^\sigma)_{n\in\mathbb{N}}$ which play a key role in the proofs of the main results. In particular, these matrices are modifications of a Lévy-Khintchine ensemble $(\mathcal{C}_n)_{n\in\mathbb{N}}$ under a certain replacement procedure which we describe below.

For $h > 0$, by (1.4) and Proposition A.1 we have that as $n \to \infty$,

$$\sum_{k=1}^{n} \pm|\mathcal{C}_n(1,k)|1_{\{|\mathcal{C}_n(1,k)|\le h\}}$$

converges in distribution to $ID(\sigma_h^2, 0, \Pi_h)$ where the $\pm$ signs are chosen using independent Rademacher variables (independent also from $\mathcal{C}_n$), and

$$\begin{aligned} \sigma_h^2 &:= \sigma^2 + \int_{|x|\le h} x^2\,\Pi(dx) \quad \text{ and} \\ \Pi_h(dx) &:= 1_{(-\infty,-h]\cup[h,\infty)}(x)\Pi(dx). \end{aligned}$$

By a diagonalization argument, we may choose a sequence of positive numbers $h_n \to 0$ such that we get the following weak convergence to a Gaussian:

$$\sum_{k=1}^{n} \pm|\mathcal{C}_n(1,k)|1_{\{|\mathcal{C}_n(1,k)|\le h_n\}} \Rightarrow \mathcal{N}(0,\sigma^2).$$

In particular, as $h_n \to 0$,

$$\begin{aligned} &\lim_{n\to\infty}\sum_{k=2}^{n} \mathbf{E}\left(|\mathcal{C}_n(1,k)|^2 1_{\{|\mathcal{C}_n(1,k)|\le h_n\}}\right) \\ (2.1)\qquad &= \lim_{n\to\infty} n\mathbf{E}\left(|\mathcal{C}_n(1,2)|^2 1_{\{|\mathcal{C}_n(1,2)|\le h_n\}}\right) = \sigma^2. \end{aligned}$$

Our replacement procedure is as follows. For all entries such that $|\mathcal{C}_n(j,k)| > h_n$ as well as for all diagonal entries $\mathcal{C}_n(j,j)$, we set $\mathcal{C}_n^\sigma(j,k) := \pm|\mathcal{C}_n(j,k)|$ where the signs $\pm$ are given by independent Rademacher variables on the upper triangle and determined on the lower triangle to preserve self-adjointness. However, the entries in positions (j,k), $j \ne k$, in the matrix $\mathcal{C}_n^\sigma$ which satisfy the condition $|\mathcal{C}_n(j,k)| \le h_n$ will remain blank for now and will be assigned values that are either 0 or σ.

We next describe how to fill in blank entries. We first need to determine the order of the rows (and columns to preserve self-adjointness) by which we fill in the blanks. Recall that $\mathcal{C}_n$ determines a geometric graph, rooted at 1, with edge-weights given by $1/\mathcal{C}_n(j,k)$ as in (1.12). Let α be the permutation of $\{1,\dots,n\}$ such that $\alpha(i)$ is the ith closest vertex from the root 1 using the distance

$$d(u,v) := \inf_{\gamma \text{ connects } u,v} \sum_{e\in\gamma} |1/\mathcal{C}_n(e)| \tag{2.2}$$

where the infimum is over all paths γ which connect vertices u and v by a sequence of edges and e stands for one of these edges. If j and k are at equal distance from the root 1, we break ties by deeming j "closer" to the root whenever $j<k$. We now fill in blank entries according to the order determined by the (random) permutation α. For instance, we fill in blanks in row 1 first since $\alpha(1)=1$ (the root is always closest to itself). Next we fill in blank entries in the row $\alpha(2)$, then row $\alpha(3)$, etc.

The procedure for filling in blank entries in row $j=\alpha(i)$ is as follows, starting with row $1=\alpha(1)$. Out of all k satisfying

$$|\mathcal{C}_n(j,k)| \le h_n, \quad k\ne j, \tag{2.3}$$

choose one uniformly at random and set this entry, in $\mathcal{C}_n^\sigma$, to σ. Set other blank entries in row j, satisfying (2.3), to zero in the matrix $\mathcal{C}_n^\sigma$. This completes the filling of row j of $\mathcal{C}_n^\sigma$, and we use the symmetry condition $\mathcal{C}_n^\sigma(j,k)=\mathcal{C}_n^\sigma(k,j)$ to fill in blank entries in the column j.

When row and column $j=\alpha(i)$ are completely filled, we repeat the procedure on row and column $\alpha(i+1)$. We continue the replacement procedure described in the previous paragraph until all blank entries have been filled; then we call $(\mathcal{C}_n^\sigma)_{n\in\mathbb{N}}$ the **modified sequence of matrices**.

3. The moment method

In this section, we use the moment method to prove a *convergence in expectation*[2] version of Theorem 1.1 in the case where there exists an almost sure bound $0<\tau<\infty$ on the entries of the Lévy-Khintchine ensemble $(\mathcal{C}_n)_{n\in\mathbb{N}}$,

$$|\mathcal{C}_n(1,2)| \le \tau \quad \text{for all } n. \tag{3.1}$$

In particular, using the associated Poisson approximation for the distribution of $\mathcal{C}_n(1,2)$ (see [Kal02, Corollary 15.16]) one sees that Π must be supported on $[-\tau,\tau]$.

Let

$$M_p(\mu) := \int_{\mathbb{R}} x^p \mu(dx)$$

be the pth moment of the measure μ. The moment method in this section consists of showing

$$\lim_{n\to\infty} M_p(\mathbf{E}\mu_{\mathcal{C}_n}) = M_p(\mathbf{E}\mu_{\mathcal{C}_\infty}), \quad \text{for all } p\in\mathbb{N}, \tag{3.2}$$

and then verifying that the moments $M_p(\mathbf{E}\mu_{\mathcal{C}_\infty})$ determine $\mathbf{E}\mu_{\mathcal{C}_\infty}$. However, the main result of this section is the following important consequence of such a verification. Recall from (1.14) the notion of a spectral measure associated to a vector.

[2]See [Tao12, Remark 2.4.1] for a definition and short discussion of this type of convergence.

Proposition 3.1 (Invariance of expected LSD under replacement procedure)**.** *If the expected limiting empirical spectral measure for a Lévy-Khintchine ensemble* $(\mathcal{C}_n)_{n\in\mathbb{N}}$ *exists and is determined by its moments, then it is equal to the limiting expected spectral measure associated to* e_1 *(the first vector of the standard basis) for any modified sequence* $(\mathcal{C}_n^\sigma)_{n\in\mathbb{N}}$.

Proof. A standard argument (see [AGZ10, Ch. 2] or [Tao12, Sec. 2.3.4] for details) shows that the p-moments are given by

$$\begin{aligned} M_p(\mathbf{E}\mu_{\mathcal{C}_n}) &= \mathbf{E}\frac{1}{n}\operatorname{tr}(\mathcal{C}_n^p) \\ &= \sum_{j_2,\dots,j_p=1}^{n} \mathbf{E}(\mathcal{C}_n(1,j_2)\mathcal{C}_n(j_2,j_3)\cdots\mathcal{C}_n(j_p,1)) \end{aligned} \tag{3.3}$$

where we have set $j_1 = 1$ by exchangeability. The ordered listings of subscript pairs

$$\big((1,j_2)(j_2,j_3),\dots,(j_p,1)\big)_{j_2,\dots,j_p=1}^{n}$$

are viewed as distinct paths of length p which start and end at 1 in the complete graph on $\{1,\dots,n\}$, with edges having orientations and with the possibility that edges are crossed multiple times. These paths are called *cycles rooted at* 1.

We now make some preliminary observations in order to rewrite (3.3) as (3.9). The expression of the pth moment in (3.9) below allows us to then prove the result.

By Proposition 4 in [Zak06], in the limit as $n\to\infty$, the only cycles that *contribute to the limiting sum* on the right side of (3.3) are "trees" in the following sense. For a given *contributing* term, if the oriented edge (j_k, j_{k+1}) is crossed $q = q(k)$ times, then it must also be crossed q times in the opposite orientation. Thus, for each k there is a corresponding $k' \neq k$ such that

$$\mathcal{C}_n(j_k, j_{k+1}) = \overline{\mathcal{C}_n(j_{k'}, j_{k'+1})}, \quad j_k = j_{k'+1},\ j_{k+1} = j_{k'}. \tag{3.4}$$

Moreover, the partition of $\{1,\dots,p\}$ which pairs each k with its corresponding k' must be a *non-crossing pair partition* (see [NS06] for details). In particular, p must be even in order to have a non-trivial moment.

If $\mathcal{C}_n(j_k, j_{k+1})$ appears $q = q(k)$ distinct times in a given term, then its conjugate (or reversed edge from $j_{k'}$ to $j_{k'+1}$) also appears $q = q(k')$ distinct times. Using independence and exchangeability, each term of the sum in (3.3) takes the form

$$\mathbf{E}|\mathcal{C}_n(1,2)|^{2q_1}\mathbf{E}|\mathcal{C}_n(1,2)|^{2q_2}\cdots\mathbf{E}|\mathcal{C}_n(1,2)|^{2q_\ell} \tag{3.5}$$

where $2q_1 + \cdots + 2q_\ell = p$.

Fix the value of j_2 and consider a cycle rooted at 1 corresponding to a term in the sum (3.3) such that $(1, j_2)$ is crossed $q = q(1)$ times in each direction for a total of $2q$ times. Removing these $2q$ edges from our cycle leaves us with several sub-cycles. These sub-cycles can be permuted and then concatenated to form two sub-cycles L and $\tilde{L}$ rooted at $L_1 := j_2$ and $\tilde{L}_1 := 1$ which avoid the edges $(1, j_2)$ and $(j_2, 1)$ (one or both of the cycles may be trivial); see Figure 5.

Write $\mathbb{L}(j_2, q)$ for the set of all pairs of cycles $(L, \tilde{L})$ which are possible, where, in particular, different permutations/concatenations leading to the same L or $\tilde{L}$ are each listed separately in $\mathbb{L}(j_2, q)$; i.e., L and $\tilde{L}$ remember their original sub-cycle structure. Also, let $s, \tilde{s}$ be the lengths of $L, \tilde{L}$ so that $s + \tilde{s} = p - 2q$, and write

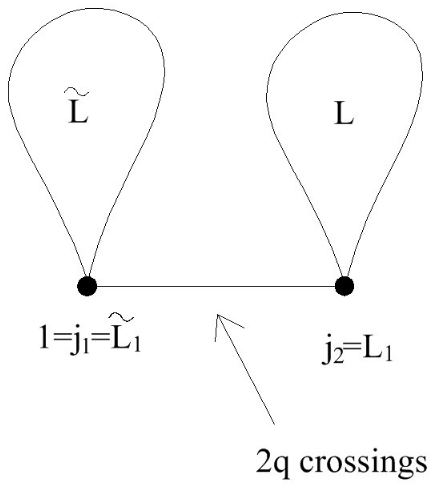

FIGURE 5

$L \equiv ((L_1, L_2), \ldots, (L_s, L_1))$ and similarly for $\tilde{L}$. Discarding some terms which do not contribute to the limiting sum, we have that (3.3) can be rewritten as

$$\sum_{q=1}^{p/2} \sum_{j_2=2}^{n} \sum_{(L,\tilde{L}) \in \mathbb{L}(j_2,q)} \mathbf{E}|\mathcal{C}_n(1,j_2)|^{2q} \mathbf{E}(\mathcal{C}_n(L_1,L_2) \cdots \mathcal{C}_n(L_s,L_1)$$

$$\mathcal{C}_n(\tilde{L}_1,\tilde{L}_2) \cdots \mathcal{C}_n(\tilde{L}_{\tilde{s}},\tilde{L}_1))$$

$$(3.6) \qquad = \sum_{q=1}^{p/2} \sum_{j_2=2}^{n} \{\mathbf{E}\left[|\mathcal{C}_n(1,j_2)|^{2q}(1_{\{|\mathcal{C}_n(1,j_2)| \le h_n\}} + 1_{\{|\mathcal{C}_n(1,j_2)| > h_n\}})\right]$$

$$\times \sum_{L,\tilde{L} \in \mathbb{L}(j_2,q)} \mathbf{E}(\mathcal{C}_n(L_1,L_2) \cdots \mathcal{C}_n(L_s,L_1) \mathcal{C}_n(\tilde{L}_1,\tilde{L}_2) \cdots \mathcal{C}_n(\tilde{L}_{\tilde{s}},\tilde{L}_1))\}.$$

Recall from (2.1) that for $\epsilon > 0$, we may find N such that $n \ge N$ implies

$$(3.7) \qquad n\mathbf{E}\left(|\mathcal{C}_n(1,2)|^2 1_{\{|\mathcal{C}_n(1,2)| \le h_n\}}\right) \le \sigma^2 + \epsilon,$$

which in turn implies

$$(3.8) \qquad n\mathbf{E}\left(|\mathcal{C}_n(1,2)|^{2q} 1_{\{|\mathcal{C}_n(1,2)| \le h_n\}}\right) \le h_n^{2q-2}(\sigma^2 + \epsilon).$$

To see this, note that a distribution satisfying (3.7) with maximum $2q$th moment is given by $\mathcal{C}_n(1,2) = \pm h_n$ with probability $\frac{\sigma^2+\epsilon}{nh_n^2}$ and $\mathcal{C}_n(1,2) = 0$ otherwise. Since $h_n \to 0$ we see that (3.8) goes to zero for $q > 1$. Multiplying out the right side of (3.6), we have that any term with a factor of $1_{\{|\mathcal{C}_n(1,j_2)| \le h_n\}}$ must have $q(1) = 1$ in order to contribute to the limiting sum. It should perhaps be noted that since we must have that $q(1) = 1$, for terms with a factor of $1_{\{|\mathcal{C}_n(1,j_2)| \le h_n\}}$, the permuting/concatenating of sub-cycles which form L and $\tilde{L}$ is not needed.

We now write

$$\mathcal{C}_n(j_k,j_{k+1}) = \mathcal{C}_n(j_k,j_{k+1})(1_{\{|\mathcal{C}_n(j_k,j_{k+1})| \le h_n\}} + 1_{\{|\mathcal{C}_n(j_k,j_{k+1})| > h_n\}})$$

for all factors in all terms of (3.3) and (3.6). For fixed $j_2 \equiv L_1$, we will categorize terms containing the factor $1_{\{|\mathcal{C}_n(1,j_2)| \le h_n\}}$ by the number of other factors in the

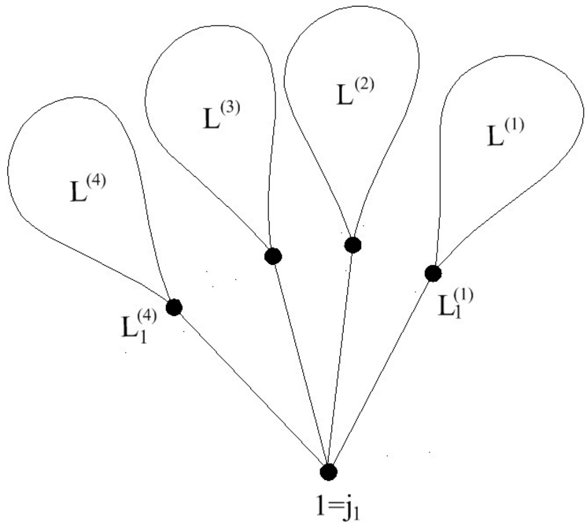

FIGURE 6

term which are of the form

$$|\mathcal{C}_n(1,j_k)|^2 1_{\{|\mathcal{C}_n(1,j_k)|\le h_n\}} \quad \text{for any} \quad k.$$

There are at most $p/2$ such factors. In particular, consider terms of (3.6) which include the factor $|\mathcal{C}_n(1,\tilde{L}_2)|^2 1_{\{|\mathcal{C}_n(1,\tilde{L}_2)|\le h_n\}}$; see Figure 6. The above procedure on our cycle rooted at 1 is repeated on the cycle $\tilde{L} =: \tilde{L}^{(1)}$, which is also rooted at 1. In other words, we fix the value of $\tilde{L}_2$ and consider cycles such that the edge $(1,\tilde{L}_2)$ is crossed exactly once in each direction. We remove these 2 edges from $\tilde{L}$ leaving us with two sub-cycles $L^{(2)}$ and $\tilde{L}^{(2)}$ rooted at $L_1^{(2)} := \tilde{L}_2^{(1)}$ and $\tilde{L}_1^{(2)} := 1$. We then repeat the procedure on the cycle $\tilde{L}^{(2)}$ to get two more sub-cycles $L^{(3)}$ and $\tilde{L}^{(3)}$ and continue this process until all edges of the form $(1,\cdot)$ or $(\cdot,1)$ are "removed". Thus, for any term containing $1_{\{|\mathcal{C}_n(1,j_2)|\le h_n\}}$ there is a corresponding list of cycles $(L^{(1)}, L^{(2)}, \ldots, L^{(M)})$. The list is of length $M \le p/2$ where M depends on the term (thus terms are categorized by their associated M value), and each cycle in the list is rooted at a different vertex in $\{2,\ldots,n\}$. Let $\mathbb{L}_M(n)$ denote the set of all possible lists of cycles of length M.

Finally, recalling that $L_1^{(1)} \equiv j_2$, the sum of all contributing terms in (3.6) can be written in the form

$$\sum_{M=0}^{p/2} \sum_{\mathbb{L}_M(n)} \prod_{i=1}^{M} \Bigg(\mathbf{E}\left[|\mathcal{C}_n(1,L_1^{(i)})|^2 1_{\{|\mathcal{C}_n(1,j_2)|\le h_n\}}\right] \\ \times \mathbf{E}\left[\prod_{i=1}^{M} \mathcal{C}_n(L_1^{(i)},L_2^{(i)})\cdots \mathcal{C}_n(L_{s^{(i)}}^{(i)},L_1^{(i)})\right]\Bigg).$$

Summing over the possible first coordinates of each cycle in the list of cycles, $L_1^{(i)} \in \{2, \ldots, n\}$, and taking the limit gives us

$$\lim_{n\to\infty} \sum_{M=0}^{p/2} \sum_{(L^{(1)},\ldots,L^{(M)})\in\mathbb{L}_M(n)} \sigma^{2M}\mathbf{E}\left[\prod_{i=1}^{M} \mathcal{C}_n(L_1^{(i)}, L_2^{(i)}) \cdots \mathcal{C}_n(L_{s^{(i)}}^{(i)}, L_1^{(i)})\right]. \tag{3.9}$$

Let $(\mathcal{C}_n^{\sigma,1})$ be matrices which are modified using only the first step of the replacement procedure, i.e., where only a single cord to infinity (from 1) has been substituted. Using the fact that

$$|\mathcal{C}_n(j_k, j_{k+1})| = |\mathcal{C}_n^{\sigma,1}(j_k, j_{k+1})| \text{ on the event } \{|\mathcal{C}_n(j_k, j_{k+1})| > h_n\},$$

a relatively straightforward calculation of $M_p(\mathbf{E}\mu_{\tilde{\mathcal{C}}_n^\sigma})$ using (3.3), also gives (3.9) by

(a) conditioning on the number of times that a given cycle rooted at 1 crosses the cord from 1 to infinity (in either direction) to be $2M$, and
(b) for a fixed set of loops $L^{(1)}, \ldots, L^{(M)}$ in (3.9) with different roots, one can identify their different roots with one single root. This single root should be thought of as the vertex at infinity which is connected to 1 by the cord in (a) above. One need only check that the two configurations of loops give the same value for the expression

$$\lim_{n\to\infty} \mathbf{E}\left[\prod_{i=1}^{M} \mathcal{C}_n(L_1^{(i)}, L_2^{(i)}) \cdots \mathcal{C}_n(L_{s^{(i)}}^{(i)}, L_1^{(i)})\right]. \tag{3.10}$$

There is a slight subtlety regarding the invariance of (3.10) under the identification of roots. The subtlety is that the dependence structure of edges crossed in (3.10) is changed under the identification of roots. However, note that we can approximate the Lévy measure by a sum of Dirac point measures, and without loss of generality, we will assume it has this form. Then, it turns out that the dependence structure of edges crossed in (3.10) does not affect the value of (3.10) since (i) the dependence structure only changes on the event that the various edges crossed have a common weight λ, and (ii) in this event, the $2q$th moment of λ times a Rademacher random variable is λ^{2q}. Thus, for example, the product of the variances of two independent λ-scaled Rademachers is exactly the fourth moment of a single λ-scaled Rademacher.

The proof of the theorem is now complete for the first step of the replacement procedure. Equivalence of moments for other steps in the replacement procedure follows similarly, and the rest of the proof is left as an exercise. □

Remark 3.2. When Π is trivial, all the q_i's in (3.5) are *all* equal to 2. This leads to the well-known fact that (3.3) is the number of Dyck words of length $2p$ which is just the pth Catalan number

$$c_p = \frac{(2p)!}{p!(p+1)!}.$$

We next have a result which relates the moments of the matrix entries to the moments of the Lévy measure. Both sets of moments are also related to the moments of the LSD using (3.5); moreover, together with the proposition below, (3.5) proves existence of the limit in (3.2).

Proposition 3.3 (Triangular array moments are related to Lévy measure moments)**.** *Suppose that* $\{\mathcal{C}(n,k), 1 \le k \le n\}_{n\in\mathbb{N}}$ *is a triangular array of random variables which are i.i.d. in each row and for which* $\sum_{k=1}^n |\mathcal{C}(n,k)|^2$ *converges weakly as* $n \to \infty$ *to an infinitely divisible law with subordinator characteristics* (σ^2, Π_s). *If the random variables are uniformly bounded,*

$$|\mathcal{C}(n,k)| \le \tau \quad \text{for all } n \text{ and } k, \tag{3.11}$$

then

$$\lim_{n\to\infty} n\mathbf{E}|\mathcal{C}(n,1)|^2 = \sigma^2 + M_1(\Pi_s)\,,$$

and for $p > 1$

$$\lim_{n\to\infty} n\mathbf{E}|\mathcal{C}(n,1)|^{2p} = M_p(\Pi_s).$$

Proof. Set $X_n := |\mathcal{C}(n,1)|^2$ with characteristic function φ_{X_n}. The characteristic function of

$$\lim_{n\to\infty} \sum_{k=1}^n |\mathcal{C}(n,1)|^2 \overset{d}{=} X$$

in (1.3) takes the form

$$\varphi_X(\theta) = \exp\left(i\theta\sigma^2 + \int_0^{\tau^2} (e^{i\theta x} - 1)\,\Pi_s(dx) \right), \tag{3.12}$$

and by convergence in distribution of the row sums and Lemma 5.8 in [Kal02],

$$\lim_{n\to\infty} n(\varphi_{X_n} - 1) = i\theta\sigma^2 + \int_0^{\tau^2} (e^{i\theta x} - 1)\,\Pi_s(dx)$$

uniformly in θ on compact subsets of $\mathbb{R}$. Since the $\{X_n\}$ are bounded and since Π_s has bounded support we may expand both sides in terms of power series and switch summations with integrals. This gives us

$$\lim_{n\to\infty} n \sum_{k\ge 1} \frac{(i\theta)^k \mathbf{E} X_n^k}{k!} = i\theta\sigma^2 + \sum_{k\ge 1} \int_0^{\tau^2} \frac{(i\theta x)^k}{k!}\,\Pi(dx) \tag{3.13}$$

uniformly on compact subsets, from which the lemma follows. ☐

To verify the "moment problem" required to use Proposition 3.1, we adapt arguments from [BG01, KSV04, Zak06]. Let Q_p be the set of $(q_1, \dots, q_\ell)$ such that $q_i \in \mathbb{N}$, $\sum_{i=1}^\ell q_i = p$, and

$$q_1 \ge q_2 \ge \cdots \ge q_\ell.$$

Also, fix a sequence of distinct colors $\{K_i\}_{i=0}^\infty$. We define $T((q_1, \dots, q_p))$ to be the number of colored rooted trees which satisfy the following:

- There are $p+1$ vertices.
- There are exactly q_i vertices of color K_i with the root being the only vertex of color K_0.
- If u and v are the same color, then the distance from u to the root is equal to the distance from v to the root.
- If u and v have the same color, then so do their parents.

Define

$$\mathcal{I}_{p,\ell} := \sum_{(q_1,\dots,q_\ell)\in Q_p} T((q_1, \dots, q_\ell)).$$

Proposition 3.4 (LSD determined by its moments). *Under assumption* (3.1),

$$M_{2p}(\mathbf{E}\mu_{\mathcal{C}_\infty}) \leq \tau^{2p} \sum_{\ell} \mathcal{I}_{p,\ell} \left(M_2(\Pi) + \Pi([-1,1]^c) + \sigma^2\right)^{\ell}, \tag{3.14}$$

and thus $\mathbf{E}\mu_{\mathcal{C}_\infty}$ *exists and is determined by its moments.*

Proof. By splitting the support of Π into $[-1,1]$ and its complement, note that $M_{2q}(\Pi) \leq M_2(\Pi) + \tau^{2q}\Pi([-1,1]^c)$. Also, without loss of generality, $\tau \geq 1$. We use Proposition 3.3 in conjunction with the argument of [Zak06, Theorem 2] (see also [BG01, Sec. 5.3] and [KSV04, Sec. IV]) to get

$$\begin{aligned}
&M_{2p}(\mathbf{E}\mu_{\mathcal{C}_\infty}) \\
&= \lim_{n\to\infty} \sum_{(q_1,\dots,q_\ell)\in Q_p} T((q_1,\dots,q_\ell)) n\mathbf{E}(|\mathcal{C}_n(1,2)|^{2q_1}) \cdots n\mathbf{E}(|\mathcal{C}_n(1,2)|^{2q_\ell}) \\
&\leq \sum_{(q_1,\dots,q_\ell)\in Q_p} T((q_1,\dots,q_\ell))(M_{2q_1}(\Pi)+\sigma^2)\cdots(M_{2q_\ell}(\Pi)+\sigma^2) \\
&\leq \tau^{2p} \sum_{(q_1,\dots,q_\ell)\in Q_p} T((q_1,\dots,q_\ell)) \left(M_2(\Pi) + \Pi([-1,1]^c) + \sigma^2\right)^{\ell} \\
&= \tau^{2p} \sum_{\ell} \mathcal{I}_{p,\ell} \left(M_2(\Pi) + \Pi([-1,1]^c) + \sigma^2\right)^{\ell}. \qquad (3.15)
\end{aligned}$$

Next, we use equation (9) in [BG01], which gives the bound

$$\mathcal{I}_{p,\ell} \leq c_p \mathcal{S}_{p,\ell} \tag{3.16}$$

(see also Proposition 10 in [Zak06]) where c_p is the pth Catalan number and

$$\mathcal{S}_{p,\ell} = \frac{1}{\ell!} \sum_{k=0}^{\ell} (-1)^{\ell-k} \binom{\ell}{k} k^{2p}$$

is a Stirling number of the second kind. By (3.15), (3.16), and Theorem 30.1 in [Bil86], $\mathbf{E}\mu_{\mathcal{C}_\infty}$ is determined by its moments if for any $R > 0$,

$$\frac{c_p}{(2p)!} \sum_{\ell=1}^{p} R^{\ell} \mathcal{S}_{p,\ell} \tag{3.17}$$

is $o(r^p)$ for some r as $p \to \infty$, and this is easily verified. For example Section 5.5 of [BG01] shows (3.17) is less than $(p^p + e^{R(p-1)})/(p!(p+1)!)$. $\square$

Remark 3.5. In [BG01], the lower bound $\mathcal{S}_{2p,\ell} \leq \mathcal{I}_{2p,\ell}$ was also established and used to show that the LSD has unbounded support (see also [Zak06, Proposition 12]). In our situation, this tells us that the Lévy-Khintchine ensembles for which the LSD has bounded support are precisely those with only a Wigner portion, i.e., those with characteristics of the form $(\sigma^2, 0, 0)$.

4. From local weak convergence to spectral convergence

In this section, to simplify things we restrict our attention to random conductance matrices $\mathcal{C}_n$ with *real* entries. The goal of this section is to present Theorem 4.2, which uses strong resolvent convergence to connect the notions of local weak convergence and convergence in distribution of the empirical spectral measures. Theorem 4.2 below is similar to [BCC11a, Theorem 2.2] (see also [BL10, BCC11b, BC12]), and its proof is an adaptation of the arguments there which treat the symmetric α-stable case:

$$(\sigma^2, 0, \Pi) = (0, 0, \operatorname{sign}(x)\alpha|x|^{-1-\alpha}dx).$$

Here we replace the α-stable Lévy measure with an arbitrary symmetric Lévy measure $\Pi(dx)$ on $\mathbb{R}\backslash\{0\}$. In particular, if one assumes self-adjointness of the limiting operator (which follows from Lemma 5.2 below), then the arguments in this section are enough to handle Theorem 1.2 in the case when $\sigma = 0$ and the entries are real.

Let us now present the precise notion of local weak convergence following the treatment in [AS04]. Let $G[\varnothing] = (V, E)$ be a $\varnothing$-rooted graph with vertex set V and edge set E both of which are at most countably infinite. Recall the distance function d defined in (2.2). The distance d naturally turns $G[\varnothing]$ into a metric space. We include $\pm\infty$ as a possible edge-weight where $\pm\infty$ is thought of as the same weight using the one-point compactification of $\mathbb{R}\backslash\{0\}$.

If $G[\varnothing]$ is connected and undirected and the edge-weight function $\mathcal{R}$ is such that for every vertex v and every $r < \infty$, the number of vertices within distance r of v is finite, then $G[\varnothing] = (V, E, \mathcal{R})$ is a **rooted geometric graph**. Henceforth all graphs will be rooted geometric graphs, and when they are rooted at the default root $\varnothing$, we may simply write G instead of $G[\varnothing]$. The set of all rooted geometric graphs is written $\mathcal{G}_\star$.

In the case that the range of $\mathcal{R}$ is positive and the underlying graph is a tree, we can interpret $\mathcal{R}$ as assigning resistances to edges. However, for technical reasons required by the proofs of our main results, *we allow $\mathcal{R}$ to take negative values.* The possibility of negative weights makes our treatment here differ slightly from [AS04]. But, using the modulus in (2.2) nevertheless permits us to reap the benefits of the metric of [AS04] on $\mathcal{G}_\star$.

Let $\mathcal{N}_{r,\varnothing}(G)$ be the r-neighborhood of $\varnothing$. This is the $\varnothing$-rooted subgraph of G formed by restricting the graph to the set of all vertices $v \in V$ such that $d(\varnothing, v) \le r$ and restricting to the set of edges that can be crossed by journeying at most distance r from the root $\varnothing$. We say r is a *continuity point* of G if there is no vertex of exact distance r from the root.

Definition 4.1 (The topology of $\mathcal{G}_\star$)**.** We say $(G_n = (V_n, E_n, \mathcal{R}_n))_{n\in\mathbb{N}}$ converges to $G = (V, E, \mathcal{R})$ in $\mathcal{G}_\star$ if for each continuity point r of G there is an n_r such that $n > n_r$ implies there exists a graph isomorphism

$$\pi_n : \mathcal{N}_{r,\varnothing}(G) \to \mathcal{N}_{r,\varnothing}(G_n)$$

which preserves the root and for which

$$\lim_{n\to\infty} \mathcal{R}_n(\pi_n^{-1}(u), \pi_n^{-1}(v)) = \mathcal{R}(u, v). \tag{4.1}$$

As noted in [AS04], the above convergence determines a topology which turns $\mathcal{G}_\star$ into a complete separable metric space. Using the usual theory of convergence in distribution, one can therefore say that a sequence of random rooted geometric graphs $(G_n)_{n\in\mathbb{N}} \subset \mathcal{G}_\star$, with distributions μ_n, converge weakly to $G \in \mathcal{G}_\star$ with distribution μ if for all bounded continuous $f : \mathcal{G}_\star \to \mathbb{R}$

$$\int_{\mathcal{G}_\star} f d\mu_n \to \int_{\mathcal{G}_\star} f d\mu. \tag{4.2}$$

Such weak convergence is called *local weak convergence.*

The following connection between local weak convergence and strong resolvent convergence was first noticed in [BL10] and [BCC11a] in the context of sparse matrices and heavy-tailed matrices, respectively (see [HO07] for related arguments).

Theorem 4.2 (Local weak convergence implies strong resolvent convergence)**.** *Let $(\mathcal{C}_{G_n})_{n\in\mathbb{N}}$, which are associated to $(G_n = (V_n, E_n, \mathcal{R}_n))_{n\in\mathbb{N}}$ as in (1.13), be essentially self-adjoint. Suppose that the graphs converge in the local weak sense to a tree $G = (V, E, \mathcal{R})$ with respect to the isomorphisms $(\pi_n)_{n\in\mathbb{N}}$ and that $\mathcal{C}_G$ is also essentially self-adjoint.*

If for each $u \in V$,

$$\lim_{\epsilon\searrow 0}\lim_{n\to\infty}\sum_{v\in V_n: v\sim\pi_n(u)} |\mathcal{C}_{G_n}(\pi_n(u), v)|^2 1_{\{|\mathcal{C}_{G_n}(\pi_n(u),v)|^2\le\epsilon\}} = 0 \ a.s., \tag{4.3}$$

then for all $z \in \mathbb{C}_+$, as $n \to \infty$:

$$\langle e_\varnothing, (\mathcal{C}_{G_n} - zI)^{-1} e_\varnothing\rangle \xrightarrow{w} \langle e_\varnothing, (\mathcal{C}_G - zI)^{-1} e_\varnothing\rangle. \tag{4.4}$$

Remark 4.3. By Proposition A.1, condition (4.3) simply says that $\sigma^2 = 0$ in (1.5).

Once one checks the local weak convergence of $(G_n[1])_{n\in\mathbb{N}}$ to a PWIT(λ_Π) and verifies self-adjointness, then the above result essentially handles the case where the Wigner component vanishes. Let us briefly outline this. First of all $\sigma = 0$ will imply condition (4.3). Next, recall that the Cauchy-Stieltjes transform (or simply Stieltjes transform) is defined as

$$S_\mu(z) := \langle \mu, (x - z)^{-1}\rangle = \int_{\mathbb{R}} \frac{\mu(dx)}{x - z}, \quad z \in \mathbb{C}\backslash\mathbb{R}. \tag{4.5}$$

Recall from (1.6) that $\mu_{\mathcal{C}_n}$ is the empirical spectral measure of $\mathcal{C}_n$. Using the fact that entries in $\mathcal{C}_n$ are i.i.d.,

$$S_{\mathbf{E}\mu_{\mathcal{C}_n}}(z) = \mathbf{E}S_{\mu_{\mathcal{C}_n}}(z) = \frac{1}{n}\mathbf{E}\mathrm{tr}(\mathcal{C}_n - zI)^{-1} = \mathbf{E}(\mathcal{C}_n - zI)^{-1}(1, 1). \tag{4.6}$$

Therefore, by (4.6), the above theorem, and a bound on the modulus of the Green's function

$$|(\mathcal{C}_n - zI)^{-1}(1, 1)| \le (\Im z)^{-1}$$

for $z \in \mathbb{C}\backslash\mathbb{R}$, we obtain convergence of $(S_{\mathbf{E}\mu_{\mathcal{C}_n}})_{n\in\mathbb{N}}$ to $S_{\mathbf{E}\mu_{\mathcal{C}_{G_\infty}}}$ where G_∞ is a PWIT(λ_Π). Lemma A.2, which tells us that the Cauchy-Stieltjes transform determines the LSD, then implies weak convergence of the expected ESDs (since $e_\varnothing$ has unit norm, the limit is a probability measure). A concentration of measure argument from [GL09], Lemma 5.1 below, extends this to a.s. weak convergence for the random ESDs.

For the proof of Theorem 4.2 we need a lemma which appears as Theorem VIII.25 in [RS80]. We state it without proof.

Lemma 4.4 (Strong resolvent convergence characterization)**.** *Suppose $\mathcal{C}_n$ and $\mathcal{C}_\infty$ are self-adjoint operators on $L^2(V)$ with a common core $\mathcal{D}$ (for all n and ∞). If*

$$\mathcal{C}_n\varphi \to \mathcal{C}_\infty\varphi \quad in \ \ L^2(V),$$

for each $\varphi \in \mathcal{D}$, then $\mathcal{C}_n$ converges to $\mathcal{C}_\infty$ in the strong resolvent sense.

Proof of Theorem 4.2. To match the setting for which we employ this theorem, let the vertex set of G_n be a subset of $\mathbb{N}$ and the vertex set of G be $\mathbb{N}^F$. By assumption, the local weak limit of $(G_n)_{n\in\mathbb{N}}$ is the tree G, with respect to the mappings

$$\pi_n : \mathbb{N}^F \to V_n \subset \mathbb{N}, \tag{4.7}$$

which are injective when restricted to some random subset of $\mathbb{N}^F$ with the same cardinality as V_n. By the Skorokhod representation theorem we will in fact assume that this weak convergence in $\mathcal{G}_\star$ is almost sure convergence on some probability space. Note that when the sequence $(\mathcal{C}_{G_n})_{n\in\mathbb{N}}$ is a sequence of $n\times n$ Lévy-Khintchine matrices, one may set $V_n = \{1,\dots,n\}$; however, in general V_n may even be infinite (in which case it is just $\mathbb{N}$).

Since $\mathbb{N}^F$ is countable we can fix some bijection with $\mathbb{N}$ and think of V_n as a subset of $\mathbb{N}^F$. In this case, the maps π_n can each be extended to (random) bijections from $\mathbb{N}^F$ to $\mathbb{N}$, and abusing notation we write π_n for these extensions. The essentially self-adjoint operators $\mathcal{C}_{G_n}$ extend to self-adjoint operators on $L^2(\mathbb{N}^F)$, using the core $\mathcal{D}_{\mathrm{fs}}$ consisting of vectors with finite support, by defining

$$\langle e_u, \mathcal{C}_{G_n} e_v\rangle := \begin{cases} \mathcal{C}_{G_n}(\pi_n(u),\pi_n(v)) & \text{if } \{\pi(u),\pi(v)\}\subset V_n, \\ 0 & \text{otherwise.}\end{cases} \tag{4.8}$$

By assumption, the closure of $\mathcal{C}_G$ is also self-adjoint using the core $\mathcal{D}_{\mathrm{fs}}$. Again abusing notation, we identify this closure with $\mathcal{C}_G$.

By local weak convergence and Skorokhod representation, we have that almost surely

$$\langle e_u, \mathcal{C}_{G_n} e_v\rangle \to \langle e_u, \mathcal{C}_G\, e_v\rangle. \tag{4.9}$$

By Lemma 4.4, we are left to show that

$$\sum_{u\in\mathbb{N}^F} |\langle e_u, \mathcal{C}_{G_n} e_v\rangle - \langle e_u, \mathcal{C}_G\, e_v\rangle|^2 \to 0$$

almost surely, as $n\to\infty$. This follows from the Vitali convergence theorem since (4.9) provides almost sure convergence and (4.3) provides uniform square integrability. $\square$

A common tool for showing local weak convergence is the following lemma about Poisson random measures which is similar to [Ste02, Lemma 4.1].

Lemma 4.5 (Convergence to a Poisson random measure)**.** *Suppose* $\{\mathcal{C}(n,k), 1\le k\le n\}_{n\in\mathbb{N}}$ *is a triangular array of real random variables which are i.i.d. in each row and for which* $\sum_{k=1}^n \mathcal{C}(n,k)$ *converges in law, as* $n\to\infty$*, to an* $ID(\sigma^2,b,\Pi)$ *random variable. Then as* $n\to\infty$*,*

$$\sum_{k=1}^n \delta_{\mathcal{C}(n,k)}$$

converge vaguely, as measures on $\mathbb{R}\backslash\{0\}$*, to a Poisson random measure* η *with intensity* $\mathbf{E}\eta = \Pi$*.*

Proof of Lemma 4.5. Note that any Lévy measure Π is also a Radon measure on $\mathbb{R}\backslash\{0\}$. Even though there is a possible singularity at 0, this is no concern since

$0 \notin \mathbb{R}\backslash\{0\}$. Therefore, by the basic convergence theorem of empirical measures to Poisson random measures (see Theorem 5.3 in [Res07]) we need only check that

$$n\mathbf{P}(\mathcal{C}_n(1,2) \in \cdot) \overset{\text{vag}}{\to} \Pi$$

vaguely as measures on $\mathbb{R}\backslash\{0\}$. This follows from Proposition A.1. □

Remark 4.6. It is instructive to recognize that the Lévy characteristics σ^2 and b bear no influence on the above lemma and consequently bear no influence on local weak convergence of the associated graphs. This is because vague convergence pushes any affect they have to the point 0 which is not in $\mathbb{R}\backslash\{0\}$. This essentially tells us that b has no effect on the LSD, which is one reason why we were allowed to set it to 0 (this statement is made rigorous by Theorem 1.2). The same is not true for σ^2 since we must have $\sigma = 0$ in order to satisfy (4.3) (uniform square integrability) and therefore to use Theorem 4.2. However, after one applies the replacement procedure, (4.3) will once again be satisfied.

The following proposition utilizes Lemma 4.5 to show local weak convergence to a PWIST. It is a variant of results in [Ald92, Sec. 3] (see also [Ald01,Ste02,BCC11a]).

Proposition 4.7 (Local weak convergence to a PWIST)**.** *Let $G_n[1]$ have conductances $\{\mathcal{C}_n^\sigma(j,k)\}_{j,k}$ which are modified Lévy-Khintchine matrices with characteristics $(\sigma^2, 0, \Pi)$ (modified as in Section 2). Then the local weak limit of $(G_n[1])_{n\in\mathbb{N}}$ is a PWIST(σ, λ_Π).*

Proof. We follow [Ald92, Sec. 3] and [BCC11a, Sec. 2.5]. For each fixed realization of the $\{\mathcal{C}_n^\sigma(j,k), 1 \le j,k \le n\}$ we consider their reciprocals, i.e., the resistances

$$\{\mathcal{R}_n^\sigma(j,k), 1 \le j,k \le n\}.$$

For any $B, H \in \mathbb{N}$, such that

$$\sum_{\ell=0}^{H} B^\ell \le n,$$

we define a rooted geometric subgraph $G_n[1]^{B,H}$ of $G_n[1]$, whose vertex set is in bijection with a B-ary tree of depth H rooted at 1. Let $V_n := \{1, \dots, n\}$. The bijection provides a partial index of vertices of $G_n[1]$ as elements in

$$J_{B,H} = \bigcup_{\ell=0}^{H} \{1, \dots, B\}^\ell \subset \mathbb{N}_0^F \tag{4.10}$$

where the indexing is given by an injective map

$$\pi_n : J_{B,H} \to V_n.$$

The map π_n easily extends to a bijection from some subset of $\mathbb{N}_0^F$ to V_n and thus can be thought of as restrictions of the maps of (4.7).

We set $I_\varnothing = \{1\}$ and set the preimage/index of the root 1 to be $\pi_n^{-1}(1) = \varnothing$. We next index the B vertices in $V_n \backslash I_\varnothing$ which have the B smallest absolute values among $\{\mathcal{R}_n^\sigma(1,k)\}_{2\le k\le n}$. The kth smallest absolute value is given the index $\varnothing k = \pi_n^{-1}(v)$, $1 \le k \le B$. As in the discussion preceding (1.11), we have written the vector $\varnothing k$ using concatenation. Breaking ties using the lexicographic order, this defines the first generation.

Now let I_1 be the union of $I_\varnothing$ and the B vertices that have been selected. If $H \ge 2$, we repeat the indexing procedure for the vertex indexed by $\varnothing 1$ (the first child of $\varnothing$) on the set $V_n \setminus I_1$. We obtain a new set $\{11, \ldots, 1B\}$ of vertices sorted by their absolute resistances. We define I_2 as the union of I_1 and this new collection. Repeat the procedure for $\varnothing 2$ on $V_n \setminus I_2$ and obtain a new set $\{21, \ldots, 2B\}$. Continuing on through $\{B1, \ldots, BB\}$, we have constructed the second generation, at depth 2, and we have indexed a total of $(B^3-1)/(B-1)$ vertices. The indexing procedure is repeated through depth H so that $(B^{H+1}-1)/(B-1)$ vertices are sorted. Call this set of vertices $V_n^{B,H} = \pi_n(J_{B,H})$. The subgraph of $G_n[1]$ generated by the vertices $V_n^{B,H}$ is denoted $G_n[1]^{B,H}$ (by "generated" we mean that we include only edges with endpoints in the specified vertex set). It is the modification of $G_n[1]$ such that any edge with at least one endpoint in the complement of $V_n^{B,H}$ is given an infinite resistance. In $G_n[1]^{B,H}$, the elements of $\{u1, \ldots, uB\}$ are the children of u.

Note that while the vertex set $V_n^{B,H}$ has a natural tree structure, $G_n[1]^{B,H}$ is actually a subgraph of a complete graph which may not be a tree.

Let $G_\infty[\varnothing]$ be a PWIST(σ, λ_Π), or a PWIT(λ_Π) if $\sigma = 0$, and write $G_\infty[\varnothing]^{B,H}$ for the finite rooted geometric graph obtained by the sorting procedure just described. Namely, $G_\infty[\varnothing]^{B,H}$ consists of the sub-tree with vertices of the form $u \in J_{B,H}$, with resistances between these vertices inherited from the infinite tree. If an edge is not present in $G_\infty[\varnothing]^{B,H}$, we may think of it as being present but having infinite resistance.

Since the conductances $\{\mathcal{C}_n^\sigma(j,k)\}$ by definition are real with a symmetric distribution, we may without loss of generality replace $\sum_{j=1}^n \pm|\mathcal{C}_n(1,j)|$ with $\sum_{j=1}^n \mathcal{C}_n(1,j)$ in (1.4). We use Lemma 4.5 on the unmodified matrices (with real and symmetrically distributed entries) to conclude that $\sum_{k=1}^n \delta_{\mathcal{C}_n(1,k)}$ converges vaguely to a Poisson random measure with intensity Π. For $h_n \to 0$, the truncation $\mathcal{C}(n,k)1_{|\mathcal{C}(n,k)| \le h_n}$ does not affect this vague convergence. Note that besides the random resistances on edges given by the Poisson random measure, there is also one more non-random resistance given by the replacement procedure (for n large enough), and the value is always $1/\sigma$. It is easily verified that the property in (4.1) is satisfied by each edge (u,v) of the tree $G_\infty[\varnothing]$.

It remains to check that for each B and H, our maps π_n are graph isomorphisms for n large enough. In other words, we must check that for each edge in $G_\infty[\varnothing]^{B,H}$ with an infinite resistance, the corresponding edges of $\left(G_n[1]^{B,H}\right)_{n\in\mathbb{N}}$ (for n large enough) must have resistances which diverge to infinity. The divergence of these resistances to infinity follows from a standard coupling argument which shows that these resistances stochastically dominate i.i.d. variables with distribution $\mathcal{R}_n(1,2)$ which clearly diverges as $n \to \infty$ (see, for example, Lemma 2.7 in [BCC11a]). $\square$

5. Proofs of the main results

In the case that a Lévy-Khintchine ensemble $(\mathcal{C}_n)_{n\in\mathbb{N}}$ is real and has characteristics of the form $(0, 0, \Pi)$, then results of Section 4 (Theorem 4.2, Proposition 4.7) imply the existence of the LSD in expectation. On the other hand, if $|\mathcal{C}_n(1,2)|$ is a.s. uniformly bounded in n, Proposition 3.4 proves the existence of the LSD in expectation.

We turn now to the general assumptions of Theorems 1.1 and 1.2. Before proving the main results, we have three preliminary lemmas. Our first preliminary lemma allows us to extend from convergence in expectation to almost sure convergence. It is a concentration of measure result first noticed in [GL09, Theorem 1] and later in [BCC11b, Lemma C.2]. We state it here without proof.

Lemma 5.1 (Concentration for ESDs)**.** *Let $\mathcal{H}_n$ be an $n \times n$ Hermitian matrix such that $\{\mathcal{H}_n(j,k), j < k\}$ are independent. For every real-valued continuous $f(x)$ going to 0 as $x \to \pm\infty$ such that $\|f\|_{TV} \le 1$, and for every $t \ge 0$,*

$$\mathbf{P}\left(\left|\int_{\mathbb{R}} f\, d\mu_{\mathcal{H}_n} - \mathbf{E}\int_{\mathbb{R}} f\, d\mu_{\mathcal{H}_n}\right| \ge t\right) \le 2\exp\left(-nt^2/2\right).$$

The next lemma verifies the self-adjointness of PWISTs required to use Theorem 4.2.

Lemma 5.2 (Self-adjointness of PWIST operators)**.** *Suppose that $G_\infty[\varnothing] = (V_\infty, E_\infty, \mathcal{R}_\infty)$ is a PWIST(σ, λ_Π). Then the associated random conductance operator $\mathcal{C}_{G_\infty}$ on $L^2(V_\infty)$, as defined in* (1.13)*, is essentially self-adjoint.*

Proof. Denote the children of the root $\varnothing$ of a PWIST(σ, λ_Π) by $\mathbb{N} = \mathbb{N}[\varnothing]$ where they are ordered according to the absolute value of the conductances on the edges where the edge to 1 has the largest absolute conductance. For $\kappa > 0$ as chosen below, define the random variable

$$\tau_\varnothing := \inf\{J : \sum_{j=J}^{\infty} |\mathcal{C}_{G_\infty}(\varnothing, j)|^2 \le \kappa\}$$

and define the i.i.d. random variables $\{\tau_v\}$ similarly by considering the conductances on $\mathbb{N}[v]$ (in place of $\mathbb{N}[\varnothing]$). By the integrability conditions on Lévy measure Π, we may choose κ large enough so that $\mathbf{E}\tau_\varnothing < 1$. We may therefore employ the proof of Proposition A.2 in [BCC11a] to show that for any PWIST, $G_\infty = (V_\infty, E_\infty, \mathcal{R}_\infty)$, there is a constant $\kappa > 0$ and a sequence of connected finite increasing subsets $(V_n)_{n\in\mathbb{N}}$ whose union is V_∞, and such that for all n and $u \in V_n$

$$\sum_{v \notin V_n : v \sim u} |\mathcal{C}_{G_\infty}(u,v)|^2 < \kappa.$$

Finally, the existence of such a κ allows us to use Lemma A.3 in [BCC11a] to conclude that any PWIST is essentially self-adjoint. Thus its closure is self-adjoint. □

The final preliminary lemma, similar to arguments in [BAG08], is used to show that the truncation in (3.1) does not effect the LSD too much. For any truncation level $\tau > 0$, let $\tau\mathcal{C}_n$ be a matrix with entries given by

$$\tau\mathcal{C}_n(j,k) := \mathcal{C}_n(j,k) 1_{\{|\mathcal{C}_n(j,k)| \le \tau\}}. \tag{5.1}$$

Lemma 5.3 (Large deviation estimate for the rank of a truncation)**.** *For every $\epsilon > 0$ and $\tau \gg 0$ (large enough depending on ϵ), there is a $\delta_{\epsilon,\tau} > 0$ such that*

$$\mathbf{P}(\operatorname{rank}(\mathcal{C}_n - \tau\mathcal{C}_n)/n \ge \epsilon) \le \exp\left(-\delta_{\epsilon,\tau} n\right).$$

Proof. Fix $\epsilon > 0$ and consider τ large enough (specified below). Define the events

$$U_{jn} := \{\text{there exists } k \text{ such that } k > j \text{ and } |\mathcal{C}_n(j,k)| > \tau\},$$
$$L_{jn} := \{\text{there exists } k \text{ such that } k < j \text{ and } |\mathcal{C}_n(j,k)| > \tau\}$$

and note that

$$\operatorname{rank}(\mathcal{C}_n - \tau\mathcal{C}_n) \le \sum_{j=1}^{n} \left(1_{U_{jn}} + 1_{L_{jn}}\right). \tag{5.2}$$

We split rows of the matrix along the diagonal to handle the dependence (due to the self-adjointness requirement) among the indicator random variables:

$$\begin{aligned}
\mathbf{P}(\operatorname{rank}(\mathcal{C}_n - \tau\mathcal{C}_n) \ge 2n\epsilon) &\le \mathbf{P}\left(\sum_{j=1}^{n} 1_{U_{jn}} \ge n\epsilon\right) + \mathbf{P}\left(\sum_{j=1}^{n} 1_{L_{jn}} \ge n\epsilon\right) \\
&\le 2\mathbf{P}\left(\sum_{j=1}^{n} 1_{U_{jn}} \ge n\epsilon\right) \\
&\le 2\mathbf{P}\left(\sum_{j=1}^{n} 1^{(j)}_{U_{1n}} \ge n\epsilon\right) \qquad (5.3)
\end{aligned}$$

where $\{1^{(j)}_{U_{1n}}\}_{j=1}^{n}$ are independent copies of $1_{U_{1n}}$. The last step follows since the independent variables $\{1_{U_{jn}}\}_{j=1}^{n}$ are each stochastically dominated by $1_{U_{1n}}$.

Since the triangular array $\{\mathcal{C}_n(1,k), 1 \le k \le n\}_{n\in\mathbb{N}}$ satisfies (1.4),

$$\lim_{n\to\infty} \mathbf{P}(U_{1n}) = 1 - \exp\{-\Pi([\tau,\infty))\},$$

so we may choose τ large enough so that

$$\sup_n \mathbf{P}(U_{1n}) = p < \epsilon.$$

The lemma follows by applying a standard large deviation estimate for i.i.d. Bernoulli(p) random variables to the right side of (5.3). □

This last lemma is used in conjunction with a metric which is compatible with weak convergence. Let

$$\|f\|_{\mathcal{L}} := \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|} + \sup_x |f(x)|.$$

Lemma 2.1 in [BAG08] says the following variant of the Dudley distance gives a topology which is compatible with weak convergence:

$$d_1(\mu,\nu) := \sup_{\|f\|_{\mathcal{L}}\le 1, f\uparrow} \left| \int f\, d\mu - \int f\, d\nu \right|. \tag{5.4}$$

Moreover, Lidskii's estimate (see equation (8) in [BAG08]) implies

$$d_1(\mu_{\mathcal{C}_n}, \mu_{\tau\mathcal{C}_n}) \le \frac{\operatorname{rank}(\mathcal{C}_n - \tau\mathcal{C}_n)}{n}. \tag{5.5}$$

Proof of Theorems 1.1 *and* 1.2. Let us first state some simplifications for the task of showing that the LSD exists as a weak limit, almost surely.

First of all, by the Borel-Cantelli lemma and Lemma 5.1, it is enough to show weak convergence of $(\mathbf{E}\mu_{\mathcal{C}_n})_{n\in\mathbb{N}}$ to $\mathbf{E}\mu_{\mathcal{C}_\infty}$. Next, by exchangeability, it is enough to show weak convergence of the expected spectral measures associated to the basis vector e_1. Finally, by Lemma A.2, it is equivalent to show convergence of the Cauchy-Stieltjes transforms of these expected spectral measures for each $z \in \mathbb{C}_+$

(the limit will be a probability measure since it is the spectral measure associated to a unit vector).

Choose a Lévy-Khintchine ensemble $(\mathcal{C}_n)_{n\in\mathbb{N}}$ and let $(\tau_m)_{m\in\mathbb{N}}$ be a sequence of positive truncation levels which go to infinity. For each truncation level τ_m, consider a new sequence of matrices $(\tau_m\mathcal{C}_n)_{n\in\mathbb{N}}$ given by (5.1). Recalling our choice of h_n from Section 2, we also consider their modifications $(\tau_m\mathcal{C}_n^\sigma)_{n\in\mathbb{N}}$ (truncation occurs before modification).

Fix m. Each modified matrix sequence $(\tau_m\mathcal{C}_n^\sigma)_{n\in\mathbb{N}}$ satisfies the hypotheses of Proposition 4.7. Thus the associated graphs have a $\mathrm{PWIST}(\sigma,\lambda_\Pi^{(m)})$ as their local weak limit as $n\to\infty$, where $\lambda_\Pi^{(m)}$ is the intensity λ_Π restricted to the set

$$(-\infty,-1/\tau_m]\cup[1/\tau_m,\infty).$$

The closure of the associated limiting operator is self-adjoint by Lemma 5.2. Moreover, by Proposition A.1 and the properties of the replacement procedure, we have for each $j\in\mathbb{N}$ that

$$\lim_{\epsilon\searrow 0}\lim_{n\to\infty}\sum_{k=1}^{n}\mathrm{Var}\left(\mathcal{C}_n^\sigma(j,k)1_{\{|\mathcal{C}_n^\sigma(j,k)|\le\epsilon\}}\right)=0, \tag{5.6}$$

which is equivalent to (4.3) since the entries $\mathcal{C}_n^\sigma(j,k)1_{\{|\mathcal{C}_n^\sigma(j,k)|\le\epsilon\}}$ have a real distribution which is symmetric for $\epsilon<\sigma$ (the truncation τ_m is unnecessary due to $1_{\{|\mathcal{C}_n^\sigma(j,k)|\le\epsilon\}}$).

By the above considerations, we may use Theorem 4.2 and the argument below (4.6) to conclude Theorem 1.2 for each sequence $(\tau_m\mathcal{C}_n^\sigma)_{n\in\mathbb{N}}$. Thus, the expected LSD of $(\tau_m\mathcal{C}_n^\sigma)_{n\in\mathbb{N}}$, denoted by $\mathbf{E}\mu_{\tau_m\mathcal{C}_\infty^\sigma}$, is the expected spectral measure at $e_\varnothing$ for the self-adjoint random conductance operator $\tau_m\mathcal{C}_\infty^\sigma$ associated to a $\mathrm{PWIST}(\sigma,\lambda_\Pi^{(m)})$.

Now take the local weak limit of the $\mathrm{PWIST}(\sigma,\lambda_\Pi^{(m)})$ graphs as $m\to\infty$. Since these graphs are truncations of a $\mathrm{PWIST}(\sigma,\lambda_\Pi)$, it is clear that their local weak limit is just a $\mathrm{PWIST}(\sigma,\lambda_\Pi)$. We may therefore apply Theorem 4.2 once more to conclude that the expected spectral measures at $e_\varnothing$ of the $\mathrm{PWIST}(\sigma,\lambda_\Pi^{(m)})$ operators converge weakly to the expected spectral measure at $e_\varnothing$ of a $\mathrm{PWIST}(\sigma,\lambda_\Pi)$ operator, which we denote by $\mathbf{E}\mu_{\mathcal{C}_\infty^\sigma}$. Thus, for every $\epsilon>0$ we can choose m large enough so that

$$d_1(\mathbf{E}\mu_{\tau_m\mathcal{C}_\infty^\sigma},\mathbf{E}\mu_{\mathcal{C}_\infty^\sigma})<\epsilon/3$$

and so that $\delta_{\epsilon,\tau_m}>0$ in Lemma 5.3.

Equation (3.5) and Propositions 3.3 and 3.4 show that the expected LSD for $(\tau_m\mathcal{C}_n)_{n\in\mathbb{N}}$ exists. Moreover, by Proposition 3.1, it is equal to $\mathbf{E}\mu_{\tau_m\mathcal{C}_\infty^\sigma}$. So we may choose n_0 large enough so that $n>n_0$ implies

$$d_1(\mathbf{E}\mu_{\tau_m\mathcal{C}_n},\mathbf{E}\mu_{\tau_m\mathcal{C}_\infty^\sigma})<\epsilon/3.$$

Lemma 5.3 and (5.5) show that we may finally choose n_1 large enough so that $n>n_1$ implies

$$d_1(\mathbf{E}\mu_{\mathcal{C}_n},\mathbf{E}\mu_{\tau_m\mathcal{C}_n})<\epsilon/3.$$

Combining the above, we have for all $n>\max(n_0,n_1)$,

$$d_1(\mathbf{E}\mu_{\mathcal{C}_n},\mathbf{E}\mu_{\mathcal{C}_\infty^\sigma})<\epsilon,$$

and so the ESDs of $(\mathcal{C}_n)_{n\in\mathbb{N}}$ converge weakly in expectation (and thus a.s.) to $\mathbf{E}\mu_{\mathcal{C}_\infty^\sigma}$, which is the expected spectral measure at $e_\varnothing$ of $\mathcal{C}_\infty^\sigma$ associated to a $\mathrm{PWIST}(\sigma,\lambda_\Pi)$.

The claim that $\mu_{\mathcal{C}_\infty}$ has bounded support if and only if Π is trivial follows from the remark at the very end of Section 3. The claim that $\mu_{\mathcal{C}_\infty}$ is symmetric follows from the fact that replacing $\mathcal{C}_n$ with $-\mathcal{C}_n$ does not change the ESD. $\square$

Proof of Proposition 1.3. The proof is an application of the resolvent identity. For details, we refer the reader to Proposition 2.1 in [Kle98] or Theorem 4.1 in [BCC11a]. The latter proof works in our setting almost word for word. $\square$

Appendix A. Some additional tools

Infinite divisibility. The following important set of criteria for convergence to an infinitely divisible law with characteristics (σ^2, b, Π) was found independently by Doeblin and Gnedenko (see Corollary 15.16 in [Kal02]). For $0 < h < 1$, define

$$\sigma_h^2 := \sigma^2 + \int_{|x|\le h} x^2\,\Pi(dx) \quad \text{and} \quad b_h := b - \int_{h<|x|} \frac{x}{1+x^2}\,\Pi(dx).$$

Also, let $\overline{\mathbb{R}}$ be the one-point compactification of $\mathbb{R}$.

Proposition A.1 (Convergence criteria for triangular arrays)**.** *Suppose* $\{\mathcal{C}(n,k), 1 \le k \le n\}_{n\in\mathbb{N}}$ *is a triangular array of random variables such that each row consists of i.i.d. random variables. The sum*

$$\sum_{j=1}^{n} \mathcal{C}(n,j)$$

converges in distribution to an $ID(\sigma^2, b, \Pi)$ *random variable if and only if for any* $0 < h < 1$ *which is not an atom of* Π*,*

- $n\mathbf{P}(\mathcal{C}(n,1) \in \cdot) \stackrel{vag}{\to} \Pi$ *on* $\overline{\mathbb{R}}\backslash\{0\}$*,*
- $n\mathbf{E}\left(|\mathcal{C}(n,1)|^2 1_{\{|\mathcal{C}(n,1)|\le h\}}\right) \to \sigma_h^2$*,*
- $n\mathbf{E}\left(\mathcal{C}(n,1) 1_{\{|\mathcal{C}(n,1)|\le h\}}\right) \to b_h$*.*

From the Cauchy-Stieltjes transform to LSDs. The use of the Cauchy-Stieltjes transform in the context of random matrices dates back to Marčenko and Pastur [MP67]. Mainly, one obtains convergence of the ESDs of the random matrices $(\mathcal{C}_n)_{n\in\mathbb{N}}$ by showing convergence of the Cauchy-Stieltjes transforms $\left(S_{\mu_{\mathcal{C}_n}}(z)\right)_{n\in\mathbb{N}}$ as defined in (4.5). The lemma given here is taken from Section 2.4 in [AGZ10].

The Cauchy-Stieltjes transform is invertible: For any open interval I such that neither endpoint is an atom of μ

$$\mu(I) = \lim_{y\to 0} \frac{1}{\pi} \int_I \Im S_\mu(x+iy)\,dx. \tag{A.1}$$

This uniquely determines the measure μ so that one then obtains the following result.

Lemma A.2 (Weak convergence via Cauchy-Stieltjes transforms)**.** *Suppose* μ_n *is a sequence of probability measures on* $\mathbb{R}$ *and for each* $z \in \mathbb{C}_+$*,* $S_{\mu_n}(z)$ *converges to* $S(z)$ *which is the Cauchy-Stieltjes transform of some probability measure* μ*. Then* μ_n *converges weakly to* μ*.*

For the proof of this lemma, see [AGZ10, Theorem 2.4.4].

Acknowledgment

This project arose out of several discussions with Shannon Starr. The author thanks him for being a sounding board and for his many helpful discussions, insights, and suggestions.

References

[AGZ10] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni, *An introduction to random matrices*, Cambridge Studies in Advanced Mathematics, vol. 118, Cambridge University Press, Cambridge, 2010. MR2760897

[Ald92] David Aldous, *Asymptotics in the random assignment problem*, Probab. Theory Related Fields **93** (1992), no. 4, 507–534, DOI 10.1007/BF01192719. MR1183889

[Ald01] David J. Aldous, *The $\zeta(2)$ limit in the random assignment problem*, Random Structures Algorithms **18** (2001), no. 4, 381–418, DOI 10.1002/rsa.1015. MR1839499

[ALS07] Luigi Accardi, Romuald Lenczewski, and Rafał Sałapata, *Decompositions of the free product of graphs*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **10** (2007), no. 3, 303–334, DOI 10.1142/S0219025707002750. MR2354364

[AS04] David Aldous and J. Michael Steele, *The objective method: probabilistic combinatorial optimization and local weak convergence*, Probability on discrete structures, Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004, pp. 1–72, DOI 10.1007/978-3-662-09444-0_1. MR2023650

[BAG08] Gérard Ben Arous and Alice Guionnet, *The spectrum of heavy tailed random matrices*, Comm. Math. Phys. **278** (2008), no. 3, 715–751, DOI 10.1007/s00220-007-0389-x. MR2373441

[BC12] Charles Bordenave and Djalil Chafaï, *Around the circular law*, Probab. Surv. **9** (2012), 1–89, DOI 10.1214/11-PS183. MR2908617

[BCC11a] Charles Bordenave, Pietro Caputo, and Djalil Chafaï, *Spectrum of large random reversible Markov chains: heavy-tailed weights on the complete graph*, Ann. Probab. **39** (2011), no. 4, 1544–1590, DOI 10.1214/10-AOP587. MR2857250

[BCC11b] Charles Bordenave, Pietro Caputo, and Djalil Chafaï, *Spectrum of non-Hermitian heavy tailed random matrices*, Comm. Math. Phys. **307** (2011), no. 2, 513–560, DOI 10.1007/s00220-011-1331-9. MR2837123

[BG01] M. Bauer and O. Golinelli, *Random incidence matrices: moments of the spectral density*, J. Statist. Phys. **103** (2001), no. 1-2, 301–337, DOI 10.1023/A:1004879905284. MR1828732

[BG05] Florent Benaych-Georges, *Classical and free infinitely divisible distributions and random matrices*, Ann. Probab. **33** (2005), no. 3, 1134–1170, DOI 10.1214/009117904000000982. MR2135315

[BGGM13] Florent Benaych-Georges, Alice Guionnet, and Camille Male, *Central limit theorems for linear statistics of heavy tailed random matrices*, Comm. Math. Phys. **329** (2014), no. 2, 641–686, DOI 10.1007/s00220-014-1975-3. MR3210147

[Bil86] Patrick Billingsley, *Probability and measure*, 2nd ed., Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986. MR830424

[BJN$^+$07] Zdzisław Burda, Jerzy Jurkiewicz, Maciej A. Nowak, Gabor Papp, and Ismail Zahed, *Free random Lévy and Wigner-Lévy matrices*, Phys. Rev. E (3) **75** (2007), no. 5, 051126, 11, DOI 10.1103/PhysRevE.75.051126. MR2361817

[BL10] Charles Bordenave and Marc Lelarge, *Resolvent of large random graphs*, Random Structures Algorithms **37** (2010), no. 3, 332–352, DOI 10.1002/rsa.20313. MR2724665

[BS01] Itai Benjamini and Oded Schramm, *Recurrence of distributional limits of finite planar graphs*, Electron. J. Probab. **6** (2001), no. 23, 13, DOI 10.1214/EJP.v6-96. MR1873300

[BS10] Zhidong Bai and Jack W. Silverstein, *Spectral analysis of large dimensional random matrices*, 2nd ed., Springer Series in Statistics, Springer, New York, 2010. MR2567175

[CB94] P. Cizeau and J. P. Bouchaud, *Theory of Lévy matrices*, Physical Review E **50** (1994), no. 3, 1810.

[DS07] R. Brent Dozier and Jack W. Silverstein, *On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices*, J. Multivariate Anal. **98** (2007), no. 4, 678–694, DOI 10.1016/j.jmva.2006.09.006. MR2322123

[EK10] Noureddine El Karoui, *On information plus noise kernel random matrices*, Ann. Statist. **38** (2010), no. 5, 3191–3216, DOI 10.1214/10-AOS801. MR2722468

[FZ97] Joshua Feinberg and A. Zee, *Non-Hermitian random matrix theory: method of Hermitian reduction*, Nuclear Phys. B **504** (1997), no. 3, 579–608, DOI 10.1016/S0550-3213(97)00502-6. MR1488584

[GL09] Adityanand Guntuboyina and Hannes Leeb, *Concentration of the spectral measure of large Wishart matrices with dependent entries*, Electron. Commun. Probab. **14** (2009), 334–342, DOI 10.1214/ECP.v14-1483. MR2535081

[HO07] Akihito Hora and Nobuaki Obata, *Quantum probability and spectral analysis of graphs*, Theoretical and Mathematical Physics, Springer, Berlin, 2007. MR2316893

[Kal02] Olav Kallenberg, *Foundations of modern probability*, 2nd ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR1876169

[Kle98] Abel Klein, *Extended states in the Anderson model on the Bethe lattice*, Adv. Math. **133** (1998), no. 1, 163–184, DOI 10.1006/aima.1997.1688. MR1492789

[KSV04] O. Khorunzhy, M. Shcherbina, and V. Vengerovsky, *Eigenvalue distribution of large weighted random graphs*, J. Math. Phys. **45** (2004), no. 4, 1648–1672, DOI 10.1063/1.1667610. MR2043849

[Küh08] Reimer Kühn, *Spectra of sparse random matrices*, J. Phys. A **41** (2008), no. 29, 295002, 21, DOI 10.1088/1751-8113/41/29/295002. MR2455271

[Kyp06] Andreas E. Kyprianou, *Introductory lectures on fluctuations of Lévy processes with applications*, Universitext, Springer-Verlag, Berlin, 2006. MR2250061

[Mal12] Camille Male, *The limiting distributions of large heavy Wigner and arbitrary random matrices*, J. Funct. Anal. **272** (2017), no. 1, 1–46, DOI 10.1016/j.jfa.2016.10.001. MR3567500

[MP67] V. A. Marčenko and L. A. Pastur, *Distribution of eigenvalues in certain sets of random matrices* (Russian), Mat. Sb. (N.S.) **72 (114)** (1967), 507–536. MR0208649

[NS06] Alexandru Nica and Roland Speicher, *Lectures on the combinatorics of free probability*, London Mathematical Society Lecture Note Series, vol. 335, Cambridge University Press, Cambridge, 2006. MR2266879

[RB88] G. J. Rodgers and A. J. Bray, *Density of states of a sparse random matrix*, Phys. Rev. B (3) **37** (1988), no. 7, 3557–3562, DOI 10.1103/PhysRevB.37.3557. MR932406

[Res07] Sidney I. Resnick, *Heavy-tail phenomena*: *Probabilistic and statistical modeling*, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2007. MR2271424

[RS80] Michael Reed and Barry Simon, *Methods of modern mathematical physics. I*: *Functional analysis*, 2nd ed., Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1980. MR751959

[Rya98] Øyvind Ryan, *On the limit distributions of random matrices with independent or free entries*, Comm. Math. Phys. **193** (1998), no. 3, 595–626, DOI 10.1007/s002200050340. MR1624843

[Sos04] Alexander Soshnikov, *Poisson statistics for the largest eigenvalues of Wigner random matrices with heavy tails*, Electron. Comm. Probab. **9** (2004), 82–91, DOI 10.1214/ECP.v9-1112. MR2081462

[Ste02] J. Michael Steele, *Minimal spanning trees for graphs with random edge lengths*, Mathematics and computer science, II (Versailles, 2002), Trends Math., Birkhäuser, Basel, 2002, pp. 223–245. MR1940139

[Tao12] Terence Tao, *Topics in random matrix theory*, Graduate Studies in Mathematics, vol. 132, American Mathematical Society, Providence, RI, 2012. MR2906465

[Wig55] Eugene P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Ann. of Math. (2) **62** (1955), 548–564, DOI 10.2307/1970079. MR0077805

[WS80] Joachim Weidmann, *Linear operators in Hilbert spaces*, Graduate Texts in Mathematics, vol. 68, Springer-Verlag, New York-Berlin, 1980. MR566954

[Zak06] Inna Zakharevich, *A generalization of Wigner's law*, Comm. Math. Phys. **268** (2006), no. 2, 403–414, DOI 10.1007/s00220-006-0074-5. MR2259200

Department of Mathematical Sciences, KAIST, Daejeon 34141, Republic of Korea
E-mail address: paul.jung@gmail.com