



# Spectrum of Lévy–Khintchine Random Laplacian Matrices

Andrew Campbell<sup>1</sup> · Sean O’Rourke<sup>2</sup>

Received: 30 November 2022 / Revised: 7 June 2023 / Accepted: 5 July 2023 /  
Published online: 26 July 2023  
© The Author(s) 2023

## Abstract

We consider the spectrum of random Laplacian matrices of the form  $L_n = A_n - D_n$  where  $A_n$  is a real symmetric random matrix and  $D_n$  is a diagonal matrix whose entries are equal to the corresponding row sums of  $A_n$ . If  $A_n$  is a Wigner matrix with entries in the domain of attraction of a Gaussian distribution, the empirical spectral measure of  $L_n$  is known to converge to the free convolution of a semicircle distribution and a standard real Gaussian distribution. We consider real symmetric random matrices  $A_n$  with independent entries (up to symmetry) whose row sums converge to a purely non-Gaussian infinitely divisible distribution, which fall into the class of Lévy–Khintchine random matrices first introduced by Jung [Trans Am Math Soc, **370**, (2018)]. Our main result shows that the empirical spectral measure of  $L_n$  converges almost surely to a deterministic limit. A key step in the proof is to use the purely non-Gaussian nature of the row sums to build a random operator to which  $L_n$  converges in an appropriate sense. This operator leads to a recursive distributional equation uniquely describing the Stieltjes transform of the limiting empirical spectral measure.

**Keywords** Symmetric random matrices · Poisson point process ·  $\alpha$ -Stable laws · Infinitely divisible laws · Heavy-tailed entries · Random Laplacian matrices · Empirical spectral measure

**Mathematics Subject Classification (2020)** Primary: 60B20

---

S. O’Rourke has been supported in part by NSF CAREER Grant DMS-2143142.

---

✉ Andrew Campbell  
andrew.campbell@ist.ac.at

Sean O’Rourke  
sean.d.orourke@colorado.edu

<sup>1</sup> Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria

<sup>2</sup> Department of Mathematics, University of Colorado at Boulder, Boulder, CO 80309, USA

## 1 Introduction

We consider the empirical spectral measure<sup>1</sup> of random Laplacian-type matrices of the form

$$L_n = A_n - D_n \quad (1.1)$$

where  $A_n = (A_{ij})_{i,j=1}^n$  is an  $n \times n$  real symmetric random matrix with independent entries up to symmetry, and  $D_n$  is a diagonal matrix with  $(D_n)_{ii} = \sum_{j=1}^n A_{ij}$ . When  $A_n$  is a Wigner matrix, i.e.  $A_n$  has independent entries up to symmetry with mean zero and variance  $\frac{1}{n}$ , the empirical spectral measure of  $A_n$  converges to Wigner's semicircle law, the empirical spectral measure of  $D_n$  converges to a standard Gaussian distribution, and it was shown in [10] that the empirical spectral measure of  $L_n$  converges to the free convolution of the semicircle law and the standard real Gaussian measure. In this paper, we will consider  $A_n$  such that the diagonal entries of  $D_n$  converge in distribution, not to the Gaussian distribution, but rather to a non-Gaussian infinitely divisible distribution. This model will include Lévy matrices, sometimes referred to as heavy-tailed Wigner matrices, where the entries of  $A_n$  are independent up to symmetry, but have infinite second moment, see Subsection 1.1 for more details. Another important example arises when  $A_n$  is the adjacency matrix of an Erdős–Rényi random graph where the expected degree of any vertex remains fixed as the number of vertices goes to infinity. These  $A_n$  fall into the class of Lévy–Khintchine matrices, a generalization of Lévy matrices defined by Jung in [24], see Subsection 1.2 for more on these matrices.

The term Laplacian comes from graph theory, where the combinatorial Laplacian of a graph with vertex set  $\{1, 2, \dots, n\}$  is defined by:

$$L_{ij} = \begin{cases} \deg(i), & \text{if } i = j \\ -1, & \text{if } i \sim j \\ 0, & \text{if } i \not\sim j, \end{cases} \quad (1.2)$$

where  $i \sim j$  if  $\{i, j\}$  is an edge in the graph and  $\deg(i)$  is the number of edges incident to a vertex  $i$ . The combinatorial Laplacian is the negative of what we refer to as the Laplacian. If the entries of  $A_n$  are almost surely nonnegative, then  $L_n$  is the infinitesimal generator of a (random) continuous time random walk, and for this reason,  $L_n$  is referred to as a Markov matrix in some of the literature. We use the term Laplacian throughout. Spectral properties of real symmetric random Laplacian matrices have been studied in [4, 10, 12, 13, 16, 17, 21–23] and for non-symmetric random Laplacian matrices in [6] when the entries of  $A_n$  are in the domain of attraction of either a real or complex Gaussian random variable. Though because of the widespread use of graph Laplacians this list is incomplete. In these light-tailed cases, the limiting spectral measure has a particularly nice free probabilistic interpretation (see [27] for

<sup>1</sup> The definition for the empirical spectral measure and other notation used throughout is established in Subsection 1.3.

an introduction to free probability and random matrices). In [10], Bryc, Dembo, and Jiang proved the following:

**Theorem 1.1** (Theorem 1.3 in [10]) *Let  $\{X_{ij} : j \geq i \geq 1\}$  be a collection of i.i.d. real random variables with  $\mathbb{E}X_{12} = 0$  and  $\mathbb{E}X_{12}^2 = 1$ ,  $X_{ij} = X_{ji}$  for  $1 \leq i \leq j$ , and let  $A_n = (X_{ij}/\sqrt{n})_{i,j=1}^n$  be a random real symmetric matrix. With probability one, the empirical spectral measure of the matrix  $L_n$  defined in (1.1) converges weakly to the free additive convolution of the semicircle and standard Gaussian measures.*

The analogous free probabilistic limit was established in [6] for non-symmetric  $A_n$ . While some of the above references study sparse Laplacian matrices, none consider random Laplacian matrices with heavy-tailed entries or sparse Laplacian matrices where the expected number of nonzero entries in a row is uniformly bounded in  $n$ .

Many of the tools and techniques we employ were developed in the study of heavy-tailed real symmetric, or Lévy, matrices by Bordenave, Caputo, and Chafaï in [7]. Lévy matrices were introduced in [14] as heavy-tailed versions of Wigner matrices. For the purposes of this paper, an important distinction between Lévy and Wigner matrices is that the row sums of a Wigner matrix converge in distribution to a Gaussian random variable, while the row sums of a Lévy matrix converge to an  $\alpha$ -stable distribution for  $0 < \alpha < 2$ . The techniques in [7] were extended by Jung in [24] to random matrices whose row sums converge in distribution to an infinitely divisible distribution.

### 1.1 Lévy Matrices

Lévy matrices are the heavy-tailed analogue of Wigner matrices, where the entries are independent up to symmetry, but fail to have two finite moments.

**Definition 1.2** A real symmetric random matrix  $X$  is a Lévy matrix if the diagonal entries are zero, the entries above the diagonal are independent and identically distributed (i.i.d.) copies of a real random variable  $\xi$ , and there exists  $\theta \in [0, 1]$  and  $\alpha \in (0, 2)$  such that

- (i)  $\lim_{t \rightarrow \infty} \frac{\mathbb{P}(\xi \geq t)}{\mathbb{P}(|\xi| \geq t)} = \theta,$
- (ii)  $\mathbb{P}(|\xi| \geq t) = t^{-\alpha} L(t)$  for all  $t \geq 1$ , where  $L$  is a slowly varying function, i.e.  $L(tx)/L(t) \rightarrow 1$  as  $t \rightarrow \infty$  for any  $x > 0$ .

The conditions in Definition 1.2 are the same conditions for  $\xi$  to be in the domain of attraction of an  $\alpha$ -stable distribution. Unlike with Wigner matrices, the natural scaling on an  $n \times n$  Lévy matrix  $X$  is not  $\sqrt{n}$ , but instead

$$a_n := \inf \left\{ x : \mathbb{P}(|X_{12}| > x) \leq n^{-1} \right\}. \tag{1.3}$$

For an  $n \times n$  Lévy matrix  $X_n$ , the matrix  $A_n$  in equation (1.1) will be defined as  $A_n := a_n^{-1} X_n$ . We will refer to  $A_n$  as a normalized Lévy matrix.

## 1.2 Lévy–Khintchine Matrices

Jung in [24] defined a generalization of Lévy matrices. Instead of assuming the entries are in the domain of attraction of an  $\alpha$ -stable distribution, the entries are in the domain of attraction of any infinitely divisible distribution.

**Definition 1.3** A sequence of real symmetric random matrices  $\{A_n\}_{n \geq 1}$  is called a *Lévy–Khintchine* random matrix ensemble with characteristics  $(\sigma^2, b, m)$  if for each  $n$ ,  $A_n = (A_{ij}^{(n)})_{i,j=1}^n$  is  $n \times n$ , the diagonal entries of  $A_n$  are 0, the non-diagonal entries are i.i.d. up to symmetry and  $\sum_{j=1}^n A_{1j}^{(n)}$  converges in distribution as  $n \rightarrow \infty$  to a random variable  $Y$  with

$$\log(\mathbb{E}e^{itY}) = -\frac{1}{2}t^2\sigma^2 + itb + \int_{\mathbb{R}} \left( e^{itx} - 1 - \frac{itx}{1+x^2} \right) dm(x) \quad (1.4)$$

for all  $t \in \mathbb{R}$ , where  $m$  is a measure on  $\mathbb{R}$  with  $m(\{0\}) = 0$  satisfying

$$\int_{\mathbb{R}} 1 \wedge |x|^2 dm(x) < \infty. \quad (1.5)$$

**Remark 1.4** It is worth noting that the distribution of  $A_{12}^{(n)}$  may change with  $n$ . However, for many important examples  $A_{12}^{(n)}$  is either a rescaling of a fixed random variable or is the product of a fixed random variable and a Bernoulli random variable where only the Bernoulli random variable is changing with  $n$ .

A random variable  $Y$  satisfying (1.4) is said to have an infinitely divisible distribution with characteristics  $(\sigma^2, b, m)$ , and (1.4) is referred to as the Lévy–Khintchine representation of  $Y$ . When  $\sigma = 0$ ,  $Y$  is called purely non-Gaussian and has an important connection to Poisson point processes with intensity measure  $m$  outlined in Propositions 3.6 and 3.7. Section 3 additionally contains background on Poisson point processes.

## 1.3 Notation

Throughout this paper, we use  $\Rightarrow$  to denote weak convergence of probability measure, convergence in distribution of random variables, and vague convergence of finite measures. For an  $n \times n$  real symmetric matrix  $M$ , the eigenvalues will always be considered in non-increasing order  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . We define the empirical spectral measure of an  $n \times n$  real symmetric matrix  $M$  to be the probability measure

$$\mu_M = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}, \quad (1.6)$$

where  $\delta_x$  is the Dirac delta measure at  $x$ .

A coupling of two probability measures  $\mu_1$  and  $\mu_2$  is a random tuple  $(X, Y)$  such that  $X$  is  $\mu_1$  distributed and  $Y$  is  $\mu_2$  distributed. The symbol  $\stackrel{d}{=}$  will be used to denote equality in distribution of random variables and  $\mathcal{L}(X)$  will be used to denote the distribution of a random variable  $X$ . For two complex-valued square integrable random variables  $\xi$  and  $\psi$ , we define the covariance between  $\xi$  and  $\psi$  as  $\text{Cov}(\xi, \psi) := \mathbb{E}[(\xi - \mathbb{E}\xi)(\psi - \mathbb{E}\psi)]$ .

Throughout we will consider Poisson point processes on  $\bar{\mathbb{R}} \setminus \{0\}$ , the one point compactification of  $\mathbb{R}$  with the origin removed, with some intensity measure  $m$ . We will consider both finite and infinite measures, so for convenience we will denote the points of this process by  $\{y_i\}_{i \geq 1}$  for general  $m$  where  $y_i = 0$  for any  $i$  greater than an appropriate (possibly identically infinite) Poisson random variable, and when considering a specific finite measure  $m$ , we will denote the points by  $\{y_i\}_{i=1}^N$  for a Poisson random variable  $N$ . See Sect. 3 for details on Poisson point processes.

For a topological space  $E$ , let  $C_K(E)$  denote the set of real-valued continuous functions on  $E$  with compact support. We will use  $\mathbb{C}_+$  to be the set of complex numbers with strictly positive imaginary part. For a probability measure  $\mu$  on  $\mathbb{R}$ , we define the function  $s_\mu : \mathbb{C}_+ \rightarrow \mathbb{C}_+$  by

$$s_\mu(z) = \int_{\mathbb{R}} \frac{1}{x - z} d\mu(x), \tag{1.7}$$

and refer to  $s_\mu$  as the Stieltjes transform of  $\mu$ .

We will use asymptotic notation ( $O, o, \Theta$ , etc.) under the assumption that  $n \rightarrow \infty$  unless otherwise stated.  $X = O(Y)$  if  $X \leq CY$  for an absolute constant  $C > 0$  and all  $n \geq C$ ,  $X = o(Y)$  if  $X \leq C_n Y$  for  $C_n \rightarrow 0$ ,  $X = \Theta(Y)$  if  $cY \leq X \leq CY$  for absolute constants  $C, c > 0$  and all  $n \geq C$ , and  $X \sim Y$  if  $X/Y \rightarrow 1$ .

## 2 Main Results

Throughout we will assume  $A_n = (A_{ij}^{(n)})_{i,j=1}^n$  is the  $n$ -th element of a Lévy–Khintchine random matrix ensemble with characteristics  $(0, b, m)$ ,  $D_n$  is a diagonal matrix with  $(D_n)_{ii} = \sum_{j=1}^n A_{ij}^{(n)}$ , and

$$L_n = A_n - D_n. \tag{2.1}$$

**Definition 2.1** Let  $\{A_n\}_{n \geq 1}$  be a Lévy–Khintchine random matrix ensemble with characteristics  $(0, b, m)$  and for each  $n \geq 1$  let  $V_{1n} \geq V_{2n} \geq \dots \geq V_{(n-1)n}$  be the order statistics of  $\{|A_{2n}^{(n)}|, |A_{3n}^{(n)}|, \dots, |A_{nn}^{(n)}|\}$ .  $\{A_n\}_{n \geq 1}$  satisfies Condition **C1** if:

- The Poisson point process with intensity measure  $m$  is almost surely summable.
- $(V_{jn})_{j \geq 1}$  is almost surely uniformly integrable in  $n$ , i.e.

$$\lim_{k \rightarrow \infty} \sup_{n > k} \sum_{i=k+1}^{n-1} V_{in} = 0, \tag{2.2}$$

almost surely.

- There exists  $\varepsilon > 0$  and  $C > 0$  such that

$$m(\{x \in \mathbb{R} : |x| \geq t\}) \leq Ct^{-\varepsilon}, \tag{2.3}$$

and

$$n\mathbb{P}\left(|A_{12}^{(n)}| \geq t\right) \leq Ct^{-\varepsilon} \tag{2.4}$$

for all  $t > 1/4$  and for every  $n \in \mathbb{N}$ .

**Remark 2.2** From Campbell’s formula (Lemma 3.3), almost sure summability of a Poisson point process with intensity measure  $m$  is equivalent to

$$\int_{\mathbb{R} \setminus \{0\}} |x| \wedge 1 dm(x) < \infty. \tag{2.5}$$

**Remark 2.3** Some interesting and important examples of random matrices satisfying condition **C1** include:

- (i)  $A_n = a_n^{-1}X_n$  for a Lévy matrix  $X_n$  with  $\alpha \in (0, 1)$  and  $a_n$  as defined in (1.3). In this case,  $m = m_\alpha$  where  $m_\alpha$  is the measure on  $\mathbb{R}$  with density

$$f(x) = \alpha|x|^{-(1+\alpha)} (\theta \mathbf{1}_{\{x>0\}} + (1 - \theta)\mathbf{1}_{\{x<0\}}),$$

for  $\alpha$  and  $\theta$  as in Definition 1.2.<sup>2</sup>

- (ii) The adjacency matrix  $A_n$  of an Erdős–Rényi random graph  $G(n, p)$  with  $np \rightarrow \lambda \in (0, \infty)$ . In this case, the row sums of  $A_n$  converge to Poisson random variables so that in (1.4)  $m = \lambda\delta_1$  and  $b = \lambda/2$ .
- (iii) The matrix  $A_n = \frac{1}{\sqrt{\lambda}}E_n \circ X_n$  where  $E_n$  is the adjacency matrix of an Erdős–Rényi random graph  $G(n, p)$  with  $np \rightarrow \lambda \in (0, \infty)$ ,  $X_n$  is chosen from the Gaussian Orthogonal Ensemble (GOE), and  $\circ$  is the Hadamard product of matrices. In this case,  $m = \lambda G_\lambda$  where  $G_\lambda$  is the centered Gaussian probability measure with variance  $\frac{1}{\lambda}$ .

The first two points of Condition **C1** will be important for handling the diagonal entries of  $L_n$ . (2.5) implies a Poisson point process with intensity measure  $m$  is almost surely summable, which is stronger than the almost sure square summability implied by (1.5). An essential component of the proofs in Sect. 6 is the convergence of the empirical point process of entries in a row of  $A_n$  to a Poisson point process with intensity measure  $m$ ; however, convergence as point processes does not necessarily imply convergence of the row sum to the sum of the Poisson point process. The uniform integrability assumption (2.2) will allow us in Sect. 6 to conclude that the row sums converge to the sum of the Poisson point process with intensity measure  $m$ . The last

<sup>2</sup> For  $\alpha \in [1, 2)$ , (2.5) and (2.2) will not hold. For this reason, we believe the case where  $\alpha \in [1, 2)$  would require different techniques than those used here.

point is a technical assumptions needed in the proof of the main theorem given below. Heuristically the last point of Condition **C1** states that the infinitely divisible random variable  $Y$  in Definition 1.3 has at least  $t^{-\varepsilon}$  tail decay, and this tail assumption holds entry-wise uniformly in  $n$ . The assumption in (2.4) is technical and used to prove tightness of the empirical spectral measures, but perhaps is not necessary, and there may be room for refinement. The choice of  $1/4$  in the final condition is arbitrary, any positive constant would be sufficient.

**Theorem 2.4** (Eigenvalue Convergence for Laplacian Lévy–Khintchine matrices) *Let  $\{A_n\}_{n \geq 1}$  be a Lévy–Khintchine random matrix ensemble with characteristics  $(0, b, m)$  all defined on the same probability space satisfying Condition **C1**, and for every  $n \in \mathbb{N}$  let  $L_n$  be defined by (2.1). Then, there exists a deterministic probability measure  $\mu_m$  depending only on  $m$  such that a.s.  $\mu_{L_n}$  converges weakly to  $\mu_m$ , as  $n \rightarrow \infty$ .*

While the random matrices satisfying Condition **C1** may appear very different for different  $m$ , a general description of  $\mu_m$  is available through its Stieltjes transform and a recursive distributional equation (RDE). A recursive distributional equation is an equation of the form

$$X \stackrel{d}{=} g((Y_1, X_1), (Y_2, X_2), \dots) \tag{2.6}$$

where  $\{X_n\}_{n=1}^\infty$  are i.i.d. copies of  $X$  and  $\{Y_n\}_{n=1}^\infty$  is some sequence of random variables independent from  $\{X_n\}_{n=1}^\infty$ . While we do not use existing results from the literature, we did find the survey [1] and the unpublished manuscript [2] helpful for better understanding RDEs and contraction arguments in proving uniqueness of solutions. We encourage the interested reader to begin there for more information on RDEs.

**Theorem 2.5** (Recursive Distributional Equation for Stieltjes Transform of  $\mu_m$ ) *Let  $\mu_m$  be the limiting deterministic measure from Theorem 2.4, and let  $s_m(z) = \int_{\mathbb{R}} \frac{1}{x-z} d\mu_m(x)$  be the Stieltjes transform of  $\mu_m$ . Then for every  $z \in \mathbb{C}_+$ ,  $s_m(z) = \mathbb{E}s_\varnothing(z)$  where  $s_\varnothing$  is the Stieltjes transform of a random probability measure. Moreover, the distribution of  $s_\varnothing$  is the unique distribution on the space of Stieltjes transforms of probability measures such that*

$$s_\varnothing(z) \stackrel{d}{=} \left( z - \sum_{j=1}^\infty \frac{y_j}{s_j(z)y_j - 1} \right)^{-1} \quad \text{for all } z \in \mathbb{C}_+, \tag{2.7}$$

where  $\{y_j\}_{j \geq 1}$  is a Poisson point process with intensity measure  $m$  and  $\{s_j\}$  is a collection of i.i.d. copies of  $s_\varnothing$  independent from the point process.

Theorems 2.4 and 2.5 give that the limiting empirical spectral measure of  $L_n$  is uniquely determined by a Poisson point process with intensity measure  $m$ . For the examples outlined in Remark 2.3, we will now give some more explicit descriptions of the corresponding point processes.

- (i) Let  $E_1, E_2, \dots$  be a sequence of independent exponential random variables with mean 1 and  $\Gamma_k = E_1 + \dots + E_k$ . Additionally let  $\varepsilon_1, \varepsilon_2, \dots$  be a sequence of i.i.d. random variables such that

$$\mathbb{P}(\varepsilon_1 = 1) = \theta = 1 - \mathbb{P}(\varepsilon_1 = -1).$$

Then (see [15] Proposition 2) the collection  $\{\varepsilon_k \Gamma_k^{-1/\alpha}\}_{k \geq 1}$  is a Poisson point process with intensity measure  $m_\alpha$ , the measure arising for Lévy matrices, example (i) in Remark 2.3.

- (ii) For the Laplacian of very sparse random graphs, discussed in Remark 2.3 (ii), the Poisson point process is quite simple. Let  $N$  be a Poisson random variable with mean  $\lambda$  and for  $k \geq 1$  define  $y_k$  by

$$y_k = \begin{cases} 1, & k \leq N, \\ 0, & k > N \end{cases} .$$

Then,  $\{y_k\}_{k \geq 1}$  is a Poisson point process with intensity measure  $\lambda \delta_1$ .

- (iii) For a very sparse GOE matrix described in example (iii) in Remark 2.3, let  $Y_1, Y_2, \dots$  be independent standard real Gaussian random variables, and let  $N$  be a Poisson random variable with mean  $\lambda$ . Define

$$y_k = \begin{cases} \frac{Y_k}{\sqrt{\lambda}}, & k \leq N \\ 0, & k > N \end{cases} .$$

Then,  $\{y_k\}_{k \geq 1}$  is a Poisson point process with intensity measure  $\lambda G_\lambda$ . This example is explored a bit further in Theorem 2.7.

RDE (2.7) can be written as:

$$s_\emptyset(z) \stackrel{d}{=} - \left( z + \sum_{j=1}^\infty y_j - \sum_{j=1}^\infty \frac{y_j^2 s_j(z)}{s_j(z) y_j - 1} \right)^{-1} . \tag{2.8}$$

If we consider a diagonal matrix  $\tilde{D}_n$  independent from  $A_n$  with independent entries  $(\tilde{D}_n)_{ii} \stackrel{d}{=} (D_n)_{ii}$  and the matrix  $\tilde{L}_n = A_n - \tilde{D}_n$ , then the work below leading up to the existence of (2.7) could be adapted in a straightforward way to arrive at a corresponding RDE for  $\tilde{L}_n$ . Specifically,

$$s_\emptyset(z) \stackrel{d}{=} - \left( z + \sum_{j=1}^\infty \tilde{y}_j - \sum_{j=1}^\infty \frac{y_j^2 s_j(z)}{s_j(z) y_j - 1} \right)^{-1} , \tag{2.9}$$

where  $\{\tilde{y}_j\}$  is an independent copy of the point process of  $\{y_j\}$ , independent of  $\{s_j\}$ . For light-tailed  $A_n$ , Theorem 1.1 gives that the limiting spectral measure of  $L_n$  is the free additive convolution of the semicircle measure and the Gaussian measure.



This is the same limiting spectral measure for  $A_n - D_n$ , when  $A_n$  is independent of  $D_n$ . In contrast, the differences between equations (2.8) and (2.9) suggest that for Lévy–Khintchine  $A_n$ , the dependence between  $A_n$  and  $D_n$  can be seen in the limiting measure  $\mu_m$ .

### 2.1 Outline

Section 3 contains a brief introduction to Poisson point processes, and important results on these processes used throughout. In Sects. 4 and 5, we define local convergence for operators on  $\ell^2(V)$  for a countable set  $V$  and use the measure  $m$  to build a random operator  $L$ . In Sect. 6, we show that  $L_n$  converges locally in distribution to  $L$ , and then, in Sect. 7 we upgrade this to convergence of the empirical spectral measures. Finally in Sect. 8, we show that the Stieltjes transform of the limiting empirical spectral measure can be described as the expected value of the unique solution to (2.7). In the appendices, we prove almost sure tightness of the collection  $\{\mu_{L_n}\}_{n \geq 1}$  and list some technical lemmas. We end this section with two corollaries of Theorem 2.5. The first is a continuity result for the map  $m \mapsto \mu_m$ . In the second, we use (2.7) to recover the free convolution of a semicircle and a standard Gaussian measure from the limiting empirical measure of very sparse random matrices.

### 2.2 Corollaries of Theorem 2.5

The first corollary of Theorem 2.5 concerns continuity of the mapping  $m \mapsto \mu_m$  where  $\mu_m$  is the limiting measure of Theorem 2.4. Uniqueness of the solution to the RDE in Theorem 2.5 is crucial to the proof of Corollary 2.6.

**Corollary 2.6** *Let  $\bar{\mathbb{R}}$  denote the one point compactification of  $\mathbb{R}$ . Let  $\{m_n\}_{n=1}^\infty$  be a collection of measures on  $\mathbb{R}$  such that*

$$\int_{\mathbb{R}} 1 \wedge |x| \, dm_n(x) < \infty,$$

for all  $n \in \mathbb{N} \cup \{\infty\}$ . Let  $\mu_{m_1}, \mu_{m_2}, \dots$  and  $\mu_{m_\infty}$  be the deterministic limiting measures described in Theorem 2.4 for a Lévy–Khintchine random matrix ensemble with characteristics  $(0, b, m_1), (0, b, m_2), \dots$  and  $(0, b, m_\infty)$ , respectively. If for any  $f \in C_K(\bar{\mathbb{R}} \setminus \{0\})$ ,

$$\int_{\mathbb{R}} f \, dm_n \rightarrow \int_{\mathbb{R}} f \, dm_\infty,$$

and for any  $\varepsilon > 0$

$$\lim_{k \rightarrow \infty} \sup_{n \geq 1} \mathbb{P} \left( \sum_{j=k}^\infty |y_j^{(n)}| > \varepsilon \right) = 0, \tag{2.10}$$

where for each  $n \in \mathbb{N}$ ,  $\{y_j^{(n)}\}_{j=1}^\infty$  is a Poisson point process with intensity measure  $m_n$ , and then,  $\mu_{m_n}$  converges weakly to  $\mu_{m_\infty}$  as  $n \rightarrow \infty$ .

**Proof** Let  $s_n$  be the Stieltjes transforms of  $\mu_{m_n}$ , and let  $r_n$  be the random Stieltjes transform solving (2.7) corresponding to  $m_n$ , so that  $s_n = \mathbb{E}r_n$ . By Lemma B.2, in order to prove Corollary 2.6 it is enough to show that pointwise  $\lim_{n \rightarrow \infty} s_n = \mathbb{E}r$ , where  $r$  solves RDE (2.7) corresponding to  $m_\infty$ . Take any subsequence  $\{s_{n_k}\}_{n_k}$  of  $\{s_n\}_n$ , with corresponding subsequence  $\{r_{n_k}\}_{n_k}$  of  $\{r_n\}_n$ . From Lemma B.3, it follows that  $\{r_{n_k}\}$  is tight in the space of analytic function on  $\mathbb{C}_+$  with the topology of uniform convergence on compact subsets, and we pass to a further subsequence  $n'_k$  converging to another random analytic function  $r(z)$ . As  $\{r_{n_k}\}$  is almost surely uniformly bounded on compact subsets it follows that  $r$  is almost surely bounded on compact subsets. For any fixed  $z \in \mathbb{C}_+$ , it follows by the dominated convergence theorem that

$$\lim_{n'_k \rightarrow \infty} s_{n'_k}(z) = \lim_{n'_k \rightarrow \infty} \mathbb{E}r_{n'_k}(z) = \mathbb{E}r(z). \tag{2.11}$$

Corollary 2.6 then follows if  $r$  is a random Stieltjes transform solution to RDE (2.7) corresponding to  $m_\infty$ .

To this end, let  $\Pi_n$  be Poisson random measures with intensity measures  $m_n$ . For a positive function  $f \in C_K(\mathbb{R} \setminus \{0\})$ ,  $1 - e^{-f(x)}$  is also a continuous function with compact support. Thus,

$$\lim_{n \rightarrow \infty} \exp\left(\int_{\mathbb{R}} 1 - e^{-f(x)} dm_n(x)\right) = \exp\left(\int_{\mathbb{R}} 1 - e^{-f(x)} dm_\infty(x)\right). \tag{2.12}$$

It follows from Propositions 3.4 and 3.5 that  $\Pi_n$  converges in distribution to  $\Pi_\infty$ . For  $n \in \mathbb{N}$  let  $\{y_j^{(n)}\}_{j \geq 1}$  be the points of the process  $\Pi_n$  and  $\{y_j\}_{j \geq 1}$  the points of the process  $\Pi_\infty$ . The points may be ordered such that for every  $j \in \mathbb{N}$ ,  $y_j^{(n)}$  converges in distribution to  $y_j$  (see Section 2 of [15] for more details). In fact, from (2.10) and Lemma 1 of [15]  $\{y_j^{(n)}\}_{j \geq 1}$  converges in distribution to  $\{y_j\}_{j \geq 1}$  in  $\ell^1(\mathbb{N})$ . Using Skorokhod’s representation theorem, we may put  $\{r_{n'_k}\}$ ,  $\{\Pi_{n'_k}\}$ ,  $\Pi_\infty$  and  $r$  on a single probability space such that all the above convergences in distributions are almost sure, and

$$\lim_{n \rightarrow \infty} \sum_{j=1}^\infty |y_j^{(n)} - y_j| = 0 \tag{2.13}$$

almost surely. For fixed  $z \in \mathbb{C}_+$ ,

$$\begin{aligned} r(z) &= \lim_{n'_k \rightarrow \infty} r_{n'_k}(z) \\ &= \lim_{n'_k \rightarrow \infty} - \left( z - \sum_{j=1}^\infty \frac{y_j^{(n'_k)}}{r_{n'_k}^{(j)}(z)y_j^{(n'_k)} - 1} \right)^{-1} \end{aligned}$$

$$\begin{aligned}
 &= - \left( z - \lim_{n'_k \rightarrow \infty} \sum_{j=1}^{\infty} \frac{y_j^{(n'_k)}}{r_{n'_k}^{(j)}(z)y_j^{(n'_k)} - 1} \right)^{-1} \\
 &= - \left( z - \sum_{j=1}^{\infty} \frac{y_j}{r_j(z)y_j - 1} \right)^{-1}, \tag{2.14}
 \end{aligned}$$

for independent copies  $r_1, r_2, \dots$  of  $r$ , where the last equality follows from (2.13). Thus,  $r$  is an analytic solution to RDE (2.7). From (2.14) and the almost sure boundedness of  $r$  on compact subsets of  $\mathbb{C}_+$  that almost surely

$$\lim_{t \rightarrow \infty} itr(it) = -1, \tag{2.15}$$

and thus  $r$  is almost surely the Stieltjes transform of a probability measure. From (2.11) and the uniqueness of the solution to RDE (2.7), it follows that for any  $z \in \mathbb{C}_+$

$$\lim_{n'_k \rightarrow \infty} s_{n'_k}(z) = s_{\infty}(z). \tag{2.16}$$

As the subsequence  $n_k$  was arbitrary it follows that  $s_n$  converge pointwise to  $s_{\infty}$  and  $\mu_{m_n}$  converges weakly to  $\mu_{m_{\infty}}$  as  $n \rightarrow \infty$ . □

Theorem 2.7 considers the  $\lambda \rightarrow \infty$  limit of example (iii) in Remark 2.3. The limiting measure is the same limiting measure found in Theorem 1.1. The works of Jiang [22] and Chatterjee and Hazra [13] established Theorem 1.1 for sparse random matrices where the expected number of nonzero entries in a row tends to infinity with the size of the matrix. Theorem 2.7, when combined with Theorem 2.4 and Remark 2.3 (iii), can then be interpreted as splitting the limit to where first  $n \rightarrow \infty$  and then the expected number of nonzero entries tends to infinity.

**Theorem 2.7** *Let  $G_{\lambda}$  denote the Gaussian probability measure with mean 0 and variance  $\frac{1}{\lambda}$ , and let  $m_{\lambda} = \lambda G_{\lambda}$ . If  $\mu_{m_{\lambda}}$  is the deterministic limiting probability measure from Theorem 2.4, then  $\mu_{m_{\lambda}}$  converges weakly to the free convolution of the semicircle distribution and the standard real Gaussian distribution, as  $\lambda \rightarrow \infty$ .*

**Proof** Denote the free convolution of a standard semicircle measure and standard Gaussian measure by  $SC \boxplus G_1$ . It is known [5] that the Stieltjes transform,  $s_{fc}$ , of  $SC \boxplus G_1$  can be defined as the unique solution to

$$s_{fc}(z) = \int_{\mathbb{R}} \frac{1}{x - z - s_{fc}(z)} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \tag{2.17}$$

satisfying  $\text{Im}(s_{fc}(z)) \geq 0$  and  $s_{fc}(z) \sim -z^{-1}$  as  $z \rightarrow \infty$ . If  $s_\lambda$  is the Stieltjes transform of  $\mu_{m_\lambda}$ , then from Theorem 2.5 we know  $s_\lambda(z) = \mathbb{E}r_\lambda(z)$  where  $r_\lambda$  satisfies the RDE

$$r_\lambda(z) \stackrel{d}{=} - \left( z - \sum_{j=1}^N \frac{y_j}{r_j(z)y_j - 1} \right)^{-1}, \tag{2.18}$$

$N \sim \text{Pois}(\lambda)$ ,  $\{y_j\}_{j=1}^\infty$  are i.i.d. Gaussian random variables with mean zero and variance  $\frac{1}{\lambda}$  and  $\{r_j\}_{j=1}^\infty$  are i.i.d. copies of  $r_\lambda$ , independent of the collection  $\{y_j\}_{j=1}^\infty$ . We will instead use the equivalent recursive distributional equation

$$r_\lambda(z) \stackrel{d}{=} - \left( z + \frac{1}{\sqrt{\lambda}} \sum_{j=1}^N y_j + \frac{1}{\lambda} \sum_{j=1}^N \frac{r_j(z)y_j^2}{1 - r_j(z)y_j/\sqrt{\lambda}} \right)^{-1}, \tag{2.19}$$

where  $\{y_j\}_{j=1}^\infty$  are i.i.d. standard real Gaussian random variables. Fix  $z \in \mathbb{C}_+$ . We first consider the sum  $S_\lambda = \frac{1}{\sqrt{\lambda}} \sum_{j=1}^N y_j$ . For  $t \in \mathbb{R}$ , define

$$\begin{aligned} \varphi_{S_\lambda}(t) &:= \mathbb{E} \exp(itS_\lambda) \\ &= \sum_{k=0}^\infty \left( e^{-\frac{t^2}{2\lambda}} \right)^k \frac{\lambda^k e^{-\lambda}}{k!} \\ &= \exp(-\lambda) \exp\left(\lambda e^{-\frac{t^2}{2\lambda}}\right) \\ &= \exp\left(-\lambda + \lambda - \frac{t^2}{2} + o\left(\frac{1}{\lambda}\right)\right), \end{aligned}$$

where here and throughout the proof asymptotic notation is as  $\lambda \rightarrow \infty$ . Thus,  $S_\lambda$  converges to a standard real Gaussian random variable as  $\lambda \rightarrow \infty$ .

We will compare the sum  $\frac{1}{\lambda} \sum_{j=1}^N \frac{r_j(z)y_j^2}{1 - r_j(z)y_j/\sqrt{\lambda}}$  to increasingly simpler sums. The first comparison is to the sum  $\frac{1}{\lambda} \sum_{j=1}^N r_j(z)y_j^2$ . Notice that

$$\begin{aligned} \left| \frac{1}{\lambda} \sum_{j=1}^N \frac{r_j(z)y_j^2}{1 - r_j(z)y_j/\sqrt{\lambda}} - \frac{1}{\lambda} \sum_{j=1}^N r_j(z)y_j^2 \right| &= \left| \frac{1}{\lambda^{3/2}} \sum_{j=1}^N \frac{r_j(z)^2 y_j^3}{1 - r_j(z)y_j/\sqrt{\lambda}} \right| \\ &\leq \frac{1}{\text{Im}(z)^2 \lambda^{3/2}} \sum_{j=1}^N \frac{|y_j|^3}{|1 - r_j(z)y_j/\sqrt{\lambda}|} \\ &\leq \frac{2}{\text{Im}(z)^2 \lambda^{3/2}} \sum_{j=1}^N |y_j|^3 \end{aligned}$$

$$+ \frac{1}{\text{Im}(z)^2 \lambda^{3/2}} \sum_{j=1}^N \frac{|y_j|^3}{|1 - r_j(z)y_j/\sqrt{\lambda}|} \mathbf{1}_{A_{j,\lambda}},$$

where  $\mathbf{1}_{A_{j,\lambda}}$  is the indicator of the event  $A_{j,\lambda} = \{|y_j| \geq \sqrt{\lambda} \text{Im}(z)/2\}$ . We will now show both pieces of this bound converge in probability to zero. From Lemma 3.3,

$$\lim_{\lambda \rightarrow \infty} \mathbb{E} \frac{1}{\lambda^{3/2}} \sum_{j=1}^N |y_j|^3 = \lim_{\lambda \rightarrow \infty} \frac{\mathbb{E}|y_1|^3}{\sqrt{\lambda}} = 0.$$

From standard tail estimates of Gaussian random variables, we have that

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \mathbb{P} \left( \sum_{j=1}^N \mathbf{1}_{A_{j,\lambda}} \neq 0 \right) &= \lim_{\lambda \rightarrow \infty} \sum_{k=0}^{\infty} \mathbb{P} \left( \sum_{j=1}^k \mathbf{1}_{A_{j,\lambda}} \neq 0 \right) e^{-\lambda} \frac{\lambda^k}{k!} \\ &\leq \lim_{\lambda \rightarrow \infty} \sum_{k=0}^{\infty} k \mathbb{P} \left( |y_1| \geq \sqrt{\lambda} \text{Im}(z)/2 \right) e^{-\lambda} \frac{\lambda^k}{k!} \\ &\leq \lim_{\lambda \rightarrow \infty} \sum_{k=0}^{\infty} k \frac{2}{\sqrt{2\pi\lambda} \text{Im}(z)} e^{-\lambda \text{Im}(z)^2/8} e^{-\lambda} \frac{\lambda^k}{k!} \\ &\leq \lim_{\lambda \rightarrow \infty} C e^{-c\lambda} \sqrt{\lambda}, \end{aligned}$$

for some positive constants  $C, c > 0$  independent of  $\lambda$ . Thus  $\frac{1}{\lambda} \sum_{j=1}^N \frac{r_j(z)y_j^2}{1 - r_j(z)y_j/\sqrt{\lambda}} - \frac{1}{\lambda} \sum_{j=1}^N r_j(z)y_j^2 \Rightarrow 0$  as  $\lambda \rightarrow \infty$ . Next we compare to the sum  $\frac{1}{\lambda} \sum_{j=1}^N (\mathbb{E}r_\lambda(z))y_j^2$ . To this end, let  $Z_j = r_j(z)y_j^2 - (\mathbb{E}r_\lambda(z))y_j^2$ , and consider the Taylor expansion of characteristic function of the real part of  $\frac{1}{\lambda} \sum_{j=1}^N Z_j$

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \mathbb{E} \exp \left( it \frac{1}{\lambda} \sum_{j=1}^N \text{Re}(Z_j) \right) &= \lim_{\lambda \rightarrow \infty} \sum_{k=0}^{\infty} \left[ \mathbb{E} \exp \left( it \frac{1}{\lambda} \text{Re}(Z_1) \right) \right]^k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \lim_{\lambda \rightarrow \infty} \exp(-\lambda + \lambda + it \mathbb{E} \text{Re}(Z_1) + O(1/\lambda)) \\ &= 1. \end{aligned}$$

An identical argument follows from the imaginary part, and we see that  $\frac{1}{\lambda} \sum_{j=1}^N Z_j$  converges in probability to zero. It is also straightforward to show  $\frac{1}{\lambda} \sum_{j=1}^N y_j^2 \Rightarrow 1$ , and thus,  $\frac{1}{\lambda} \sum_{j=1}^N \mathbb{E}(r_\lambda(z))y_j^2 - \mathbb{E}r_\lambda(z)$  converges in probability to zero. These three comparisons lead to

$$\frac{1}{\lambda} \sum_{j=1}^N \frac{r_j(z)y_j^2}{1 - r_j(z)y_j/\sqrt{\lambda}} - \mathbb{E}r_\lambda(z) = \frac{1}{\lambda} \sum_{j=1}^N \frac{r_j(z)y_j^2}{1 - r_j(z)y_j/\sqrt{\lambda}} - \frac{1}{\lambda} \sum_{j=1}^N r_j(z)y_j^2$$

$$\begin{aligned}
 &+ \frac{1}{\lambda} \sum_{j=1}^N r_j(z) y_j^2 - \frac{1}{\lambda} \sum_{j=1}^N \mathbb{E}(r_\lambda(z)) y_j^2 \\
 &+ \mathbb{E}(r_\lambda(z)) \left( \frac{1}{\lambda} \sum_{j=1}^N y_j^2 - 1 \right),
 \end{aligned}$$

which converges in distribution to 0. Since this limit is a constant, we may conclude that jointly

$$\left( \frac{1}{\sqrt{\lambda}} \sum_{j=1}^N y_j, \frac{1}{\lambda} \sum_{j=1}^N \frac{r_j(z) y_j^2}{1 - r_j(z) y_j / \sqrt{\lambda}} - \mathbb{E}r_\lambda(z) \right) \Rightarrow (Y, 0), \tag{2.20}$$

where  $Y$  is a standard Gaussian random variable.

Let  $\{\lambda_n\}_{n=1}^\infty$  be an arbitrary increasing sequence of positive real numbers going to infinity and let  $\{\lambda_{n_k}\}$  be an arbitrary subsequence. From Lemma B.3,  $\{r_{\lambda_{n_k}}\}_{n_k}$  is tight as a family of random analytic functions on  $\mathbb{C}_+$  with the topology of uniform convergence on compact subsets, and thus, there exists a further subsequence  $\lambda_{n_{k'}}$  such that  $r_{\lambda_{n_{k'}}}(z) \rightarrow \tilde{r}(z)$  for some random analytic function  $\tilde{r}$ . Fix  $z \in \mathbb{C}_+$ , it follows from the dominated convergence theorem that  $\mathbb{E}r_{\lambda_{n_{k'}}}(z) \rightarrow \mathbb{E}\tilde{r}(z) =: r(z)$  for some deterministic limit  $r(z)$ . As  $z \in \mathbb{C}_+$  was arbitrary, it follows from the above convergence in distribution and the continuous mapping theorem that

$$\begin{aligned}
 r(z) &= \lim_{n_{k'} \rightarrow \infty} \mathbb{E}r_{\lambda_{n_{k'}}}(z) \\
 &= - \lim_{n_{k'} \rightarrow \infty} \mathbb{E} \left( z + \frac{1}{\sqrt{\lambda_{n_{k'}}}} \sum_{j=1}^N y_j + \frac{1}{\lambda_{n_{k'}}} \sum_{j=1}^N \frac{r_{\lambda_{n_{k'}}}(z) y_j^2}{1 - r_{\lambda_{n_{k'}}}(z) y_j / \sqrt{\lambda_{n_{k'}}}} \right)^{-1} \\
 &= \mathbb{E} \frac{-1}{z + Y + r(z)} \\
 &= \int_{\mathbb{R}} \frac{1}{x - z - r(z)} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx,
 \end{aligned}$$

pointwise on  $\mathbb{C}_+$ . Thus,  $r(z) = s_{fc}(z)$  along every one of these further subsequences of  $\{\lambda_{n_k}\}$ , and  $s_\lambda(z) = \mathbb{E}r_\lambda(z) \rightarrow s_{fc}(z)$ . By Lemma B.2, this pointwise convergence of the Stieltjes transforms implies  $\mu_{m_\lambda}$  converges weakly to  $SC \boxplus G_1$  as  $\lambda \rightarrow \infty$ .  $\square$

The matrix  $X_n$  in Remark 2.3 (iii) has Gaussian entries, and for convenience, we stated Theorem 2.7 for the corresponding measure  $\lambda G_\lambda$ . However, the proof can be adapted in a straightforward way to the analogous measures corresponding to  $X_n$  from Remark 2.3 (iii) having entries with mean zero, variance  $\frac{1}{\lambda}$ , and three finite moments.

### 3 Poisson Point Processes and Infinitely Divisible Distributions

This section contains a brief overview of Poisson point processes and their relation to infinitely divisible distributions. It includes results used throughout the paper, whose proofs can be found in [15, 25, 28, 29].

We will assume throughout the section that  $E = \bar{\mathbb{R}} \setminus \{0\}$  with the relative topology, where  $\bar{\mathbb{R}}$  is the one point compactification of  $\mathbb{R}$ . Many of the results below can be extended to higher-dimensional Euclidean space and other sufficiently nice topological spaces. It is also worth noting that we consider  $\bar{\mathbb{R}} \setminus \{0\}$  for the purpose of defining the appropriate notion of a compact subset of  $\mathbb{R}$ , i.e. one which is closed and bounded away from 0. However, many of the results below could be stated on  $\mathbb{R}$ .

Denote by  $\mathcal{M}(E)$  the set of simple point Radon measures

$$\mu = \sum_{x \in S} \delta_x, \quad (3.1)$$

where  $S$  is a multiset, and  $\mu((-\infty, -r) \cup (r, \infty)) < \infty$  for any  $r > 0$ . Denote by  $\mathcal{H}(E)$  the set of supports corresponding to measures in  $\mathcal{M}(E)$ , where the support of a simple point measure  $\mu$  is taken to be the multiset  $S$ . The elements of  $\mathcal{H}(E)$  are called configurations.

Intuitively, a point process should be a random collection of points on some space. From this intuition, we want to define a point processes to be random variables taking values in  $\mathcal{H}(E)$ ; however, to do so one would need to define an appropriate  $\sigma$ -algebra on  $\mathcal{H}(E)$ . This is done through the space  $\mathcal{M}(E)$ , as spaces of measures can be topologized in a straightforward way using an appropriate class of test functions. From this topology, one can then take the Borel  $\sigma$ -algebra to make  $\mathcal{M}(E)$  a measurable space which our random variables can take values in. To this end, we say a sequence of measure  $\mu_n$  in  $\mathcal{M}(E)$  converge vaguely to a measure  $\mu$  if for any continuous compactly supported  $f$  on  $E$

$$\int_E f d\mu_n \rightarrow \int_E f d\mu.$$

We refer to the topology on  $\mathcal{M}(E)$  defined through vague convergence as the *vague topology*. In fact,  $\mathcal{M}(E)$  with the vague topology is a Polish space, and thus complete and metrizable.

The vague topology on  $\mathcal{M}(E)$  can be pushed forward to configurations  $\mathcal{H}(E)$  by considering the bijection,  $F$ , from simple point Radon measures to configurations defined by  $F(\mu) = S$  in (3.1). Then, vague convergence  $\mu_n \rightarrow \mu$  corresponds to point-wise convergence of the configurations in some labeling of the elements of the multisets  $F(\mu_n)$ , i.e. there exists labeling  $F(\mu_n) = \{x_n^{(1)}, x_n^{(2)}, \dots\}$ , where  $x_n^{(1)}, x_n^{(2)}, \dots$  need not be distinct, such that for all  $k \in \mathbb{N}$ ,  $x_n^{(k)} \rightarrow x^{(k)}$ .

With  $\mathcal{M}(E)$  appropriately topologized, we define a simple point process  $N$  as a measurable mapping from a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  to  $(\mathcal{M}(E), \mathcal{B}(\mathcal{M}(E)))$ , where  $\mathcal{B}(\mathcal{M}(E))$  is the Borel  $\sigma$ -algebra defined by the vague topology.

**Definition 3.1** Let  $m$  be a Borel measure on  $E$ . A point process  $N$  is a Poisson point process with intensity measure  $m$  if we have the following:

1. For any Borel set  $A$ ,  $N(A)$  is a Poisson random variable with mean  $m(A)$ .<sup>3</sup>
2. If  $A_1, \dots, A_k$  are pairwise disjoint Borel sets, then  $N(A_1), \dots, N(A_k)$  are independent random variables.

Lemma 3.2 is a technical result for Poisson point processes with intensity measure satisfying (2.5). We do not believe that the result is novel; however, we are unable to find a reference in the point process literature. The proof below is a modification of the proof for Lemma A.4 in [7].

**Lemma 3.2** *Let  $m$  be a measure on  $\mathbb{R} \setminus \{0\}$  such that*

$$\int_{\mathbb{R}} 1 \wedge |x| dm(x) < \infty, \tag{3.2}$$

and let  $\{y_k\}_{k=1}^N$  be a Poisson point process with intensity measure  $m$  where  $N = \infty$  a.s. if  $m(\mathbb{R} \setminus \{0\}) = \infty$ . If  $\tau_\kappa = \inf \left\{ t \geq 0 : \sum_{k=t+1}^N y_{(k)}^2 \leq \kappa \right\}$  where  $y_{(1)} \geq y_{(2)} \geq \dots$  is a non-increasing ordering of  $\{y_k\}_{k=1}^N$ , then  $\mathbb{E}\tau_\kappa < \infty$  for all  $\kappa > 0$  and  $\mathbb{E}\tau_\kappa \rightarrow 0$  as  $\kappa \rightarrow \infty$ .

**Proof** The fact that  $\tau_\kappa$  is almost surely finite follows from the integrability condition on  $m$ . Additionally  $\mathbb{P}(\tau_\kappa = 0) = \mathbb{P}\left(\sum_{k=1}^N y_k^2 \leq \kappa\right)$  and clearly converges to 1 as  $\kappa \rightarrow \infty$ . Thus, it is sufficient to prove  $\mathbb{E}\tau_\kappa < \infty$  for all  $\kappa > 0$ . Let  $0 < r_t < 1$  be some monotonically decreasing function of  $t$  such that  $r_t \rightarrow 0$  as  $t \rightarrow \infty$ ,  $S_t = \sum_{k=1}^N y_k^2 \mathbf{1}_{\{|y_k| \leq r_t\}}$ , and  $\Pi = \sum_{k=1}^N \delta_{y_j}$ . Define the event  $A_t = \{\Pi((-\infty, -r_t) \cup (r_t, \infty)) \geq t\}$ . On  $A_t^c$  the collection of points summed over in the definition of  $\tau_\kappa$  is a strict subset of the collection of points summed over in the definition of  $S_t$ . Then,

$$\begin{aligned} \{\tau_\kappa > t\} &\subseteq (\{\tau_\kappa > t\} \cap A_t^c) \cup (\{\tau_\kappa > t\} \cap A_t) \\ &\subseteq (\{S_t \geq \kappa\} \cap A_t^c) \cup A_t \\ &\subseteq \{S_t \geq \kappa\} \cup A_t. \end{aligned}$$

We will now show for appropriate  $r_t$  the probabilities of the events  $\{S_t \geq \kappa\} \cup \{\Pi((-\infty, -r_t) \cup (r_t, \infty)) \geq t\}$  are summable in  $t \in \mathbb{N}$ .

For a Poisson random variable  $X$  with mean  $\lambda$ ,  $\mathbb{P}(X \geq t) \leq \exp(-t \log(\frac{t}{\lambda e}))$ . Letting  $E_t = \mathbb{E}\Pi((-\infty, -r_t) \cup (r_t, \infty)) = m((-\infty, -r_t) \cup (r_t, \infty)) < \infty$  we have

$$\mathbb{P}(\Pi((-\infty, -r_t) \cup (r_t, \infty)) \geq t) \leq \exp\left(-t \log\left(\frac{t}{E_t e}\right)\right). \tag{3.3}$$

For the event  $\{S_t \geq \kappa\}$  notice that for every  $\theta > 0$ ,

$$\mathbb{P}(S_t \geq \kappa) \leq \exp(-\theta\kappa)\mathbb{E}\exp(\theta S_t),$$

<sup>3</sup> In the case  $m(A) = \infty$ , we say  $N(A)$  is a Poisson random variable which is almost surely  $\infty$ .



and from Campbell’s formula, Lemma 3.3,

$$\begin{aligned} \mathbb{E} \exp(\theta S_t) &= \exp \left( \int_{-r_t}^{r_t} (e^{\theta x^2} - 1) dm(x) \right) \\ &\leq C \exp \left( \theta \int_{-r_t}^{r_t} x^2 dm(x) \right). \end{aligned}$$

Letting  $\theta = \theta_t = \left( \int_{-r_t}^{r_t} x^2 dm(x) \right)^{-1}$  the above gives us

$$\mathbb{P}(\tau_\kappa > t) \leq \exp \left( -t \log \left( \frac{t}{E_t e} \right) \right) + C' \exp(-\theta_t \kappa). \tag{3.4}$$

Notice the integrability assumption on  $m$  implies

$$\varepsilon m((-\infty, -\varepsilon) \cup (\varepsilon, \infty)) \leq C \tag{3.5}$$

for all  $\varepsilon \in (0, 1)$  and some constant  $C$ . From the integrability assumption, we also have that

$$\int_{-1}^1 |x| dm(x) = I \tag{3.6}$$

for some  $0 < I < \infty$ . From this, we see that

$$\int_{-r_t}^{r_t} x^2 dm(x) \leq r_t I, \tag{3.7}$$

and  $\theta_t \geq (I r_t)^{-1}$ . Taking  $r_t = \frac{c}{t}$  for some  $c > 0$  gives  $\theta_t \geq (cI)^{-1}t$ . For an appropriate choice of  $c > 0$  (3.5) and the definition of  $E_t$  imply that  $E_t \leq t/2e$ . Thus, from (3.4) we get

$$\mathbb{E} \tau_\kappa = \sum_{t=0}^{\infty} \mathbb{P}(\tau_\kappa \geq t) < \infty. \tag{3.8}$$

This completes the proof. □

The next lemma, known as Campbell’s formula, is a fundamental result in the theory of point processes and gives a description of the functions belonging almost surely to  $L^1(\Pi)$  for a Poisson random measure  $\Pi$ .

**Lemma 3.3** (Campbell’s Formula, Sect. 3.2 in [26]) *Let  $\Pi$  be a Poisson point process on  $E$  with intensity measure  $m$ . Let  $u : E \rightarrow \mathbb{R}$  be a measurable function. Then with probability 1,*

$$S = \int_E u(x) d\Pi(x) < \infty \tag{3.9}$$

if and only if

$$\int_E 1 \wedge |u(x)| \, dm(x) < \infty. \tag{3.10}$$

If either of the above integrals are finite then

$$\mathbb{E} \exp(\theta S) = \exp \left( \int_E e^{\theta u(x)} - 1 \, dm(x) \right) \tag{3.11}$$

for any  $\theta \in \mathbb{C}$  for which the integral on the right-hand side is finite. Moreover,

$$\mathbb{E} \int_E u(x) \, d\Pi(x) = \int_E u(x) \, dm(x), \tag{3.12}$$

whenever  $u \geq 0$  or  $\int |u(x)| \, dm(x) < \infty$ .

The next two lemmas, which can be found in Chapter 5 of [29], characterize Poisson point processes and convergence of point processes in terms of (3.11).

**Proposition 3.4** (Theorem 5.1 in [29]) *A point process  $\Pi$  on  $E$  is a Poisson point process with intensity measure  $m$  if and only if*

$$\mathbb{E} \exp \left( - \int_E u(x) \, d\Pi(x) \right) = \int_E (e^{-u(x)} - 1) \, dm(x),$$

for every bounded positive measurable function  $u$ .

**Proposition 3.5** (Theorem 5.2 in [29]) *Let  $\Pi_1, \Pi_2, \dots$  and  $\Pi_\infty$  be simple point processes on  $E$ . Then  $\Pi_n \Rightarrow \Pi_\infty$  as  $n \rightarrow \infty$  if and only if*

$$\mathbb{E} \exp \left( - \int_E u(x) \, d\Pi_n(x) \right) \rightarrow \mathbb{E} \exp \left( - \int_E u(x) \, d\Pi_\infty(x) \right),$$

as  $n \rightarrow \infty$  for every bounded positive measurable function  $u$ .

The next two propositions outline the connection between Poisson point processes and triangular arrays converging to infinitely divisible distributions. For  $0 < h < 1$ , define

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

**Proposition 3.6** (Corollary 15.16 in [25]) *Suppose  $\{X_{ni} : 1 \leq i \leq n\}_{n \geq 1}$  is a triangular array of random variables such that each row consists of i.i.d. random variables. Then, the sum*

$$\sum_{i=1}^n X_{ni},$$

converges in distribution to an infinitely divisible random variable with characteristic  $(\sigma^2, b, m)$  as  $n \rightarrow \infty$  if and only if for every  $0 < h < 1$  which is not an atom of  $m$

- $n\mathbb{P}(X_{n1} \in \cdot) \Rightarrow m(\cdot)$  on  $\overline{\mathbb{R}} \setminus \{0\}$ ,
- $n\mathbb{E} \left[ X_{n1}^2 \mathbf{1}_{\{|X_{n1}| \leq h\}} \right] \rightarrow \sigma_h^2$ , and
- $n\mathbb{E} \left[ X_{n1} \mathbf{1}_{\{|X_{n1}| \leq h\}} \right] \rightarrow b_h$ ,

as  $n \rightarrow \infty$ .

**Proposition 3.7** (Theorem 5.3 in [29]) *Suppose  $\{X_{ni} : 1 \leq i \leq n\}_{n \geq 1}$  is a triangular array of random variables on  $\overline{\mathbb{R}} \setminus \{0\}$  such that each row consists of i.i.d. random variables. Let  $N$  be a Poisson point process with intensity measure  $m$ . Then,*

$$\sum_{i=1}^n \delta_{X_{ni}} \Rightarrow N$$

as  $n \rightarrow \infty$  if and only if

$$n\mathbb{P}(X_{n1} \in \cdot) \Rightarrow m(\cdot)$$

as  $n \rightarrow \infty$ .

### 4 Operators on $\ell^2(V)$

Let  $V$  be a countable set and let  $\ell^2(V)$  denote the Hilbert space defined by the inner product

$$\langle \phi, \psi \rangle := \sum_{u \in V} \bar{\phi}_u \psi_u, \quad \phi_u = \langle \delta_u, \phi \rangle,$$

where  $\delta_u$  is the unit vector supported on  $u \in V$ . Let  $\mathcal{D}(V)$  denote the dense subset of  $\ell^2(V)$  of vectors with finite support. Let  $(w_{uv})_{u,v \in V}$  be a collection of real numbers with  $w_{uv} = w_{vu}$  such that for all  $u \in V$ ,

$$\sum_{v \in V} |w_{uv}|^2 < \infty.$$

We then define a symmetric linear operator  $A$  with domain  $\mathcal{D}(V)$  by

$$\langle \delta_u, A\delta_v \rangle = \langle \delta_v, A\delta_u \rangle = w_{uv}. \tag{4.1}$$

**Definition 4.1** (Local Convergence) *Suppose  $(A_n)$  is a sequence of bounded operators on  $\ell^2(V)$  and  $A$  is a linear operator on  $\ell^2(V)$  with domain  $\mathcal{D}(A) \supset \mathcal{D}(V)$ . For any  $u, v \in V$ , we say that  $(A_n, u)$  converges locally to  $(A, v)$ , and write*

$$(A_n, u) \rightarrow (A, v),$$

if there exists a sequence of bijections  $\sigma_n : V \rightarrow V$  such that  $\sigma_n(v) = u$  and, for all  $\phi \in \mathcal{D}(V)$ ,

$$\sigma_n^{-1} A_n \sigma_n \phi \rightarrow A \phi,$$

in  $\ell^2(V)$ , as  $n \rightarrow \infty$ .

Here we use  $\sigma_n$  for the bijection on  $V$  and the corresponding linear isometry defined in the obvious way. This notion of convergence is useful to random matrices for two reasons. First, we will make a choice on how to define the action of an  $n \times n$  matrix on  $\ell^2(V)$ , and the bijections  $\sigma_n$  help ensure the choice of location for the support of the matrix does not matter. Second, local convergence also gives convergence of the resolvent operator at the distinguished points  $u, v \in V$ . This comes down to the fact that local convergence is strong operator convergence, up to the isometries. See [8] for details.

**Theorem 4.2** (Theorem 2.2 in [7]) *If  $(A_n)_{n=1}^\infty$  and  $A$  are self-adjoint operators such that  $(A_n, u)$  converges locally to  $(A, v)$  for some  $u, v \in V$ , then, for all  $z \in \mathbb{C}_+$ ,*

$$\langle \delta_u, (A_n - z)^{-1} \delta_u \rangle \rightarrow \langle \delta_v, (A - z)^{-1} \delta_v \rangle \tag{4.2}$$

as  $n \rightarrow \infty$ .

To apply this to random operators, we say that  $(A_n, u) \rightarrow (A, v)$  in distribution if there exists a sequence of random bijections  $\sigma_n$  such that  $\sigma_n^{-1} A_n \sigma_n \phi \rightarrow A \phi$  in distribution for every  $\phi \in \mathcal{D}(V)$ .

### 5 Poisson Weighted Infinite Tree

Let  $\rho$  be a positive Radon measure on  $\mathbb{R} \setminus \{0\}$ . **PWIT**( $\rho$ ) is the random infinite weighted rooted tree defined as follows. The vertex set of the tree is identified with  $\mathbb{N}^f := \bigcup_{k \in \mathbb{N} \cup \{0\}} \mathbb{N}^k$  by indexing the root as  $\mathbb{N}^0 = \emptyset$ , the offspring of the root as  $\mathbb{N}$  and, more generally, the offspring of some  $v \in \mathbb{N}^k$  as  $(v1), (v2), \dots \in \mathbb{N}^{k+1}$ . Define  $T$  as the tree on  $\mathbb{N}^f$  with edges between parents and offspring. Let  $\{\Xi_v\}_{v \in \mathbb{N}^f}$  be independent realizations of a Poisson point process with intensity measure  $\rho$ . Let  $\Xi_\emptyset = \{y_1, y_2, \dots\}$  be ordered such that  $|y_1| \geq |y_2| \geq \dots$  with the convention  $y_i = 0$  for all  $i$  large enough<sup>4</sup> if  $\rho(\mathbb{R} \setminus \{0\}) < \infty$ , and assign the weight  $y_i$  to the edge between  $\emptyset$  and  $i$ , assuming such an ordering is possible. More generally assign the weight  $y_{vi}$  to the edge between  $v$  and  $vi$  where  $\Xi_v = \{y_{v1}, y_{v2}, \dots\}$  and  $|y_{v1}| \geq |y_{v2}| \geq \dots$ , again with the convention  $y_{vi} = 0$  for all  $i$  larger than  $\Xi_v(\mathbb{R} \setminus \{0\})$  if  $\rho(\mathbb{R} \setminus \{0\}) < \infty$ .

For a measure  $m$  on  $\mathbb{R} \setminus \{0\}$  satisfying (1.5) and a realization of **PWIT**( $m$ ) define the linear operator  $A$  on  $\mathcal{D}(\mathbb{N}^f)$  by the formulas

$$\langle \delta_v, A \delta_{vk} \rangle = \langle \delta_{vk}, A \delta_v \rangle = y_{vk} \tag{5.1}$$

<sup>4</sup> If  $\rho(\mathbb{R} \setminus \{0\}) < \infty$  then the number of points in  $\mathbb{R} \setminus \{0\}$  is a Poisson random variable. By large enough we mean larger than this random variable.

and  $\langle \delta_v, A\delta_u \rangle = 0$  otherwise. From (1.5) one can see that the points in  $\Xi_v$  are almost surely square summable for every  $v \in \mathbb{N}^f$ , and thus  $A$  is a well defined linear operator on  $\mathcal{D}(\mathbb{N}^f)$ , though is possibly unbounded on  $\ell^2(\mathbb{N}^f)$ .

### 5.1 Poisson Weighted Infinite Tree with Loops

The Poisson weighted infinite tree has been utilized in [7–9, 11, 24] to study the empirical spectral distribution of heavy-tailed random matrices by showing the random matrices converge to the operator defined by (5.1) for an appropriate measure  $m$ . One key feature of those matrices is the diagonal elements are negligible when compared to the largest entries in a row or column. This will not be the case for the Laplacian matrix  $L_n$ , thus we will need to define an operator on a slightly modified graph.

Let  $m$  be a measure on  $\mathbb{R} \setminus \{0\}$  such that

$$\int_{\mathbb{R} \setminus \{0\}} |x| \wedge 1 dm(x) < \infty. \tag{5.2}$$

Define the Poisson weighted infinite tree with loops **PWITL**( $m$ ) as the random weighted graph with vertex set  $\mathbb{N}^f$  and edge set  $E \cup \bigcup_{v \in \mathbb{N}^f} \{v, v\}$  where  $E$  is the edge set of **PWIT**( $m$ ). The weights on edges in  $E$  of **PWITL**( $m$ ) are the weights on edges in  $E$  of **PWIT**( $m$ ), while the weight on a loop  $\{v, v\}$  is

$$y_{vv} := -y_{ul} - \sum_{k=1}^{\infty} y_{vk}, \tag{5.3}$$

where  $ul = v$  if  $v$  is not  $\emptyset$  and the weight on  $\{\emptyset, \emptyset\}$  is

$$y_{\emptyset\emptyset} := - \sum_{k=1}^{\infty} y_k. \tag{5.4}$$

(5.2) is enough to guarantee  $y_{vv}$  is a well-defined random variable, see Lemma 3.3. Define the operator  $L$  by

$$\langle \delta_v, L\delta_{vk} \rangle = \langle \delta_{vk}, L\delta_v \rangle = y_{vk} \quad \text{and} \quad \langle \delta_v, L\delta_v \rangle = y_{vv}, \tag{5.5}$$

and  $\langle \delta_v, L\delta_u \rangle = 0$  otherwise. In which case we say  $L$  is the operator associated to **PWITL**( $m$ ).

We will show the sequence  $\{(L_n, 1)\}_{n \geq 1}$  converges locally in distribution to  $(L, \emptyset)$  where  $L$  is the linear operator on  $\ell^2(\mathbb{N}^f)$  associated to the **PWITL**( $m$ ).

### 5.2 Self-Adjointness

In this section, we review and apply a criteria established by Bordenave, Caputo, and Chafaï in [7] for unbounded operators to be essentially self-adjoint. There are

two minor issues which prevent immediately applying their results to the operator  $L$  associated to **PWITL**( $m$ ). First is they consider operators with skeletons which are trees, and not trees with loops. This is easy to overcome. The second obstacle is in the application of the criteria they consider only point processes associated to  $\alpha$ -stable distributions and not more general infinitely divisible distributions. This is overcome by Lemma 3.2.

**Proposition 5.1** (Lemma A.3 in [7]) *Let  $A$  be a linear operator on  $\ell^2(\mathbb{N}^f)$  defined by (4.1). We say  $u \sim v$  if  $u = v$ ,  $u = vk$ , or  $v = uk$  for some  $k \in \mathbb{N}$ . Assume  $w_{uv} = 0$  if  $u \not\sim v$ . Suppose there exists a constant  $\kappa > 0$  and sequence of finite connected subsets  $S_n \subset \mathbb{N}^f$ , such that  $S_n \subset S_{n+1}$ ,  $\mathbb{N}^f = \bigcup_{n \in \mathbb{N}} S_n$ , and for every  $n$  and  $v \in S_n$ ,*

$$\sum_{u \notin S_n : u \sim v} |w_{uv}|^2 \leq \kappa. \tag{5.6}$$

Then  $A$  is essentially self-adjoint.

**Proof** Proposition 5.1 is not stated identically to Lemma A.3 in [7]; however, the only added assumption is that vertices are connected to themselves, so that the graph of the skeleton of  $A$  is not a tree. The step in the proof given in [7] which uses the tree structure is the fact that if  $v \in S_n$ ,  $u \sim v$ , and  $v' \in S_n \setminus \{v\}$  then  $u \not\sim v'$  which is also true for a tree with loops. □

**Proposition 5.2** (Proposition A.2 in [7]) *Let  $m$  be a measure on  $\mathbb{R} \setminus \{0\}$  satisfying (1.5). Let  $\{\Xi_v\}_{v \in \mathbb{N}^f}$  be a collection of Poisson point process on  $\mathbb{R} \setminus \{0\}$  with intensity measure  $m$ . Let  $\Xi_\emptyset = \{y_1, y_2, \dots\}$  be ordered such that  $|y_1| \geq |y_2| \geq \dots$ , and  $\Xi_v = \{y_{v1}, y_{v2}, \dots\}$  be ordered such that  $|y_{v1}| \geq |y_{v2}| \geq \dots$  with the convention the  $y_{vk}$  or  $y_k$  are eventually zero if  $m(\mathbb{R} \setminus \{0\}) < \infty$ . Additionally let  $\{y_{vv}\}_{v \in \mathbb{N}^f}$  be a collection of real random variables. Define the symmetric linear operator  $A$  on  $\ell^2(\mathbb{N}^f)$  by*

$$\langle \delta_v, A\delta_{vk} \rangle = \langle \delta_{vk}, A\delta_v \rangle = y_{vk}, \quad \text{and} \quad \langle \delta_v, A\delta_v \rangle = y_{vv},$$

and  $\langle \delta_u, A\delta_v \rangle = 0$  otherwise. Then, with probability 1,  $A$  is essentially self-adjoint.

**Proof** Proposition 5.2 may not initially appear to be Proposition A.2 in [7]; however, the proofs are essentially identical. The only minor difference is replacing Lemma A.4 in [7] with Lemma 3.2. □

### 6 Local Convergence for the Laplacian of Lévy–Khintchine Matrices

For an  $n \times n$  matrix  $M$ , extend  $M$  to a bounded operator on  $\ell^2(\mathbb{N}^f)$  as follows. For  $1 \leq i, j, \leq n$ , let  $\langle \delta_i, M\delta_j \rangle = M_{ij}$ , and  $\langle \delta_u, M\delta_v \rangle = 0$  otherwise.

**Theorem 6.1** *Let  $L_n$  be the matrix defined by (2.1), where  $\{A_n\}_{n \geq 1}$ , a Lévy–Khintchine random matrix ensemble satisfying **CI**. Let  $L$  be the linear operator on  $\ell^2(\mathbb{N}^f)$  associated with **PWITL**( $m$ ). Then, in distribution,  $(L_n, 1) \rightarrow (L, \emptyset)$ , as  $n \rightarrow \infty$ .*

The rest of this section is devoted to the proof of Theorem 6.1. Before considering  $(L_n, 1)$ , we begin by showing  $(A_n, 1)$  converges to  $(L + D, \emptyset)$  where  $D$  is a diagonal operator. This follows from the work of Jung in [24]. We include the proof to establish notation and for the convenience of the reader. We define a network as a graph with edge weights taking values in some normed space. To begin, let  $G_n$  be the complete network, without loops, on  $\{1, \dots, n\}$  whose weight on edge  $\{i, j\}$  equals  $\xi_{ij}^n$  for some collection  $(\xi_{ij}^n)_{1 \leq i < j \leq n}$  of random variables taking values in some normed space. Now consider the rooted network  $(G_n, 1)$  with the distinguished vertex 1. For any realization  $(\xi_{ij}^n)$ , and for any  $B, H \in \mathbb{N}$  such that  $(B^{H+1} - 1)/(B - 1) \leq n$ , we will define a finite rooted subnetwork  $(G_n, 1)^{B,H}$  of  $(G_n, 1)$  whose vertex set coincides with a  $B$ -ary tree of depth  $H$ . To this end, we partially index the vertices of  $(G_n, 1)$  as elements in

$$J_{B,H} := \bigcup_{l=0}^H \{1, \dots, B\}^l \subset \mathbb{N}^f,$$

the indexing being given by an injective map  $\sigma_n$  from  $J_{B,H}$  to  $V_n := \{1, \dots, n\}$ . We set  $I_\emptyset := \{1\}$  and the index of the root  $\sigma_n^{-1}(1) = \emptyset$ . The vertex  $v \in V_n \setminus I_\emptyset$  is given the index  $(k) = \sigma_n^{-1}(v)$ ,  $1 \leq k \leq B$ , if  $\xi_{1,v}^n$  has the  $k$ -th largest norm value among  $\{\xi_{1j}^n, j \neq 1\}$ , ties being broken by lexicographic order<sup>5</sup>. This defines the first generation, and let  $I_1$  be the union of  $I_\emptyset$  and this generation. If  $H \geq 2$  repeat this process for the vertex labeled (1) on  $V_n \setminus I_1$  to order  $\{\xi_{(1)j}^n\}_{j \in V_n \setminus I_1}$  to get  $\{11, 12, \dots, 1B\}$ . Define  $I_2$  to be the union of  $I_1$ , and this new collection. Repeat again for (2), (3),  $\dots$ ,  $(B)$  to get the second generation and so on. Call this vertex set  $V_n^{B,H} = \sigma_n J_{B,H}$ .

For a realization  $T$  of **PWITL**( $m$ ), recall we assign the weight  $y_{vk}$  to the edge  $\{v, vk\}$  and the weight  $y_{vv}$  to the edge  $\{v, v\}$ . Then  $(T, \emptyset)$  is a rooted network. Call  $(T, \emptyset)^{B,H}$  the finite rooted subnetwork obtained by restricting  $(T, \emptyset)$  to the vertex set  $J_{B,H}$ , and the edge set without the loops. If an edge is not present in  $(T, \emptyset)^{B,H}$  assign the weight 0. We say a sequence  $(G_n, 1)^{B,H}$ , for fixed  $B$  and  $H$ , converges in distribution, as  $n \rightarrow \infty$ , to  $(T, \emptyset)^{B,H}$  if the joint distribution of the weights converges weakly.

Let  $\xi_{ij}^n = L_{ij} = A_{ij}^{(n)}$ , where  $L_{ij}$  is the  $ij$ -th entry of  $L_n$  for  $1 \leq i < j \leq n$ . We aim to show with the choice of weights  $(\xi_{ij}^n)_{1 \leq i < j \leq n}$  that for fixed  $B, H$   $(G_n, 1)^{B,H}$  converges weakly to  $(T, \emptyset)^{B,H}$ .

Order the elements of  $J_{B,H}$  lexicographically, i.e.  $\emptyset < 1 < 2 < \dots < B < 11 < 12 < \dots < B \dots B$ . For  $v \in J_{B,H}$  let  $\mathcal{O}_v$  denote the offspring of  $v$  in  $(G_n, 1)^{B,H}$ . By construction  $I_\emptyset = \{1\}$  and  $I_v = \sigma_n(\bigcup_{w < v} \mathcal{O}_w)$ , where  $w < v$  must be strict in this union. Thus at every step of the indexing procedure we order the weights of neighboring edges not already considered at a previous step. Thus for all  $v$ ,

$$(\xi_{\sigma_n(v),j}^n)_{j \notin I_v} \stackrel{d}{=} (\xi_{1,j}^n)_{1 < j \leq n - |I_v|}.$$

<sup>5</sup> To help keep track of notation in this section, note that  $v = (w) \in V_n$  if  $w \in J_{B,H}$  and  $\sigma_n(w) = v$ .

Note that by independence, Proposition 3.7 still holds if you take the sum of Dirac measures at the random variables over  $\{1, \dots, n\} \setminus I$  for any fixed finite set  $I$ . Thus, by Proposition 3.7 the weights from a fixed parent to its offspring in  $(G_n, 1)^{B,H}$  converge weakly to those of  $(T, \varnothing)^{B,H}$ . By independence we can extend this to joint convergence. Recall  $(G_n, 1)^{B,H}$  is a complete graph and not a tree with loops. Thus, it remains to show the edges in  $(G_n, 1)^{B,H}$  which were not considered in the sorting procedure converge to 0. This was shown for heavy-tailed weights in [7] and for more general Lévy–Khintchine weights in [24].

Let  $L$  be the operator associated to **PWITL**( $m$ ). For fixed  $B, H$  let  $\sigma_n^{B,H}$  be the map  $\sigma_n$  above associated with  $(G_n, 1)^{B,H}$ , and arbitrarily extend  $\sigma_n^{B,H}$  to a bijection on  $\mathbb{N}^f$ , where  $V_n$  is considered in the natural way as a subset of the offspring of  $\varnothing$ . From the Skorokhod representation theorem, we may assume  $(G_n, 1)^{B,H}$  converges almost surely to  $(T, \varnothing)^{B,H}$ . Thus, there are sequences  $B_n, H_n$  tending to infinity and  $\hat{\sigma}_n := \sigma_n^{B_n, H_n}$  such that for any pair  $v, w \in \mathbb{N}^f$  with  $w \neq v$ ,  $\xi_{\hat{\sigma}_n(v), \hat{\sigma}_n(w)}^n$  converges almost surely to

$$\begin{cases} y_{vk}, & \text{if } w = vk \text{ for some } k \\ y_{wk}, & \text{if } v = wk \text{ for some } k \\ 0, & \text{otherwise.} \end{cases}$$

Thus for any  $u, v \in \mathbb{N}^f$  with  $u \neq v$

$$\langle \delta_u, \hat{\sigma}_n^{-1} L_n \hat{\sigma}_n \delta_v \rangle = \langle \delta_u, \hat{\sigma}_n^{-1} A_n \hat{\sigma}_n \delta_v \rangle \rightarrow \langle \delta_u, (L + D) \delta_v \rangle = \langle \delta_u, L \delta_v \rangle \tag{6.1}$$

almost surely. We now consider the diagonal elements. Let  $u \in \mathbb{N}^f, B = H = k$  for some  $k \in \mathbb{N}$  such that  $u \in J_{k,k}$ . From the above we know almost surely

$$\sum_{v \in J_{k,k}} \xi_{(u),(v)}^n \rightarrow \sum_{v \in J_{k,k}, v \sim u} y_v. \tag{6.2}$$

Assume  $v \notin J_{k,k}$ , then  $\xi_{(u),(v)}^n \rightarrow 0$  almost surely. By the uniform summability condition of **C1** we have almost surely

$$\sum_{v \notin J_{k,k}} \xi_{(u),(v)}^n \rightarrow 0. \tag{6.3}$$

As  $k$  was arbitrarily large we have that almost surely for any  $v \in \mathbb{N}^f$

$$\langle \delta_v, \hat{\sigma}_n^{-1} L_n \hat{\sigma}_n \delta_v \rangle \rightarrow \langle \delta_v, L \delta_v \rangle. \tag{6.4}$$

From linearity it suffices to show for every  $v \in \mathbb{N}^f$  that  $\hat{\sigma}_n^{-1} L_n \hat{\sigma}_n \delta_v \rightarrow L \delta_v$ , i.e.

$$\sum_{u \in \mathbb{N}^f} \left[ \langle \delta_u, \hat{\sigma}_n^{-1} L_n \hat{\sigma}_n \delta_v \rangle - \langle \delta_u, L \delta_v \rangle \right]^2 \rightarrow 0. \tag{6.5}$$



We have shown  $\langle \delta_u, \hat{\sigma}_n^{-1} L_n \hat{\sigma}_n \delta_v \rangle \rightarrow \langle \delta_u, L \delta_v \rangle$  almost surely for every  $u \in \mathbb{N}^f$ , thus (6.5) holds if  $\{\langle \delta_u, \hat{\sigma}_n^{-1} L_n \hat{\sigma}_n \delta_v \rangle\}_{u \in \mathbb{N}^f}$  is uniformly square-summable. This follows from the uniform summability of **C1**. This completes the proof of Theorem 6.1.

We will need the following extension of Theorem 6.1.

**Theorem 6.2** *Let  $L_n$  be the matrix defined by (2.1) for  $\{A_n\}_{n \geq 1}$ , a Lévy–Khintchine random matrix ensemble satisfying **C1**. If  $L$  and  $L'$  are two independent copies of the linear operator on  $\ell^2(\mathbb{N}^f)$  associated to **PWITL**( $m$ ), then, in distribution,  $(L_n \oplus L_n, (1, 2)) \rightarrow (L \oplus L', (\emptyset, \emptyset))$  as  $n \rightarrow \infty$ .*

**Proof** Using Proposition 2.6 in [7] and the arguments above, we can construct isometries  $\sigma_n$  on  $\ell^2(\mathbb{N}^f) \oplus \ell^2(\mathbb{N}^f)$  such that for any  $v \in \mathbb{N}^f$ ,  $\sigma_n^{-1}(L_n \oplus L_n)\sigma_n(\delta_v, 0) \rightarrow L \oplus L'(\delta_v, 0)$  and  $\sigma_n^{-1}(L_n \oplus L_n)\sigma_n(0, \delta_v) \rightarrow L \oplus L'(0, \delta_v)$  in  $\ell^2(\mathbb{N}^f) \oplus \ell^2(\mathbb{N}^f)$  almost surely. The result then follows by linearity.  $\square$

### 7 Resolvent Convergence and the Proof of Theorem 2.4

**Theorem 7.1** *Let  $s_{L_n}(z)$  be the Stieltjes transform of  $\mu_{L_n}$ , and let  $s_\emptyset(z)$  be the Stieltjes transform of the measure  $\mu_\emptyset$  defined by*

$$\langle \delta_\emptyset, f(L)\delta_\emptyset \rangle = \int_{\mathbb{R}} f d\mu_\emptyset \tag{7.1}$$

for any continuous bounded function  $f : \mathbb{R} \rightarrow \mathbb{C}$ , where  $f(L)$  is defined by the continuous functional calculus. Then,

$$\lim_{n \rightarrow \infty} \mathbb{E}s_{L_n}(z) = \mathbb{E}s_\emptyset(z) \tag{7.2}$$

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

**Proof** For  $z \in \mathbb{C}_+$ , we define the operators

$$R_n(z) = (L_n - z)^{-1}, \tag{7.3}$$

and

$$R(z) = (L - z)^{-1}. \tag{7.4}$$

Additionally for  $u, v \in \mathbb{N}^f$ , we define the functions  $R_n(z)_{uv}, R(z)_{uv} : \mathbb{C}_+ \rightarrow \mathbb{C}$  by

$$R_n(z)_{uv} := \langle \delta_u, R_n(z)\delta_v \rangle, \quad \text{and} \quad R(z)_{uv} := \langle \delta_u, R(z)\delta_v \rangle. \tag{7.5}$$

From Proposition 5.2,  $L$  is self-adjoint with probability 1. Thus from Theorem 4.2 and Theorem 6.1

$$R_n(z)_{11} = \langle \delta_1, (L_n - z)^{-1}\delta_1 \rangle \Rightarrow \langle \delta_\emptyset, (L - z)^{-1}\delta_\emptyset \rangle = R(z)_{\emptyset\emptyset}. \tag{7.6}$$

For every  $z \in \mathbb{C}_+$ ,  $R_n(z)_{11}$  and  $R(z)_{\emptyset\emptyset}$  are bounded, thus

$$\lim_{n \rightarrow \infty} \mathbb{E}R_n(z)_{11} = \mathbb{E}R(z)_{\emptyset\emptyset}. \tag{7.7}$$

By definition  $s_{\emptyset}(z) = R(z)_{\emptyset\emptyset}$ , while

$$s_{L_n}(z) = \frac{1}{n} \operatorname{tr}(L_n - z)^{-1} = \frac{1}{n} \sum_{i=1}^n R_n(z)_{ii}. \tag{7.8}$$

It is clear from the matrix of cofactors method of inversion  $R_n(z)_{ii} \stackrel{d}{=} R_n(z)_{jj}$  for every  $i, j \in [n]$ . Thus,

$$\begin{aligned} \mathbb{E}s_{L_n}(z) &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}R_n(z)_{ii} \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}R_n(z)_{11} \\ &= \mathbb{E}R_n(z)_{11}. \end{aligned}$$

This completes the proof. □

### 7.1 Proof of Theorem 2.4

We are now ready to complete the proof of Theorem 2.4. From Lemma A.1,  $\{\mu_{L_n}\}_{n \geq 1}$  is almost surely tight. Consider the Stieltjes transform  $s_{L_n}$  of  $\mu_{L_n}$ . From tightness and Lemma B.2, it is enough to prove that almost surely there exists a probability measure with Stieltjes transform  $s$ , such that for any subsequence  $\{n_k\}$

$$\lim_{n_k \rightarrow \infty} s_{L_{n_k}}(z) = s(z), \tag{7.9}$$

for all  $z \in \mathbb{C}_+$ . We know from Theorem 7.1 that for all  $z \in \mathbb{C}_+$

$$\lim_{n \rightarrow \infty} \mathbb{E}s_{L_n}(z) = \mathbb{E}s_{\emptyset}(z). \tag{7.10}$$

We now upgrade this to almost surely convergence of  $s_{L_n}(z)$  to  $\mathbb{E}s_{\emptyset}(z)$ . For  $z \in \mathbb{C}_+$

$$\mathbb{E}|s_{L_n}(z) - \mathbb{E}s_{\emptyset}(z)| \leq \mathbb{E}|s_{L_n}(z) - \mathbb{E}s_{L_n}(z)| + |\mathbb{E}s_{L_n}(z) - \mathbb{E}s_{\emptyset}(z)|.$$

For  $z \in \mathbb{C}_+$

$$\mathbb{E}|s_{L_n}(z) - \mathbb{E}s_{L_n}(z)| = \mathbb{E} \left| \frac{1}{n} \sum_{i=1}^n (R_n(z)_{ii} - \mathbb{E}R_n(z)_{ii}) \right|, \tag{7.11}$$

and by the exchangeability of the matrix entries

$$\begin{aligned} \mathbb{E} \left[ \left| \frac{1}{n} \sum_{i=1}^n R_n(z)_{ii} - \mathbb{E}R_n(z)_{ii} \right|^2 \right] &= \frac{1}{n} \mathbb{E} |R_n(z)_{11} - \mathbb{E}R_n(z)_{11}|^2 \\ &\quad + \frac{n-1}{n} \text{Cov}(R_n(z)_{11}, R_n(z)_{22}) \\ &\leq \frac{1}{n \text{Im}(z)^2} \\ &\quad + \frac{n-1}{n} \text{Cov}(R_n(z)_{11}, R_n(z)_{22}). \end{aligned}$$

From Theorems 4.2 and 6.2, we know  $R_n(z)_{11}$  and  $R_n(z)_{22}$  are asymptotically independent random variables bounded uniformly in  $n$ , and thus asymptotically uncorrelated. From this, we get

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left| \frac{1}{n} \sum_{i=1}^n R_n(z)_{ii} - \mathbb{E}R_n(z)_{ii} \right|^2 \right] = 0,$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E}|s_{L_n}(z) - \mathbb{E}s_{\emptyset}(z)| = 0. \tag{7.12}$$

Taking  $\mu_m = \mathbb{E}\mu_{\emptyset}$  completes the proof of Theorem 2.4.

### 8 Proof of Theorem 2.5

We will follow the approach of [7] and take advantage of the tree structure on  $\mathbb{N}^f$  to arrive at (2.7) before proving uniqueness. Let  $L$  be the operator associated to **PWITL**( $m$ ), we have already seen that  $s_m(z) = \mathbb{E}s_{\emptyset}(z)$  where

$$s_{\emptyset}(z) = \langle \delta_{\emptyset}, (L - z)^{-1} \delta_{\emptyset} \rangle. \tag{8.1}$$

We now decompose the operator  $L$  as

$$L = C + \bigoplus_{k=1}^{\infty} L_k \tag{8.2}$$

where

$$\begin{aligned} \langle \delta_k, C\delta_{\emptyset} \rangle &= \langle \delta_k, L\delta_{\emptyset} \rangle \\ \langle \delta_k, C\delta_k \rangle &= -y_k \\ \langle \delta_{\emptyset}, C\delta_{\emptyset} \rangle &= \langle \delta_{\emptyset}, L\delta_{\emptyset} \rangle \end{aligned} \tag{8.3}$$

and for every  $k \in \mathbb{N}$ ,  $L_k$  is supported on  $k\mathbb{N}^f = \{kv \in \mathbb{N}^f : v \in \mathbb{N}^f\}$ . Note  $\langle \delta_v, C\delta_u \rangle = 0$  for any other combination of  $u, v \in \mathbb{N}^f$ . Under this decomposition  $\{L_k\}_{k \geq 1}$  is a collection of i.i.d. random operators each equal in distribution, up to an isometry, to  $L$ . For convenience, define the operator  $\tilde{L}$  by

$$\tilde{L} := \bigoplus_{k=1}^{\infty} L_k, \tag{8.4}$$

and the operators  $R(z) := (L - z)^{-1}$  and  $\tilde{R}(z) := (\tilde{L} - z)^{-1}$  for  $z \in \mathbb{C}_+$ . From (8.2) we get the resolvent identity

$$\tilde{R}(z)CR(z) = \tilde{R}(z) - R(z). \tag{8.5}$$

Additionally denote by  $R_{uv}(z) := \langle \delta_u, R(z)\delta_v \rangle$  and  $\tilde{R}_{uv}(z) := \langle \delta_u, \tilde{R}(z)\delta_v \rangle$ . Note  $\tilde{R}_{\emptyset\emptyset}(z) = -z^{-1}$ ,  $\tilde{R}_{kl}(z) = 0$  for all  $k, l \in \mathbb{N}$  with  $k \neq l$ , and  $\tilde{R}_{\emptyset k}(z) = 0 = \tilde{R}_{k\emptyset}(z)$  for all  $k \in \mathbb{N}$ .

From (8.5), one immediately gets

$$\langle \delta_k, \tilde{R}(z)CR(z)\delta_{\emptyset} \rangle = -R_{k\emptyset}(z). \tag{8.6}$$

It also follows that

$$\begin{aligned} \langle \delta_k, \tilde{R}(z)CR(z)\delta_{\emptyset} \rangle &= \left\langle \delta_k, \tilde{R}(z)C \sum_{v \in \mathbb{N}^f} R_v(z)\delta_v \right\rangle \\ &= \left\langle \delta_k, \tilde{R}(z)R_{\emptyset\emptyset}(z)y_{\emptyset\emptyset}\delta_{\emptyset} \right\rangle + \left\langle \delta_k, \tilde{R}(z) \sum_{l \in \mathbb{N}} R_{\emptyset\emptyset}(z)y_l\delta_l \right\rangle \\ &\quad + \left\langle \delta_k, \tilde{R}(z) \sum_{j \in \mathbb{N}} R_{j\emptyset}(z)y_j\delta_{\emptyset} \right\rangle - \left\langle \delta_k, \tilde{R}(z) \sum_{j \in \mathbb{N}} R_{j\emptyset}(z)y_j\delta_j \right\rangle \\ &= 0 + \tilde{R}_{kk}(z)y_k R_{\emptyset\emptyset}(z) + 0 - \tilde{R}_{kk}(z)y_k R_{k\emptyset}(z). \end{aligned}$$

Rearranging we arrive at

$$R_{k\emptyset}(z) = \frac{\tilde{R}_{kk}(z)y_k R_{\emptyset\emptyset}(z)}{\tilde{R}_{kk}y_k - 1}. \tag{8.7}$$

A similar computation for  $\langle \delta_{\emptyset}, \tilde{R}(z)CR(z)\delta_{\emptyset} \rangle$  gives

$$\tilde{R}_{\emptyset\emptyset}(z) - R_{\emptyset\emptyset}(z) = \tilde{R}_{\emptyset\emptyset}(z)y_{\emptyset\emptyset}R_{\emptyset\emptyset}(z) + \sum_{j=1}^{\infty} \tilde{R}_{\emptyset\emptyset}(z)y_j R_{j\emptyset}(z). \tag{8.8}$$

Combining (8.7) and (8.8) gives

$$\begin{aligned} \tilde{R}_{\emptyset\emptyset}(z) &= R_{\emptyset\emptyset}(z) + \tilde{R}_{\emptyset\emptyset}(z)y_{\emptyset\emptyset}R_{\emptyset\emptyset}(z) + \sum_{j=1}^{\infty} \tilde{R}_{\emptyset\emptyset}(z)y_j R_{j\emptyset}(z) \\ &= R_{\emptyset\emptyset}(z) + \tilde{R}_{\emptyset\emptyset}(z)y_{\emptyset\emptyset}R_{\emptyset\emptyset}(z) + \sum_{j=1}^{\infty} \tilde{R}_{\emptyset\emptyset}(z)y_j \frac{\tilde{R}_{jj}(z)y_j R_{\emptyset\emptyset}(z)}{\tilde{R}_{jj}y_j - 1} \\ &= R_{\emptyset\emptyset}(z) \left[ 1 + \tilde{R}_{\emptyset\emptyset}(z)y_{\emptyset\emptyset} + \sum_{j=1}^{\infty} \tilde{R}_{\emptyset\emptyset}(z)y_j \frac{\tilde{R}_{jj}(z)y_j}{\tilde{R}_{jj}y_j - 1} \right], \end{aligned}$$

which, along with  $\tilde{R}_{\emptyset\emptyset}(z) = -z^{-1}$ , implies

$$R_{\emptyset\emptyset} = - \left( z - y_{\emptyset\emptyset} - \sum_{j=1}^{\infty} \frac{\tilde{R}_{jj}(z)y_j^2}{\tilde{R}_{jj}(z)y_j - 1} \right)^{-1}. \tag{8.9}$$

Noting  $y_{\emptyset\emptyset} = -\sum_{j=1}^{\infty} y_j$  gives (2.7). Note that for  $j \in \mathbb{N}$   $\tilde{R}_{jj}(z)$  depends only on  $z$  and  $L_j$ , and hence  $\{\tilde{R}_{jj}(z)\}_{j \in \mathbb{N}}$  is a collection of i.i.d. random variables independent of  $\{y_j\}_{j \in \mathbb{N}}$ .

### 8.1 Uniqueness

In this section, we prove uniqueness of the solution to (2.7) from Theorem 2.5. While the argument is technical, the core is a contraction approach. We will show the map  $T$  defined below in (8.10) would contract, in an appropriate metric, two fixed points belonging to a nice subset of all probability measures on the space of Stieltjes transforms. We then extend this result to any two potential fixed points by moving from this metric to a functional separating distinct points.

Let  $\mathcal{S}$  be the set of Stieltjes transforms of probability measures on  $\mathbb{R}$  and  $\mathcal{P}(\mathcal{S})$  be the set of probability measures on  $\mathcal{S}$ . Define  $T : \mathcal{P}(\mathcal{S}) \rightarrow \mathcal{P}(\mathcal{S})$  as follows: for  $\mu \in \mathcal{P}(\mathcal{S})$

$$T(\mu) := \mathcal{L} \left( - \left( z - \sum_{j=1}^N \frac{y_j}{s_j(z)y_j - 1} \right)^{-1} \right), \tag{8.10}$$

where  $\{s_j\}$  are i.i.d. with distribution  $\mu$ ,  $\{y_j\}_{j=1}^{\infty}$  is a Poisson point process with a fixed intensity measure  $m$  independent of the collection  $\{s_j\}_{j \geq 1}$ ,  $N$  is a Poisson random variable with mean  $m(\mathbb{R})$  such that  $y_j = 0$  if  $j > N$ , and  $\mathcal{L}(X)$  is the law of a random variable  $X$ . Thus the distribution of  $s_{\emptyset}$  is a fixed point of  $T$  and we aim to show it is the unique fixed point. The notation of distance for which  $T$  contracts fixed

points will involve the infimum over all couplings of these fixed point measures. Let  $\mu_1, \mu_2 \in \mathcal{P}(\mathcal{S})$  be two fixed points of  $T$  and let  $(s(z), r(z))$  be an arbitrary coupling of  $\mu_1$  and  $\mu_2$ . Additionally let  $\mu_r$  and  $\mu_s$  be the random probability measures on  $\mathbb{R}$  defined uniquely by

$$s(z) = \int_{\mathbb{R}} \frac{1}{x - z} d\mu_s(x) \text{ and } r(z) = \int_{\mathbb{R}} \frac{1}{x - z} d\mu_r(x),$$

for all  $z \in \mathbb{C}_+$ . For now, we will assume there exists  $M \in \mathbb{N}$  such that almost surely  $\mu_r([-M, M]) \geq \frac{1}{2}$  and  $\mu_s([-M, M]) \geq \frac{1}{2}$ . This assumption will be removed later. As  $r$  and  $s$  are analytic functions on the upper half plane, we will consider them only on the box

$$\mathcal{C}_M = \left\{ z \in \mathbb{C} : |\operatorname{Re}(z)| \leq \frac{1}{2}, f_m(M) \leq \operatorname{Im}(z) \leq f_m(M) + 1 \right\}, \tag{8.11}$$

where  $f_m$  is a positive increasing function on  $\mathbb{N}$  such that  $f_m(M) \rightarrow \infty$  as  $M \rightarrow \infty$ , which will be chosen later to satisfy (8.17) below. Note for  $z \in \mathcal{C}_M$ , the assumption on  $\mu_r$  and  $\mu_s$  imply  $\operatorname{Im}(r(z)), \operatorname{Im}(s(z)) \geq \frac{1}{2} \frac{\operatorname{Im}(z)}{(M + \frac{1}{2})^2 + \operatorname{Im}(z)^2}$ . Let  $\{(r_j, s_j)\}_{j=1}^N$  be i.i.d. copies of  $(s(z), r(z))$ . Define the random functions  $\tilde{s}$  and  $\tilde{r}$ , pointwise on  $\mathbb{C}_+$  and the sample space, by

$$\tilde{s}(z) = - \left( z - \sum_{j=1}^N \frac{y_j}{s_j(z)y_j - 1} \right)^{-1} \text{ and } \tilde{r}(z) = - \left( z - \sum_{j=1}^N \frac{y_j}{r_j(z)y_j - 1} \right)^{-1} \tag{8.12}$$

where  $\{y_j\}_{j=1}^\infty$  is a Poisson point process with intensity measure  $m$  independent of the collection  $\{(r_j, s_j)\}_{j=1}^\infty$ . If  $\mu_1$  and  $\mu_2$  are fixed points of  $T$ , then  $(\tilde{r}, \tilde{s})$  is a coupling of  $\mu_1$  and  $\mu_2$ . We show that

$$\mathbb{E} \sup_{z \in \mathcal{C}_M} |\tilde{r}(z) - \tilde{s}(z)| \leq \frac{4}{5} \mathbb{E} \sup_{z \in \mathcal{C}_M} |r(z) - s(z)|$$

for an appropriate choice of  $f_m(M)$  independent of the coupling  $(r, s)$ . First note

$$\begin{aligned} \mathbb{E} \sup_{z \in \mathcal{C}_M} |\tilde{r}(z) - \tilde{s}(z)| &\leq \mathbb{E} \sup_{z \in \mathcal{C}_M} \frac{1}{\operatorname{Im}(z)^2} \sum_{j=1}^N \frac{|r_j(z) - s_j(z)| y_j^2}{|r_j(z)y_j - 1| |s_j(z)y_j - 1|} \\ &\leq \mathbb{E} \sum_{j=1}^N \sup_{z \in \mathcal{C}_M} \frac{1}{\operatorname{Im}(z)^2} \frac{|r_j(z) - s_j(z)| y_j^2}{|r_j(z)y_j - 1| |s_j(z)y_j - 1|}. \end{aligned} \tag{8.13}$$

To handle the denominator, we will consider separately the points where  $\operatorname{Re}(r_j(z)y_j)$  is small and the few points where  $\operatorname{Im}(r_j(z)y_j)$  is large. Let  $\hat{m}$  be equal to  $m$  with support restricted to  $[-f_m(M)/2, f_m(M)/2]$  and  $\tilde{m} := m - \hat{m}$ . Decompose the point process  $\{y_j\}_{j=1}^N$  into two independent Poisson point processes  $\{\hat{y}_j\}_{j=1}^{\hat{N}}$  and  $\{\tilde{y}_j\}_{j=1}^{\tilde{N}}$

with intensity measures  $\hat{m}$  and  $\tilde{m}$ , respectively. We will divide the sum in (8.13) into two sums over these point processes. To begin note for  $z \in \mathcal{C}_M, |s_j(z)\hat{y}_j - 1||r_j(z)\hat{y}_j - 1| \geq 1/4$  and thus

$$\begin{aligned} & \mathbb{E} \sum_{j=1}^{\hat{N}} \sup_{z \in \mathcal{C}_M} \frac{1}{\text{Im}(z)^2} \frac{|r_j(z) - s_j(z)|\hat{y}_j^2}{|r_j(z)\hat{y}_j - 1||s_j(z)\hat{y}_j - 1|} \\ & \leq \mathbb{E} \sum_{j=1}^{\hat{N}} \sup_{z \in \mathcal{C}_M} \frac{4}{\text{Im}(z)^2} |r_j(z) - s_j(z)|\hat{y}_j^2 \\ & \leq \frac{4}{f_m(M)^2} \mathbb{E} \sum_{j=1}^{\hat{N}} \sup_{z \in \mathcal{C}_M} |r_j(z) - s_j(z)|\hat{y}_j^2 \\ & = \frac{4}{f_m(M)^2} \mathbb{E} \sup_{z \in \mathcal{C}_M} |r_1(z) - s_1(z)|I_{M,m}, \end{aligned}$$

where  $I_{M,m} = \int_{-f_m(M)/2}^{f_m(M)/2} y^2 dm(y)$  and the last equality follows from Lemma 3.3 and independence. From (2.3)

$$\begin{aligned} \int_{-f_m(M)/2}^{f_m(M)/2} y^2 dm(y) & \leq \int_{-1}^1 y^2 dm(y) + \int_{1 \leq |x| \leq \sqrt{f_m(M)}} y^2 dm(y) \\ & \quad + \int_{\sqrt{f_m(M)} \leq |x| \leq f_m(M)/2} y^2 dm(y) \\ & \leq C' + C' f_m(M) + \frac{f_m(M)^2}{4} m(\{x : |x| \geq \sqrt{f_m(M)}\}) \\ & \leq C f_m(M)^{2-\varepsilon/2} \end{aligned}$$

where  $\varepsilon > 0$  is from (2.3), and  $C, C' > 0$  are constants which depend only on the measure  $m$ . Thus,

$$\mathbb{E} \sum_{j=1}^{\hat{N}} \sup_{z \in \mathcal{C}_M} \frac{1}{\text{Im}(z)^2} \frac{|r_j(z) - s_j(z)|\hat{y}_j^2}{|r_j(z)\hat{y}_j - 1||s_j(z)\hat{y}_j - 1|} \leq \frac{C'}{f_m(M)^{\varepsilon/2}} \mathbb{E} \sup_{z \in \mathcal{C}_M} |r_1(z) - s_1(z)|. \tag{8.14}$$

To handle the other sum first note  $|s_j(z)\tilde{y}_j - 1||r_j(z)\tilde{y}_j - 1| \geq \tilde{y}_j^2 C_{z,M}^{-1}$  where

$$C_{z,M} = \frac{4 \left( \left( M + \frac{1}{2} \right)^2 + \text{Im}(z)^2 \right)^2}{\text{Im}(z)^2}.$$

Then,

$$\begin{aligned}
 & \mathbb{E} \sum_{j=1}^{\tilde{N}} \sup_{z \in \mathcal{C}_M} \frac{1}{\text{Im}(z)^2} \frac{|r_j(z) - s_j(z)| \tilde{y}_j^2}{|r_j(z) \tilde{y}_j - 1| |s_j(z) \tilde{y}_j - 1|} \\
 & \leq \mathbb{E} \sum_{j=1}^{\tilde{N}} \sup_{z \in \mathcal{C}_M} \frac{C_{z,M}}{\text{Im}(z)^2} |r_j(z) - s_j(z)| \\
 & \leq \frac{C_{i(f_m(M)+1),M}}{f_m(M)^2} \sum_{j=1}^{\tilde{N}} \sup_{z \in \mathcal{C}_M} |r_j(z) - s_j(z)| \\
 & = \frac{C_{i(f_m(M)+1),M}}{f_m(M)^2} \mathbb{E} \tilde{N} \mathbb{E} \sup_{z \in \mathcal{C}_M} |r_1(z) - s_1(z)|, \tag{8.15}
 \end{aligned}$$

where the final equality follows from Lemma 3.3. Finally combining (8.14) and (8.15) gives

$$\mathbb{E} \sup_{z \in \mathcal{C}_M} |\tilde{r}(z) - \tilde{s}(z)| \leq \left( \frac{C'}{f_m(M)^{\epsilon/2}} + \frac{C_{i(f_m(M)+1),M}}{f_m(M)^2} \mathbb{E} \tilde{N} \right) \mathbb{E} \sup_{z \in \mathcal{C}_M} |r_1(z) - s_1(z)|. \tag{8.16}$$

Notice this coefficient is independent of the coupling and depends only on  $M$ ,  $f_m$  and  $m$ . From the definition of  $C_{z,M}$ , we have that  $C_{z+i,M} / \text{Im}(z)^2 \rightarrow 4$  as  $\text{Im}(z) \rightarrow \infty$ . We also have that  $\mathbb{E} \tilde{N} = m((-\infty, -f_m(M)/2) \cup (f_m(M)/2, \infty)) \leq C f_m(M)^{-\epsilon}$ . We choose  $f_m$  to be such that

$$\left( \frac{C'}{f_m(M)^{\epsilon/2}} + \frac{C_{i(f_m(M)+1),M}}{f_m(M)^2} \mathbb{E} \tilde{N} \right) \leq \frac{4}{5}, \tag{8.17}$$

for each  $M \in \mathbb{N}$ . As the left-hand side of (8.17) is decreasing in  $f_m(M)$ ,  $f_m$  may be chosen to be increasing and unbounded.

Next we remove that assumption that, for some  $M$ , almost surely  $\mu_r$  and  $\mu_s$  have half their mass in  $[-M, M]$ . For a positive, increasing, unbounded function  $f_m$  on  $\mathbb{N}$ , we define the function  $d_{f_m} : \mathcal{P}(\mathcal{S})^2 \rightarrow [0, \infty)$  by

$$d_{f_m}(\mu_1, \mu_2) = \inf_{(r,s) \in \mathcal{C}(\mu_1, \mu_2)} \mathbb{E} \sum_{M=1}^{\infty} \sup_{z \in \mathcal{C}_M} |r(z) - s(z)| / 2^M, \tag{8.18}$$

and the function  $d'_{f_m} : \mathcal{P}(\mathcal{S})^2 \rightarrow [0, \infty)$  by

$$d'_{f_m}(\mu_1, \mu_2) = \inf_{(r,s) \in \mathcal{C}(\mu_1, \mu_2)} \mathbb{E} \sum_{M=1}^{\infty} \sup_{z \in \mathcal{C}_M} |r(z) - s(z)| \mathbf{1}_{A_M} / 2^M,$$



where  $\mathbf{1}_{A_M}$  is the indicator function of the event

$$A_M = \left\{ \mu_r([-M, M]) \geq \frac{1}{2} \text{ and } \mu_s([-M, M]) \geq \frac{1}{2} \right\},$$

$\mathcal{C}_M$  is the set defined by (8.11), and  $\mathcal{C}(\mu_1, \mu_2)$  is the set of all couplings of  $\mu_1$  and  $\mu_2$  for  $\mu_1, \mu_2 \in \mathcal{P}(\mathcal{S})$ . It is straightforward to check that  $\rho : \mathcal{S} \times \mathcal{S} \rightarrow [0, \infty)$  defined by

$$\rho(s, r) = \sum_{M=1}^{\infty} \sup_{z \in \mathcal{C}_M} |r(z) - s(z)|/2^M \tag{8.19}$$

is a metric on  $\mathcal{S}$ , and thus,  $d_{f_m}$  is the 1<sup>st</sup>-Wasserstein metric on  $\mathcal{P}(\mathcal{S})$  (see [18] Chapter 11 for details). Let  $\mu_1$  and  $\mu_2$  be two fixed points of  $T$ . Let  $(s, r)$  be a coupling of  $\mu_1$  and  $\mu_2$  such that

$$\mathbb{E} \sum_{M=1}^{\infty} \sup_{z \in \mathcal{C}_M} |r(z) - s(z)|\mathbf{1}_{A_M}/2^M \leq \frac{10}{9}d'_{f_m}(\mu_1, \mu_2), \tag{8.20}$$

and let  $\tilde{s}$  and  $\tilde{r}$  be built from i.i.d. copies of  $(s, r)$  as in (8.12). Using the specific coupling  $(\tilde{s}, \tilde{r})$ , (8.16), and (8.20) we get

$$\begin{aligned} d'_{f_m}(\mu_1, \mu_2) &\leq \mathbb{E} \sum_{M=1}^{\infty} \sup_{z \in \mathcal{C}_M} |\tilde{r}(z) - \tilde{s}(z)|\mathbf{1}_{A_M}/2^M \\ &\leq \frac{4}{5} \mathbb{E} \sum_{M=1}^{\infty} \sup_{z \in \mathcal{C}_M} |r(z) - s(z)|\mathbf{1}_{A_M}/2^M \\ &\leq \frac{8}{9}d'_{f_m}(\mu_1, \mu_2), \end{aligned}$$

and thus  $d'_{f_m}(\mu_1, \mu_2) = 0$ .

If  $d'_{f_m}$  was a metric, it would be immediate that  $\mu_1 = \mu_2$ ; however, it is not clear this is the case. The only property of a metric needed is that  $d'_{f_m}$  separates distinct points in  $\mathcal{P}(\mathcal{S})$ , and thus, we conclude the proof using the following lemma.

**Lemma 8.1** *Fix a positive, increasing, unbounded function  $f_m : \mathbb{N} \rightarrow \mathbb{R}$ .  $d'_{f_m}(\mu_1, \mu_2) = 0$  if and only if  $d_{f_m}(\mu_1, \mu_2) = 0$  for the metric  $d_{f_m}$  defined in (8.18).*

**Proof** Assume  $d'_{f_m}(\mu_1, \mu_2) = 0$ , fix  $\varepsilon > 0$ , and note there exists  $M_0 \in \mathbb{N}$  such that for any  $M \geq M_0$  and any coupling  $(r, s)$  one has

$$\mathbb{P}(A_M) \geq 1 - \varepsilon. \tag{8.21}$$

We have that

$$\inf_{(r,s) \in \mathcal{C}(\mu_1, \mu_2)} \mathbb{E} \sup_{z \in \mathcal{C}_{M_0}} |r(z) - s(z)| \mathbf{1}_{A_{M_0}} = 0, \tag{8.22}$$

and thus we can find a sequence of couplings  $\{(r_n, s_n)\}_{n=1}^\infty$  such that

$$\lim_{n \rightarrow \infty} \mathbb{E} \sup_{z \in \mathcal{C}_{M_0}} |r_n(z) - s_n(z)| \mathbf{1}_{A_{M_0}^n} = 0, \tag{8.23}$$

where  $A_{M_0}^n = \{\mu_r^n([-M, M]) \geq \frac{1}{2} \text{ and } \mu_s^n([-M, M]) \geq \frac{1}{2}\}$  and  $\mu_r^n$  and  $\mu_s^n$  are the random probability measures associated to  $r_n$  and  $s_n$ . Let  $M_1$  be such that  $f_m(M_1) > \frac{1}{2\varepsilon}$ , and hence for any  $M \geq M_1$

$$\sup_{z \in \mathcal{C}_M} |r(z) - s(z)| \leq \varepsilon, \tag{8.24}$$

for any Stieltjes transforms  $r$  and  $s$ .

We will now extend the convergence in (8.23) to the supremum over the larger compact set  $\tilde{\mathcal{C}} = \cup_{j=1}^{M_1} \mathcal{C}_j$ . The  $L^1$ -convergence of the random variables  $\sup_{z \in \mathcal{C}_{M_0}} |r_n(z) - s_n(z)| \mathbf{1}_{A_{M_0}^n}$  in (8.23) to zero implies convergence in probability to zero. Thus, we can find a subsequence converging almost surely to zero, and without loss of generality, we denote this subsequence  $\{\sup_{z \in \mathcal{C}_{M_0}} |r_n(z) - s_n(z)| \mathbf{1}_{A_{M_0}^n}\}_{n=1}^\infty$ . Let  $G = \{\lim_{n \rightarrow \infty} \sup_{z \in \mathcal{C}_{M_0}} |r_n(z) - s_n(z)| \mathbf{1}_{A_{M_0}^n} = 0\}$ , and decompose  $G$  into  $G_1 = \{\omega \in G : \mathbf{1}_{A_{M_0}^n}(\omega) = 1 \text{ i.o.}\}$  and  $G_0 = G \setminus G_1$ . Clearly on  $G_0$  the random variables  $\sup_{z \in \tilde{\mathcal{C}}} |r_n(z) - s_n(z)| \mathbf{1}_{A_{M_0}^n}$  are eventually identically 0. For  $\omega \in G_1$ , we consider the further subsequence  $\{n_k\}$  such that  $\mathbf{1}_{A_{M_0}^{n_k}}(\omega) = 1$  for all  $k$ . For this outcome  $\omega$ , we have  $\{(r_{n_k} - s_{n_k})(\omega)\}_{n=1}^\infty$  is a sequence of complex analytic functions on  $\mathbb{C}_+$ , uniformly bounded on compact subsets of  $\mathbb{C}_+$ , converging uniformly to 0 on a set with an accumulation point. Thus applying the Vitali convergence theorem for analytic functions, Lemma B.1, we get that  $\sup_{z \in \tilde{\mathcal{C}}} |(r_{n_k}(z) - s_{n_k}(z))|(\omega) \rightarrow 0$  as  $n_k \rightarrow \infty$ . From the above and the bounded convergence theorem, we get

$$\lim_{n \rightarrow \infty} \mathbb{E} \sup_{z \in \tilde{\mathcal{C}}} |r_n(z) - s_n(z)| \mathbf{1}_{A_{M_0}^n} = 0. \tag{8.25}$$

Combining (8.21), (8.24), and (8.25), we obtain

$$\begin{aligned} d_{f_m}(\mu_1, \mu_2) &= \inf_{(r,s) \in \mathcal{C}(\mu_1, \mu_2)} \mathbb{E} \sum_{M=1}^\infty \sup_{z \in \mathcal{C}_M} |r(z) - s(z)| / 2^M \\ &\leq \inf_{(r,s) \in \mathcal{C}(\mu_1, \mu_2)} \mathbb{E} \sum_{M=1}^\infty \left( \sup_{z \in \mathcal{C}_M} |r(z) - s(z)| \mathbf{1}_{A_{M_0}} / 2^M + \frac{2}{f_m(1)} \mathbb{P}(A_{M_0}^c) / 2^M \right) \end{aligned}$$

$$\begin{aligned}
 &< \inf_{(r,s) \in \mathcal{C}(\mu_1, \mu_2)} \mathbb{E} \sum_{M=1}^{M_1} \sup_{z \in \tilde{\mathcal{C}}} |r(z) - s(z)| \mathbf{1}_{A_{M_0}} / 2^M + \varepsilon + \frac{2}{f_m(1)} \varepsilon \\
 &\leq \inf_{(r,s) \in \mathcal{C}(\mu_1, \mu_2)} \mathbb{E} \sup_{z \in \tilde{\mathcal{C}}} |r(z) - s(z)| \mathbf{1}_{A_{M_0}} + \left(1 + \frac{2}{f_m(1)}\right) \varepsilon \\
 &= \left(1 + \frac{2}{f_m(1)}\right) \varepsilon.
 \end{aligned}$$

As  $\varepsilon > 0$  was arbitrary we have  $d_{f_m}(\mu_1, \mu_2) = 0$ . For the other direction, note

$$\begin{aligned}
 d'_{f_m}(\mu_1, \mu_2) &= \inf_{(r,s) \in \mathcal{C}(\mu_1, \mu_2)} \mathbb{E} \sum_{M=1}^{\infty} \sup_{z \in \mathcal{C}_M} |r(z) - s(z)| \mathbf{1}_{A_M} / 2^M \\
 &\leq \inf_{(r,s) \in \mathcal{C}(\mu_1, \mu_2)} \mathbb{E} \sum_{M=1}^{\infty} \sup_{z \in \mathcal{C}_M} |r(z) - s(z)| / 2^M \\
 &= d_{f_m}(\mu_1, \mu_2),
 \end{aligned}$$

for any  $\mu_1, \mu_2$ . □

**Acknowledgements** The first author thanks Yizhe Zhu for pointing out reference [30]. We thank David Renfrew for comments on an earlier draft. We thank the anonymous referee for a careful reading and helpful comments.

**Funding** Open access funding provided by Institute of Science and Technology (IST Austria).

**Data Availability** All data generated or analyzed during this study are included in this published article.

## Declarations

**Conflict of interest** The authors have no competing interests to declare that are relevant to the content of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix A: Tightness of $\mu_{L_n}$

**Lemma A.1** *Let  $\{A_n\}_{n \geq 1}$  be a Lévy–Khintchine random matrix ensemble with characteristics  $(0, b, m)$  satisfying **CI** and  $L_n$  be the matrix defined by (2.1). Then there*

exists  $r > 0$  such that almost surely

$$\limsup_{n \rightarrow \infty} \int_{\mathbb{R}} |t|^r d\mu_{L_n}(t) < \infty,$$

and thus almost surely  $\{\mu_{L_n}\}_{n \geq 1}$  is tight.

**Proof** This is essentially an extension of the argument in Lemma B.3 of [7] to Lévy–Khintchine matrices and with the matrix  $-D_n$  added. Applying Lemma B.5 and noting  $|\lambda_i(A)| = s(A)$  for some singular value of a Hermitian matrix  $A$ , one gets for any  $0 \leq k \leq n - 1$

$$|\lambda_{1+k}(L_n)| \leq |\lambda_{1+\lfloor k/2 \rfloor}(A_n)| + |\lambda_{1+\lceil k/2 \rceil}(-D_n)|.$$

Thus,

$$\int_{\mathbb{R}} |t|^r d\mu_{L_n}(t) \leq 8 \left( \int_{\mathbb{R}} |t|^r d\mu_{A_n}(t) + \int_{\mathbb{R}} |t|^r d\mu_{D_n}(t) \right). \tag{A.1}$$

$D_n$  is a diagonal matrix so

$$\int_{\mathbb{R}} |t|^r d\mu_{D_n}(t) = \frac{1}{n} \sum_{k=1}^n \left| \sum_{j \in \{1, \dots, n\} \setminus \{k\}} A_{kj}^{(n)} \right|^r \leq \frac{1}{n} \sum_{k=1}^n \left| \sum_{j \in \{1, \dots, n\} \setminus \{k\}} |A_{kj}^{(n)}| \right|^r. \tag{A.2}$$

Assuming  $0 \leq r \leq 2$  and applying the Schatten Bound, Lemma B.4, to  $A_n$  we get

$$\int_{\mathbb{R}} |t|^r d\mu_{A_n}(t) \leq \frac{1}{n} \sum_{k=1}^n \left| \sum_{j \in \{1, \dots, n\} \setminus \{k\}} |A_{kj}^{(n)}|^2 \right|^{r/2}. \tag{A.3}$$

We will prove that almost surely

$$\limsup_{n \rightarrow \infty} \int_{\mathbb{R}} |t|^r d\mu_{D_n}(t) < \infty. \tag{A.4}$$

The proof for  $\mu_{A_n}$  follows with only minor changes. Both follow the arguments of Lemma B.1 in [7]. Define the random variable  $Y_{n,k}$  by

$$Y_{n,k} := \left| \sum_{j \in \{1, \dots, n\} \setminus \{k\}} |A_{kj}^{(n)}| \right|^r,$$

for all  $k \in \{1, \dots, n\}$ . From the proof of Lemma B.1 in [7], we see it is enough to show

$$\sup_{n \geq 1} \mathbb{E}Y_{n,1}^4 < \infty. \tag{A.5}$$

Define for any  $0 \leq a < b$

$$S_{n,a,b} := \sum_{j=2}^n |A_{1j}^{(n)}| \mathbf{1}_{\{|A_{1j}^{(n)}| \in [a,b)\}}.$$

Then  $Y_{n,1}^4 = S_{n,0,\infty}^{4r} = (S_{n,0,1/2} + S_{n,1/2,\infty})^{4r}$  and

$$\mathbb{E}Y_{n,1}^4 \leq 2^{4r-1} \left( \mathbb{E}S_{n,0,1/2}^{4r} + \mathbb{E}S_{n,1/2,\infty}^{4r} \right).$$

If we further assume  $4r < 1$ , we can apply Jensen’s inequality to get

$$\mathbb{E}S_{n,0,1/2}^{4r} \leq (\mathbb{E}S_{n,0,1/2})^{4r}.$$

Applying Proposition 3.6 to the triangular array  $\{|A_{1j}^{(n)}|\}$  we get that

$$(\mathbb{E}S_{n,0,1/2})^{4r} \rightarrow \hat{b}_{1/2}^{4r}$$

as  $n \rightarrow \infty$ . For the larger entries of the row

$$\mathbb{E}S_{n,1/2,\infty}^{4r} \leq \mathbb{E}[(N_n M_n)^{4r}]$$

where

$$M_n := \max_{2 \leq j \leq n} |A_{1j}^{(n)}|,$$

and

$$N_n := |\{2 \leq j \leq n : |A_{1j}^{(n)}| \geq 1/2\}|,$$

where  $|G|$  is the cardinality of a set  $G$ . Again, using Proposition 3.6 we get for sufficiently large  $n$  and any  $k \in \mathbb{N}$

$$\begin{aligned} \mathbb{P}(N_n = k) &\leq \binom{n}{k} \mathbb{P}(|A_{12}^{(n)}| \geq 1/2)^k \\ &= \frac{\binom{n}{k}}{n^k} \left( n \mathbb{P}(|A_{12}^{(n)}| \geq 1/2) \right)^k \\ &\leq \frac{\binom{n}{k}}{n^k} c^k \end{aligned}$$

$$\leq \frac{c^k}{k!}$$

where  $c = m((-\infty, 1/2] \cup [1/2, \infty)) + 1$ , and we see

$$\mathbb{P}(N_n \geq k) \leq e^c - \sum_{j=1}^{k-1} \frac{c^j}{j!} \leq \frac{C c^k}{k!}.$$

Thus  $\sup_{n \geq 1} \mathbb{E}N_n^\eta < \infty$  for any  $\eta > 0$ . For  $M_n$  note that from (2.4)

$$\mathbb{P}(M_n \geq t) \leq n\mathbb{P}(|A_{12}^{(n)}| \geq t) \leq Ct^{-\varepsilon}.$$

It follows that  $\sup_{n \geq 1} \mathbb{E}M_n^\gamma < \infty$  for any  $0 < \gamma < \varepsilon$ . Returning to  $S_{n,1/2,\infty}$  and applying Hölder’s inequality, we get

$$\sup_{n \geq 1} \mathbb{E}S_{n,1/2,\infty}^{4r} \leq \sup_{n \geq 1} \mathbb{E}[(N_n M_n)^{4r}] \leq \sup_{n \geq 1} (\mathbb{E}N_n^{4rp})^{1/p} (\mathbb{E}M_n^{4rq})^{1/q} < \infty$$

for small enough  $q$ , completing the proof. □

### Appendix B: Additional lemmas

**Lemma B.1** (Vitali’s convergence theorem for analytic functions, Lemma 2.14 in [3]) *Let  $f_1, f_2, \dots$  be analytic in  $D$ , a connected open set of  $\mathbb{C}$ , satisfying  $|f_n(z)| \leq M$  for every  $n$  and  $z \in D$ , and  $f_n(z)$  converges as  $n \rightarrow \infty$  for each  $z$  in a subset of  $D$  having an accumulation point in  $D$ . Then there exists a function  $f$  analytic in  $D$  for which  $f_n(z) \rightarrow f(z)$  for all  $z \in D$ . Moreover on any set bounded by a contour interior to  $D$ , the convergence is uniform.*

Though Stieltjes transforms are not uniformly bounded on  $\mathbb{C}_+$ , it is straightforward to apply Theorem B.1 to them by considering first  $\mathbb{C}_{+,m} = \{z \in \mathbb{C} : \text{Im}(z) > 1/m\}$  and letting  $m \rightarrow \infty$ .

**Lemma B.2** (Theorem B.9 in [3]) *Assume that  $\{\mu_n\}$  is a sequence of functions probability measure, with Stieltjes transforms  $\{s_n\}$ . Then,*

$$\lim_{n \rightarrow \infty} s_n(z) = s(z) \tag{B.1}$$

for all  $z \in \mathbb{C}_+$  if and only if there exists a positive measure  $\mu$  with Stieltjes transform  $s$  such that  $\mu_n$  converges to  $\mu$  vaguely.

For the following lemma, we use the notation of [30]. For a connected open domain  $D \subset \mathbb{C}$ , let  $\mathcal{H}(D)$  be the space of analytic functions on  $D$  equipped with the topology of uniform convergence on compact subsets of  $D$ .

**Lemma B.3** (Proposition 2.5 in [30]) *Let  $X_1, X_2, X_3, \dots$  be a sequence of random analytic functions on a connected open set  $D \subset \mathbb{C}$ , with probability distribution measures  $\mu_{X_1}, \mu_{X_2}, \dots$  on  $\mathcal{H}(D)$ . If for every  $K \subset D$  compact  $\{\max_{z \in K} |X_n(z)|\}_{n \geq 1}$  is a tight sequence of random variables, then  $\{\mu_{X_n}\}_{n \geq 1}$  is tight in the space of probability measures on  $\mathcal{H}(D)$ .*

**Lemma B.4** (Schatten Bound, see proof of Theorem 3.32 in [31]) *Let  $A$  be an  $n \times n$  complex Hermitian matrix with rows  $R_1, \dots, R_n$ . Then for every  $0 < r \leq 2$ ,*

$$\sum_{k=1}^n |\lambda_k(A)|^r \leq \sum_{k=1}^n \|R_k\|_2^r,$$

where  $\|\cdot\|_2$  is the Euclidean norm on  $\mathbb{C}^n$ .

**Lemma B.5** (See [20], Chapter 3) *If  $A$  and  $B$  are  $n \times n$  complex matrices and  $s_1(A) \geq s_2(A) \geq \dots \geq s_n(A)$  and  $s_1(B) \geq s_2(B) \geq \dots \geq s_n(B)$  are the singular values of  $A$  and  $B$  then*

$$\begin{aligned} s_1(AB) &\leq s_1(A)s_1(B) \text{ and } s_1(A+B) \leq s_1(A) + s_1(B), \\ \max_{1 \leq k \leq n} |s_k(A+B) - s_k(A)| &\leq s_1(B), \\ s_{i+j-1}(A+B) &\leq s_i(A) + s_j(B) \end{aligned}$$

for  $1 \leq i, j \leq n$  and  $i + j \leq n + 1$ .

**Lemma B.6** (See [8] Lemma C.1 and [19] Theorem VIII.9.2) *Let  $Z$  be a positive random variable such that for every  $t > 0$ ,*

$$\mathbb{P}(Z \geq t) = L(t)t^{-\alpha}$$

for some slowly varying function  $L$  and some  $\alpha \in (0, 2)$ . Then, for every  $p > \alpha$ ,

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}[Z^p \mathbf{1}_{\{Z \leq t\}}]}{c(p)L(t)t^{p-\alpha}} = 1,$$

where  $c(p) := \alpha/(p - \alpha)$ .

## References

1. Aldous, D.J., Bandyopadhyay, A.: A survey of max-type recursive distributional equations. *Ann. Appl. Probab.* **15**(2), 1047–1110 (2005). <https://doi.org/10.1214/105051605000000142>
2. Alsmeyer, G.: *Random recursive equations and their distributional fixed points*. Available from <https://www.uni-muenster.de/Stochastik/lehre/SS11/StochRekGleichungen/book.pdf>, (2012)
3. Bai, Z., Silverstein, J.W.: *Spectral Analysis of Large Dimensional Random Matrices*. Springer Series in Statistics, 2nd edn. Springer, New York (2010)
4. Bandeira, A.S.: Random Laplacian matrices and convex relaxations. *Found. Comput. Math.* **18**(2), 345–379 (2018). <https://doi.org/10.1007/s10208-016-9341-9>

5. Biane, P.: On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.* **46**(3), 705–718 (1997). <https://doi.org/10.1007/s10208-016-9341-9>
6. Bordenave, C., Caputo, P., Chafaï, D.: Spectrum of Markov generators on sparse random graphs. *Commun. Pure Appl. Math.* **67**(4), 621–669 (2014). <https://doi.org/10.1007/s10208-016-9341-9>
7. Bordenave, C., Caputo, P., Chafaï, D.: Spectrum of large random reversible Markov chains: heavy-tailed weights on the complete graph. *Ann. Probab.* **39**(4), 1544–1590 (2011). <https://doi.org/10.1007/s10208-016-9341-9>
8. Bordenave, C., Caputo, P., Chafaï, D.: Spectrum of non-Hermitian heavy tailed random matrices. *Commun. Math. Phys.* **307**(2), 513–560 (2011). <https://doi.org/10.1007/s00220-011-1331-9>
9. Bordenave, C., Caputo, P., Chafaï, D., Piras, D.: Spectrum of large random Markov chains: Heavy-tailed weights on the oriented complete graph. *Random Matrices Theor. Appl.* **6**(02), 175006 (2017). <https://doi.org/10.1007/s00220-011-1331-9>
10. Bryc, W., Dembo, A., Jiang, T.: Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.* **34**(1), 1–38 (2006). <https://doi.org/10.1007/s00220-011-1331-9>
11. Campbell, A., O’Rourke, S.: Spectrum of heavy-tailed elliptic random matrices. *Electr. J. Probab.* **27**, 1–56 (2022). <https://doi.org/10.1214/22-EJP849>
12. Chakrabarty, A., Hazra, R.S., den Hollander, F., Sfragara, M.: Spectra of adjacency and Laplacian matrices of inhomogeneous Erdős-Rényi random graphs. *Random Matrices Theor. Appl.* **10**(1), 34 (2021)
13. Chatterjee, A., Hazra, R.S.: Spectral properties for the Laplacian of a generalized Wigner matrix. *Random Matrices Theor. Appl.* **11**(3), 66 (2022)
14. Cizeau, P., Bouchaud, J.P.: Theory of Lévy matrices. *Phys. Rev. E* **50**, 1810–1822 (1994). <https://doi.org/10.1103/PhysRevE.50.1810>
15. Davydov, Y., Egorov, V.: On convergence of empirical point processes. *Statist. Probab. Lett.* **76**(17), 1836–1844 (2006). <https://doi.org/10.1016/j.spl.2006.04.030>
16. Ding, X.: On some spectral properties of large block Laplacian random matrices. *Statist. Probab. Lett.* **99**, 61–69 (2015). <https://doi.org/10.1016/j.spl.2006.04.030>
17. Ding, X., Jiang, T.: Spectral distributions of adjacency and Laplacian matrices of random graphs. *Ann. Appl. Probab.* **20**(6), 2086–2117 (2010). <https://doi.org/10.1214/10-AAP677>
18. Dudley, R.M.: *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics, vol. 74. Cambridge University Press, Cambridge (2002)
19. Feller, W.: *An Introduction to Probability Theory and its Applications*, vol. II, 2nd edn. Wiley, New York (1971)
20. Horn, R., Johnson, C.R.: *Topics in Matrix Analysis*. Cambridge University Press, Cambridge (1994)
21. Huang, J., Landon, B.: Spectral statistics of sparse Erdős-Rényi graph Laplacians. *Ann. Inst. Henri Poincaré Probab. Stat.* **56**(1), 120–154 (2020). <https://doi.org/10.1214/19-AIHP957>
22. Jiang, T.: Empirical distributions of Laplacian matrices of large dilute random graphs. *Random Matrices Theor. Appl.* **1**(3), 20 (2012). <https://doi.org/10.1142/S2010326312500049>
23. Jiang, T.: Low eigenvalues of Laplacian matrices of large random graphs. *Probab. Theor. Related Fields* **153**(3–4), 671–690 (2012). <https://doi.org/10.1142/S2010326312500049>
24. Jung, P.: Lévy-Khintchine random matrices and the poisson weighted infinite skeleton tree. *Trans. Am. Math. Soc.* (2014). <https://doi.org/10.1090/tran/6977>
25. Kallenberg, O.: *Foundations of modern probability*. In: *Probability and its Applications*. Springer-Verlag, New York (2002)
26. Kingman, J.F.C.: *Poisson processes*, Oxford Studies in Probability, vol. 3. Oxford University Press, New York (1993)
27. Mingo, J.A., Speicher, R.: *Free Probability and Random Matrices*. In: *Fields Institute Monographs*, vol. 35. Springer, New York (2017)
28. Resnick, S.I.: *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering, 1st edn. Springer-Verlag, New York (1987)
29. Resnick, S.I.: *Heavy-Tailed Phenomena*. Springer Series in Operations Research and Financial Engineering, 1st edn. Springer-Verlag, New York (2007)
30. Shirai, T.: *Limit theorems for random analytic functions and their zeros*, Functions in number theory and their probabilistic aspects, RIMS Kôkyûroku Bessatsu, B34, Res. Inst. Math. Sci. (RIMS), Kyoto, pp. 335–359, (2012)
31. Zhan, X.: *Matrix Inequalities*. Lecture Notes in Mathematics, vol. 1790. Springer-Verlag, New York (2002)



---

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.