Scuola dottorale in scienze Matematiche e Fisiche
Sezione Matematica
XXVII ciclo del corso di Dottorato

Spectrum of non Hermitian random Markov matrices with heavy tailed weights

Candidato
Dott. Daniele Piras

Relatore
Prof. Pietro Caputo

Coordinatore
Prof. Luigi Chierchia

Contents

Introduction

This thesis concerns the convergence of the empirical spectral distribution of random matrices, that is the probability measure concentrated on the spectrum $\{\lambda_1(M), ..., \lambda_n(M)\}$ of a complex $n \times n$ matrix M. Namely we define the empirical spectral distribution μ_M as

$$\mu_M := \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(M)}.$$

We will consider matrices M whose entries will be random variables, as a consequence μ_M will be a random probability measure.

The study of the spectrum of random matrices goes back to the $'50$ when the Hungarian physicist Eugene Wigner, proved that the empirical spectral distribution (ESD) of a sequence of Hermitian random matrices, whose entries are independent random variables with unitary variance, up to rescaling weakly converges to the probability measure

$$\mu_{sc}(dx) = \frac{1}{2\pi} \sqrt{4 - x^2}\, \mathbb{1}_{|x|<2} dx.$$

The law μ_{sc} is the so called Wigner semi-circular law, see figure 1.

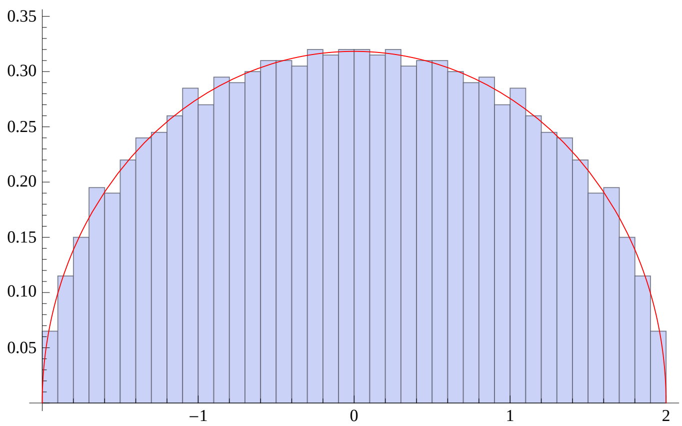


Figure 1: *Histogram of the spectrum of a 1000×1000 Wigner matrix with standard Gaussian entries, plotted with the density of* μ_{sc}.

In the same years Wigner conjectured that the limit spectral distribution of a sequence of non Hermitian matrices with unitary variance independent random entries, un to rescaling, is the uniform law on the unitary disc of the complex plane, see figure 2.

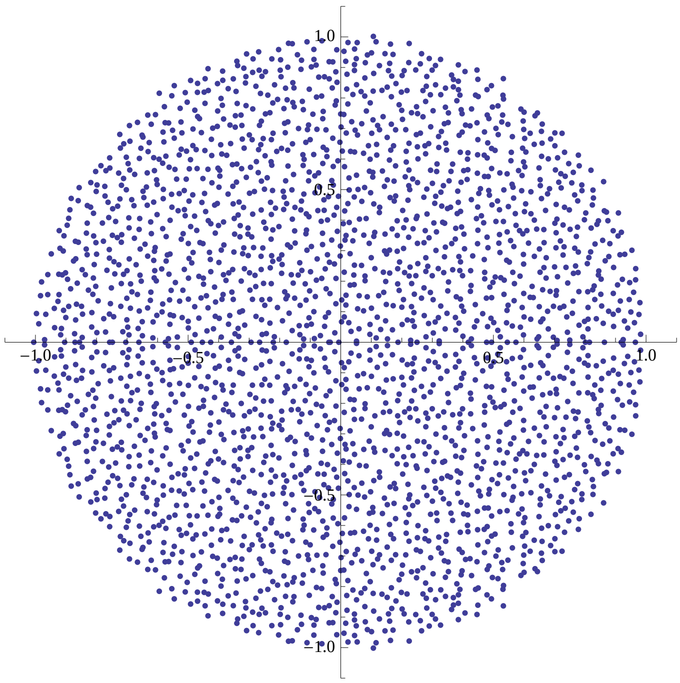


Figure 2: *Plot of the spectrum of a 2500×2500 matrix with i.i.d. standard Gaussian entries.*

The proof of the conjecture has a long history, it indeed has been proved by Tao and Vu in 2009, after more than 40 years of partial results. We just mention very few of them, for a more detailed sequence of the results we refer to [13, Section 2].

The first piece of proof for the circular law theorem is by Metha in [19], who in 1967, using the work of Ginibre [15], proved the result for the expected empirical distribution for complex Gaussian entries. In 1997 Bai, in [4], is the first to obtain results in the universal case, using the work of Vyacheslav Girko, but assuming stronger hypothesis on the law of the entries, such as bounded density and finite sixth moment. Finally in 2009, Tao and Vu proved the original conjecture, in [27].

The proof is based on the Girko Hermitization method, a trick to pull back the non-Hermitian problem to Hermitian matrices. We give more details in chapter 1, but the core of the method is to work with the spectrum of the singular values of the non-Hermitian matrix and to take advantage of their logarithmic bond with the spectrum of the eigenvalues of M, see Theorem 1.2.1 and Lemma 1.2.3.

In 2008 Ben Arous and Guionnet, proved a heavy tailed counterpart of the Wigner theorem, see also Zakharevich [30]. They indeed proved an existence result for the limiting spectral distribution of a sequence of Hermitian matrices whose entries are independent with common law in the domain of attraction of the α-stable law, with $\alpha \in (0,2)$. As in the finite second moment scenario, the limiting spectral measure does not depend on the law of the entries of the matrix, but only on the parameter α. Remarkable work in this regime is also by Belinschi, Dembo and Guionnet [5]. These works established rigorously a number of prediction made by physicists Bouchaudand and Fizeau [14]

In 2010 Bordenave, Caputo, and Chafaï [9], with a new and independent approach based on the objective method introduced by Aldous and Steele in [2], give an alternative proof of the convergence in the heavy tailed setting. They prove that the heavy tailed matrix, suitably rescaled, locally converges to an infinite poissonian weighted tree called PWIT. The heavy tails regime is in many ways more difficult than the bounded variance regime. Indeed we do not have an explicit expression for the limiting distribution. Nevertheless the approach of [9], is powerful enough to give some properties of the limiting spectral measure, by means of recursive analysis on the limiting tree.

In 2012 the same authors in [10], prove an analogous of the Circular law, for non-Hermitian matrices with i.i.d. heavy tailed entries, using the same approach from [9] combined with the Hermitization techniques by Tao and Vu.

Very little is known if the entries of the matrix are not independent.

An interesting problem with non-independent entries is obtained by considering Markov matrices, i.e. matrices with non-negative entries and row sum equal to 1. In this case is natural to associate the random matrix with the corresponding weighted random graph, and to interpret the elements of the matrix as the transition probabilities of the random walk on the graph.

Suppose $U_{i,j}$ is a collection of i.i.d. random variables, and define the Markov matrix $X_n = (X_{i,j})_{i,j=1}^n$,

$$X_{i,j} = \frac{U_{i,j}}{\rho_i} \qquad \rho_i = \sum_{j=1}^n U_{i,j}.$$

By construction, the spectrum of X_n is a subset of $\{z \in \mathbb{C} : |z| \leq 1\}$. The convergence of ESD of P has been investigated in recent works by Bordenave, Caputo, and Chafaï. When the variables $\{U_{i,j}\}$ have finite variance $\sigma^2 \in (0,+\infty)$ and unitary mean, $\mu_{\frac{\sqrt{n}}{\sigma}X_n}$ behaves as the ESD of $\frac{U_n}{\sqrt{n\sigma^2}}$ where $U_n = (U_{i,j})$ is the non normalized matrix. Namely it converges to the circular law when U_n is non-Hermitian, see [11] and figure 3, and to the Wigner semi-circular law when U_n is Hermitian, see [8].

This analysis can be extended to other models with non-independent entries such as random Markov generators and zero sum matrices, see the work of Bordenave, Ca-

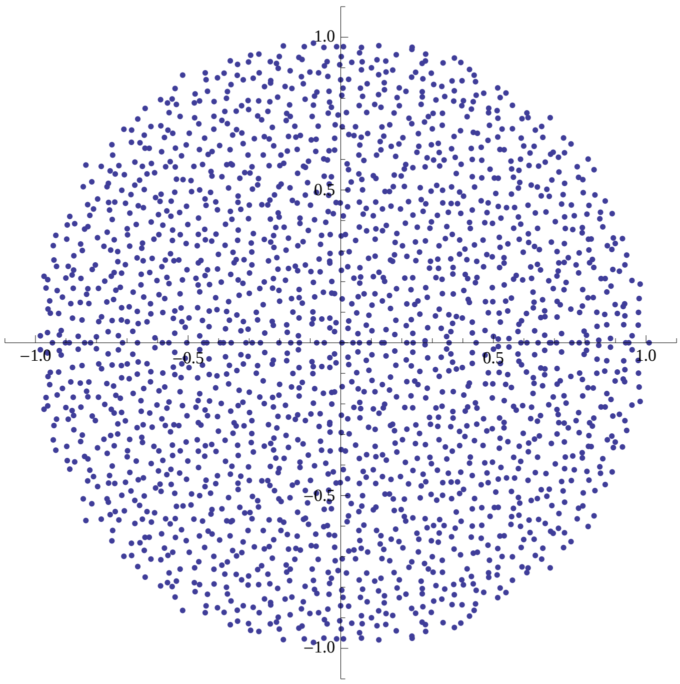


Figure 3: *Plot of the spectrum of a 2000×2000 rescaled random non-Hermitian Markov matrix, with $U_{i,j}$ are i.i.d exponential random variables with mean 1.*

puto, and Chafaï [12] and Tao [25]. In [9], Bordenave, Caputo, and Chafaï, study the markovian case when U_n is a symmetric heavy tailed random matrix. Of remarkable interest is the case $\alpha \in (0,1)$. In this regime the ESD of the matrix X_n, without any scaling factor, converges to a non trivial measure concentrated on the unitary disc of the complex plane. In this work we study the non-Hermitian version of this heavy tailed matrix, that is the case where U_n has i.i.d. heavy tailed entries, with $\alpha \in (0,1)$. Our main result concerns the convergence of the ESD of the associated Markov matrix X_n, to a non trivial measure supported in the unitary disc of the complex plane, depending only on the parameter α. In contrast with the Hermitian case, this limiting distribution should have a remarkable concentration on a disc with radius $r < 1$, see figures 4,5,6 and 7 for a plot of the eigenvalues, and figure 8 and 9 for a radial plot of the density of the eigenvalues. They represent simulation of μ_{X_n}, where X_n is a $n \times n$ random Markov matrices, and the non normalized matrix U_n has i.i.d. entries $U_{i,j} \sim V^{-1/\alpha}$, with $V \sim Unif(0,1)$, for various values of α and n, as reported in captions.

This thesis consists of 4 chapters. In chapter 1 we introduce some of the more remarkable technical tools we need in the following chapters. In order to relieve the reading we dilute those technical tools in a survey of some results cited above. The last section of chapter 1, is a more detailed introduction to our work.

In chapter 2, we prove the local convergence of a non-Hermitian random Markov matrix X_n with entries in the domain of attraction of an α-stable law with $\alpha \in (0,1)$ to a modified PWIT whose generations have alternated distribution, this generalizes the analysis of [10].

In chapter 3, we use the resolvent convergence implied by the local convergence, to prove the convergence of the ESD of the singular values of X_n, see theorem 1.6.2, to a measure with finite exponential moments, see proposition 3.3.3.

In chapter 4, we use the Girko's Hermitization method to obtain the convergence of the spectrum of the eigenvalues of X_n, see theorem 1.6.1, to a non trivial measure, see section 4.6.

The techniques used in this work follow very closely, with few exceptions, the works [8, 9, 10, 11].

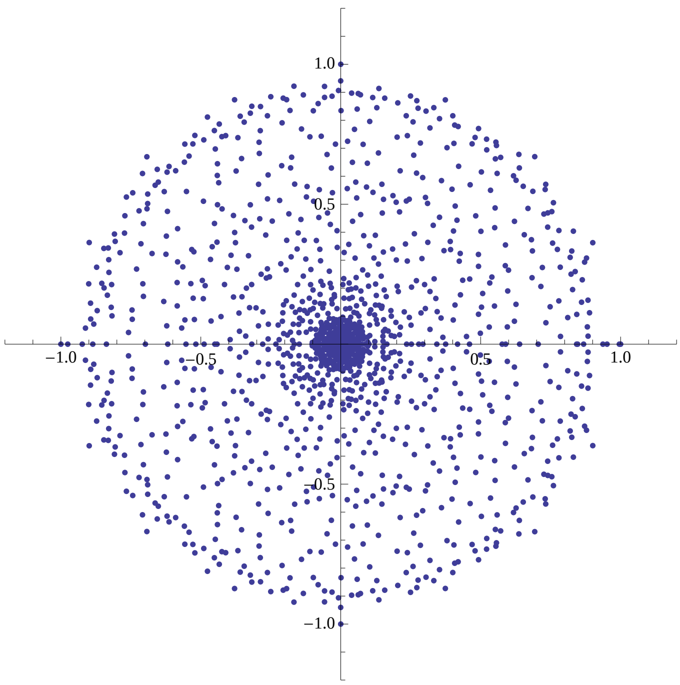


Figure 4: 2000×2000 $\alpha = 0,1$

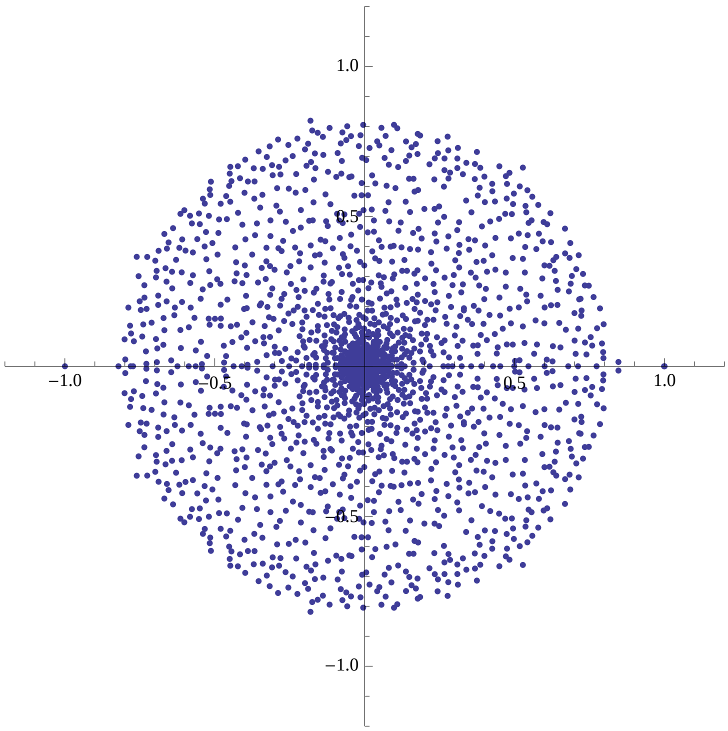


Figure 5: 2000×2000 $\alpha = 0,3$

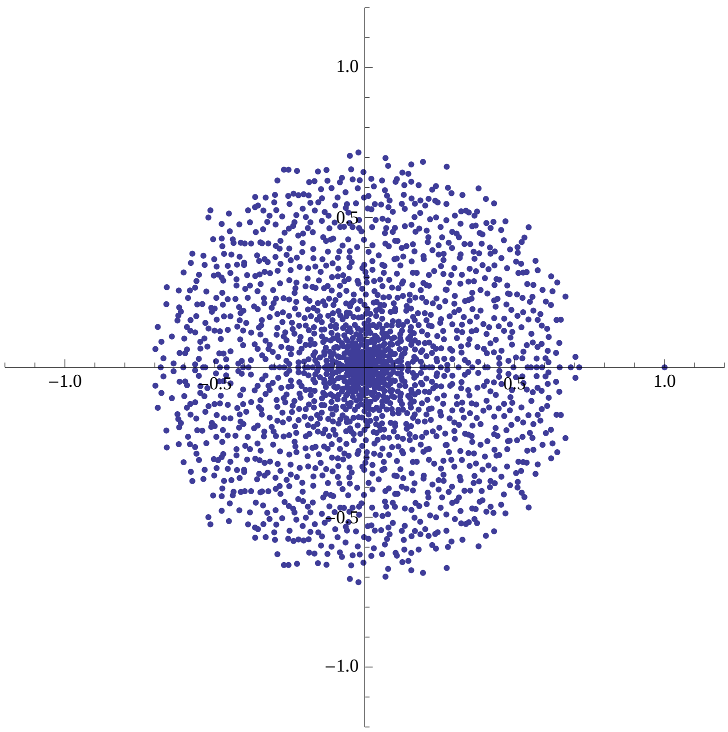


Figure 6: 2000×2000 $\alpha = 0,5$

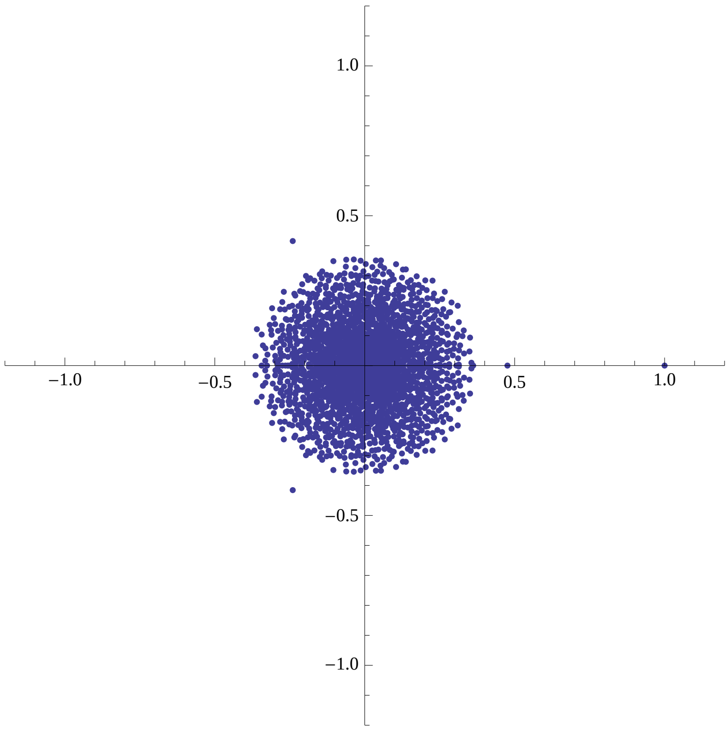


Figure 7: 2000×2000 $\alpha = 0,9$

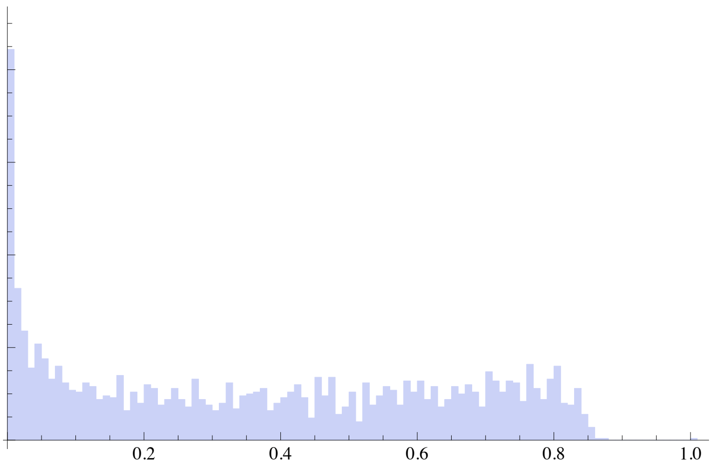


Figure 8: *Radial plot of the density for* $n = 2500$ *and* $\alpha = 0,3$

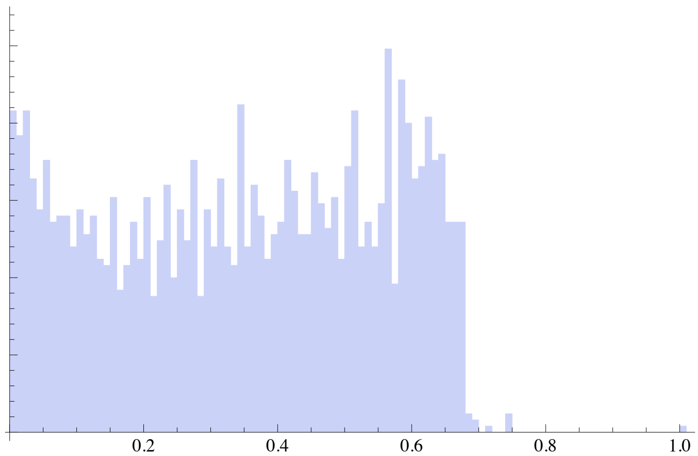


Figure 9: *Radial plot of the density for* $n = 2500$ *and* $\alpha = 0,5$

Chapter 1

Preliminaries

Consider a matrix $M \in \mathcal{M}_n(\mathbb{C})$, the set of $n \times n$ complex matrices. Call $\lambda_1(M), ..., \lambda_n(M)$ its eigenvalues counting multiplicity, we define the empirical spectral distribution (ESD) of M as

$$\mu_M = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(M)}. \tag{1.0.1}$$

This defines a unitary mass measure on $\mathbb{C}$, μ_M is a probability measure defined on the complex plane. Moreover, if we let the entries of the matrix M be random variables, μ_M becomes a random probability measure defined on the complex plane. We are interested in its asymptotic behavior as the order n of the matrix diverges to infinity, when the entries of M are random variables.

The aim of this chapter is to introduce some fundamental technical tools we need through the following chapters. Starting from the finite variance case, with the Wigner theorem for Hermitian matrices and the circular law theorem for non Hermitian matrices, and continuing with the results for heavy tailed matrices of Bordenave, Caputo and Chafaï, referred to as BCC from now on, we present some results we cited in the introduction, since the technicalities we need for the proofs of those theorems are the same we need in the following chapters.

1.1 Semi-circular Law

Consider a matrix $X_n = (X_{i,j})$ such that $X_{i,j} = \overline{X}_{j,i}$, for $1 \leq i < j$, and where $\{X_{i,j}\}_{1 \leq i < j \leq n}$ is a collection of i.i.d. random variables with common law $\mathcal{P}$ on $\mathbb{C}$, and $\{X_{i,i}\}_{i=1}^{n}$ is a collection of i.i.d random variables with common law $\mathcal{Q}$ on $\mathbb{R}$. Then $X_n = (X_{i,j})_{1 \leq i,j \leq n}$ is a random Hermitian matrix, also called Wigner matrix. If the variables $X_{i,j}$ have finite variance, namely $\mathrm{Var}(X_{1,2}) = \mathbb{E}\left[\,|X_{1,2}|^2\,\right] - \mathbb{E}\left[\,|X_{1,2}|\,\right]^2 = \sigma^2 < +\infty$, then the ESD of $(n\sigma^2)^{-1/2}X_n$, converges to a probability measure defined on the interval $[-2,2]$ of the real line, called *semi-circular* law. We will call this law μ_{sc}. The semi-circular law is the probability measure

$$\mu_{sc}(dx) = \frac{1}{2\sigma\pi}\sqrt{4-x^2}\mathbb{1}_{|x|\leq 2}dx. \tag{1.1.1}$$

The result was proved by Wigner in 1955, and is contained in [29]. It can be considered the starting point of the random matricx theory.

Theorem 1.1.1. *Consider a random Wigner matrix X_n, such that $Var(X_{1,2}) = \sigma^2 < +\infty$, then almost surely,*

$$\mu_{(n\sigma^2)^{-1/2}X_n} \xrightarrow[n\to+\infty]{(w)} \mu_{sc}.$$

We write $\mu_n \xrightarrow[n\to+\infty]{(w)} \mu_0$, when the sequence of measure μ_n weakly converges to the measure μ_0 in the classical sense, i.e. $\mu_n(h) \xrightarrow[n\to+\infty]{} \mu_0(h)$ for any bounded and continuous function h.

Many different proofs of this result have been given, see [3, Chapter 2]. We give the idea of the proof using the Resolvent method, since it gives us the opportunity to introduce the Cauchy-Stieltjes transform of a measure, the Resolvet of a matrix and how they are linked in random matrix theory.

1.1.1 Cauchy-Stieltjes transform, resolvent and pointed spectral measure

We first introduce the Cauchy-Stiltjes transform of a real measure and the resolvent of a matrix, then, given some properties of these objects, we will see how they are linked to each other, and how, through this link, one can prove the Wigner theorem.

Cauchy-Stieltjes transform

For a finite measure μ supported on $\mathbb{R}$, define the Cauchy-Stieltjes transform as

$$g_\mu(z) := \int \frac{1}{x-z}\mu(dx) \tag{1.1.2}$$

for $z \in \mathbb{C}^+ := \{w \in \mathbb{C} : Im(z) > 0\}$. If μ has bounded support we have

$$g_\mu(z) = \sum_{n\geq 0} z^{-n-1}\int x^n\mu(dx). \tag{1.1.3}$$

If μ is a probability measure, then $g_\mu(z)$ is an analytic function from $\mathbb{C}^+ \to \mathbb{C}^+$, its modulus is bounded by $Im(z)^{-1}$ and it characterizes the measure μ, in the sense that we can deduce μ from $g_\mu(z)$, see [3, Section 2.4]. Moreover the convergence of the

Cauchy-Stieltjes transform is equivalent to the weak convergence as stated in the next result proved in e.g. [7].

Theorem 1.1.2. *Let μ and $(\mu_n)_{n\geq 1}$ be a sequence of real probability measures, then the following are equivalent.*

i) $\mu_n \xrightarrow[n\to+\infty]{(w)} \mu$.

ii) $g_{\mu_n}(z) \xrightarrow[n\to+\infty]{} g_\mu(z)$ *for all* $z \in \mathbb{C}^+$.

iii) There exists a set $D \subset \mathbb{C}^+$ *with an accumulation point such that for all* $z \in D$, $g_{\mu_n}(z) \xrightarrow[n\to+\infty]{} g\mu(z)$.

The Cauchy-Stieltjes transform of the semi-circular law satisfies the following fixed point equation for all $z \in \mathbb{C}^+$,

$$g_{\mu_{sc}}(z) = -(z + g_{\mu_{sc}}(z))^{-1}, \tag{1.1.4}$$

see e.g. [7].

Resolvent

Let M be a $n \times n$ complex Hermitian matrix, take $z \in \mathbb{C}^+$, then, if $\mathbb{1}_n$ is the $n \times n$ identity matrix, $M - z\mathbb{1}_n$ is invertible, one defines the resolvent of M as the function $R : \mathbb{C}^+ \to \mathcal{M}_n(\mathbb{C})$ such that

$$R(z) := (M - z\mathbb{1}_n)^{-1}.$$

If $(v_k)_{k=1}^n$ is an orthonormal basis of eigenvectors of M, which exists by spectral theorem, we can decompose the resolvent matrix as

$$R(z) = \sum_{k=1}^n \frac{1}{\lambda_k(M) - z} v_k v_k^*.$$

Using this decomposition we can observe that $R(z)$ is a normal matrix, namely $R(z)R(z)^* = R(z)^*R(z)$. Moreover $R(z)$ is bounded, $\|R(z)\|_{2\to 2} \leq Im(z)^{-1}$ and $z \mapsto R(z)$ is an analytic function on $\mathbb{C}^+$.

1.1.2 Pointed spectral measure

In order to introduce the bond between Cauchy-Stieltjes transform and Resolvent we need to introduce the pointed spectral measure. Let ξ be a vector of $\mathbb{C}^n$ with unitary ℓ^2 norm, and $(v_k)_{k=1}^n$ an orthonormal basis of $\mathbb{C}^n$ given by the eigenvectors of M. We can define the real probability measure

$$\mu_M^\xi = \sum_{k=1}^n |\langle v_k, \xi\rangle|^2 \delta_{\lambda_k(M)}. \tag{1.1.5}$$

We call this measure, spectral measure pointed at ξ. This measure could analogously be defined by the unique probability measure μ_M^ξ such that

$$\int x^k \mu_M^\xi(dx) = \left\langle \xi, M^k \xi \right\rangle,$$

for any integer $k \geq 1$. Note that for the special case $\xi = (\frac{1}{\sqrt{n}}, ..., \frac{1}{\sqrt{n}})$ we find again the ESD of M, $\mu_M = \mu_M^{(\frac{1}{\sqrt{n}},...,\frac{1}{\sqrt{n}})}$. Also, if $(e_1, ..., e_n)$ is the canonical base of $\mathbb{C}^n$, for $\xi = (\xi_i)_{i=1}^n = (\langle \xi, e_i \rangle)_{i=1}^n$,

$$\mu_M^\xi = \sum_{i=1}^n |\xi_i|^2 \mu_M^{e_i},$$

and

$$\mu_M = \frac{1}{n} \sum_{i=1}^n \mu_M^{e_i}.$$

Now consider an Hermitian matrix $M \in \mathcal{H}_n(\mathbb{C})$, the subset of $\mathcal{M}_n(\mathbb{C})$ of the Hermitian matrices, the link between the Chauchy-Stieltjes transform and Resolvent is that,

$$\langle \xi, R(z)\xi \rangle = \int \frac{1}{x - z} \mu_M^\xi(dx) = g_{\mu_M^\xi}(z).$$

where μ_M^ξ is the pointed spectral measure we defined in (1.1.5). Moreover, considering μ_M, the ESD of M,

$$g_{\mu_M}(z) = \int \frac{1}{x - z} \mu_M(dx) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\lambda_i(M) - z} = \frac{1}{n} \mathrm{Tr}(R(z)).$$

Let us remark that the measure μ_M, is actually a random probability measure. We can deal with this extra randomness introducing concentration inequalities.

1.1.3 Concentration Inequalities

Concentration inequalities are a powerful tool in probability theory, and more generally in measure theory. What essentially concentration inequalities state is that, under suitable hypothesis, a random object is, with very high probability, close to a constant. The random object is generally a function, on which we require mild hypothesis, of a family of random variables and the constant is the expected values of the function under the probability measure.

The first result we present is the one we will use the most, see [10, Lemma C.2]. It refers to random matrix with independent half-rows, as the Hermitian random matrix X_n from the Wigner theorem. Recall the BV-norm of a function $f : \mathbb{R} \to \mathbb{R}$, vanishing at ∞, i.e. $\lim_{x \to +\infty} f(x) = 0$, is defined as

$$\|f\|_{BV} = \sup_{(x_k)_{k \in \mathbb{Z}}} \sum_k |f(x_{k+1}) - f(x_k)|,$$

with $x_{k+1} \geq x_k$ for any k.

Theorem 1.1.3 (Concentration of ESD with independent half-rows)**.** *Let M be an Hermitian random matrix. For $1 \leq k \leq n$ define the variables $M_k := (M_{k,j})_{j=1}^k \in \mathbb{R}^k$. If the variables $(M_k)_{k=1}^n$ are independent, then for any $f : \mathbb{R} \to \mathbb{R}$ such that $\|f\|_{BV} \leq 1$ and every $t \geq 0$,*

$$\mathbb{P}\left(\left|\int f d\mu_M - \mathbb{E}\left[\int f d\mu_M\right]\right| \geq t\right) \leq 2\exp\left(-\frac{nt^2}{8}\right)$$

The proof is based on the Azuma-Hoeffding's inequality, see e.g. [18].

There is a huge family of very useful concentration inequalities based on Logarithmic-Sobolev type inequality, which we will not present, but for which we refer to [7, Subsection 3.3.2]. What we want to present instead is the following concentration inequality by Talagrand in [24], which we will use in Chapter 4.

Theorem 1.1.4 (Talagrand concentration inequility)**.** *Let K be a convex subset of $\mathbb{R}$ with diameter $D = \sup_{x,y \in K} |x - y|$. Consider a convex Lipschitz real valued function defined on K^n, with Lipschitz constant $\|f\|_{Lip}$. Let $\mathbb{P} = P_1 \otimes \cdots \otimes P_n$ be a product measure on K_n and let $M_{\mathbb{P}}(f)$ be the median of f under $\mathbb{P}$. Then for any $t > 0$*

$$\mathbb{P}\left(|f - M_{\mathbb{P}}(f)| \geq t\right) \leq 4e^{-\frac{t^2}{4D^2\|f\|_{Lip}}}$$

Note that if $E_{\mathbb{P}}(f)$ is the mean of f under the measure $\mathbb{P}$ of the theorem,

$$\begin{aligned} |M_{\mathbb{P}}(f) - E_{\mathbb{P}}(f)| &\leq \int_0^{+\infty} \mathbb{P}\left(|f - M_{\mathbb{P}}(f)| \geq t\right) dt \\ &\leq 4\int_0^{+\infty} \exp\left\{-\frac{t^2}{4D^2\|f\|_{Lip}}\right\} dt \\ &= 4\sqrt{\pi} D \|f\|_{Lip}. \end{aligned}$$

We may then deduce an equality involving the mean of f under $\mathbb{P}$.

1.1.4 Rough sketch of proof of Theorem 1.1.1

The idea of the proof of theorem 1.1.1, using the resolvent method, is to show the convergence of $g_{\mu_{\frac{X_n}{\sigma\sqrt{n}}}}$ to $g_{\mu_{sc}}$. Fix $\sigma^2 = 1$. The first step is to consider $\mathbb{E}\left[g_{\mu_{\frac{X_n}{\sqrt{n}}}}\right]$, since $g_{\mu_{\frac{X_n}{\sqrt{n}}}}$ concentrates around its mean, by theorem 1.1.3. Then, a simple calculation gives

$$\mathbb{E}\left[g_{\mu_{\frac{X_n}{\sqrt{n}}}}\right] = \frac{1}{n}\mathbb{E}\left[\operatorname{Tr} R(z)\right]$$

Where $R(z)$ is the resolvent of $X_n/\sqrt{n}$. Now, since $\{(R(z))_{i,i}\}_{i=1}^n$ are variables with the same distribution, one can conclude that

$$\frac{1}{n}\mathbb{E}\left[\,\mathrm{Tr}R(z)\,\right]=\mathbb{E}\left[\,(R(z))_{1,1}\,\right].$$

Now, using concentration arguments, properties of the resolvent, and the fact that all the non diagonal entries of X_n have the same law, one can show that, if one call $\lim_{n\to\infty}\mathbb{E}\left[g_{\mu_{\frac{X_n}{\sqrt{n}}}}\right] := g_{\mu_\infty}$, then g_{μ_∞} satisfies the same fixed point equation as the Cauchy-Stieltjes transform, namely

$$g_{\mu_\infty}=-\left(z+g_{\mu_\infty}\right)^{-1}.$$

Since the Cauchy-Stieltjes transform of probability measure uniquely characterize a distribution, $\mu_\infty=\mu_{sc}$. Now the convergence of Cahucy-Stieltjes transform is equivalent to weak convergence of measure, then

$$\mu_{\frac{X_n}{\sqrt{n}}}\xrightarrow[n\to+\infty]{(w)}\mu_{sc},$$

see [7] for a more detailed proof.

1.2 Circular law

In this section we analyze the limiting behavior of a non-Hermitian matrix. For a matrix $M\in\mathcal{M}_n(\mathbb{C})$, recall its spectrum is $\{\lambda_1(M),...,\lambda_n(M)\}$, where the eigenvalues are order in such a way that $|\lambda_1(M)|\geq\cdots\geq|\lambda_n(M)|$. We can consider the singular values, defined by

$$s_i(M):=\lambda_i(\sqrt{MM^*})\qquad\text{for } i=1,...,n,$$

where $M^*=\overline{M}^T$ is the conjugate-transpose matrix of M, then $\{s_1(M),...,s_n(M)\}\subset[0,+\infty)$. Again the singular values are such that $s_1(M)\geq\cdots\geq s_n(M)$. We can define an analogous of the ESD for the singular values, as

$$\nu_M:=\frac{1}{n}\sum_{k=1}^{n}\delta_{s_k(M)}.$$

Note that the spectrum of the singular values of M^* and M^T, equals the spectrum of the singular values of M. Also, it will be useful to notice that the $2n\times 2n$ Hermitian matrix

$$\begin{pmatrix}0 & M\\ M^* & 0\end{pmatrix}\tag{1.2.1}$$

has spectrum $\{\pm s_1(M),...,\pm s_n(M)\}$. The spectrum of the singular values and the spectrum of the eigenvalues are linked by Weyl's inequalities, see [28].

Theorem 1.2.1 (Weyl's inequalities)**.** *For any $M \in \mathcal{M}_n(\mathbb{C})$, and every $1 \leq k \leq n$, if $|\lambda_1(M)| \geq \cdots \geq |\lambda_n(M)|$, and $s_1(M) \geq \cdots \geq s_n(M)$, are the spectrum of the eigenvalues and singular values respectively, then*

$$\prod_{i=1}^{k} s_i(M) \geq \prod_{i=1}^{k} |\lambda_i(M)|.$$

Moreover

$$\prod_{i=n-k+1}^{n} s_i(M) \leq \prod_{i=n-k+1}^{n} |\lambda_i(M)|.$$

Then when $k = n$ we obtain the equality,

$$\prod_{i=1}^{n} s_i(M) = \prod_{i=1}^{n} |\lambda_i(M)|.$$

Also, using majorization techniques, one may deduce from Weyl's inequalities, that for every real-valued function f, such that $t \mapsto f(e^t)$ is increasing and convex on $[s_n(M), s_1(M)]$, for every $1 \leq k \leq n$,

$$\sum_{i=1}^{k} f(|\lambda_i(M)|) \leq \sum_{i=1}^{k} f(s_i(M)).$$

In particular for $k = n$ and $f(x) = x^2$,

$$\sum_{i=1}^{n} |\lambda_i(M)|^2 \leq \sum_{i=1}^{n} s_i(M)^2 = \mathrm{Tr}(MM^*) = \sum_{i,j=1}^{n} |M_{i,j}|^2. \tag{1.2.2}$$

For the proofs and more details we refer to [13]. Consider now a random matrix X_n, whose entries are a collection of i.i.d. random variables $(X_{i,j})_{i,j=1}^n$, with common law $\mathcal{P}$ on $\mathbb{C}$. One can prove an analogous of theorem 1.2.2 for the non-Hermitian matrix X_n.

Theorem 1.2.2. *Consider a matrix $X_n = (X_{i,j})_{i,j=1}^n$ where $\{X_{i,j}\}_{1 \leq i,j \leq n}$ is a collection of i.i.d. complex random variables with common law $\mathcal{P}$, such that $Var(X_{1,1}) = 1$. Then*

$$\mu_{\frac{X_n}{\sqrt{n}}} \xrightarrow[n \to +\infty]{(w)} \mathcal{C}_1$$

where $\mathcal{C}_1$ is the uniform law on the unit disc of the complex plane, with density $z \mapsto \frac{1}{\pi}\mathbb{1}_{|z| \leq 1}$, for $z \in \mathbb{C}$.

Some special case of this theorem can be directly proved, e.g. for the case $X_{1,1} \sim N(0, \frac{1}{2}\mathbb{1}_2)$ see [13, Theorem 3.5]. Anyways for the general case, the key ingredient for the proof is to take back the problem to Hermitian matrices, through a technique called Hermitization.

1.2.1 Logarithmic potential and Hermitization

We define the logarithmic potential of a probability measure μ defined on $\mathbb{C}$ as

$$U_\mu(z) = -\int \log|z-w|\mu(dw),$$

and we define $\mathcal{P}(\mathbb{C})$ the set of probability measure for which this integral is finite. Then $U_\mu(z)$ is a function defined on $\mathbb{C} \to [-\infty, +\infty]$. One can compute the Logarithmic potential of the Circular law $\mathcal{C}_1$, see e.g. [23],

$$U_{\mathcal{C}_1}(z) = \begin{cases} -\log|z| & \text{if } |z| > 1 \\ \\ \frac{1}{2}(1-|z|^2) & \text{if } |z| \leq 1 \end{cases}. \tag{1.2.3}$$

Note that the logarithmic potential uniquely determines the measure, in the sense that for every $\mu, \nu \in \mathcal{P}(\mathbb{C})$ if $U_\mu(z) = U_\nu(z)$ for almost every $z \in \mathbb{C}$, then $\mu = \nu$, see [13, Lemma 4.1].

Consider now a matrix $M \in \mathcal{M}_n(\mathbb{C})$, and let us compute the Logarithmic potential of its ESD,

$$U_{\mu_M(z)} = -\int_{\mathbb{C}} \log|w-z|\mu_M(dw) = -\frac{1}{n}\log|\det(M-z)|.$$

For $z \notin \{\lambda_1(M), ..., \lambda_n(M)\}$,

$$-\frac{1}{n}\log|\det(M-z)| = -\frac{1}{n}\log(\sqrt{(M-z)(M-z)^*}) = -\int_0^{+\infty} \log(t)\nu_{M-z}(dt).$$

The two spectra are then bonded. This relation is the heart of the Hermitization method. For a sequence of non Hermitian matrix $(M_n)_{n\geq 1}$, this method let us deduce the convergence of μ_{M_n}, by the convergence of ν_{M_n}, and the convergence of ν_{M_n} is a problem of convergence of spectrum Hermitian matrices. The price to pay is, in first place, the introduction of the complex variable z. Also if one wants to take the limit in equation

$$U_{\mu_{M_n}}(z) = -\int_0^{+\infty} \log(t)\nu_{M_n-z}(dt),$$

the weak convergence of $(\nu_{M_n-z})_{n\geq 1}$ is not sufficient, since $\log(\cdot)$ is not bounded on $[0, +\infty)$. Thus one has to require the uniform integrability of $\log(\cdot)$ with respect to the sequence of measures $(\nu_{M_n-z})_{n\geq 1}$. Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is uniformly integrable for a sequence of probability measure $(\pi_n)_{n\geq 1}$ if

$$\lim_{t\to+\infty} \sup_{n\geq 1} \int_{|f|>t} |f(x)|\pi_n(dx) = 0.$$

We use the following property, if $\pi_n \xrightarrow[n\to+\infty]{(w)} \pi$ for some π and f is continuous and uniformly integrable for $(\pi_n)_{n\geq 1}$, then f is integrable with respect to π and

$$\lim_{n\to+\infty} \int f(x)\pi_n(dx) = \int f(x)\pi(dx).$$

This method has been used in almost every work related to non Hermitian random matrices. The original idea goes back to Girko, we will then refer to this method, as Girko Hermitization method.

Lemma 1.2.3 (Hermitization)**.** *Let $(M_n)_{n\geq 1}$ be a sequence of complex random matrices, where $M_n \in \mathcal{M}_n(\mathbb{C})$ for every $n \geq 1$. Suppose that there exists a family of non random probability measure ν_z, $z \in \mathbb{C}$, supported in $[0, +\infty)$ such that, for almost all $z \in \mathbb{C}$, almost surely*

i) $\nu_{M_z - z\mathbb{1}_n} \xrightarrow[n\to+\infty]{(w)} \nu_z$.

ii) $\log(\cdot)$ *is uniformly integrable for* $(\nu_{M_n - z})_{n\geq 1}$

Then there exists a probability measure $\mu \in \mathcal{P}(\mathbb{C})$ such that

j) almost surely $\mu_{M_n} \xrightarrow[n\to+\infty]{(w)} \mu$.

jj) for almost all $z \in \mathbb{C}$

$$U_\mu(z) = -\int_0^{+\infty} \log(s)\nu_z(ds).$$

For a proof based on the work of Tao and Vu, we refer to [13]. The hypothesis of the lemma can be weakened, mostly we can require less than uniform integrability for the $\log(\cdot)$, we refer to [13, Remark 4.4]. Anyways if $i)$ and $ii)$, both hold in probability for almost all $z \in \mathbb{C}$, then $j)$ and $jj)$ hold with the convergence in probability in $j)$, see [13, Lemma 4.3].

1.2.2 Small singular values

The proof of theorem 1.2.2, is based on the Hermitization lemma. Part $i)$ is, as already noticed, a problem of convergence of Hermitian matrices, which we can deal with using the approach for the proof of theorem 1.1.1.

To prove the uniform integrability of part $ii)$, one proves a stronger condition, namely the existence of a positive p such that

$$\limsup_n \int s^p \nu_{n^{-1/2}X_n - z}(ds) < +\infty \tag{1.2.4}$$

and

$$\limsup_n \int s^{-p} \nu_{n^{-1/2}X_n - z}(ds) < +\infty. \tag{1.2.5}$$

The finiteness of integral (1.2.4), for $p \le 2$, follows by the law of large numbers, consider the special case $z = 0$,

$$\int s^2 \nu_{n^{-1/2}X_n}(ds) = \frac{1}{n^2}\sum_{i=1}^n s_i(X)^2 = \frac{1}{n^2}\mathrm{Tr}(XX^*) = \frac{1}{n^2}\sum_{i,j=1}^n |X_{i,j}|^2 \xrightarrow[n\to+\infty]{} \mathbb{E}\left[\,|X_{1,1}|^2\,\right].$$

The hard part is to prove the finiteness of the integral (1.2.5). The turning point is the following lemma by Tao and Vu in [27], to lower bound the small singular values of $n^{-1/2}X_n - z\mathbb{1}$.

Lemma 1.2.4 (Count of small singular values)**.** *There exists $c_0 > 0$ and $0 < \gamma < 1$ such that, almost surely, for n big enough and $n^{1-\gamma} \le i \le n-1$ and all $M \in \mathcal{M}_n\mathbb{C}$,*

$$s_{n-i}(n^{-1/2}X_n + M) \ge c_0 \frac{i}{n}.$$

The key ingredient for the proof is the following lemma from [27], which relates the singular values, to distance of some opportune vector spaces.

Lemma 1.2.5 (Tao-Vu negative second moment)**.** *If M is a full rank $n' \times n$ complex matrix, with $n' \le n$ and rows $R_1, ..., R_{n'}$. Set $R_{-i} = span\{R_j : j \ne i\}$ then*

$$\sum_{i=1}^{n'} s_i^{-2}(M) = \sum_{i=1}^{n'} \mathrm{dist}(R_i, R_{-i})^{-2}.$$

The smallest singular value $s_n(n^{-1/2}X_n - z\mathbb{1})$, deserves a special mention. One indeed needs a polynomial lower bound on the least singular value in order to bound the integral of s^{-p} with respect to $\nu_{n^{-1/2}X_n - z\mathbb{1}}$. Such an estimate is not trivial to prove. Anyway, assuming further hypothesis on the law of the entries of X_n, such as absolute continuity with bounded density, the proof is rather simple, and is based on the properties of the convolution of densities, see e.g. our proof of lemma 4.2.3. One can prove the same result in complete generality, but with much more work, see [26].

Lemma 1.2.6 (Polynomial lower bound on least singular value)**.** *For every $a, d > 0$, there exists $b > 0$ such that if M is a deterministic complex $n \times n$ matrix with $s_1(M) \le n^d$, then*

$$\mathbb{P}\left(s_n(X_n + M) \le n^{-b}\right) \le n^{-a}.$$

In particular there exists $b > 0$, which may depends on d, such that almost surely, for n big enough,

$$s_n(X_n + M) \ge n^{-b}.$$

With these results, we can now give the desired bound on (1.2.5), set $Y_n(z) = n^{-1/2}X_n - z$,

$$\begin{aligned}\int s^{-p}\nu_{Y_n(z)}(ds) &= \frac{1}{n}\sum_{i=1}^{n}(s_i(Y_n(z)))^{-p}\\ &\le \frac{1}{n}\sum_{i=1}^{n-\lfloor n^{1-\gamma}\rfloor}(s_i(Y_n(z)))^{-p} + \frac{1}{n}\sum_{i=n-\lfloor n^{1-\gamma}\rfloor+1}^{n} s_n(Y_n(z)))^{-p}\\ &\le c_0^{-p}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{i}{n}\right)^{-p} + 2n^{-\gamma+bp},\end{aligned}$$

now the first term of the right hand side is a Riemann sum for $\int_0^1 s^{-p}ds$, which converges for $0 < p < 1$. Then for $0 < p < \min(\gamma/b, 1)$, we have the bound and we can apply the Hermitization lemma to establish the convergence of $\mu_{\frac{X_n}{\sqrt{n}}}$ to a measure μ, whose logarithmic potential is

$$U_\mu(z) = -\int_0^{+\infty}\log(t)\nu_z(dt)$$

for all $z \in \mathbb{C}$. Since ν_z does not depend on the law of $X_{1,1}$, it follows that also μ does not depends on the law of the entries of X_n, then μ is the circular law, the same as the Gaussian case, see e.g. [13, Chapter 3]. It is also possible to explicitly compute the integral $\int_0^{+\infty}\log(s)\nu_z(ds)$, to prove that it matches (1.2.3), the logarithmic potential of the uniform law on the unit disc of the complex plane, see [20].

1.3 Heavy tailed Hermitian random matrices

The above scenario can be drastically perturbed when we consider random matrices whose entries have heavy tails at infinity. For any $\alpha > 0$, we define $\mathbb{H}_\alpha$, as the class of laws supported in $[0, +\infty)$, with regularly varying tail of index α, meaning that for every $t > 0$,

$$G(t) := \mathcal{L}(t, +\infty) = t^{-\alpha}L(t),$$

where $L(t)$ is a function slowly variating at infinity, i.e. for any $x > 0$,

$$\lim_{t\to+\infty}\frac{L(xt)}{L(t)} = 1.$$

Define $a_n = \inf\{t > 0 : nG(t) \le 1\}$. Then $nG(a_n) = nL(a_n)a_n^{-\alpha} \xrightarrow[n\to+\infty]{} 1$, and for all $t > 0$,

$$nG(a_n t) \xrightarrow[n\to+\infty]{} t^{-\alpha}. \tag{1.3.1}$$

It is known that a_n has regular variation at ∞ with index $1/\alpha$, so that $a_n = n^{1/\alpha}\ell(n)$, for some function $\ell(n)$ slowly varying at ∞, see Resnick [21] for more details. We call

$\mathbb{H}_\alpha^*$ the subset of $\mathbb{H}_\alpha$ of the laws such that $L(t) \to c > 0$ when $t \to +\infty$, this let us take $a_n = c^{1/\alpha} n^{1/\alpha}$ in equation (1.3.1).

The case $\alpha > 2$ corresponds to the Wigner theorem. We now consider random matrices $X_n = (X_{i,j})$ with i.i.d. entries, up to requiring $X_n^* = X_n$, such that $U_{i,j} = |X_{i,j}|$ is in $\mathbb{H}_\alpha$, for $\alpha \in (0,2)$. A random variable Y is in the domain of attraction of an α-stable law, if and only if the law $\mathcal{L}$ of $|Y|$ is in $\mathbb{H}_\alpha$ for $\alpha \in (0,2)$, and if it exists the limit

$$\theta = \lim \frac{\mathbb{P}\left(X_{i,j} > t\right)}{\mathbb{P}\left(|X_{i,j}| > t\right)} \in [0,1]. \tag{1.3.2}$$

For X_n, rescaled by a_n, the following result holds.

Theorem 1.3.1 (Symmetric i.i.d matrix, $\alpha \in (0,2)$). *For every $\alpha \in (0,2)$ there exists a symmetric probability distribution μ_α on $\mathbb{R}$ depending only on α, such that almost surely,*

$$\mu_{a_n^{-1} X_n} \xrightarrow[n \to +\infty]{(w)} \mu_\alpha.$$

The limiting spectral measure μ_α has bounded density and heavy tail of index α.

This result was first proved by Ben Arous and Guionnet in [6], then, with a different method, by Caputo, Bordenave and Chafaï in [9]. Their key idea is to exhibit the local convergence of the sequence of random matrices $a_n^{-1} X_n$ to a self-adjoint operator defined as the adjacency matrix of an infinite rooted tree with random edge weights, the so called Poisson infinite weighted tree (PWIT) introduced by David Aldous in [1].

1.3.1 Local convergence and Poisson weighted infinite tree

The basic idea for the proof of theorem 1.3.1, is to apply the resolvent method in this context. It needs a bit of extra work, but once one sees the matrix X_n as the adjacency matrix of an undirected weighted graph, one can prove that $a_n^{-1} X_n$ converges locally to a PWIT, and using known spectral theory one can transfer the convergence to the resolvents.

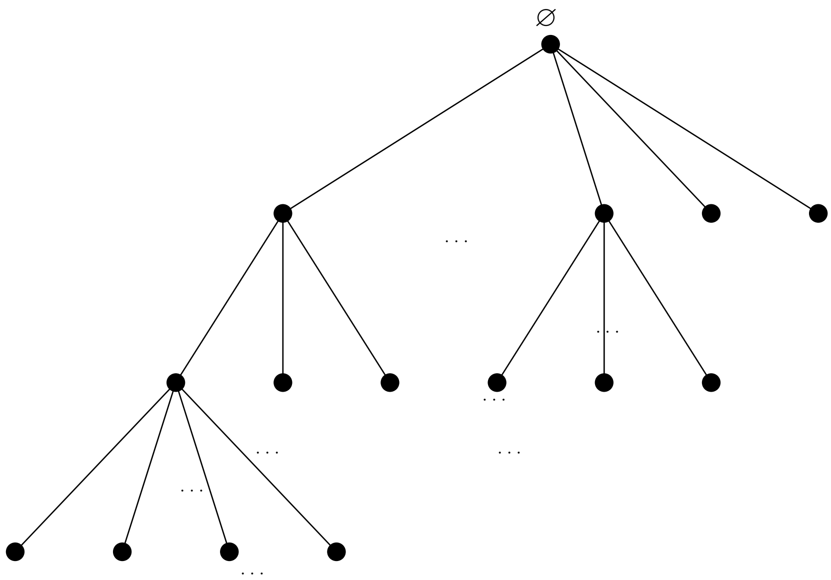
We will present more details on the local convergence and its bond with the resolvent convergence in chapter 2. Iif A_n is the adjacency matrix of the weighted graph $G = (V_n, E)$, rooted in $v \in V_n = \{1, 2, ..., n\}$, then A_n can be interpreted as on operator on a suitable Hilbert space. Moreover if A_n converges to an operator A on the same space, and both A_n and A are self adjoint, then also the respective resolvents converge. The right notion of convergence will be the local convergence.

In this regime, the operator A_n will be $a_n^{-1} X_n$, for which we will show the convergence to an operator to be defined, in the Hilbert space $\ell^2(V)$, where V is the the vertex set of the PWIT, and the scalar product is the usual

$$\langle \varphi, \psi \rangle = \sum_{v \in V} \varphi_v \bar{\psi}_v, \quad \varphi_v = \langle \delta_v, \varphi \rangle,$$

where $\varphi, \psi \in \mathbb{C}$, and δ_v denote the unit vector with support v.

Let us now introduce the Poisson weighted infinite tree. The PWIT(ν) is a random rooted tree, with vertex set identified with $\mathbb{N}^f := \cup_{k \geq 1} \mathbb{N}^k$, where $\mathbb{N}^0 = \{\varnothing\}$, is the root. The root's offsprings are indexed by the elements of $\mathbb{N}$, and, in general, the offspring os some vertex $v \in \mathbb{N}^k$ are indexed $(v1), (v2), ... \in \mathbb{N}^{k+1}$. To the edges of the tree we assign marks according to a collection $\{\Xi_v\}_{v \in \mathbb{N}^f}$, of independent realizations of a Poisson process of intensity ν on $\mathbb{R}$. Starting from the root $\varnothing$, we sort $\Xi_\varnothing = \{y_1, y_2, ...\}$ in such a way that $|y_1| \leq |y_2| \leq \cdots$, and assign the mark y_i to the offspring of the offspring labeled i. Recursively, for any $v \in \mathbb{N}^f$, sort Ξ_v, and assign the mark y_{vi} to the offspring labeled vi. Since Ξ_v has $\nu(\mathbb{R})$ elements, in average, as convention, if $\nu(\mathbb{R}) < +\infty$ we assign the mark ∞ to the remaining edges.



Let us then introduce the limiting operator of $a_n^{-1} X_n$. Let θ be as in equation (1.3.2), then define the positive Borel measure on the real line ℓ_θ as,

$$\ell_\theta(dx) = \theta \mathbb{1}_{x>0} dx + (1-\theta) \mathbb{1}_{x<0} dx,$$

and consider a realization of $PWIT(\ell_\theta)$. Denote the mark from vertex $v \in \mathbb{N}^k$ to $vk \in \mathbb{N}^{k+1}$ by y_{vk}. Note that almost surely

$$\sum_{k \geq 1} |y_{vk}|^{-2/\alpha} < +\infty,$$

since almost surely $\lim_{k\to\infty} k^{-1}|y_{vk}| = 1$. Define $\mathcal{D}$ the dense set in $\ell^2(V)$ of vectors of finite support, we may then define a linear operator $T : \mathcal{D} \to \ell^2(V)$ by letting, for $v, w \in \mathbb{N}^f$,

$$T(v,w) = \langle \delta_v, T\delta_w \rangle = \begin{cases} \mathrm{sign}(y_w)|y_w|^{-1/\alpha} & \text{if } w = vk \text{ for some integer } k \\ \mathrm{sign}(y_v)|y_v|^{-1/\alpha} & \text{if } v = wk \text{ for some integer } k \\ 0 & \text{otherwise} \end{cases} \tag{1.3.3}$$

The operator T is symmetric. It is also self-adjont, see [9, Proposition A.2].

The next step is to prove the local convergence of $(a_n^{-1}X_n, 1)$ to $(T, \varnothing)$. We will not go into details, since we present a similar argument in chapter 2. For the details of the heavy-tailed hermitian case see the proof of [9, Theorem 2.3.i)]. It is based on the fact that the order statistics of the row vectors of $a_n^{-1}X_n$ converges to a Poisson point process of intensity $\alpha x^{-\alpha-1}dx$, so that, if we consider for example vertex 1, and define $V_1 \geq \cdots \geq V_n$ the ordered statistics of $(X_{11}, ..., X_{1,n})$, then

$$a_n^{-1}(V_1, ..., V_n) \xrightarrow[n\to+\infty]{d} (\gamma_1, \gamma_2, ...) \tag{1.3.4}$$

where $(\gamma_i)_{i\geq 1}$ is an ordered Poisson point process of intensity $\alpha x^{-\alpha-1}dx$, see [17] or [9, Lemma 2.4].

The convergence in equation (1.3.4), is to be interpreted as follows: for any fixed $k \geq 1$ the joint law of $a_n^{-1}(V_1, ..., V_k)$ converges weakly to $(\gamma_1, \gamma_2, ..., \gamma_k)$.

We can now compute the Cauchy-Stieltjes transform of $\mu_{a_n^{-1}X_n}$. In particular, for $z \in \mathbb{C}^+$,

$$g_{\mu_{a_n^{-1}X_n}}(z) = \int \frac{1}{x-z}\mu_{a_n^{-1}X_n}(dx) = \frac{1}{n}\sum_{i=1}^{n}(R_{a_n^{-1}X_n}(z))_{ii},$$

if $(R_{a_n^{-1}X_n}(z)) = (a_n^{-1}X_n - z\mathbb{1})^{-1}$ is the resolvent of $a_n^{-1}X_n$. Again, by concentration inequalities we can focus on the expected value of the random measure $\mu_{a_n^{-1}X_n}$, and by exchangeability of the variables,

$$\mathbb{E}\left[g_{\mu_{a_n^{-1}X_n}}(z) \right] = \mathbb{E}\left[(R_{a_n^{-1}X_n}(z))_{11} \right].$$

We can then exploit the bond between local convergence and resolvent convergence, and affirm

$$\mathbb{E}\left[(R_{a_n^{-1}X_n}(z))_{11} \right] = \mathbb{E}\left[\left\langle \delta_1, (a_n^{-1}X_n - z\mathbb{1})^{-1}\delta_1 \right\rangle \right] \xrightarrow[n\to+\infty]{} \mathbb{E}\left[\left\langle \delta_\varnothing, (T-z)^{-1}\delta_\varnothing \right\rangle \right]. \tag{1.3.5}$$

Define $\mathbb{E}\left[\left\langle \delta_\varnothing, (T-z)^{-1}\delta_\varnothing \right\rangle \right] =: \mathbb{E}[h(z)]$, and note that $\mathbb{E}[h(z)] = \mathbb{E}\left[g_{\mu_\varnothing} \right] = g_{\mathbb{E}[\mu_\varnothing]}$, is the expected value of the pointed spectral measure with respect to the vector $\delta_\varnothing$, associated to the self-adjoint operator T.

The limiting measure μ_α is not explicitly known, but $h(z)$ satisfies a recursive distributional equation, from which one can deduce some properties of μ_α. Namely, by recursive properties of $\mathbb{N}^f$,the vertex set of the PWIT, one can prove the distributional equality

$$h(z) \stackrel{d}{=} -\left(z + \sum_{k\geq 1} \xi_k h_k(z)\right)^{-1} \tag{1.3.6}$$

where $(h_k(z))$ are i.i.d. random variables with the same law as $h(z)$, and $\{\xi_k\}_{k\geq 1}$ is an independent Poisson point process with intensity $\frac{\alpha}{2}x^{-1-\alpha/2}dx$. For the proof we refer to [9, Thoerem 4.1].

In particular from equation (1.3.6), one can prove the following properties of the limiting spectral distribution μ_α:

i) μ_α is absolutely continuous on $\mathbb{R}$ with bounded density

ii) The density of μ_α at 0 is equal to

$$\frac{1}{\pi}\Gamma\left(1+\frac{2}{\alpha}\right)\left(\frac{\Gamma(1-\frac{\alpha}{2})}{\Gamma(1+\frac{\alpha}{2})}\right)^{-1/\alpha}$$

iii) μ_a is heavy tailed and, as $t \to +\infty$,

$$\mu_\alpha(t,+\infty) \sim \frac{1}{2}t^{-\alpha}.$$

1.4 Non Hermitian Heavy tailed random matrices

In 2012 BCC, proved an analogous of theorem 1.2.2 for non Hermitian matrices X_n with i.i.d entries $(X_{i,j})_{i,j=1}^n$, whose law is absolutely continuous with bounded density, and in $\mathbb{H}_\alpha$, $\alpha \in (0,2)$, see [10].

Theorem 1.4.1. *There exists a probability measure μ_α on $\mathbb{C}$ depending only on α such that, almost surely*

$$\mu_{a_n^{-1}X_n} \xrightarrow[n\to+\infty]{(w)} \mu_\alpha.$$

The measure μ_α has bounded density and finite moments of any order.

The approach for the proof relies on Girko's Hermitization method, so that an intermediate step is to prove the convergence for the ESD of singular values of the matrix $a_n^{-1}X_n - z\mathbb{1}$.

Theorem 1.4.2. *For all $z \in \mathbb{C}$ there exists a probability measure $\nu_{\alpha,z}$ on $[0,+\infty)$ depending only on α and z such that, almost surely*

$$\nu_{a_n^{-1}X_n - z\mathbb{1}} \xrightarrow[n\to+\infty]{(w)} \nu_{\alpha,z}.$$

As in the light tailed case, this is a problem of convergence of the spectrum of Hermitian matrices, see [10, Theorem 1.2]. The approach is via bipartization.

1.4.1 Bipartization of a matrix

Given a complex $n \times n$ matrix $A \in \mathcal{M}_n(\mathbb{C})$, we define its bipartized version B, in $\mathcal{M}_n(\mathcal{M}_2(\mathbb{C})) \simeq \mathcal{M}_{2n}(\mathbb{C})$, as

$$B = (B_{ij})_{i,j=1}^n \qquad \text{where} \qquad B_{ij} = \begin{pmatrix} 0 & A_{ij} \\ (A^*)_{ij} & 0 \end{pmatrix} = \begin{pmatrix} 0 & A_{ij} \\ \overline{A_{ji}} & 0 \end{pmatrix}.$$

Since $(B^*)_{ij} = \overline{B_{ji}}$, as an element of $\mathcal{M}_{2n}(\mathbb{C}), B$ is Hermitian.

In graph term, we can identify the non Hermitian matrix A as the weight matrix of an oriented network on the vertex set $\{1, ..., n\}$, with weight A_{ij} on the oriented edge (i, j). The bipartized version of A, the Hermitian matrix B, is the weight matrix of a non oriented network on $\{-1, ..., -n, 1, ..., n\}$, with weight A_{ij} on the non oriented edge $\{i, -j\}$.

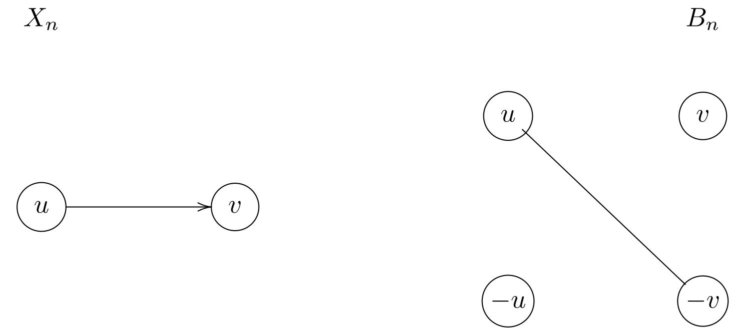


Figure 1.1: *The edge (i, j) maps into $\{i, -j\}$ with the bipartization.*

Through the permutation σ, defined as

$$\sigma(i) = \begin{cases} 2i - 1 & \text{if } 1 \leq i \leq n \\ 2(i - n) & \text{if } n + 1 \leq i \leq 2n, \end{cases}$$

applied to the row vectors, we can rearrange the entries of the matrix B as

$$\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$$

We will denote the first n rows and columns of the bipartized matrix as $\{-1, ..., -n\}$ and the last n as $\{1, ..., n\}$. With this notation there is no edge of type either $\{i, j\}$ or $\{-i, -j\}$. The non-oriented graph is indeed a bipartized graph. Since we will use Girko's Hermitization method to prove the convergence of the empirical spectral distribution of eigenvalues of the matrix X_n, we will deal with the matrix $X_n - z\mathbb{1}$. We call its bipartized version $B_n(z)$. Define

$$\mathbb{H}_+ = \left\{ U(z,\eta) = \begin{pmatrix} \eta & z \\ \bar{z} & \eta \end{pmatrix} \,:\, z \in \mathbb{C},\, \eta \in \mathbb{C}_+ \right\} \subset \mathcal{M}_2(\mathbb{C}),$$

and $U \otimes \mathbb{1}_n \in \mathcal{M}_n(\mathcal{M}_2(\mathbb{C}))$ as $(U \otimes \mathbb{1}_n)_{ij} = \delta_{ij} U$, then $B_n(z) = (B_n - U(z,0) \otimes \mathbb{1}_n)$, where B_n is the bipartized version of X_n. As an element of $\mathcal{M}_{2n}$ the resolvent of the the matrix $B_n(z)$ has the form,

$$R_{B_n(z)}(\eta) = (B_n(z) - \eta \mathbb{1}_{2n})^{-1} = (B_n - U(z,\eta) \otimes \mathbb{1}_n)^{-1} = R_{B_n}(U). \tag{1.4.1}$$

Where $R_{B_n}(U)$ the resolvent matrix defined in $\mathcal{M}_n(\mathcal{M}_2(\mathbb{C}))$. For $1 \leq i,j \leq n$ $(R_{B_n}(U))_{ij} \in \mathcal{M}_2(\mathbb{C})$. When $i = j$, the matrix

$$(R_{B_n}(U))_{ii} = \begin{pmatrix} a_i(z,\eta) & b_i(z,\eta) \\ b_i'(z,\eta) & c_i(z,\eta) \end{pmatrix},$$

has bounded entries, see [10, Lemma 2.2].

For a complex $n \times n$ matrix A define the symmetrized version of ν_A as

$$\check{\nu}_A = \frac{1}{n} \sum_{i=1}^{n} (\delta_{\sigma_i(A)} + \delta_{-\sigma_i(A)}).$$

once one has rearranged the entries of the bipartized matrix, it is easy to notice that,

$$\mu_B = \check{\nu}_A \tag{1.4.2}$$

if B is the bipartized version of A, see [10, Theorem 2.1]. Thus the analysis of the spectrum of singular values of the non Hermitian matrix A, is reduced to the study of the spectrum of the Hermitian matrix B.

In [10], the authors develop all the results needed for the convergence to bipartized matrices. In particular it continues to hold the fact that the local convergence implies the convergence of the resolvents, but now the resolvents have to be considered of the bipartized operators. Also for the non Hermitian matrices the local convergence is to a Poisson weighted infinite tree, but the measure now has density $2\ell_\theta$. This can be explained roughly by a simple observation. Focus on a single vertex in the non bipartized matrix, e.g. vertex 1. We have two vectors of weights, $(X_{12}, ..., X_{1,n})$ and $(X_{21}, ..., X_{n1})$ if we do not consider the loop X_{11}. The component of the first vector are the weights of the outgoing edges from 1, the second are the weights of the edges incoming to 1. Since the order statistics of both those vectors converges to an ordered Poisson point process of intensity $\alpha x^{-\alpha-1}dx$ if rescaled by a factor a_n, by thinning property of Poisson process, this is equivalent to have a Poisson process of double intensity, which we then split into two process of halved intensity. We refer to [10, subsection 2.6], for further details.

Call then A the limiting operator defined on PWIT$(2\ell_\theta)$. We again can transfer the local convergence to resolvents of the bipartized operators, provided the self-adjointness, which is proved in [10, Proposition 2.8]. So that, if B_n is the bipartized version of $a_n^{-1}X_n$ and B is the bipartized operator of A, one has

$$(R_{B_n}(U))_{1,1} := \begin{pmatrix} a_1(z,\eta) & b_1(z,\eta) \\ b_1'(z,\eta) & c_1(z,\eta) \end{pmatrix} \xrightarrow[n\to+\infty]{(w)} (R_B(U))_{\varnothing,\varnothing} =: \begin{pmatrix} a(z,\eta) & b(z,\eta) \\ b'(z,\eta) & c(z,\eta) \end{pmatrix}. \tag{1.4.3}$$

Since B is almost surely self-adjoint, it implies it exists a measure $\nu_{\varnothing,z}$, the pointed spectral measure on vector $\delta_\varnothing$, such that

$$a(z,\eta) := (R_B(U))_{\varnothing,\varnothing} = \int \frac{1}{x-\eta}\nu_{\varnothing,z}(dx) = g_{\nu_{\varnothing,z}}(\eta).$$

By concentration inequalities we can focus on expected values, and compute $g_{\mathbb{E}\left[\check{\nu}_{a_n^{-1}X_n-z}\right]}(\eta)$ which is equal to $\mathbb{E}\left[a_1(z,\eta)\right]$ from equation (1.4.3), see [10, Theorem 2.1], so that again by (1.4.3),

$$\lim_{n\to\infty}\mathbb{E}\left[\check{\nu}_{a_n^{-1}X_n-z}\right] = \lim_{n\to\infty}\mathbb{E}\left[a_1(z,\eta)\right] = \mathbb{E}\left[a(z,\eta)\right],$$

since, again, the resolvent is a bounded and analytic function of the spectrum. Thus $\check{\nu}_{a_n^{-1}X_n-z}$ converges to a measure $\nu_{\alpha,z} = \mathbb{E}\left[\nu_{\varnothing,z}\right]$. For the uniform integrability of $\log(\cdot)$ for the sequence of measure $(\nu_{a_n^{-1}X_n-z})_{n\geq 1}$, one has essentially to readapt to the heavy tailed setting, the proof of the light tailed case. We will not present any details now, since we will go through similar arguments in chapter 4, for further details we refer to [10, section 3]. This gives an idea of the proof of theorem 1.4.1.

As $h(z)$ for the Hermitian case, also the random variable $a(z,\eta)$ satisfies some recursive distributional equation, through which it is possible to give some properties for the limiting spectral distribution μ_α. Namely we have,

$$a(z,\eta) \stackrel{d}{=} \frac{\eta + \sum_{k\geq 1}\xi_k a_k}{|z|^2 - (\eta + \sum_{k\geq 1}\xi_k a_k)(\eta + \sum_{k\geq 1}\xi_k' a_k')},$$

where $(a_k)_{k\geq 1}$ and $(a_k')_{k\geq 1}$ are i.i.d. copies of $a(z,\eta)$ and $(\xi_k)_{k\geq 1}$, $(\xi_k')_{k\geq 1}$ are independent poisson Process on $[0,+\infty)$ of intensity $\frac{\alpha}{2}x^{-1-\alpha/2}dx$.

One can then deduce some properties of μ_α, see [10, Theorem 1.3].

1.5 Markov Matrices

As already noticed in the introduction, Markov matrices are an interesting instance of the problem of the convergence of the spectrum of random matrices with non-independent entries. Starting from a collection of random variables $\{U_{i,j}\}_{i,j\geq 1}^{n}$, we can define the random Markov matrix

$$X_n = (X_{i,j})_{i,j=1}^{n} = \left(\frac{U_{i,j}}{\rho_i}\right)_{i,j=1}^{n} \qquad \rho_i = \sum_{j=1}^{n} U_{i,j}$$

When the entries of the random matrix have light tails, due to a law of large numbers for ρ_i, the work of BCC shows that, up to rescaling, the behavior of the Markov matrix X_n is the same as non normalized matrix, see [11, 8].

We then directly go through the analysis of the spectrum of random Markov matrices with heavy tailed entries.

1.5.1 Hermitian Heavy tailed Markov matrices

In [9], Bordenave, Caputo, and Chafaï study the case of Markov matrices, with heavy tailed entries. When $\alpha \geq 1$, it still holds a law of large numbers for the normalization ρ_i, and one can reduce the problem to the non-normalized case. We focus on the case $\alpha \in (0,1)$. In this regime the entries of $U_n = (U_{i,j})_{i,j=1}^{n}$ have infinite mean, and there is no law of large numbers. Anyways, if one considers the first row of the matrix U_n, $U^{(1)} = (U_{1,1}, ..., U_{1,n})$, for the vector of its order statistics $(V_1, ..., V_n)$, still holds (1.3.4), and

$$a_n^{-1}(V_1, ..., V_n) \xrightarrow[n\to+\infty]{(w)} (\gamma_1, \gamma_2, ...).$$

Where $\{\gamma_i\}_{i\geq 1}$ is an ordered Poisson process of intensity $\alpha x^{-\alpha-1}dx$. Since almost surely $\lim_{n\to+\infty} \gamma_n / n^{-1/\alpha} = 1$, and

$$\sum_{n\geq 1} n^{-1/\alpha} < +\infty$$

for $\alpha \in (0,1)$, one can prove that for the normalization holds the following convergence result,

$$a_n^{-1}\rho_i = a_n^{-1}\sum_{i=1}^{n} U_{1,i} = a_n^{-1}\sum_{i=1}^{n} V_i \xrightarrow[n\to+\infty]{} \sum_{i\geq 1}\gamma_i,$$

see [9, Lemma 2.4]. Combining this latter with (1.3.4), one can prove that

$$\left(\frac{V_1}{\rho_1}, \cdots, \frac{V_n}{\rho_1}\right) \xrightarrow[n\to+\infty]{(w)} \left(\frac{\gamma_1}{\sum_{i\geq 1}\gamma_i}, \frac{\gamma_2}{\sum_{i\geq 1}\gamma_i}, ...\right).$$

The law of the vector on the right hand side, is called Poisson -Dirichlet law of index α. As in the non-normalized case, we can use this convergence result to establish the local convergence of the Markov matrix X_n to an operator defined on the PWIT, which now weights edges with Poisson-Dirichlet weights. Namely, consider a realization of a PWIT(ℓ_1), where ℓ_1 is the Lebesgue measure on $[0,+\infty)$, and define

$$\rho(v) = y_v^{-1/\alpha} + \sum_{k\geq 1} y_{vk}^{-1/\alpha},$$

where $y_\varnothing = 0$. Then we can define the linear operator K, on the dense subset of $\mathbb{N}^f$ of vectors with finite support $\mathcal{D}$, as follow,

$$K(v,w) = \langle \delta_v, K\delta_w \rangle = \begin{cases} \frac{y_w^{-1/\alpha}}{\rho(v)} & \text{if } w = vk \text{ for some integer } k \\ \frac{y_v^{-1/\alpha}}{\rho(v)} & \text{if } v = wk \text{ for some integer } k \\ 0 & \text{otherwise} \end{cases} \tag{1.5.1}$$

Note that K is not symmetric, but it becomes symmetric in the weighted Hilbert space $\ell^2(V,\rho)$, defined by the scalar product

$$\langle \varphi, \psi \rangle_\rho := \sum_{u\in V} \rho(u)\bar{\varphi}_u \psi_u.$$

Moreover on the same Hilbert space K is a bounded self-adjoint operator, since by Schwartz's inequility

$$\begin{aligned} \langle K\varphi, K\varphi \rangle_\rho^2 &= \sum_{u\in V} \rho(u) |\sum_{v\in V} K(u,v)\varphi_v|^2 \\ &\leq \sum_{u\in V} \rho(u) \sum_{v\in V} K(u,v)|\varphi_v|^2 \\ &= \sum_{v\in V} \rho(v)|\varphi_v|^2 = \langle \varphi, \varphi \rangle_\rho^2. \end{aligned}$$

To work on the unweighted Hilbert space we actually have to consider the operator

$$S(u,v) = \sqrt{\frac{\rho(v)}{\rho(w)}} K(v,w) = \frac{T(v,w)}{\sqrt{\rho(v)\rho(w)}} \tag{1.5.2}$$

for $v, w \in \mathbb{N}^f$, and T is as in (1.3.3). The map $\delta_v \mapsto \sqrt{\rho(v)}\delta_v$, induces a linear isometry between $\ell^2(V,\rho)$ and the unweighted Hilbert space $\ell^2(V)$. In this way, one can prove that S is the local limiting operator of the matrix $P_n = (P_{i,j})_{i,j=1}^n$ and $P_{i,j} = U_{i,j}/\sqrt{\rho_i \rho_j}$. For the proof see, [9, Theorem 2.3 (iii)]. Note that by (1.5.2), the spectrum of S is contained in $[-1,1]$. Call

$$R^{(n)}(z) = (P_n - z\mathbb{1}_n)^{-1} \quad \text{and} \quad R(z) = (S-z)^{-1}$$

the resolvents of P_n and S respectively. For $l \in \mathbb{N}$ set

$$p_l := \left\langle \delta_\varnothing, S^l \delta_\varnothing \right\rangle.$$

Note that $p_l = \rho(\varnothing)^{-1}\langle \delta_\varnothing, K\delta_\varnothing\rangle$ is the probability that the random walk on the PWIT associated to K comes back to the root after l steps, starting from the root. In particular $p_{2n+1} = 0$ for any $n \geq 0$. Set $p_0 = 1$, and let $\mu_\varnothing$ done the spectral measure of S pointed at the vector $\delta_\varnothing$. Equivalently, $\mu_\varnothing$ is the spectral measure of K pointed at the $\ell^2(V,\rho)$ normalized vector $\hat{\delta}_\varnothing = \delta_\varnothing/\sqrt{\rho(\varnothing)}$. In particular,

$$p_l = \int x^l \mu_\varnothing(dx).$$

Since all odd moments equal 0, $\mu_\varnothing$ is symmetric. For any $z \in \mathbb{C}^+$ we have

$$\langle \delta_\varnothing, R(z)\delta_\varnothing\rangle = \int_{-1}^{1} \frac{1}{x-z}\mu_\varnothing(dx) = g_{\mu_\varnothing}.$$

By construction

$$\frac{1}{n}Tr(R^{(n)}(z)) = \int_{-1}^{1} \frac{1}{x-z}\mu_K(dx) = g_{\mu_K}.$$

By exchangeability and linearity we have

$$\mathbb{E}\left[\frac{1}{n}Tr(R^{(n)}(z))\right] = \mathbb{E}\left[R^{(n)}(z)_{1,1}\right] = \mathbb{E}\left[\, g_{\mu_K}\,\right] = g_{\mathbb{E}[\,\mu_K\,]}.$$

Now by local convergence, we may infer that

$$\lim_{n\to+\infty} g_{\mathbb{E}[\,\mu_K\,]} = \lim_{n\to+\infty} \mathbb{E}\left[R^{(n)}(z)_{1,1}\right] = \mathbb{E}\left[\,\langle \delta_\varnothing, R(z)\delta_\varnothing\rangle\,\right] = g_{\mathbb{E}[\,\mu_\varnothing\,]}.$$

Now one has to prove a concentration-type result, as

$$\lim_{n\to+\infty} |g_{\mu_K} - g_{\mathbb{E}[\,\mu_K\,]}| = 0$$

for any $z \in \mathbb{C}^+$, see [9], for details. This gives an idea of the proof of the following result, provided to set $\widetilde{\mu}_\alpha = \mathbb{E}\left[\,\mu_\varnothing\,\right]$.

Theorem 1.5.1. *For every $\alpha \in (0,1)$, there exists a symmetric probability distribution $\widetilde{\mu}_\alpha$ supported on $[-1,1]$, depending only on α, such that almost surely,*

$$\mu_K \xrightarrow[n\to+\infty]{(w)} \widetilde{\mu}_\alpha$$

1.6 Non Hermitian random Markov matrices with heavy tailed weights

We investigate the convergence of the spectrum of a non Hermitian Markov matrix, with heavy tailed entries. For entries in $\mathbb{H}^*_\alpha$, $\alpha \in (0,1)$, with bounded density, if X_n is the normalized matrix

$$X_n = (X_{ij})_{i,j=1}^n = \left(\frac{U_{ij}}{\sum_k U_{ik}}\right)_{i,j=1}^n = \left(\frac{U_{ij}}{\rho_i}\right)_{i,j=1}^n, \tag{1.6.1}$$

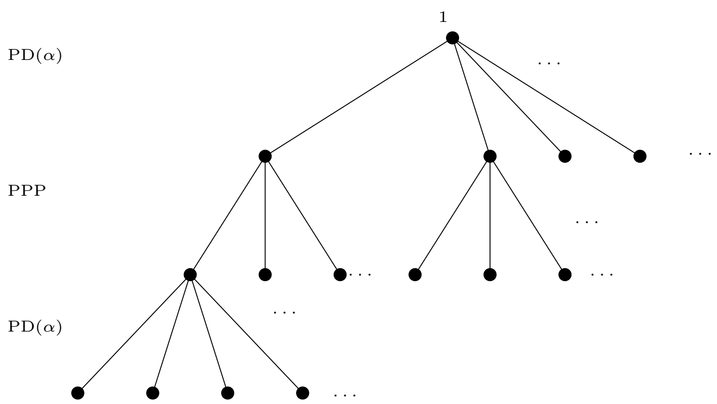
we prove the next result.

Theorem 1.6.1. *If the law $\mathcal{L}$ of $\{U_{ij}\}_{i,j=1}^{n} \in \mathbb{H}_\alpha^*$ has bounded density, then there exists a probability measure μ_α on $\{z \in \mathbb{C} : |z| \leq 1\}$, depending only on α, such that for X_n as in* (1.6.1),

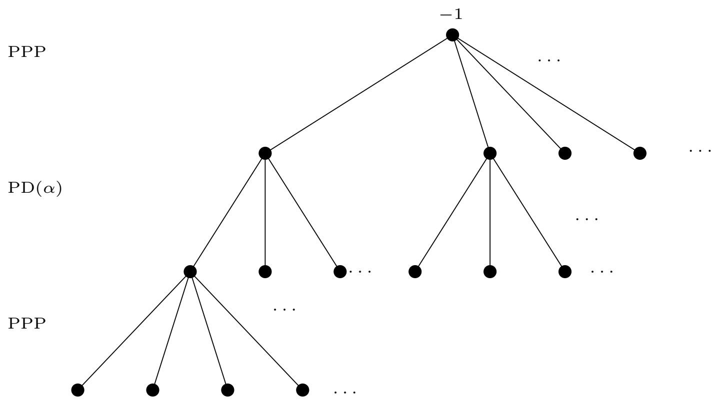
$$\mu_{X_n} \xrightarrow[n\to+\infty]{(w)} \mu_\alpha$$

We want to apply Girko's hermitization method. Therefore we first prove the convergence of the ESD of the singular values of $X_n - z\mathbb{1}_n$, to a measure $\nu_{z,\alpha}$, then we prove the uniform integrability in probability of the function $\log(\cdot)$, with respect to the sequence $(X_n - z\mathbb{1}_n)_{n\geq 1}$.

The first step in the proof of the convergence of $\nu_{X_n - z}$, is to investigate the local convergence of $B_n(z)$, the bipartized version of $X_n - z\mathbb{1}$, since we want to take advantage of the bond between local convergence and resolvent convergence.Focus on vertex 1 of X_n, and consider, $R_1 := (X_{11}, ..., X_{1,n})$ and $C_1 := (X_{11}, ..., X_{n1})$. Those vectors are the first row and the first column of the matrix X_n respectively. For both vectors consider the order statistics $\hat{R}_1$ and $\hat{C}_1$. Unlike the i.i.d heavy tails case, they now have different limiting behavior. Indeed $\hat{R}_1$ behaves as in the Hermitian case and it converges to a $PD(\alpha)$. The column vector $\hat{C}_1$ has a different limiting behavior, we will prove it converges to a function of a Poisson process of intensity $\alpha x^{-\alpha-1}dx$. Then the local limit in this scenario is an infinite weight tree as the PWIT, but it has generations with alternated distribution. As in the following figure.



Moreover the alternation depends on which vertex we pick to be the root. So that if the root is in one of the n vertices $\{-1, ..., -n\}$ we introduced for bipartizing X_n, the alternation is switched.



Namely we have two different local limits, depending on which vertex we decide to be the root. This implies that the limiting spectral measure of the singular values will be an average of this two different limiting spectral distribution, so that in the case $z = 0$, if $\mathcal{A}_+$ and $\mathcal{A}_-$ are the two limiting operators, we have

$$\mathbb{E}\left[\check{\nu}_{X_n}\right] = \mathbb{E}\left[\mu_{B_n(0)}\right] \xrightarrow[n\to+\infty]{(w)} \frac{1}{2}\mathbb{E}\left[\mu_{(\mathcal{A}_+,\varnothing)}\right] + \frac{1}{2}\mathbb{E}\left[\mu_{(\mathcal{A}_-,\varnothing)}\right].$$

We are still able to transport the convergence to resolvent since both limiting operators are essentially self-adjoint, see proposition 3.1.1. We prove the following result.

Theorem 1.6.2. *For any $z \in \mathbb{C}$ and $\alpha \in (0,1)$ there exists a measure $\nu_{z,\alpha}$, depending only on z and α, such that*

$$\nu_{X_n - \mathbb{1}z} \xrightarrow[n\to+\infty]{(w)} \nu_{z,\alpha}$$

While μ_{X_n} is concentrated on $\{w \in \mathbb{C} : |w| \leq 1\}$, here $\nu_{X_n} - z$ can have unbounded support. Simulations however suggest a very light tail of the distribution, see e.g. figure 1.2.

We indeed prove a finiteness result for the exponential moment of the limiting spectral distribution of the singular values, see proposition 3.3.3.

We then readapt results from [10] and [11] to prove the uniform integrability of the $\log(\cdot)$ for the sequence $(\nu_{X_n - z\mathbb{1}_n})$ and conclude the prove of theorem 1.6.1.

We will also observe that that the logarithmic potential of μ_α is not infinite, so that μ_α is not a Dirac delta in 0, and that the second moment of μ_α, is strictly less than 1, so that μ_α is not supported in $\{z \in \mathbb{C} : |z| = 1\}$.

As suggested by figures 4, 5, 6 and 7, the measure μ_α should exhibit some interesting concentration phenomenon within a disc of radius $z < 1$.

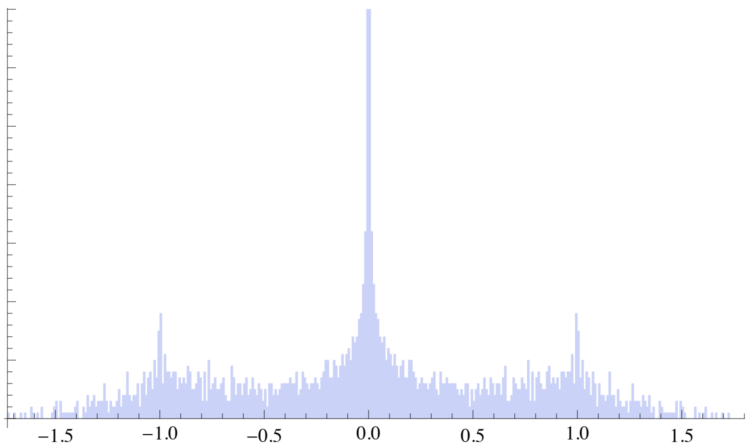


Figure 1.2: *Histogram of* $\check{\nu}_{X_n}$ *with* $n = 1000$ *and* $\alpha = 0.5$*. The entries of the non-normalized matrix are i.i.d variables distributed as* $U^{-1/\alpha}$*, where* U *has the uniform distribution in* $[0, 1]$.

Chapter 2

Local Convergence

This chapter is dedicated to the study of the local structure of non Hermitian random Markov matrices with heavy tailed weights. We will look at the matrix X_n as an adjacency matrix of a weighted graph.

2.1 Local weak convergence to PWIT

This section explores the local convergence. We first introduce the notion of local convergence, then look at convergence, as a vector of random variables, of the rows of the matrix B_n, and finally we explore the local convergence of B_n.

2.1.1 Local operator convergence

Let V be a countable set, consider the Hilbert space $\ell^2(V)$, with the scalar product

$$\langle \varphi, \psi \rangle = \sum_{v \in V} \varphi_v \bar{\psi}_v, \quad \varphi_v = \langle \delta_v, \varphi \rangle$$

where $\varphi, \psi \in \mathbb{C}^V$ and δ_v is the unit vector with support v. Let $\mathcal{D}(V)$ be the dense subset of $\ell^2(V)$, of the vectors with finite support.

Definition 2.1.1 (Local convergence). *Suppose $(A_n)_{n\geq 1}$ is a sequence of bounded operators on $\ell^2(V)$ and A is a linear operator on the same space with domain $D(A) \supset \mathcal{D}(V)$. For any $u, v \in V$ we say that (A_n, u) converges locally to (A, v) and write*

$$(A_n, u) \xrightarrow[n\to+\infty]{loc} (A, v),$$

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

$$\sigma_n^{-1} A_n \sigma_n \varphi \xrightarrow[n\to+\infty]{} A\varphi$$

in $\ell^2(V)$.

Assume in addition that A is closed and $\mathcal{D}(V)$ is a core for A, the local convergence is the standard strong convergence of operators in $\ell^2(V)$, up to a re-indexing of V which preserves a distinguished element. The nice property of local convergence is its bond with strong convergence of the resolvents, as stated in the next theorem from [10].

Theorem 2.1.2 (From local convergence to resolvent)**.** *Assume that (A_n) and A satisfy the condition of definition 2.1.1, and $(A_n, u) \xrightarrow[n\to+\infty]{loc} (A, v)$, for some $u, v \in V$. Let B_n be the self-adjoint bipartized operator of A_n. If the bipartized operator B of A is self-adjoint, and $\mathcal{D}(V)$ is a core for B, then, for all $U \in \mathbb{H}_+$,*

$$\langle \delta_u, R_{B_n}(U)\delta_u \rangle \xrightarrow[n\to+\infty]{} \langle \delta_v, R_B(U)\delta_v \rangle, \tag{2.1.1}$$

where $R_B(U) = (B(z) - \eta)^{-1}$ is the resolvent of $B(z)$.

This results shall be applied to random operators on $\ell^2(V)$, which satisfy the conditions of Definition 2.1.1 with probability one. In this case we say that $(A_n, u) \xrightarrow[n\to+\infty]{loc} (A, v)$ in distribution, if there exists a random bijection σ_n as in Definition 2.1.1 such that $\sigma_n^{-1} A_n \sigma_n \varphi \xrightarrow[n\to+\infty]{d} A\varphi$, for all $\varphi \in \mathcal{D}(V)$. A random vector $\varphi_n \in \ell^2(V)$ converges in distribution to $\varphi \in \ell^2(V)$ if for all bounded continuous $f : \ell^2(V) \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}\,[\, f(\varphi_n)\,] = \mathbb{E}\,[\, f(\varphi)\,]\,.$$

2.1.2 Single row convergence

There are two types of row vectors of B_n. Fix $v \in \{1, 2, ..., n\}$, e.g. $v = 1$. The weights of the outgoing edges correspond to the first row of the matrix X_n, the vector

$$(X_{11}, X_{12}, ..., X_{1n}) = \frac{1}{\rho_1}(U_{11}, U_{12}, ..., U_{1n}).$$

Its convergence has already been explored by Bordenave, Caputo, and Chafaï in [9]. If $V_1 \geq V_2 \geq \cdots \geq V_n$ correspond to the order statistics of the vector $(U_{11}, U_{12}, ..., U_{1n})$ then

$$\frac{1}{\rho_1}(V_1, V_2, ..., V_n) \xrightarrow[n\to+\infty]{d} \left(\frac{\gamma_1}{\sum_{k\geq 1}\gamma_k}, \frac{\gamma_2}{\sum_{k\geq 1}\gamma_k}, ... \right) \stackrel{d}{=} PD(0, \alpha),$$

where $\{\gamma_i\}_{i\geq 1}$ is an ordered Poisson point process of intensity $\alpha x^{-\alpha-1}dx$, and the vector has distribution Poisson-Dirichlet of parameter α, denoted $PD(0, \alpha)$, see [9, Lemma2.4].

On the other hand, if $v \in \{-1, -2, ..., -n\}$, the weights on the edges correspond to the first column of the matrix X_n. Fix $v = -1$, the vector is

$$(X_{11}, X_{21}, ..., X_{n1}) = \left(\frac{U_{11}}{\rho_1}, \frac{U_{21}}{\rho_2}, ..., \frac{U_{n1}}{\rho_n} \right).$$

The setting here is slightly different, indeed the first column is a vector of i.i.d. variables. We will prove its convergence to a function of a Poisson point process.

Proposition 2.1.3. *Let $\hat{X}^{(n)}$ be the vector of the order statistic of $(X_{11}, ..., X_{n1})$, then*

$$\hat{X}^{(n)} \xrightarrow[n\to+\infty]{d} \left(\frac{\xi_1}{c(\alpha)+\xi_1}, \frac{\xi_2}{c(\alpha)+\xi_2}, ... \right), \tag{2.1.2}$$

where $\{\xi_i\}_{i\geq 1}$ is an ordered Poisson Process of intensity $\alpha x^{-\alpha-1}dx$, and $c(\alpha)$ is an absolute constant depending only on α,

$$c(\alpha) = \int_0^{+\infty} s^{-\alpha} \mu(ds),$$

where μ is the law of the one sided stable distribution of index α.

To prove Proposition 2.1.3, we will use the next result from [21], for the proof we refer to [21, Theorem 5.3].

Lemma 2.1.4 (Poisson point process with Radon intensity measure)**.** *Let $\xi_1^n, \xi_2^n, ...$ be sequences of i.i.d. random variables on $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ such that*

$$n\mathbb{P}\left(\xi_1^n \in \cdot\right) \xrightarrow[n\to+\infty]{(w)} \nu, \tag{2.1.3}$$

where ν is a Radon measure in $\mathbb{R}$. Then for any finite set $I \subset \mathbb{N}$ the random measure

$$\sum_{i\in\{1,...,n\}\setminus I} \delta_{\xi_i^n}$$

converges weakly as $n \to +\infty$ to $PPP(\nu)$, the Poisson point process on $\mathbb{R}$ with intensity law ν, for the usual vague topology on Radon measure.

For the notion of vague convergence we refer to [21, Section 3.5].

Lemma 2.1.5. *Let A be a nonnegative random variable in $\mathbb{H}_\alpha^*$. Define $B(n) = a_n^{-1} \sum_{i=1}^n A_i$, where A_i are i.i.d. copies of A, independent of A. Then there exists an absolute constant $c(\alpha) > 0$ depending only on α, such that for any $t > 0$*

$$\lim_{n\to\infty} n\mathbb{P}\left(A > ta_n B(n)\right) = c(\alpha) t^{-\alpha}, \qquad t > 0 \tag{2.1.4}$$

Proof. Let μ_n denote the law of $B(n)$, so that, for any fixed $t > 0$ one has

$$n\mathbb{P}\left(A > ta_n B(n)\right) = \int_0^\infty n\mathbb{P}\left(A > sta_n\right) \mu_n(ds) = c^{-1} t^{-\alpha} \int_0^\infty s^{-\alpha} L(a_n ts) \mu_n(ds), \tag{2.1.5}$$

since, as the law of the entries is in $\mathbb{H}_\alpha^*$, a_n of equation (1.3.1) is of the form $a_n = c^{1/\alpha} n^{1/\alpha}$. Recall that μ_n converges weakly to μ, the law of $B := \sum_{i=1}^\infty \gamma_i$, where $\{\gamma_i\}$ denotes the Poisson point process of intensity $\alpha x^{-\alpha-1}dx$ on the interval $[0, \infty)$, so that μ is the law of the one sided stable distribution of index α. Thus, we need to show that for any fixed $t > 0$:

$$\lim_{n\to\infty} c^{-1} \int_0^\infty s^{-\alpha} L(a_n ts)\mu_n(ds) = c_1, \tag{2.1.6}$$

for some positive constant c_1.

We fix $\varepsilon > 0$ and start by showing that

$$\lim_{n\to\infty} c^{-1} \int_\varepsilon^\infty s^{-\alpha} L(a_n ts)\mu_n(ds) = \int_\varepsilon^\infty s^{-\alpha}\mu(ds), \tag{2.1.7}$$

Indeed, $|c^{-1}L(a_n ts) - 1| \leq u(n)$, for some function $u(n) \to 0$ as $n \to \infty$ uniformly in $s \geq \varepsilon$, and $\int_\varepsilon^\infty s^{-\alpha}\mu_n(ds) \to \int_\varepsilon^\infty s^{-\alpha}\mu(ds)$ by weak convergence. This implies (2.1.7).

Next, fix some constant K so large that $1/2 \leq c^{-1}L(x) \leq 2$ for all $x \geq K(t \wedge 1)$. Let us show that

$$\lim_{n\to\infty} \int_0^{Ka_n^{-1}} s^{-\alpha} L(a_n ts)\mu_n(ds) = 0. \tag{2.1.8}$$

This follows from the

$$L(a_n ts) = s^\alpha t^\alpha a_n^\alpha \mathbb{P}\left(A > a_n ts\right) \leq s^\alpha t^\alpha a_n^\alpha.$$

Indeed, with this bound one has

$$\int_0^{Ka_n^{-1}} s^{-\alpha} L(a_n ts)\mu_n(ds) \leq t^\alpha a_n^\alpha \int_0^{Ka_n^{-1}} \mu_n(ds). \tag{2.1.9}$$

Since $\int_0^{Ka_n^{-1}} \mu_n(ds)$ is the probability that $B(n) \leq Ka_n^{-1}$ one has

$$\mathbb{P}\left(B(n) \leq Ka_n^{-1}\right) = \mathbb{P}\left(a_n^{-1}\sum_{i=1}^n A_i \leq Ka_n^{-1}\right) \leq \mathbb{P}\left(\max_{i=1,\ldots,n} A_i \leq K\right) = \mathbb{P}\left(A \leq K\right)^n. \tag{2.1.10}$$

From the definition of K it follows that $\mathbb{P}\left(A < K\right) \leq 1-u$, for $u = u(K) = (c/2)^\alpha K^{-\alpha} > 0$. Thus $\int_0^{Ka_n^{-1}} \mu_n(ds) \leq (1-u)^n$ decays exponentially to zero. This, together with (2.1.9), implies (2.1.8).

Finally, we need to consider the integral

$$\mathcal{I}(\varepsilon, n) := \int_{Ka_n^{-1}}^\varepsilon s^{-\alpha}\mu_n(ds). \tag{2.1.11}$$

Since $s \geq K/a_n$ here one has $c^{-1}L(a_n ts) \leq 2$ by definition of K. Therefore

$$\mathcal{I}(\varepsilon, n) \leq 2\int_{Ka_n^{-1}}^\varepsilon s^{-\alpha}\mu_n(ds)$$

We want to show that

$$\limsup_{n\to\infty} \int_{Ka_n^{-1}}^\varepsilon s^{-\alpha}\mu_n(ds) \leq q(\varepsilon)\,, \tag{2.1.12}$$

for some $q(\varepsilon) \to 0$ as $\varepsilon \to 0$.

Let Z denote the one-sided stable distribution with index α. As in Lemma 3.5 of [10] one has that there exist $\eta > 0$ and $p \in (0,1)$ such that A dominates stochastically ηDZ where D is a Bernoulli(p) variable independent of Z. Thus $B(n)$ stochastically dominates $\bar{B}(n) := a_n^{-1}(D_1 Z_1 + \cdots D_n Z_n)$ where D_i are iid copies of D and Z_i are iid copies of Z. If $\bar{\mu}_n$ denotes the law of $\bar{B}(n)$ then it is given by the mixture $\bar{\mu}_n = \sum_{k=0}^n p(k,n)\nu_k$ where $p(k,n) = \binom{n}{k}p^k(1-p)^{n-k}$, and ν_k is the law of $a_n^{-1}(Z_1 + \cdots + Z_k)$. Recall that $Z_1 + \cdots + Z_k$ has the same law of $k^{1/\alpha}Z$, so that ν_k is the law of $\left(\frac{k}{n}\right)^{1/\ a} Z$. Let E be the event that $D_1 + \cdots + D_n \geq pn/2$. We can therefore estimate

$$\int_{Ka_n^{-1}}^{\varepsilon} s^{-\alpha}\mu_n(ds) \leq (K/a_n)^{-\alpha}\mathbb{P}\left(E^c\right) + \sum_{k=pn/2}^{n} p(k,n) \int_{Ka_n^{-1}}^{\varepsilon} s^{-\alpha}\nu_k(ds). \tag{2.1.13}$$

$\mathbb{P}\left(E^c\right)$ decays to zero exponentially in n by Chernoff bound,

$$\mathbb{P}\left(E^c\right) = \mathbb{P}\left(D_1 + \cdots + D_n \leq \frac{pn}{2}\right), \tag{2.1.14}$$

see e.g. [18] .Thus the first term above vanishes in the limit $n \to \infty$. We now consider the second term. For any $k \in [pn/2, n]$ one has that ν_k is the law of λZ, for a constant $\lambda = (k/n)^{1/\alpha}$ bounded above and below uniformly in n, since $\lambda \in [(p/2)^{1/\alpha}, 1]$. If ν_k is the law of λZ as above and ν denotes the law of Z then using a change of variables one has

$$\limsup_{n\to\infty} \sup_{k\in[pn/2,n]} \int_{Ka_n^{-1}}^{\varepsilon} s^{-\alpha}\nu_k(ds) \leq \sup_{\lambda\in[(p/2)^{1/\alpha},1]} \lambda^{-\alpha} \int_0^{\varepsilon/\lambda} x^{-\alpha}\nu(dx) \leq q(\varepsilon),$$

for some $q(\varepsilon) \to 0$. Here we are using the easily established fact that $\int_0^{\varepsilon} x^{-\alpha}\nu(dx) \to 0$ as $\varepsilon \to 0$ (this follows because $\nu(0,x]$ behaves as $e^{-x^{-1/\alpha}}$ for small x). Inserting this in the previous expressions, we have obtained

$$\limsup_{n\to\infty} \int_{Ka_n^{-1}}^{\varepsilon} s^{-\alpha}\mu_n(ds) \leq \mathbb{P}\left(E\right) q(\varepsilon) \leq q(\varepsilon)\,. \tag{2.1.15}$$

This ends the proof of (2.1.12).

Putting together (2.1.7),(2.1.8) and (2.1.12), we have shown that for any fixed $t > 0$ one has

$$\lim_{n\to\infty} nt^{\alpha}\mathbb{P}\left(A > ta_n B(n)\right) = c(\alpha) = \int_0^{\infty} s^{-\alpha}\mu(ds)\,. \tag{2.1.16}$$

□

We now have all the ingredients for the proof of Proposition 2.1.3.

Proof of Proposition 2.1.3. The j-th column vector has form

$$\left(\frac{U_{ij}}{\rho_i}\right)_{i=1}^n = \left(\frac{U_{ij}}{U_{i1}+U_{i2}+\cdots+U_{in}}\right)_{i=1}^n$$

for i.i.d. variables $U_{i1},...,U_{in} \in \mathbb{H}_\alpha^*$, for any $i = 1,...,n$. Fix $j = 1$. We can rearrange the vector as

$$\left(\frac{U_{i1}}{U_{i1}+U_{i2}+\cdots+U_{in}}\right)_{i=1}^n = \left(\frac{1}{\left(1+\frac{U_{i2}+\cdots+U_{in}}{U_{i1}}\right)}\right)_{i=1}^n .$$

Let us now focus on the vector $\left(\frac{U_{i1}}{U_{i2}+\cdots+U_{in}}\right)_{i=1}^n$. Define $B_n^{(i)} := a_n^{-1}(U_{i2}+\cdots+U_{in})$. By lemma 2.1.5, the ordered statistics of the vector

$$\left(\frac{U_{i1}}{U_{i2}+\cdots+U_{in}}\right)_{i=1}^n = \left(\frac{U_i}{a_n B_n^{(i)}}\right)_{i=1}^n ,$$

verify condition (2.1.3) of Lemma 2.1.4, for $\left(\xi_i^{(n)}\right)_{i\geq 1} := \left(\frac{U_i}{a_n B_i^{(n)}}\right)_{i\geq 1}$, with $\nu(t,+\infty) = c(\alpha)t^{-\alpha}$, where $c(\alpha)$ is an absolute constant depending only on α, and as a consequence, the ordered statistics converge to the vector

$$\left(\frac{\gamma_1}{c(\alpha)}, \frac{\gamma_2}{c(\alpha)}, ...\right), \tag{2.1.17}$$

where $\{\gamma_i\}_{i\geq 1}$ is the ordered Poisson process of intensity $\alpha x^{-\alpha-1}$. Thus if $(V_1,...,V_n)$ is the vector of the ordered statistics of $\left(\frac{U_{i1}}{a_n B_n^{(i)}}\right)_{i=1}^n$, the vector of the ordered statistics of $\left(\frac{U_{i1}}{\rho_i}\right)_{i=1}^n$ is

$$\left(\frac{1}{1+\frac{1}{V_i}}\right)_{i=1}^n .$$

by monotonicity and continuity for $x \in (0,+\infty)$ of the function,

$$x \longmapsto \frac{1}{1+\frac{1}{x}}$$

follows the statement. □

For the row vectors one has a result of almost sure uniform square integrability, as proved in [9, Lemma 2.4]. We prove an analogous result for the column vector.

Lemma 2.1.6. *Consider the family of i.i.d. random variables in $\mathbb{H}_\alpha^*$, $\{A_{ij}\}_{i,j\geq 1}$, and define for every $n \geq 1$, $B_n^{(i)} := a_n^{-1}(A_{i2}+\cdots+A_{in})$. Then,*

$$\lim_{k\to+\infty} \sup_n \sum_{j=k+1}^n \left(\frac{A_{j1}}{a_n B_n^{(j)}}\right)^2 = 0 \tag{2.1.18}$$

Proof. By proposition 2.1.3, if $G(t) := \mathbb{P}\left(\frac{A_{11}}{a_n B_n^{(1)}} > t\right)$, we have that

$$nG(a_n t) \xrightarrow[n\to+\infty]{} c(\alpha)t^{-\alpha}. \tag{2.1.19}$$

Define $G^{-1}(y) = \inf\{w > 0 : G(w) \leq y\}$ for $y \in (0,1)$. If $\hat{V} = (\hat{V}_1, ..., \hat{V}_n)$ is the vector of the ordered statistics of the vector $\left(\frac{A_{11}}{a_n B_n^{(1)}}, \frac{A_{21}}{a_n B_n^{(2)}}, ..., \frac{A_{n1}}{a_n B_n^{(n)}}\right)$, then following the approach of [17],

$$\hat{V} \stackrel{d}{=} \left(G^{-1}\left(\frac{\gamma_1}{\gamma_{n+1}}\right), G^{-1}\left(\frac{\gamma_2}{\gamma_{n+1}}\right), ..., G^{-1}\left(\frac{\gamma_n}{\gamma_{n+1}}\right)\right), \tag{2.1.20}$$

where $\gamma_k = \sum_{i=1}^k E_i$, and $\{E_i\}_{i\geq 1}$ is a collection of i.i.d. exponential random variable with mean 1. Now, by equation (2.1.4) for any $\delta > 0$ we can find an integer n_0 such that

$$a_n^{-1} Y_k = a_n^{-1} G^{-1}(\gamma_k/\gamma_{n+1}) \leq (n\gamma_k/(1+\delta)\gamma_{n+1})^{-1/\alpha}$$

for $n \geq n_0$. Since $n/\gamma_{n+1} \to 1$, a.s. the expression above is a.s. bounded by $2(1+\delta)^{1/\alpha}\gamma_k^{-\alpha}$, the claim follows by the a.s. summability of $\gamma_k^{-1/\alpha}$. □

2.1.3 Local convergence to a modified PWIT

We now investigate the limiting local structure of the bipartite graph identified by B_n. Define the weighted rooted network (G_n, v), induced by the weights given by the matrix weights B_n, obtained by distinguishing the vertex labeled v. For any $L, P \in \mathbb{N}$ such that $1 + L + \cdots + L^P \leq n$, we want to define $(G_n, v)^{L,P}$ a subnetwork of (G_n, v), which vertex set is identified with the vertex set of a L-ary tree of depth H, rooted in v.

Fix $v = 1$. For any fixed realization of the marks $\{U_{ij}\}_{i,j=1}^n$, we partially order the vertices of $(G_n, 1)$ as elements of

$$J_{L,P} = \bigcup_{k=0}^{P} \{1, 2, ..., L\}^k \subset \mathbb{N}^f.$$

The indices will be given by a map

$$\sigma_n : J_{L,P} \to V_n \tag{2.1.21}$$

Set $I_\varnothing = \{1\}$, and the index of the root 1 is $\sigma_n^{-1}(1) = \varnothing$. To a vertex v in $V_n \setminus I_\varnothing$ is given the index (k) if it is the k-th largest value in $\mathcal{N}_1 = \{U_{j,1} : j \neq 1\}_{j=1}^n$, the neighborhood of 1, for $1 \leq k \leq L$. This defines the first generation. Define I_1 as the union of $I_\varnothing$ and the L selected vertices. If $P \geq 2$, we repeat the indexing procedure, starting from (1), the first child of $\varnothing$, on the set $V_n \setminus I_1$, We obtain a new set indexed $\{11, ..., 1L\}$. Define I_{11} as the union of I_1 and the L selected vertices. This procedure is repeated until depth P, when $(L^{P+1}-1)/(L-1)$ vertices are indexed. Call this set of vertices $V_n^{L,P} = \sigma_n J_{L,P}$. Even though $V_n^{L,P}$ has the structure of the vertex set of a tree, Its edges set still hascircuits and loops. In the next proposition we prove that in the limit it actually converges to a tree. In the sense that all the edges $\{u,v\} \in J_{L,P} \times J_{L,P}$ that do not belong to the tree vanish. For the sake of clarity, define

$$E^{L,P} = \{\{u,v\} \in J_{L,P} \times J_{L,P} : \nexists k \in \{1,...,L\} \text{ such that } u = vk \text{ or } v = uk\}. \quad (2.1.22)$$

Proposition 2.1.7. *The edges in $E^{L,P}$ vanish in the limit, if*

$$n\mathbb{P}\left(\frac{U_{12}}{a_n} \in \cdot\right) \xrightarrow[n\to+\infty]{(w)} \nu,$$

where ν is a Radon measure on $\mathbb{R}$ with no mass at 0.

We need the next lemma for the proof.

Lemma 2.1.8. *Let $E^{L,P}$ be as in* (2.1.22)*, then the vector*

$$\left\{\frac{U_{i,j}}{a_n} : \{i,j\} \in E^{L,P}\right\},$$

is stochastically dominated by i.i.d. random variables distributed as $U_{1,2}/a_n$.

Proof. The claim is implied by the following. Let $Y_1, ..., Y_m$ be i.i.d positive random variables. And let $m = n_1 + \cdots + n_p$, for positive $p, n_1, .., n_p$, so that the variables are divided into k blocks $I_1, ..., I_p$, such that for any $j = 1, .., p$, $|I_j| = n_j$. Fix now some integers $0 \le \kappa_j \le n_j$ for any j, and call $q_1^j, ..., q_{k_j}^j$, the random indices of the k_j largest values of the variables of the j-th block. Call this random set of indices J^j, so that $J^j = \{q_1^j, ..., q_{k_j}^j\}$. Call $J = \cup_{i=1}^p J^i$, where $J^i = \emptyset$ if $k_i = 0$. Define $\widetilde{Y} = \{Y_i : i \notin J\}$. We now prove that $\widetilde{Y}$ is stochastically dominated by $m - |J|$ i.i.d. copies of Y_1.

Construct the following coupling. Extract a realization $y_1, ..., y_m$ of $Y_1, ..., Y_m$, and isolate the m blocks. Consider the vector

$$\mathcal{Z} = (z_1^1, ..., z_{n_1-k_1}^1, z_1^2, ..., z_{n_2-k_2}^2, ..., z_{n_p-k_p}^p),$$

obtained by extracting uniformly at random $n_1 - k_1$ values from $y_1, ..., y_{n_1}$, $n_2 - k_2$ form $y_{n_1+1}, ..., y_{n_2}$ and so on. Now we construct the vector $\mathcal{V}$. For the first block take $v_i^1 = z_i^1$, for any $i = 1, ..., n_1 - k_1$, whenever $i \in I_1 \setminus J_1$ was picked for the vector $\mathcal{Z}$. Assign the remaining values through an independent uniform permutation of the variables $y_i, i \in I_i \setminus J^i$, not picked for the vector $\mathcal{Z}$. Repeat this procedure for any block. By construction, coordinate-wise $\mathcal{Z} \ge \mathcal{V}$. The proof ends noticing that $\mathcal{V}$ is distributed as $\widetilde{Y}$, while $\mathcal{Z}$ is distributed as a vector of $m - |J|$ i.i.d. copies of Y_1. $\square$

proof of Proposition 2.1.7. By lemma 2.1.8, if $(U_1, ..., U_{|E^{L,P}|})$ is a vector of i.i.d. variables with the same law as $U_{i,j}$, then by union bound,

$$\begin{aligned}\mathbb{P}\left(\max_{\{i,j\}\in E^{L,P}} \frac{U_{i,j}}{a_n} > t\right) &\leq \mathbb{P}\left(\max_{k=1,...,|E^{L,P}|} \frac{U_k}{a_n} > t\right) \\ &\leq |E^{L,P}|\mathbb{P}\left(\frac{U_1}{a_n} > t\right) \\ &\leq \frac{|E^{L,P}|}{n} n\mathbb{P}\left(\frac{U_1}{a_n} > t\right) \xrightarrow[n\to+\infty]{} 0\end{aligned}$$

By settings $n\mathbb{P}\left(U_1 > ta_n\right) = nG(a_n t) \xrightarrow[n\to+\infty]{} t^{-\alpha}$, then for any $t > 0$,

$$\mathbb{P}\left(\max_{\{i,j\}\in E^{L,P}} \frac{U_{i,j}}{a_n} > t\right) \leq t^{-\alpha}\frac{|E^{L,P}|}{n} \xrightarrow[n\to+\infty]{} 0.$$

□

We now observe that, ρ_i stochastically dominates ρ_1, which normalized by a_n, converges in distribution to a positive random variable (the one-sided stable law of index α). Then proposition 2.1.7, implies

$$\mathbb{P}\left(\max_{\{i,j\}\in E^{L,P}} \frac{U_{i,j}}{\rho_i} > t\right) \xrightarrow[n\to+\infty]{} 0.$$

The distribution of the limiting tree depends on the vertex picked to be the root. Consider the special case $P = 1$, namely we only look at the first generation. For any $1 \leq L \leq n-1$, the rooted network $(G_n, 1)^{L,1}$ converges to a tree of depth 1, as proved in proposition 2.1.7. The weights on the edges converge to the first L maxima of a $PD(\alpha)$. If we map $v = -1$ into the root, the weights will converge to the first L maxima of the vector (2.1.2), see proposition 2.1.3.

We generalize the convergence in the next proposition, but first let us define the limiting objects. Consider a realization of PWIT$(\alpha x^{-\alpha-1}dx)$. For any $v \in \mathbb{N}^f$, call $d(v) = \text{dist}(\varnothing, v)$, the graph distance of v from $\varnothing$ on the infinite tree, and denote y_{vk} the mark from v to vk, for $k \in \mathbb{N}$. Define,

$$\rho(v) := \frac{y_v}{c(\alpha) + y_v} + \sum_{i\geq 1} y_{vi}$$

if $v = \varnothing$, set $\rho(\varnothing) := \sum_{i\geq 1} y_i$. We define two random operators on $\mathcal{D}(\mathbb{N}^f)$, for any $u \in \mathbb{N}^f$, and $k \in \mathbb{N}$, as

$$\langle \delta_u, \mathcal{A}_+\delta_{uk}\rangle = \langle \delta_{uk}, \mathcal{A}_+\delta_u\rangle = \begin{cases} \dfrac{y_{uk}}{\rho(u)} & \text{if } d(u) \equiv 0 (\text{mod } 2) \\ \\ \dfrac{y_{uk}}{c(\alpha) + y_{uk}} & \text{if } d(u) \equiv 1 (\text{mod } 2) \\ \\ 0 & \text{otherwise} \end{cases} \tag{2.1.23}$$

$$\langle \delta_u, \mathcal{A}_- \delta_{uk} \rangle = \langle \delta_{uk}, \mathcal{A}_- \delta_u \rangle = \begin{cases} \dfrac{y_{uk}}{c(\alpha) + y_{uk}} & \text{if } d(u) \equiv 0 (\text{mod } 2) \\ \\ \dfrac{y_{uk}}{\rho(u)} & \text{if } d(u) \equiv 1 (\text{mod } 2) \\ \\ 0 & \text{otherwise} \end{cases} \tag{2.1.24}$$

where $c(\alpha)$, is an absolute constant depending only on α, whose explicit representation is

$$c(\alpha) = \int_0^{+\infty} x^{-\alpha} \mu(dx), \tag{2.1.25}$$

where μ is the law of the one sided stable distribution of index α, see proof of Lemma 2.1.5.

Proposition 2.1.9 (Local weak convergence to a tree). *Let G_n be the complete network on $\{-1, ..., -n\} \cup \{1, ..., n\}$, whose mark on edge (i, j) equals $(B_n)_{i,j}$. Then for all integers L, P as $n \to \infty$, in distribution,*

$$(G_n, 1)^{L,P} \to (\mathcal{A}_+, \varnothing)^{L,P} \qquad (G_n, -1)^{L,P} \to (\mathcal{A}_-, \varnothing)^{L,P},$$

Where $(\mathcal{A}_\pm, \varnothing)^{L,P}$ is the subtree with vertices in $V^{L,P}$ and marks inherited from the infinite tree.

Proof. Let us focus on the convergence of $(G_n, 1)^{L,P} \xrightarrow[n \to +\infty]{loc} (\mathcal{A}_+, \varnothing)^{L,P}$. By proposition 2.1.7, of $(G_n, 1)^{L,P}$ converges to a L-ary tree of depth P. We now focus on the convergence of the marks. We order the element of $J_{L,P}$ in lexicographic order, i.e. $\varnothing \prec 1 \prec \cdots \prec B \cdots B$. For $v \in J_{L,P}$, let $\mathcal{O}_v$ denote the set of the offspring of v in $(G_n, 1)^{L,P}$. By construction $I_\varnothing = \{1\}$ and $I_v = \sigma_n(\cup_{w \prec v} \mathcal{O}_w)$. The indexing procedure at every step sorts the marks of neighboring edges that have not been explored at an earlier step, then by construction the offspring of the tree are independent. Thus the marks from a parent to his offspring in $(G_n, 1)^{L,P}$ behave as independent vectors. Namely if the parent is at even distance from the root (0 is even), then the marks converges to a Poisson Dirichlet process, otherwise if the parent is at odd distance from the root, the marks to its offsprings behave as (2.1.2). Thus the marks converge weakly to those in $(\mathcal{A}_+, \varnothing)^{L,P}$. $\square$

We now improve the convergence to the infinite tree. To prove the local weak convergence we extend B_n to an operator on $\mathcal{D}(\mathbb{N}^f)$ setting $\langle \delta_i, B_n \delta_j \rangle = B_{i,j}$, for $i, j \in \{-1, ..., -n\} \cup \{1, ..., n\}$ otherwise 0. The proof is a compound of the results of [9], and [10], due to the double nature of the limiting operators.

Theorem 2.1.10. *Let B_n be the bipartized version of X_n, $\mathcal{A}_+$ as in* (2.1.23), *and $\mathcal{A}_-$ as in* (2.1.24) *then*

$$(B_n, 1) \xrightarrow[n\to+\infty]{loc} (\mathcal{A}_+, \varnothing) \qquad (B_n, -1) \xrightarrow[n\to+\infty]{loc} (\mathcal{A}_-, \varnothing)$$

Proof. We again focus on the convergence to $\mathcal{A}_+$. By proposition 2.1.9, for any $L, P \in \mathbb{N}$ such that $1 + L + \cdots + L^P \leq n$, the rooted network $(G_n, 1)^{L,P}$ locally converges,

$$(G_n, 1)^{L,P} \xrightarrow[n\to+\infty]{loc} (\mathcal{A}_+, \varnothing)^{L,P}, \tag{2.1.26}$$

for$\mathcal{A}_+$ as in (2.1.23). If $\sigma_n^{L,P}$ is the bijection defined in (2.1.21), we can extend $\sigma_n^{L,P}$ to a bijection on the whole set $\mathbb{N}^f$. By Skorohod representation theorem we may assume that (2.1.26) is an a.s. convergence, and we can find diverging sequences L_n and P_n and a sequence of bijections $\widetilde{\sigma}_n := \sigma_n^{L_n,P_n}$, such that $1 + L_n + \cdots + L_n^{P_n} \leq 2n$ and such that $U_{\widetilde{\sigma}_n(u),\widetilde{\sigma}_n(v)} \xrightarrow[n\to+\infty]{} \langle \delta_u, \mathcal{A}_+\delta_v \rangle$, if $v = vk$, 0 otherwise. Now, for any $v \in \mathbb{N}^f$, we have to prove

$$\sum_u \left(\left\langle \delta_u, (\widetilde{\sigma}_n^{-1} B_n \widetilde{\sigma}_n)\delta_v \right\rangle - \langle \delta_u, \mathcal{A}_+\delta_v \rangle \right)^2 \xrightarrow[n\to+\infty]{} 0. \tag{2.1.27}$$

We already proved that, for any u, $U_{\widetilde{\sigma}_n(u),\widetilde{\sigma}_n(v)} \xrightarrow[n\to+\infty]{} \langle \delta_u, \mathcal{A}_+\delta_v \rangle$, plus we have the uniform square integrability results from [9, Lemma 2.4], for the row-type generation, and Lemma 2.1.6, for the colum-type generation, up to noticing that, for every i,

$$\lim_{k\to\infty} \sup_n \sum_{j=k+1}^{n} \frac{U_{ji}}{U_{j1} + U_{j2} + \cdots U_{jn}} \leq \lim_{k\to\infty} \sup_n \sum_{j=k+1}^{n} \frac{U_{ji}}{a_n B_n^{(j)}},$$

where $B_n^{(i)} = (U_{i2} + \cdots + U_{in})$.

□

Chapter 3

Singular values

The aim of this chapter is to prove the convergence of the empirical spectral distribution of the singular values of the matrix $X_n - z\mathbb{1}$, via theorem 2.1.2. We also prove a result for the finiteness of the exponential moment of the limiting spectral distribution of the singular values.

3.1 Resolvent convergence

In order to apply theorem 2.1.2, we shall check the self-adjointeness of our bipartized operators. Our operators are symmetric, densely-defined but unbounded, we are then interested in their unique self-adjoint extension, which, with a slight abuse of notation, we will denote as the operators, once we prove they are essentially self-adjoint.

3.1.1 Self-adjointness of limit operators

Equations (2.1.23) and (2.1.24) define two a.s. self adjoint operators, as proved in the next proposition.

Proposition 3.1.1 (Self-adjointness of limit operators)**.** *For any $z \in \mathbb{C}$ the operators $\mathcal{A}_+$ and $\mathcal{A}_-$ defined in* (2.1.23) *and* (2.1.24) *are essentially self-adjoint with probability one.*

In order to prove proposition 3.1.1, we recall two lemmas from [10].

Lemma 3.1.2. *Let $V = \mathbb{N}^f$ denote the vertex set of the PWIT, and let $\mathcal{D}$ be the space of the finitely supported vectors. Write $u \sim v$ if $u = vk$ or $v = uk$ for some $k \in \mathbb{N}$. Let $A : \mathcal{D} \to \ell^2(V)$ denote the symmetric linear operatore defined by*

$$\langle \delta_u, A\delta_v \rangle = w_{u,v} = \bar{w}_{v,u},$$

and such that $w_{u,v} = 0$ *whenever* u *and* v *are not neighbors. Suppose there exists a constant* $\kappa > 0$ *and a sequence of connected finite subsets* $(S_n)_{n\geq 1}$ *in* V*, such that* $S_n \subset S_{n+1}$*,* $\bigcup_n S_n = V$ *and for every* $v \in S_n$*,*

$$\sum_{u \notin S_n : u \sim v} |w_{uv}|^2 < \kappa. \tag{3.1.1}$$

Then A *is essentially self-adjoint.*

Lemma 3.1.3. *Let* $\kappa > 0$*,* $0 < \alpha < 2$ *and let* $0 < x_1 < x_2 < \cdots$ *be a Poisson Process of intensity 1 on* $(0, +\infty)$*. Define* $\tau = \inf\{t \in \mathbb{N} : \sum_{k=t+1}^{\infty} x_k^{-2/\alpha} \leq \kappa\}$*. Then* $\mathbb{E}[\tau]$ *is finite for any* κ *and it goes to 0 as* κ *goes to* ∞*.*

Proof of Proposition 3.1.1. We prove the proposition for $\mathcal{A}_+$, through a stochastic domination argument. For any realization of $\mathcal{A}_+$ define $\mathbb{A}$ as

$$\langle \delta_u, \mathbb{A}\delta_{uk} \rangle = \langle \delta_{uk}, \mathbb{A}\delta_u \rangle = \begin{cases} \dfrac{y_{uk}}{\sum_{i\geq 1} y_{ui}} & \text{if } d(u) \equiv 0 (\text{mod } 2) \\[2ex] \dfrac{y_{uk}}{c(\alpha)} & \text{if } d(u) \equiv 1 (\text{mod } 2) \\[2ex] 0 & \text{otherwise} \end{cases} \tag{3.1.2}$$

It follows, for any $u, v \in \mathbb{N}^f$

$$\mathbb{P}\left(\mathbb{A}(u, v) \geq \mathcal{A}_+(u, v)\right) = 1, \tag{3.1.3}$$

where $\mathbb{A}(u, v) = \langle \delta_u, \mathbb{A}, \delta_v \rangle$, and $c(\alpha)$ is the absolute constant depending only on α defined in equation (2.1.25). Since for any index $i \in \mathbb{N}^f$, $y_i \geq 0$, we have, with probability one, for any $u, v \in \mathbb{N}^f$ such that $d(u) \equiv 0(\text{mod } 2)$ and $d(v) \equiv 1(\text{mod } 2)$,

$$\frac{y_{uk}}{y_u + \sum_{j\geq 1} y_{uj}} \leq \frac{y_{uk}}{\sum_{j\geq 1} y_{uj}} \qquad \text{and} \qquad \frac{y_{vk}}{c(\alpha) + y_{vk}} \leq \frac{y_{vk}}{c(\alpha)}.$$

Thus $\mathbb{A}$ stochastically dominates $\mathcal{A}_+$. The condition (3.1.1) concerns the finiteness of the sum of the operator weights on a collection of subsets of the vertices of the tree. By (3.1.3), if the sum is finite for $\mathbb{A}$ then it is finite for $\mathcal{A}_+$.

For any $v \in V^+$ define

$$\tau_v^+(\kappa) = \inf\left\{ t \geq 0 : \sum_{p=t+1}^{\infty} y_{vp}^2 \leq \kappa \right\}.$$

By construction $(\tau_v^+(\kappa))_{v \in V^+}$ is a collection of i.i.d. variables. Fix now $\bar{\kappa}$ such that $\mathbb{E}[\tau_v^+(\kappa)] < 1$. Such a κ exists by Lemma 3.1.3.

Similarly for any $v \in V^-$ define $\tau_v^-(\kappa)$. The collection $(\tau_v^-(\kappa))_{v\in V^-}$ is again i.i.d. by construction. Since the Poisson-Dirichlet generations sum up to 1, it holds and analogous of Lemma 3.1.3, and we can again fix a $\bar{\kappa}$ such that $\mathbb{E}\left[\tau_v^-(\kappa)\right] < 1$. Call now

$$\tau = \min_{v\in V^+, u\in V^-} \{\tau_v^+(\kappa), \tau_u^-(\kappa)\}.$$

The proof ends as in [10, Proposition 2.8]. Consider an i.i.d. collection of variables indexed by the vertices set V, distributed as τ, $\{t_v\}_{v\in V}$. Fix the κ such that $\mathbb{E}\left[\tau_v\right] < 1$. Now put a green mark on the on all vertices such that $\tau_v \geq 1$, and red otherwise. Define the subforest T^g of T, where we put an edge between v and vk if v is green and $1 \leq k \leq \tau_v$. Then if the root is red, we set $S_1 = \{\varnothing\}$, if the root is green we consider $T^g_\varnothing = (V^g_\varnothing, E^g_\varnothing)$, the subtree of $T^g_\varnothing$ containing the root. Due to the choice of κ, $T^g_\varnothing$ is almost sure finite. Consider $L^g_\varnothing$, the set of the leaves of $T^g_\varnothing$, namely the set of the vertices in $V^g_\varnothing$ such that for all $1 \leq k \leq \tau_v$, vk is red. Set $S_1 = V^g_\varnothing \bigcup_{v\in L^g_\varnothing} \{vk : 1 \leq k \leq \tau_v\}$. Now we define the outer boundary of $\{\varnothing\}$ as $\partial_\tau\{\varnothing\} = \{1, ..., \tau_\varnothing\}$ and for $v = (i_1, ..., i_k) \in \mathbb{N}^f \setminus \{\varnothing\}$ we set $\partial_\tau\{v\} = \{(i_1, ..., i_{k-1}, i_{k+1})\} \cup \{(i_1, ..., i_k, 1), ..., (i_1, ..., i_k, \tau_v)\}$. For a connected set S the outer boundary is

$$\partial_\tau S = \left(\bigcup_{v\in S} \partial_\tau\{v\}\right) \setminus S.$$

Now for each vertex $u_1, ..., u_k \in \partial_\tau S_1$, we repeat the above procedure to the rooted subtrees $T_{u_1}, ..., T_{u_k}$. We set $S_2 = S_1 \bigcup_{1\leq i\leq k} \bigcup_{v\in L^g_{u_i}} \{vk : 1 \leq k \leq \tau_v\}$. Iteratively, we may thus define an increasing connected sequence (S_n) of vertices with the properties required for corollary 3.1.2. □

3.2 Singular values

In this section we give the proof of the weak convergence of the empirical spectral distribution. Due to concentration properties, it will be sufficient to prove the convergence of the expectation of the random measure $\nu_{X_n - z\mathbb{1}}$.

3.2.1 Proof Theorem 1.6.2

Recall the structure of the matrix

$$X_n = (X_{ij})_{i,j=1}^n = \left(\frac{U_{ij}}{\sum_j U_{ij}}\right)_{i,j=1}^n = \left(\frac{U_{ij}}{\rho_i}\right)_{i,j=1}^n.$$

where $\{U_{i,j}\}_{i,j=1}^n$ is a i.i.d. collection of variables with law in $\mathbb{H}^*_\alpha$.

Proof of theorem 1.6.2. Rearranging the entries of the matrix

$$B(z) = \left(X_n - \begin{pmatrix} 0 & z \\ \bar{z} & 0 \end{pmatrix} \otimes \mathbb{1}_n \right),$$

we note that it is similar to,

$$\begin{pmatrix} 0 & (X_n - z\mathbb{1}) \\ (X_n - z\mathbb{1})^* & 0 \end{pmatrix}.$$

The spectrum of this matrix is $\{\pm\sigma_1(X_n - z\mathbb{1}), ..., \pm\sigma_n(X_n - z\mathbb{1})\}$, as already observed in [10, Theorem 2.1], then $\mu_{B_n(z)} = \check{\nu}_{X_n - z\mathbb{1}}$. Define

$$R_n(U)_{kk} = ((B_n - U(z,\eta) \otimes \mathbb{1}_n)^{-1})_{kk} = \begin{pmatrix} a_k(z,\eta) & b_k(z,\eta) \\ b'_k(z,\eta) & c_k(z,\eta) \end{pmatrix}.$$

We have

$$\mathrm{Tr}(R_n(U)) = \sum_{k=1}^{n} (a_k(z,\eta) + c_k(z,\eta)) = \sum_{k=1}^{n} (\sigma_k(X_n - z\mathbb{1}) - \eta)^{-1} + (-\sigma_k(X_n - z) - \eta)^{-1}.$$

By Theorem 2.1.10, $(B_n, 1)$ locally converges to $(\mathcal{A}_+, \varnothing)$. By Proposition 3.1.1, $\mathcal{A}_+$ is an essentially self-adjoint operator, as a consequence, by Theorem 2.1.2, we have the convergence of the resolvents. Namely, for any $i \in \{1, ..., n\}$,

$$\mathbb{E}\left[R_n(U)_{ii} \right] \xrightarrow[n\to+\infty]{} \mathbb{E}\left[R_{\mathcal{A}_+}(U)_{\varnothing\varnothing} \right].$$

Analogously, for any index in $\{-1, ..., -n\}$,

$$\mathbb{E}\left[R_n(U)_{-i-i} \right] \xrightarrow[n\to+\infty]{} \mathbb{E}\left[R_{\mathcal{A}_-}(U)_{\varnothing\varnothing} \right].$$

Moreover, by the essential self-adjointness of both $\mathcal{A}_+$ and $\mathcal{A}_-$, it follows that there exist measures $\check{\nu}^+_{\varnothing,z}$ and $\check{\nu}^-_{\varnothing,z}$ such that

$$R_{\mathcal{A}_\pm}(U)_{\varnothing\varnothing} = \int \frac{\check{\nu}^\pm_{\varnothing,z}}{x - \eta} = m_{\check{\nu}^\pm_{\varnothing,z}}(\eta).$$

Call $\check{\nu}_{z,\alpha} = \frac{1}{2}\mathbb{E}\left[\nu^+_{\varnothing,z} \right] + \frac{1}{2}\mathbb{E}\left[\nu^-_{\varnothing,z} \right]$, then

$$\begin{aligned} m_{\mathbb{E}[\check{\nu}_{X_n - z\mathbb{1}}]}(\eta) &= m_{\mathbb{E}\left[\mu_{B_n(z)}\right]}(\eta) \\ &= \frac{1}{2n}\mathbb{E}\left[\mathrm{Tr}(R_{B_n}(U)) \right] \\ &= \frac{1}{2}\mathbb{E}\left[R_n(U)_{11} \right] + \frac{1}{2}\mathbb{E}\left[R_n(U)_{-1-1} \right] \\ &\xrightarrow[n\to+\infty]{} \frac{1}{2}\mathbb{E}\left[R_{\mathcal{A}_+}(U)_{\varnothing\varnothing} \right] + \frac{1}{2}\mathbb{E}\left[R_{\mathcal{A}_-}(U)_{\varnothing\varnothing} \right] \\ &= m_{\frac{1}{2}\mathbb{E}\left[\nu^+_{\varnothing,z}\right] + \frac{1}{2}\mathbb{E}\left[\nu^-_{\varnothing,z}\right]}(\eta) = m_{\check{\nu}_{\alpha,z}}(\eta). \end{aligned}$$

Thus we proved $\mathbb{E}\left[\check{\nu}_{X_n - z\mathbb{1}} \right] \xrightarrow[n\to+\infty]{} \mathbb{E}\left[\check{\nu}_{\varnothing,z} \right] = \check{\nu}_{z,\alpha}$. Since the matrix B_n is Hermitian, we are in the hypothesis of concentration theorem 1.1.3. We can then exploit the bound of lemma and Borel-Cantelli first lemma to upgrade the convergence to almost sure. $\square$

3.3 Moments

We now estimate the moments of the limiting spectral distribution of the singular values $\nu_{z,\alpha}$. To this end we shall use Bennett's inequality to bound the moments of the random variable $\Gamma_1^{(n)} := \sum_{i=1}^n X_{i,1}$, for reasons that will be clear later. Let us now give a prove of the Bennett's inequality.

Theorem 3.3.1 (Bennett's inequality)**.** *Let Z be a random variable such that, $\mathbb{E}\left[\, Z \,\right] = 0$, $\mathbb{E}\left[\, Z^2 \,\right] = \sigma^2$, $|Z| < M$, where M is a positive constant. Then if $Z_1, .., Z_n$ are independent copies of Z, and $t \geq 0$,*

$$\mathbb{P}\left(\sum_{i=1}^n Z_i \geq t\right) \leq \exp\left\{-\frac{n\sigma^2}{M^2}\varphi\left(\frac{tM}{n\sigma^2}\right)\right\}, \tag{3.3.1}$$

where $\varphi(x) = (1+x)\log(1+x) - x$.

Proof. By Cernoff bound, for any $s > 0$,

$$\mathbb{P}\left(\sum_{i=1}^n Z_i \geq t\right) \leq e^{-st} \exp \prod_{i=1}^n \mathbb{E}\left[\, e^{sZ_i} \,\right].$$

Now,

$$\begin{aligned}
\mathbb{E}\left[\, e^{sZ} \,\right] &= \mathbb{E}\left[\sum_{k=0}^\infty \frac{(sZ)^k}{k!}\right] \\
&= 1 + \sum_{k=2}^\infty \frac{\mathbb{E}\left[\, (sZ)^k \,\right]}{k!} \\
\text{(Hölder) } &\leq 1 + \sum_{k=2}^\infty \frac{s^k M^{k-2}\mathbb{E}\left[\, Z^2 \,\right]}{k!} \\
&= 1 + \frac{\sigma^2}{M^2}\sum_{k=2}^\infty \frac{(sM)^k}{k!} \\
&= 1 + \frac{\sigma^2}{M^2}(e^{sM} - 1 - sM) \\
&\leq \exp\left\{\frac{\sigma^2}{M^2}(e^{sM} - 1 - sM)\right\}
\end{aligned}$$

Then,

$$\mathbb{P}\left(\sum_{i=1}^n Z_i \geq t\right) \leq \exp\left\{\frac{n\sigma^2}{M^2}(e^{sM} - 1 - sM) - st\right\}$$

Minimizing over $s > 0$ the right hand side, we get (3.3.1). □

The variable $\Gamma_1^{(n)}$ is the sum of the elements of the first column of the matrix X_n, a vector of i.i.d. random variables. In order to apply Bennet inequality we need to compute $\mathbb{E}[X_{11}]$ and $\mathrm{Var}(X_{11})$. By symmetry $\mathbb{E}[X_{11}] = n^{-1}$. Since for every $i, j = 1, ..., n$, $\mathbb{E}[X_{11}] = \mathbb{E}[X_{i,j}]$, moreover $\sum_{j=1}^{n} X_{j,k} = 1$, then $\mathbb{E}[X_{i,j}] = n^{-1}$, for every $i, j = 1, ..., n$.

For the variance we give a lower bound, which is sufficient for our purpose. Define the event $E_1^{(n)} = \{X_{11} = \max_{i=1,...,n} X_{1i}\}$, it has probability $1/n$. Since $\mathbb{E}\left[(X_{1,1} - \mathbb{E}[X_{1,1}])^2\right] = \mathbb{E}\left[X_{1,1}^2\right] - \frac{1}{n^2}$, we bound $\mathbb{E}\left[X_{1,1}^2\right]$,

$$\begin{aligned}\mathbb{E}\left[X_{1,1}^2\right] &\geq \mathbb{E}\left[X_{1,1}^2; E_1^{(n)}\right]\\ &= \mathbb{P}\left(E_1^{(n)}\right)\mathbb{E}\left[\max_{i=1,...,n} X_{1,i}^2\right]\end{aligned}$$

By Proposition 2.1.3,

$$\max_{i=1,...,n}\{X_{i,1}\} \xrightarrow[n\to+\infty]{(w)} \frac{\gamma_1}{c(\alpha)+\gamma_1},$$

where $\{\gamma_i\}_{i\geq 1}$, is an ordered Poisson process of intensity $\alpha x^{-\alpha-1}dx$, then

$$\mathbb{E}\left[\max_{i=1,...,n} X_{1,i}^2\right] \xrightarrow[n\to+\infty]{} \mathbb{E}\left[\left(\frac{\gamma_1}{c(\alpha)+\gamma_1}\right)^2\right] = \kappa(\alpha) < 1.$$

As a consequence there exists a $n \gg 1$, such that $\mathbb{E}\left[\max_{i=1,...,n} X_{1,i}^2\right] \geq \kappa(\alpha)/2$, and $\mathbb{E}\left[X_{1,i}^2\right] \geq \kappa(\alpha)/2n$. Thus for n big enough,

$$\mathbb{E}\left[(X_{1,1} - \mathbb{E}[X_{1,1}])^2\right] = \mathbb{E}\left[X_{1,1}^2\right] - \frac{1}{n^2} \geq \frac{\kappa(\alpha)}{2n} - \frac{1}{n^2} \geq \frac{\kappa(\alpha)}{n4}.$$

Call $\{Z_i\}_{i=1}^n := \{X_{1,i} - n^{-1}\}_{i=1}^n$. To estimate the moments of $\Gamma_1^{(n)}$ we apply Bennett's inequality to the collection $\{Z_i\}$. Note that for every $i = 1, ..., n$, $|Z_i| \leq 1$. For any $k > 1$, we have

$$\begin{aligned}\mathbb{E}\left[|\Gamma_i^{(n)} - 1|^k\right] &= \int_0^{+\infty} \mathbb{P}\left(|\Gamma_i^{(n)} - 1|^k \geq t\right) dt\\ &\leq \int_0^{+\infty} \mathbb{P}\left(\Gamma_i^{(n)} \geq 1 - t^{1/k}\right) dt + \int_0^{+\infty} \mathbb{P}\left(\Gamma_i^{(n)} \geq 1 + t^{1/k}\right) dt\\ &= \int_0^{1} \mathbb{P}\left(\Gamma_i^{(n)} \geq 1 - t^{1/k}\right) dt + \int_0^{+\infty} \mathbb{P}\left(\Gamma_i^{(n)} \geq 1 + t^{1/k}\right) dt\\ &\leq 1 + \int_0^{+\infty} \mathbb{P}\left(\Gamma_i^{(n)} \geq 1 + t^{1/k}\right) dt\end{aligned}$$

By Bennett's inequality, applied to the variables Z_i, with $M = 1$, $\mathbb{E}[Z_i] = 0$, and $\sigma^2 \geq \kappa(\alpha)/n4$, we have

$$\int_0^{+\infty} \mathbb{P}\left(\Gamma_i^{(n)} \geq 1 + t^{1/k}\right) dt \leq \int_0^{+\infty} \exp\left\{-\frac{\kappa(\alpha)}{4}\phi\left(\frac{4t^{1/k}}{\kappa(\alpha)}\right)\right\},$$

where $\phi(t)$ is as in theorem 3.3.1. Note that $\phi'(t) = \log(t+1)$, so that $\phi(t)$ is an increasing monotone function for $t > 0$.

$$\begin{aligned}\mathbb{E}\left[|\Gamma_i^{(n)} - 1|^k\right] &\leq 1 + \int_0^{+\infty} \exp\left\{-\frac{\kappa(\alpha)}{4}\varphi\left(\frac{4t^{1/k}}{\kappa(\alpha)}\right)\right\} dt \\ &= 1 + \int_0^{+\infty} kt^{k-1} e^{-\frac{\kappa(\alpha)}{4}\varphi\left(\frac{4t}{\kappa(\alpha)}\right)} dt < +\infty.\end{aligned}$$

Since the right hand side does not depend on n, we have the uniform bound

$$\liminf_n \mathbb{E}\left[|\Gamma_i^{(n)} - 1|^k\right] \leq 1 + \int_0^{+\infty} kt^{k-1} e^{-\frac{\kappa(\alpha)}{4}\varphi\left(\frac{4t}{\kappa(\alpha)}\right)} dt.$$

We can now bound the moments of both operators $\mathcal{A}_+$ and $\mathcal{A}_-$, defined in equation (2.1.23) and (2.1.24) respectively.

Proposition 3.3.2. *For any $k \geq 1$*

$$\mathbb{E}\left[\left\langle \delta_\varnothing, \mathcal{A}_\pm^{2k} \delta_\varnothing \right\rangle\right] \leq \mathcal{C}_{2k} \mathbb{E}\left[\Gamma^k\right],$$

where $\mathcal{C}_{2k} = (k+1)^{-1}\binom{2k}{k}$ is the number of Dyck paths of length $2k$, and Γ is the random variable

$$\Gamma = \sum_{i=1}^{+\infty} \frac{\gamma_i}{c(\alpha) + \gamma_i},$$

with $\{\gamma_i\}_{i\geq 1}$ ordered Poisson process of intensity $\alpha x^{-\alpha-1}dx$, and $c(\alpha)$ defined in (2.1.25).

Proof. Define the operators

$$\langle \delta_u, \widetilde{\mathcal{A}}_+ \delta_{uk} \rangle = \langle \delta_{uk}, \widetilde{\mathcal{A}}_+ \delta_u \rangle = \begin{cases} \dfrac{y_{uk}}{\sum_{i\geq 1} y_{ui}} & \text{if } d(u) \equiv 0 (\text{mod } 2) \\ \\ \dfrac{y_{uk}}{c(\alpha) + y_{uk}} & \text{if } d(u) \equiv 1 (\text{mod } 2) \\ \\ 0 & \text{otherwise} \end{cases} \tag{3.3.2}$$

$$\langle \delta_u, \widetilde{\mathcal{A}}_- \delta_{uk} \rangle = \langle \delta_{uk}, \widetilde{\mathcal{A}}_- \delta_u \rangle = \begin{cases} \dfrac{y_{uk}}{c(\alpha) + y_{uk}} & \text{if } d(u) \equiv 0 (\text{mod } 2) \\ \\ \dfrac{y_{uk}}{\sum_{i\geq 1} y_{ui}} & \text{if } d(u) \equiv 1 (\text{mod } 2) \\ \\ 0 & \text{otherwise} \end{cases} \tag{3.3.3}$$

We will from now on consider those operators, since

$$\mathbb{E}\left[\left\langle \delta_\varnothing, \mathcal{A}_\pm^{2k} \delta_\varnothing \right\rangle\right] \leq \mathbb{E}\left[\left\langle \delta_\varnothing, \widetilde{\mathcal{A}}_\pm^{2k} \delta_\varnothing \right\rangle\right], \tag{3.3.4}$$

see proof of proposition 3.1.1. Define $O^k_\varnothing$ as the set of all path starting from $\varnothing$ and returning to $\varnothing$ in $2k$ steps. For a path $\eta \in O^k_\varnothing$, $\eta = (\varnothing, v_1, v_2, ...v_{2k-2}, v_{2k-1}, \varnothing)$, with $v_i \in \mathbb{N}^f$, define $W(\eta)$, as the product of the weights of the edges along the path, say:

$$W(\eta) = \omega_{\varnothing,v_1}\omega_{v_1,v_2}\cdots\omega_{v_{2k-2},v_{2k-1}}\omega_{v_{2k-1},\varnothing}.$$

Since both operators $\mathcal{A}_+$ and $\mathcal{A}_-$ are defined on a tree, in particular on an acyclic graph, the weights of $W(\eta)$ will be pairwise identical, and the path weight can be rewritten as

$$W(\eta) = \omega^2_{\varnothing,v_1}\omega^2_{v_1,v_2}\cdots\omega^2_{v_{k-2},v_{k-1}}\omega^2_{v_{k-1},v_k}.$$

for some $v_1, ..., v_k \in \mathbb{N}^f$. Since $0 \le \omega_{i,j} \le 1$ for any $i, j \in \mathbb{N}^f$,

$$W(\eta) \le \sqrt{W(\eta)} = \omega_{\varnothing,v_1}\omega_{v_1,v_2}\cdots\omega_{v_{k-2},v_{k-1}}\omega_{v_{k-1},v_k}.$$

Then

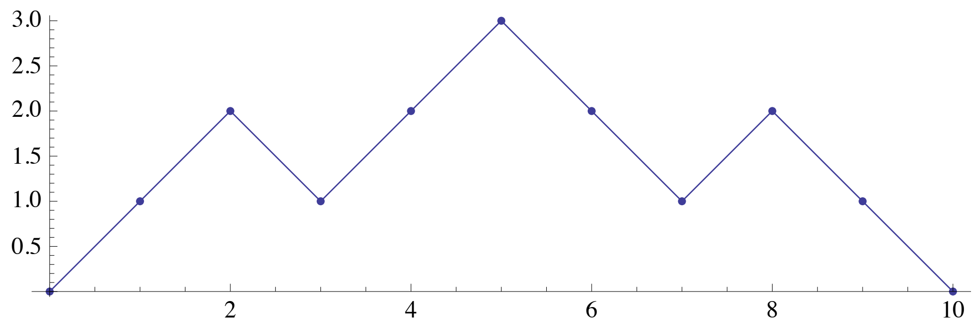
$$\mathbb{E}\left[\left\langle \delta_\varnothing, \widetilde{\mathcal{A}}^{2k}_\pm \delta_\varnothing \right\rangle\right] \le \sum_{\eta \in O^k_\varnothing} \sqrt{W(\eta)}.$$

We should now enumerate, for any $k \ge 1$, the paths in $O^k_\varnothing$.

Let us give an example. Fix $k = 5$, and consider the path

$$\eta = (\varnothing, 1, 11, 1, 12, 121, 12, 1, 13, 1, \varnothing),$$

and plot its distance from $\varnothing$ at each step.



The result is a Dyck path. However, the same Dyck path can be associated to (infinitely) many other paths $\eta \in O^k_\varnothing$, namely when the path takes the step from 12 to 121, the vertex 121 could have been replaced by any other vertex of its generation, and the Dyck path would have been the same. Then, in order to count the total weight of all paths η associated to the Dyck path of the example, the first step is to sum over $D_{12} = \{12k : k \in \mathbb{N}\}$, the descendants of 12, obtaining

$$\sum_{i_1,i_2,i_3,i_5}\sum_{i_4\in D_{12}} \omega_{\varnothing,i_1}\omega_{i_1,i_2}\omega_{i_1,i_3}\omega_{i_3,i_4}\omega_{i_1,i_5} = \sum_{i_1,i_2,i_3,i_5} \omega_{\varnothing,i_1}\omega_{i_1,i_2}\omega_{i_1,i_3}\omega_{i_1,i_4}\left(\sum_{i_4\in D_{i_3}} \omega_{i_3,i_5}\right),$$

if $D_v := \{vk : k \in \mathbb{N}\}$. For $v \in \mathbb{N}^f$, call $\Omega_v = \sum_{k\geq 1} \omega_{vk}$. Since we are summing inside the expectation, and generations are independent, we have

$$\mathbb{E}\left[\,\omega_{\varnothing,i_1}\omega_{i_1,i_2}\omega_{i_1,i_3}\omega_{i_1,i_4}\Omega_{i_3}\,\right] = \mathbb{E}\left[\,\omega_{\varnothing,i_1}\omega_{i_1,i_2}\omega_{i_1,i_3}\omega_{i_1,i_4}\,\right]\mathbb{E}\left[\,\Omega_{i_3}\,\right].$$

Note now that the variable Ω can be distributed either as Γ or as the sum of the component of a $PD(\alpha)$. Furthermore observe that we obtain the same graphic (Dyck path) replacing any vertex with any other of its generation, what it matters is the sequence of the steps. The total weight of the paths associated to the Dyck path of the example is given by the sum,

$$\sum_{i_1\in D_\varnothing}\ \sum_{i_2,i_3,i_4\in D_{i_1}}\ \sum_{i_5\in D_{i_3}} \omega_{\varnothing,i_1}\omega_{i_1,i_2}\omega_{i_1,i_3}\omega_{i_1,i_4}\omega_{i_3,i_5}.$$

Once we take the expectations, it becomes

$$\mathbb{E}\left[\sum_{i_1\in D_\varnothing}\ \sum_{i_2,i_3,i_4\in D_{i_1}}\ \sum_{i_5\in D_{i_3}} \omega_{\varnothing,i_1}\omega_{i_1,i_2}\omega_{i_1,i_3}\omega_{i_1,i_4}\omega_{i_3,i_5}\right] = \mathbb{E}\left[\,\Omega\,\right]\mathbb{E}\left[\,\Omega^3\,\right]\mathbb{E}\left[\,\Omega\,\right].$$

Now notice that if Ω is a Poisson-Dirichlet type generation, for any $n \geq 1$, $\mathbb{E}\left[\,\Omega^n\,\right] = 1$. Moreover, since $\mathbb{E}\left[\,\Gamma\,\right] = 1$ and for any integers $0 < n_1 \leq n_2$, $\mathbb{E}\left[\,\Gamma^{n_1}\,\right] \leq \mathbb{E}\left[\,\Gamma^{n_2}\,\right]$, independently of the distribution of the generation we have $\mathbb{E}\left[\,\Omega^n\,\right] \leq \mathbb{E}\left[\,\Gamma^n\,\right]$. Then by Hölder inequality

$$\mathbb{E}\left[\,\Omega\,\right]\mathbb{E}\left[\,\Omega^3\,\right]\mathbb{E}\left[\,\Omega\,\right] \leq \mathbb{E}\left[\,\Gamma\,\right]\mathbb{E}\left[\,\Gamma^3\,\right]\mathbb{E}\left[\,\Gamma\,\right] \leq \mathbb{E}\left[\,\Gamma^5\,\right]^{1/5}\mathbb{E}\left[\,\Gamma^5\,\right]^{3/5}\mathbb{E}\left[\,\Gamma^5\,\right]^{1/5} = \mathbb{E}\left[\,\Gamma^5\,\right].$$

For a generic Dyck path π, of length $2k$, we will prove that the total weight of the paths η associated to π (we will write $\eta \sim \pi$) does not exceed $\mathbb{E}\left[\,\Gamma^k\,\right]$.

Since by Jensen inequality,

$$\mathbb{E}\left[\,\Gamma^{a_1}\,\right]^{a_2}\mathbb{E}\left[\,\Gamma^{b_1}\,\right]^{b_2} \leq \mathbb{E}\left[\,\Gamma^{a_1a_2+b_1b_2}\,\right],$$

steps taken over independent progenies, are less expensive than steps in the same progeny, provided that the steps are taken at same root distance.

Therefore, given any Dyck path π, we can upper bound the expected total weight of its associated paths η, which we will denote as $\eta \sim \pi$, simply keeping track of how many times it touches progenies at a given distance from $\varnothing$. Call d_i the number of times the path touches progenies at distance i from $\varnothing$. Since the length of the π is $2k$, $d_1 + d_2 + \cdots + d_k = k$, and

$$\mathbb{E}\left[\sum_{\eta\sim\pi}\sqrt{W(\eta)}\right] \leq \mathbb{E}\left[\,\Gamma^{d_1}\,\right]\mathbb{E}\left[\,\Gamma^{d_2}\,\right]\cdots\mathbb{E}\left[\,\Gamma^{d_k}\,\right],$$

again by Hölder inequality,

$$\prod_{i=1}^{k} \mathbb{E}\left[\Gamma^{d_i}\right] \leq \prod_{i=1}^{k} \mathbb{E}\left[\Gamma^{k}\right]^{\frac{d_i}{k}} = \mathbb{E}\left[\Gamma^{k}\right].$$

In conclusion, call the Dyck paths of length $2k$, $\pi_1, \pi_2, ..., \pi_{\mathcal{C}_{2k}}$. We have

$$\begin{aligned}
\mathbb{E}\left[\left\langle \delta_\varnothing, \widetilde{\mathcal{A}}_{\pm}^{2k}\delta_\varnothing \right\rangle\right] &\leq \mathbb{E}\left[\sum_{\eta\in O_\varnothing^k} \sqrt{W(\eta)}\right] \\
&\leq \sum_{i=1}^{\mathcal{C}_{2k}} \mathbb{E}\left[\sum_{\eta\sim\pi_i} \sqrt{W(\eta)}\right] \\
&\leq \mathcal{C}_{2k}\mathbb{E}\left[\Gamma^k\right].
\end{aligned}$$

□

The bound on the moments of both $\mathcal{A}_+$ and $\mathcal{A}_-$ are powerful enough to let us deduce a bound on the exponential moment of the limiting spectral distribution of the eigenvalues.

Proposition 3.3.3. *For any $\lambda \in \mathbb{R}$,*

$$\int e^{\lambda t}\nu_{\alpha,z}(dt) < +\infty.$$

Proof. By straightforward computation,

$$\begin{aligned}
\int e^{\lambda t}\nu_{\alpha,z}(dt) &= \int \sum_{k=0}^{+\infty} \frac{(\lambda t)^k}{k!}\nu_{\alpha,z}(dt) \\
&= \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} \int t^k \nu_{\alpha,z}(dt) \\
&= \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} \int t^k \mathbb{E}\left[\nu_{\varnothing,z}\right](dt) \\
&= \sum_{k=0}^{+\infty} \frac{\lambda^{2k}}{(2k)!}\frac{1}{2}\left(\mathbb{E}\left[\left\langle \delta_\varnothing, \mathcal{A}_+^{2k}\delta_\varnothing\right\rangle\right] + \mathbb{E}\left[\left\langle \delta_\varnothing, \mathcal{A}_-^{2k}\delta_\varnothing\right\rangle\right]\right) \\
&\leq \sum_{k=0}^{+\infty} \frac{\lambda^{2k}}{(2k)!}\frac{1}{2}\left(\mathcal{C}_{2k}\mathbb{E}\left[\Gamma^k\right] + \mathcal{C}_{2k}\mathbb{E}\left[\Gamma^k\right]\right) \\
&= \sum_{k=0}^{+\infty} \frac{\lambda^{2k}}{(2k)!}\frac{1}{k+1}\frac{(2k)!}{k!k!}\mathbb{E}\left[\Gamma^k\right] \\
&\leq \sum_{k=0}^{+\infty} \frac{\lambda^{2k}}{(k+1)!k!}\mathbb{E}\left[\Gamma^k\right]
\end{aligned}$$

Since by Bennett's inequality we have bounds for the moments of $|\Gamma - 1|$, we observe that

$$\begin{aligned}
\mathbb{E}\left[\Gamma^k\right] &\le \mathbb{E}\left[|\Gamma|^k\right] \\
&\le \mathbb{E}\left[(|\Gamma - 1| + 1)^k\right] \\
&= \mathbb{E}\left[\sum_{i=0}^{k}\binom{k}{i}|\Gamma - 1|^i\right] \\
&\le \mathbb{E}\left[|\Gamma - 1|^k\right]\sum_{i=0}^{k}\binom{k}{i} = 2^k\mathbb{E}\left[|\Gamma - 1|^k\right].
\end{aligned}$$

Now by Skorokhod representation theorem we may assume $\Gamma^{(n)}$ converges to Γ almost surely, and by Fatou's lemma

$$\mathbb{E}\left[\Gamma^k\right] \le \liminf_n \mathbb{E}\left[(\Gamma^{(n)})^k\right].$$

Thus

$$\begin{aligned}
\int e^{\lambda t}\nu_{\alpha,z}(dt) &\le \sum_{k=0}^{+\infty}\frac{\lambda^{2k}}{(k+1)!k!}\mathbb{E}\left[\Gamma^k\right] \\
&\le \liminf_n \sum_{k=0}^{+\infty}\frac{\lambda^{2k}2^k}{(k+1)!k!}\mathbb{E}\left[|\Gamma^{(n)} - 1|^k\right] \\
&\le \sum_{k=0}^{+\infty}\frac{\lambda^{2k}2^k}{(k+1)!k!}\left(1 + \int_0^{+\infty} kt^{k-1}\exp\left\{-\frac{\kappa(\alpha)}{4}\varphi\left(\frac{4t}{\kappa(\alpha)}\right)\right\}dt\right) \\
&\le e^{2\lambda^2} + 2\lambda^2\int_0^{+\infty}\left(\sum_{k=0}^{+\infty}k\frac{(2\lambda^2 t)^{k-1}}{(k+1)!k!}\right)\exp\left\{-\frac{\kappa(\alpha)}{4}\varphi\left(\frac{4t}{\kappa(\alpha)}\right)\right\}dt \\
&\le e^{2\lambda^2} + 2\lambda^2\int_0^{+\infty}\exp\left\{2\lambda^2 t - \frac{\kappa(\alpha)}{4}\varphi\left(\frac{4t}{\kappa(\alpha)}\right)\right\}dt < +\infty.
\end{aligned}$$

□

Chapter 4

Eigenvalues

We will prove the convergence of the empirical spectral distribution of the eigenvalues via Girko's Hermitization method, which will require further hypothesis for the law of the matrix entries, such as the bounded density.

4.1 Tightness

Since X_n is a Markov matrix, the empirical spectral distribution μ_{X_n} has support in the unitary disc of the complex plane $\{z \in \mathbb{C} : |z| \leq 1\}$, consequently it is tight. We say a sequence of measure π_n defined on $\mathbb{C}$ is tight, if for any $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset \mathbb{C}$ such that $\pi_n(K_\varepsilon) > 1 - \varepsilon$.

In the next lemma we prove the tightness fir the empirical spectral distribution of the singular values.

Lemma 4.1.1 (Tightness)**.** *For all $z \in \mathbb{C}$ and any $\delta > 0$, for any n,*

$$\int_{\mathbb{R}} t^2 \nu_{X_n - z\mathbb{1}}(dt) \leq (1+\delta) + (1+\delta^{-1})|z|^2,$$

then $(\nu_{X_n - z\mathbb{1}})_{n\geq 1}$ is tight.

Proof. By a property of the singular values, if N and M are complex $n \times n$ matrices with spectrum of singular values $\sigma_1(N) \geq \cdots \sigma_n(N)$ and $\sigma_1(M) \geq \cdots \sigma_n(M)$ respectively, then

$$\max_{1\leq i\leq n} |\sigma_i(M) - \sigma_i(N)| \leq \sigma_1(M - N),$$

see e.g. [16]. Applying this property to the matrices $M = X_n - z\mathbb{1}$ and $N = X_n$, for every $k = 1, .., n$, holds

$$\begin{aligned}\sigma_k(X_n - z\mathbb{1}) - \sigma_k(X_n) &\leq \max_{i=1,..,n} |\sigma_i(X_n - z\mathbb{1}) - \sigma_i(X_n)| \\ &\leq \sigma_1(X_n - z\mathbb{1} - X_n) \\ &= \sigma_1(-z\mathbb{1}) = |z|.\end{aligned}$$

So that,

$$\sigma_k(X_n - z\mathbb{1}) \leq \sigma_k(X_n) + |z|.$$

Then, for any $n \in \mathbb{N}$ and $z \in \mathbb{C}$, and $\delta > 0$,

$$\begin{aligned}\int_{\mathbb{R}} t^2 \nu_{X_n - z\mathbb{1}}(dt) &= \frac{1}{n}\sum_{k=1}^{n}(\sigma_k(X_n - z\mathbb{1}))^2 \\ &\leq \frac{1}{n}\sum_{k=1}^{n}(\sigma_k(X_n) + |z|)^2 \\ &\leq \frac{1}{n}\sum_{k=1}^{n}\left[(1+\delta)\sigma_k^2(X_n) + (1+\delta^{-1})|z|^2\right] \\ &= (1+\delta^{-1})|z|^2 + \frac{1+\delta}{n}\sum_{k=1}^{n}\sigma_k^2(X_n)\end{aligned}$$

Now,

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}\sigma_k^2(X_n) &= \frac{1}{n}\mathrm{Tr}(X^*X) \\ &= \frac{1}{n}\sum_{i,j=1}^{n}(X^*)_{ij}X_{ji} \\ &= \frac{1}{n}\sum_{i,j=1}^{n}X_{ji}^2 \\ &\leq \frac{1}{n}\sum_{i,j=1}^{n}X_{ji} = 1.\end{aligned}$$

□

4.2 Invertibility

This section is dedicated to prove a bound for the smallest singular value of the matrix $X_n - z\mathbb{1}$. First let us recall the Rudelson-Vershynin row bound, see [22].

Lemma 4.2.1 (Rudelson-Vershynin row bound)**.** *Let M be a complex $n \times n$ matrix with rows $R_1, ..., R_n$, then*

$$n^{-1/2}\min_{i=1,..,n}\mathrm{dist}(R_i, R_{-i}) \leq s_n(M) \leq \min_{i=1,..,n}\mathrm{dist}(R_i, R_{-i}),$$

where R_{-i} is the vector space span$(R_j; j \neq i)$.

In our setting $X_n - z\mathbb{1}$ is non-invertible for $z = 1$, since X_n is a Markov matrix. Anyway if z is at positive distance from 1, we can bound the smallest eigenvalue $s_n(X_n - z\mathbb{1})$.We now readapt a result for light tailed matrices from [11]. The first step is to prove the following lemma.

Lemma 4.2.2. *For any $z \in \mathbb{C}$ define the $n \times n$ complex matrix,*

$$A_z := \mathbb{1}_n - z \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}.$$

Fix $\delta > 0$. If $|1 - z| > \delta$ then

$$s_n(A_z) \geq \frac{\delta}{(1 + \delta + \sqrt{n}|z|)}$$

Proof. Since $|1-z| > \delta > 0$, A_z is invertible. The rank if the symmetric matrix $A_z A_z^* - \mathbb{1}_n$ is at most 2, call indeed

$$M_1 := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}.$$

We have

$$\begin{aligned} A_z A_z^* - \mathbb{1}_n &= (\mathbb{1}_n - zM_1)(\mathbb{1}_n - zM_1)^* - \mathbb{1}_n \\ &= (\mathbb{1}_n - zM_1)(\mathbb{1}_n - \bar{z}M_1^T) - \mathbb{1}_n \\ &= \mathbb{1}_n - \bar{z}M_1^T - zM_1 + |z|^2 M_1 M_1^T - \mathbb{1}_n \\ &= -\bar{z}M_1^T - zM_1 + |z|^2 M_1 M_1^T \end{aligned}$$

Since

$$M_1 M_1^T = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

We have

$$-\bar{z}M_1^T - zM_1 + |z|^2 M_1 M_1^T = \begin{pmatrix} |z|^2 - z - \bar{z} & |z|^2 - \bar{z} & \cdots & |z|^2 - \bar{z} \\ |z|^2 - z & |z|^2 & \cdots & |z|^2 \\ \vdots & \vdots & \ddots & \vdots \\ |z|^2 - z & |z|^2 & \cdots & |z|^2 \end{pmatrix}.$$

Then A_z has at least $n-2$ singular values equal to 1, and in particular $s_n(A_z) \leq 1 \leq s_1(A_z)$. Then A_z is lower triangular with eigenvalues $1-z, 1, ..., 1$, by Weyl inequality

$$|1-z| = \prod_{i=1}^{n} |\lambda_i(A_z)| = \prod_{i=1}^{n} s_i(A_z) = s_1(A_z)s_n(A_z). \tag{4.2.1}$$

We also have that

$$s_1^2(A_z) + s_n^2(A_z) + (n-2) = \mathrm{Tr}(A_z A_z^*) = |1-z|^2 + (n-1)(1+|z|^2),$$

which gives,

$$s_1^2(A_z) + s_n^2(A_z) = 1 + (n-1)|z|^2 + |1-z|^2. \tag{4.2.2}$$

Then $s_1^2(A_z)$ and $s_n^2(A_z)$ are the solution of the equation

$$X^2 - (1 + (n-1)|z|^2 + |1-z|^2)X + |1-z|^2 = 0.$$

Take $n \gg 1$, so that we do not have complex conjugate solution, then we can

$$s_1^2(A_z) = \frac{1 + (n-1)|z|^2 + |1-z|^2 + \sqrt{(1 + (n-1)|z|^2 + |1-z|^2)^2 - 4|1-z|^2}}{2}$$

and

$$s_n^2(A_z) = \frac{1 + (n-1)|z|^2 + |1-z|^2 - \sqrt{(1 + (n-1)|z|^2 + |1-z|^2)^2 - 4|1-z|^2}}{2}.$$

Now, by (4.2.1),

$$\begin{aligned}
s_n^2(A_z) &= \frac{|1-z|^2}{s_1^2(A_z)} \\
&= \frac{2|1-z|^2}{1 + (n-1)|z|^2 + |1-z|^2 + \sqrt{(1 + (n-1)|z|^2 + |1-z|^2)^2 - 4|1-z|^2}} \\
&\geq \frac{|1-z|^2}{1 + (n-1)|z|^2 + |1-z|^2} \\
&= \frac{1}{\frac{1}{|1-z|^2} + 1 + \frac{|z|^2(n-1)}{|1-z|^2}} \\
&\geq \frac{1}{\frac{1}{\delta^2} + 1 + \frac{|z|^2(n-1)}{\delta^2}} \\
&= \frac{\delta^2}{1 + \delta^2 + |z|^2(n-1)}
\end{aligned}$$

Thus

$$s_n(A_z) \geq \sqrt{\frac{\delta^2}{1 + \delta^2 + |z|^2(n-1)}} \geq \frac{\delta}{1 + \delta + \sqrt{n}|z|} \tag{4.2.3}$$

□

Lemma 4.2.3 (Invertibility). *For any $\delta > 0$, if $z \in \mathbb{C}$ is such that $|1-z| > \delta$ and $|z| < \delta^{-1}$, then there exists $r(\delta) > 0$ such that, a.s.*

$$\liminf_{n\to\infty} n^r s_n(X_n - z\mathbb{1}_n) = +\infty.$$

Proof. Fix $a > 0$ and $z \in \mathbb{C}$ with $|1-z| > \delta$, $|z| < \delta^{-1}$. We can rewrite $X_n = D_n U_n$, where

$$D_n = \operatorname{diag}(\rho_1^{-1}, ..., \rho_n^{-1}) \qquad \text{and} \qquad U_n = \begin{pmatrix} U_{11} & \cdots & U_{1n} \\ \vdots & \ddots & \vdots \\ U_{n1} & \cdots & U_{nn} \end{pmatrix}.$$

Thus we can decompose $X_n - z\mathbb{1}_n$ as follow,

$$X_n - z\mathbb{1}_n = D_n Y_n \quad \text{where} \quad Y_n := U_n - zD_n^{-1}.$$

Define the event

$$\mathcal{A}_n = \bigcap_{i=1}^{n} \{\rho_i \leq n^c\}.$$

We have, if $n \gg 1$,

$$\begin{aligned}
\mathbb{P}\left(\mathcal{A}_n^C\right) &= \mathbb{P}\left(\bigcup_{i=1}^{n} \{\rho_i > n^\eta\}\right) \\
&\leq n\mathbb{P}\left(\rho_1 > n^\eta\right) \\
&\leq n\mathbb{P}\left(n \max_{i=1,..,n} U_i > n^\eta\right) \\
&= n(1 - \mathbb{P}\left(\max_{i=1,..,n} U_i \leq n^{\eta-1}\right)) \\
&= n(1 - (1 - \mathbb{P}\left(U_1 > n^{\eta-1}\right))^n) \\
&= n(1 - (1 - L(n^{\eta-1})n^{-\alpha(\eta-1)})^n) \\
&\leq n(1 - (1 - 2cn^{-\alpha(\eta-1)})^n) \\
&\leq 2n(1 - \exp\{-2cn^{-\alpha(\eta-1)+1}\}) \\
&= 2n(1 - 1 + 2cn^{-\alpha(\eta-1)+1} + o(n^{-\alpha(\eta-1)+1})) \\
&\leq 8cn^{-\alpha(\eta-1)+2}.
\end{aligned}$$

Then, for $\eta > (2+\alpha)/\alpha + 1$, $\mathbb{P}\left(\mathcal{A}_n^C\right) \leq n^{-a}$, for $n \gg 1$. Since $s_n(D) = \min_i \rho_i^{-1}$ and for any complex $n \times n$ matrix M and N, $s_n(MN) \geq s_n(M)s_n(N)$, on the event $\mathcal{A}_n$ we have

$$\{s_n(X_n - z) \leq t\} \subset \{s_n(D_n)s_n(Y) \leq t\} \subset \{s_n(Y) \leq tn^{-\eta}\},$$

for every $t > 0$. Now, for every $b' > 0$, one may select $b > 0$ and set $t = t(n) = n^{-b}$ such that $t(n)n^{-\eta} \leq n^{-b'}$, for $n \gg 1$. Thus, on the event $\mathcal{A}_n$, for $n \gg 1$,

$$\mathcal{X}_n := \{s_n(X_n - z\mathbb{1}_n) \leq n^{-b}\} \subset \{s_n(Y) \leq n^{-b'}\} := \mathcal{Y}_n.$$

Consequently, for any $b' > 0$, there exists a $b > 0$ such that for $n \gg 1$,

$$\mathbb{P}(\mathcal{X}_n) = \mathbb{P}(\mathcal{X}_n \cap \mathcal{A}_n) + \mathbb{P}\left(\mathcal{X}_n \cap \mathcal{A}_n^C\right) \leq \mathbb{P}(\mathcal{Y}_n) + \mathbb{P}\left(\mathcal{A}_n^C\right) \leq \mathbb{P}(\mathcal{Y}_n) + n^{-a}.$$

Then we have to prove that, for some $b' > 0$, depending on δ, for $n \gg 1$,

$$\mathbb{P}(\mathcal{Y}_n) = \mathbb{P}\left(s_n(Y) \leq n^{-b'}\right) \leq n^{-a}. \tag{4.2.4}$$

Define A_z as in lemma 4.2.2. For every $1 \leq k \leq n$, let P_k be the $n \times n$ permutation matrix for the transposition $(1, k)$. Note that $P_1 = \mathbb{1}$. For every column vector e_i of the canonical base of $\mathbb{R}_n$,

$$(P_k A_z P_k)e_i = \begin{cases} e_i & \text{if } i \neq k, \\ e_k - z(e_1 + \cdots + e_n) & \text{if } i = k. \end{cases}$$

Now, if $R_1, ..., R_n$ and $R'_1, ..., R'_n$ are the rows of the matrices U and Y, then

$$Y = \begin{pmatrix} R'_1 \\ \vdots \\ R'_n \end{pmatrix} = \begin{pmatrix} R_1 P_1 A_z P_1 \\ \vdots \\ R_n P_n A_z P_n \end{pmatrix}.$$

Define the vector space $R'_{-i} := \text{span}\{R'_j : j \neq i\}$ for every $1 \leq i \leq n$. From Rudelson-Vershynin row bound, Lemma 4.2.1,

$$\min_{i=1,..,n} \text{dist}(R'_i, R'_{-i}) \leq \sqrt{n} s_n(Y).$$

As a consequence, by the union bound, for any $u \geq 0$,

$$\begin{aligned}
\mathbb{P}\left(\sqrt{n} s_n(Y) \leq u\right) \mathbb{P}\left(\min_{i=1,...,n} \text{dist}(R'_i, R'_{-i}) \leq u\right) \\
&\leq \mathbb{P}\left(\exists i = 1, ..., n \,:\, \text{dist}(R'_i, R'_{-i}) \leq u\right) \\
&\leq \sum_{i=1}^{n} \mathbb{P}\left(\text{dist}(R'_i, R'_{-i}) \leq u\right) \\
&\leq n \max_{i=1,...,n} \mathbb{P}\left(\text{dist}(R'_i, R'_{-i}) \leq u\right)
\end{aligned}$$

The law of the random variable $\text{dist}(R'_i, R'_{-i})$ does not depend on i. We fix $i = 1$. Let V' be a unit normal vector to R'_{-1}. Such a vector is not unique, we just pick one, and this defines a random variable on the unit sphere $\mathbb{S}^{n-1} := \{x \in \mathbb{C} : \|x\|_2 = 1\}$. Since $V' \in (R'_{-1})^{\perp}$ and $\|V'\|_2 = 1$,

$$|R'_1 \cdot V'| \leq \text{dist}(R'_1, R'_{-1}) = \min_{v' \in R^{\perp}_{-1}} |R'_1 \cdot v'|.$$

Let ν be the distribution of V' on $\mathbb{S}^{n-1}$. Since V' and R_1' are independent, for any $u \geq 0$,

$$\mathbb{P}\left(\operatorname{dist}(R_1', R_{-1}') \leq u\right) \leq \mathbb{P}\left(|R_1' \cdot V'| \leq u\right) = \int_{\mathbb{S}^{n-1}} \mathbb{P}\left(|R_1' \cdot v'| \leq u\right) \nu(dv').$$

Fix $v' \in \mathbb{S}^{n-1}$, then $R_1' \cdot v' = R_1 v$ where $v := P_1 A_z P_1 v' = A_z v'$. Lemma 4.2.2, provides the bound, if $|z| \leq \delta^{-1}$,

$$\|v\|_2 = \|A_z v'\| \geq \min_{x \in \mathbb{S}^{n-1}} \|A_z x\|_2 = s_n(A_z) \geq (1 + \frac{1}{\delta} + \frac{1}{\delta^2}\sqrt{n})^{-1} := c(\delta, \sqrt{n})^{-1}.$$

But $\|v\|_2 \geq c(\delta, \sqrt{n})^{-1}$ implies, that there exists $j_0 \in \{1, ..., n\}$ such that $|v_{j_0}|^{-1} \leq \sqrt{n} c(\delta, \sqrt{n})$. Therefore

$$|Re(v_{j_0})| \leq \sqrt{2n} c(\delta, \sqrt{n}) \qquad \text{or} \qquad |Im(v_{j_0})| \leq \sqrt{2n} c(\delta, \sqrt{n}).$$

Suppose $|Re(v_{j_0})| \leq \sqrt{2n} c(\delta, \sqrt{n})$. Observe that

$$\mathbb{P}\left(|R_1' \cdot v'| \leq u\right) = \mathbb{P}\left(|R_1 \cdot v| \leq u\right) \leq \mathbb{P}\left(|Re(v_{j_0})| \leq u\right).$$

The real random variable $Re(R_1 \cdot v)$ is a sum of independent real random variables and one of them is $U_{1j_0} Re(v_{j_0})$, which is absolutely continuous with a density bounded above by $Bc(\delta, \sqrt{n})\sqrt{2n}$, where B is the bound of the density of U_{11}. Consequently by a basic property of convolutions of probability measures, the real random variable $Re(R_1 \cdot v)$ is absolutely continuous with a density φ bounded above by $Bc(\delta, \sqrt{n})\sqrt{2n}$, and therefore,

$$\mathbb{P}\left(|Re(v_{j_0})| \leq u\right) = \int_{-u}^{u} \varphi(s) ds \leq Bc(\delta, \sqrt{n})\sqrt{2n} 2u \leq c(\delta) B n u.$$

Finally, to end the proof we have

$$\mathbb{P}\left(\sqrt{n} s_n(Y) \leq u\right) \leq c(\delta) B n^2 u.$$

Thus, (4.2.4), holds with $b' = d + 1/2$, taking $u = n^{-d}$ such that $c(\delta) B n^2 n^{-d} \leq n^{-a}$ for $n \gg 1$. $\square$

4.3 Distance of a row from a vector space.

Tao and Vu's negative second moment lemma, bond singular values to distance of the row vectors to some vector spaces of not too big dimension. Here we readapt to our setting some results from [10], with the further observation that, since any complex $n \times n$ matrix M, has the same spectrum of the singular values as M^* and M^T, we will prove next result for $(X_n - z)^*$, exploiting this fact to bound the singular values.

Proposition 4.3.1. *Let $0<\gamma<1/4$ and R be a row of $a_n(X_n-z)^*$. There exists $\delta>0$ depending on α and γ such that or all d-dimensional subspace W of $\mathbb{C}^n$ with $n-d\geq n^{1-\gamma}$ one has*

$$\mathbb{P}\left(\operatorname{dist}(R,W)\leq n^{\frac{1-2\gamma}{\alpha}}\right)\leq e^{n^{-\delta}}. \tag{4.3.1}$$

Before going into the proof of Proposition 4.3.1, recall the concentration inequality by Talagrand, in theorem 1.1.4.

Proof of proposition 4.3.1. Assume R is the first row of $a_n(X_n-z\mathbb{1})$, then $R=a_n(X_n^{(1)}-ze_1)$, if $X_n^{(1)}$ is the first row of the matrix X_n^* (or the first column of X_n). We have

$$\operatorname{dist}(R,W)\geq\operatorname{dist}(a_nX_n^{(1)}-ze_1,\operatorname{span}(e_1,W))=\operatorname{dist}(X_n^{(1)},W_1),$$

if we set $W_1=\operatorname{span}(W,e_1)$. Note that $d\leq\dim W_1\leq d+1$. Recall R is a column type vector, then

$$a_n^{-1}R=\left(\frac{U_{i1}}{U_{i1}+\cdots+U_{in}}\right)_{i=1}^{n}=\left(\frac{U_{i1}}{U_{i1}+a_nB_n^{(i)}}\right)_{i=1}^{n}$$

if one defines $B_n^{(i)}:=a_n^{-1}(U_{i2}\cdots+U_{in})$.Call

$$(Z_i)_{i=1}^n:=a_n\left(\frac{U_{i1}}{U_{i1}+a_nY_i}\right)_{i=1}^{n}, \tag{4.3.2}$$

note then that for any $\varepsilon>0$, and any $i=1,...,n$,

$$\begin{aligned}
\mathbb{P}(Z>t)&=\mathbb{P}\left(a_n\frac{U}{U+a_nB_n^{(i)}}>t\right)=\mathbb{P}\left(\frac{a_nB_n^{(i)}}{U}<a_nt^{-1}-1\right)\\
&\leq\mathbb{P}\left(\frac{a_nB_n^{(i)}}{U}<a_nt^{-1}\right)\leq\mathbb{P}\left(\frac{U}{B_n^{(i)}}>t\right)\\
&\leq\mathbb{P}\left(\frac{U}{B_n^{(i)}}>t\,,\,B_n^{(i)}\geq n^{-\varepsilon}\right)+\mathbb{P}\left(B_n^{(i)}<n^{-\varepsilon}\right)\\
&\leq\mathbb{P}\left(U>tn^{-\varepsilon}\right)+2e^{-n^{\varepsilon\alpha}} &(4.3.3)\\
&\leq L(tn^{\varepsilon})t^{-\alpha}n^{\varepsilon\alpha}+2e^{-n^{\varepsilon\alpha}} &(4.3.4)
\end{aligned}$$

for a positive ε. Equaiton (4.3.3) holds since $B_n^{(i)}=a_n^{-1}\sum_{k=2}^nU_{ik}$, and for any positive x,

$$\begin{aligned}
\mathbb{P}\left(B_n^{(i)}<x\right)&=\mathbb{P}\left(a_n^{-1}\sum_{k=2}^{n}U_{ik}<x\right)\\
&\leq\mathbb{P}\left(\max_{k=1,..,n}U_{ik}<xa_n\right)\\
&=(1-\mathbb{P}(U_{i2}>a_nx))^n\\
&=(1-L(a_nx)(a_nx)^{-\alpha})^n\\
&\leq 2e^{-x^{-\alpha}}.
\end{aligned}$$

Then if $n \gg 1$, $\mathbb{P}\left(B_n^{(i)} < n^{\varepsilon}\right) \leq 2e^{-n^{\varepsilon\alpha}}$. Define $\mathcal{J} = \{i \,:\, Z_i \leq b_n\}$.

$$\begin{aligned}\mathbb{P}\left(|\mathcal{J}| < n - \sqrt{n}\right) =& \mathbb{P}\left(\sum_{i=1}^{n} \mathbb{1}_{Z_i > b_n} \geq \sqrt{n}\right)\\ &\leq e^{-\sqrt{n}}\left(\mathbb{E}\left[\, e^{\mathbb{1}_{Z_1 > b_n}}\,\right]\right)^n\\ &\leq e^{-\sqrt{n}}\left(1 + e\mathbb{P}\left(Z_1 \geq b_n\right)\right)^n\\ &\leq e^{-\sqrt{n} + c_0 b_n^{-\alpha} n^{1+\varepsilon\alpha}} \leq e^{-n^{\delta}}\end{aligned}$$

for a positive δ, where c_0 is a positive constant, if one takes $b_n = a_n n^{-2\gamma/\alpha}$, for ε small enough. It follows we can reduce to prove the statement under the condition $\{|\mathcal{J}| \geq n - \sqrt{n}\}$, moreover we can consider any fixed $\mathcal{I} \subset \{1, ..., n\}$ such that $|\mathcal{I}| \geq n - \sqrt{n}$. Assume $\mathcal{I} = \{1, 2, ..., n'\}$ with $n' \geq n - \sqrt{n}$. Let be $I = \text{span}(e_i\,;\, i \in \mathcal{I})$ and π_I the orthogonal projection on I. Define $W_2 = \pi_I(W_1) = \pi_I(\text{span}(W_1, e_1))$. By Grassmann inequality

$$\dim(W_2) \geq \dim(W_1 \cap I) \geq \dim W_1 + \dim I - \dim(W_1 \cup I) \geq d - \sqrt{n}.$$

Define

$$W' = \text{span}(W_2, \mathbb{E}\left[\,\pi_I(Z)\,|\,\mathcal{J} = \mathcal{I}\,\right]) \qquad Y = Z - \mathbb{E}\left[\,\pi_I(Z)\,|\,\mathcal{J} = \mathcal{I}\,\right])$$

Define P the matrix of the orthogonal projection on $(W')^{\perp}$ in $\mathbb{C}^{n'}$. Then

$$\mathbb{E}\left[\,\text{dist}^2(Y, W')|\mathcal{J} = \mathcal{I}\,\right] = \mathbb{E}\left[\,\|PY\|^2|\mathcal{J} = \mathcal{I}\,\right] \tag{4.3.5}$$

$$= \mathbb{E}\left[\,Y_1^2|\mathcal{J} = \mathcal{I}\,\right]\text{tr}P \tag{4.3.6}$$

By lemma $\mathbb{E}\left[\,Y_1^2|\mathcal{J} = \mathcal{I}\,\right] \geq \mathbb{E}\left[\,Z_1^2|\mathcal{J} = \mathcal{I}\,\right] = \mathbb{E}\left[\,Z_1^2|\mathcal{J} = \mathcal{I}\,\right] \geq \frac{c(\alpha)}{2n} b_n^{2-\alpha}$. Since $\text{tr}P \geq n' - \dim W' \geq \frac{1}{2}(n - d)$,

$$\mathbb{E}\left[\,\text{dist}^2(Y, W')|\mathcal{J} = \mathcal{I}\,\right] \geq \frac{c(\alpha)}{4} b_n^{2-\alpha} \frac{n-d}{n} \geq n^{(1-2\gamma)\frac{2}{\alpha} + \gamma - \varepsilon},$$

for $\varepsilon > 0$, small enough. Under $\mathbb{P}\left(\cdot\,|\mathcal{J} = \mathcal{I}\right)$, $b_n^{-1}(Y_1, ..., Y_n)$ is a vector of centered and independent variables on $\{z \in \mathbb{C}\,;\,|z| < 1\}^{n'}$, thus by Talagrand concentration inequality, applied to $F(x) = \text{dist}(x, W')$,

$$\mathbb{P}\left(|\text{dist}(Y, W') - M\text{dist}(Y, W')| > r\,|\mathcal{J} = \mathcal{I}\right) \leq 4e^{-\frac{r^2}{4b_n^2}}.$$

□

We will need and analogous of Lemma C.1 for the truncated moments of [10].

Lemma 4.3.2. *If Z is as in equation (4.3.2), for any $p \geq 1$, there exist positive constants c_1, c_2 such that,*

$$c_1 t^{p-\alpha} \leq \mathbb{E}\,[\, Z^p \,|\, Z \leq t \,] \leq c_2 t^{p-\alpha}.$$

Proof. Since

$$\lim_{t \to +\infty} \frac{\mathbb{E}\,[\, Z^p \,|\, Z \leq t \,]}{\mathbb{E}\,[\, Z^p \mathbb{1}_{Z \leq t} \,]} = 1,$$

we will show that there exist two constants $c_1, c_2 > 0$, both independent of t, such that

$$c_1 t^{p-\alpha} \leq \mathbb{E}\,[\, Z^p \mathbb{1}_{Z \leq t} \,] \leq c_2 t^{p-\alpha}.$$

($\leq$) For the upper bound we just use the fact that $U > 0$, and equation (4.3.4).

$$\begin{aligned}
\mathbb{E}\,[\, Z^p \, \mathbb{1}_{Z \leq t} \,] = \mathbb{E}\,[\, Z^p \, \mathbb{1}_{Z^p \leq t^p} \,] &= \int_0^{t^p} \mathbb{P}\,(Z^p > x)\, dx \\
&= \int_0^{t^p} \mathbb{P}\left(a_n \frac{U}{U + a_n B_n} > x^{1/p} \right) dx \\
&\leq \int_0^{t^p} x^{-\alpha/p} dx \\
&= \frac{p}{p-\alpha} t^{p-\alpha}
\end{aligned}$$

($\geq$)

$$\begin{aligned}
\mathbb{E}\,[\, Z^p \, \mathbb{1}_{Z \leq t} \,] &= \int_0^{t^p} \mathbb{P}\,(Z^p > x)\, dx \\
&= \int_0^{t^p} \mathbb{P}\left(\frac{U}{U + a_n B_n} > \frac{x^{1/p}}{a_n} \right) dx \\
&= \int_0^{t^p} \mathbb{P}\left(\frac{a_n B_n}{U} < \frac{a_n}{x^{1/p}} - 1 \right) dx \\
&\geq \int_0^{t^p} \mathbb{P}\left(\frac{a_n B_n}{U} < \frac{a_n}{2x^{1/p}} \right) dx \\
&= \int_0^{t^p} \mathbb{P}\left(\frac{U}{a_n B_n} > \frac{2x^{1/p}}{a_n} \right) dx \\
&\geq \frac{c(\alpha)}{2n} \int_0^{t^p} \left(\frac{2x^{1/p}}{a_n} \right)^{-\alpha} dx \\
&\geq \frac{c(\alpha)}{2^{1+\alpha}} \int_0^{t^p} x^{-\alpha/p} dx \\
&= \frac{c(\alpha)}{2^{1+\alpha}} \frac{p}{p-\alpha} t^{p-\alpha}
\end{aligned}$$

□

Proposition 4.3.3. *Assume* $0 < \gamma < \alpha/4$*, there exists an event* E *such that,*

$$\mathbb{E}\left[\,\mathrm{dist}^{-2}(R,W);E\,\right] \le c\,(n-d)^{-\frac{2}{\alpha}} \qquad \mathbb{P}\left(E^C\right) \le c\,n^{-(1-2\gamma)/\alpha}$$

Proof. As in the proof of proposition 4.3.1, define the vector subspaces $W_1 = \mathrm{span}(W, e_1)$, so that

$$\mathrm{dist}(R,W) \ge \mathrm{dist}(a_n\left(\frac{U_{11}}{\rho_1},...,\frac{U_{n1}}{\rho_n}\right), W_1) = \mathrm{dist}((Z_1,...,Z_n), W').$$

If we define $Z_i = a_n\frac{U_{1i}}{U_{1i}+a_nB_n^{(i)}}$ for $i = 1,...,n$, and $Z := (Z_1,...,Z_n)$. Define the set $\mathcal{I} := \{i \in \{1,...,n\}$ such that $Z_i \le \sqrt{a_n}\}$. We have

$$\mathbb{P}\left(|\mathcal{I}| < n - n^{1/2+\varepsilon}\right) \le e^{-n^\delta} \tag{4.3.7}$$

for a positive δ, see proof of proposition 4.3.1. Then is sufficient to prove that for any set $I \subset \{1,...,n\}$ with $|I| \ge n - n^{1/2+\varepsilon}$,

$$\mathbb{E}\left[\,\mathrm{dist}^{-2}(R,W);E_I\,|\,I=\mathcal{I}\,\right] \le \kappa(n-d)^{-2/\alpha},$$

for some event E_I such that $\mathbb{P}\left((E_I)^c\,|\,I=\mathcal{I}\right) \le n^{-(1-2\gamma)/\alpha}$. The we will set $E := E_I \cap \{|\mathcal{I}| > n - n^{1/2+\varepsilon}\}$. Without loss of generality we can assume $I = \{1,2,..,n'\}$ with $n' \ge n - n^{1/2+\varepsilon}$. Let $\pi_I(\cdot)$ be the orthogonal projection onto $\mathrm{span}(e_i : i = 1,...,n')$. If $W_2 := \pi_I(W_1)$, set

$$W' := \mathrm{span}(W_2, \mathbb{E}\left[\,\pi_I(Z)\,|\,I=\mathcal{I}\,\right]).$$

Note that $d - n^{1/2+\varepsilon} \le \dim W' \le d+2$. Call $Y = Z - \mathbb{E}\left[\,\pi_I(Z)\,|\,I=\mathcal{I}\,\right]$, so that Y is a vector of centered random variables under $\mathbb{P}\left(\cdot|I=\mathcal{I}\right)$. Then

$$\mathrm{dist}(R,W) \ge \mathrm{dist}(Z,W_1) \ge \mathrm{dist}(Y,W').$$

Denote P the matrix of the orthogonal projection to $(W')^\perp$ in $\mathbb{C}^{n'}$. By construction,

$$\mathbb{E}\left[\,\mathrm{dist}Y,W'\,\right] = \mathbb{E}\left[\,\sum_{i,j=1}^n Y_iP_{ij}Y_j\,|\,I=\mathcal{I}\,\right] = \mathbb{E}\left[\,Y^2\,|\,I=\mathcal{I}\,\right]\mathrm{tr}P.$$

Define $S = \sum_{i=1}^n P_{ii}Y_i^2$, where $P_{ii} = (e_i, Pe_1) \in [0,1]$, $\sum_i P_{ii} = \mathrm{tr}P$ and $n-(d+1) \le \mathrm{tr}P \le n-d$.

$$\begin{aligned}\mathbb{E}\left[\,(\mathrm{dist}^2(Y,W')-S)^2\,|\,I=\mathcal{I}\,\right] &= \mathbb{E}\left[\,\Big(\sum_{i\neq j}^n Y_iP_{ij}Y_j\Big)^2\,|\,I=\mathcal{I}\,\right]\\ &\le 2\mathbb{E}\left[\,Y_1^2\,|\,I=\mathcal{I}\,\right]^2\mathrm{tr}P^2 \le 2\mathbb{E}\left[\,Z_1^2\,|\,I=\mathcal{I}\,\right]^2\mathrm{tr}P^2\\ &\le c\,a_n^2\frac{n-d}{n}.\end{aligned}$$

by Lemma 4.3.2. Now let ζ be a random variable with one-side α stable distribution with $\alpha \in (0,1)$. By [10, Lemma 3.6], if we define the event

$$\Gamma = \left\{ \sum_{i=1}^{n} P_{ii} Z_i^2 \geq \varepsilon (n-d)^{\frac{2}{\alpha}} \zeta \right\},$$

then $\mathbb{P}\left(\Gamma^C\right) \leq e^{-n^\delta}$, for some positive δ and ε. Since for any $a, b \in \mathbb{R}$ holds $(a-b)^2 \geq \alpha^2/2 - b^2$, then $S \geq \frac{1}{2} S_a - S_b$ with

$$S_a = \sum_{i=1}^{n} P_{ii} Z_i^2 \qquad S_b = \sum_{i=1}^{n} P_{ii} \mathbb{E}\left[Z_i \,|\, I = \mathcal{I} \right]^2 = (n-d)\frac{a_n^2}{n^2}.$$

Let G_1 be the event $S_a \geq 3S_b$. By hypothesis $\gamma \leq \alpha/4$, then there exists a positive ε,

$$\mathbb{P}\left(G_1^c \cap \Gamma \,|\, I = \mathcal{I}\right) \leq \mathbb{P}\left(\zeta \leq c\,(n-d)^{-\frac{2}{\alpha}}(n-d)\frac{a_n^2}{n^2} \,|\, I = \mathcal{I}\right) \leq \mathbb{P}\left(\zeta \leq c\, n^{-\varepsilon} \,|\, I = \mathcal{I}\right).$$

Since $\mathbb{E}\left[\zeta^{-m}\right]$ is finite for any positive m, by Markov's inequality $\mathbb{P}\left(Z \leq t\right) \leq t^{-m}\mathbb{E}\left[\zeta^{-m}\right]$ .Thus for any $p > 0$, there exist a positive constant κ_p such that

$$\mathbb{P}\left(G_1^c \cap \Gamma\right) \leq \kappa_p n^{-p}.$$

Set $G = G_1 \cap \Gamma$. On the event G $S \geq \frac{1}{6} S_a \geq \frac{\varepsilon}{6}(n-d)^{\frac{1}{\beta}}\zeta = \frac{\varepsilon}{6}(n-d)^{\frac{2}{\alpha}}\zeta$, then for some constant $c_1 > 0$,

$$\mathbb{E}\left[S^{-2}; G\right] \leq c_1 (n-d)^{-4/\alpha}\mathbb{E}\left[\zeta^{-2}\right] \leq c_2 (n-d)^{-4/\alpha}.$$

Thus $\mathbb{E}\left[S^{-2}; G \,|\, I = \mathcal{I}\right] = O((n-d)^{-4/\alpha})$. Since $a \geq b/2$ implies $|a-b| \geq |b|/2$, using Markov's and Cauchy-Schwarz's inequalities

$$\begin{aligned}
\mathbb{P}\left(\mathrm{dist}^2(Y, W') \leq S/2; G \,|\, I = \mathcal{I}\right) &\leq \mathbb{P}\left(\left|\frac{\mathrm{dist}^2(Y, W') - S}{S}\right| \geq \frac{1}{2}; G \,|\, I = \mathcal{I}\right) \\
&\leq 2\mathbb{E}\left[\frac{|\mathrm{dist}^2(Y, W') - S|}{S}; G \,|\, I = \mathcal{I}\right] \\
&\leq 2\sqrt{\mathbb{E}\left[|\mathrm{dist}^2(Y, W') - S|^2; G \,|\, I = \mathcal{I}\right]\mathbb{E}\left[S^{-2}; G \,|\, I = \mathcal{I}\right]}.
\end{aligned}$$

If we set $G_2 = \{\mathrm{dist}^2(Y, W') \geq S/2\}$, then $\mathbb{P}\left(G_2^c \cap G \,|\, I = \mathcal{I}\right) = O(a_n n^{-1/2}(n-d)^{1/2-\alpha/2})$ that is $\mathbb{P}\left(G_2^c \cap G\right) = O(n^{-(1-2\gamma)/\alpha})$ under our hypothesis on γ. Furthermore, by Cauchy-Schwarz inequality

$$\mathbb{E}\left[\mathrm{dist}^{-2}; G_2 \cap G \,|\, I = \mathcal{I}\right] \leq \mathbb{E}\left[S^{-1}; G \,|\, I = \mathcal{I}\right] = O((n-d)^{-2/\alpha}).$$

Define $E = G_2^c \cap G$.

$$\begin{aligned}
\mathbb{P}\left(E^c\right) &\leq \mathbb{P}\left(\Gamma^c\right) + \mathbb{P}\left(\Gamma \cap G_1^c\right) + \mathbb{P}\left(\Gamma \cap G_1 \cap G_2^c\right) \\
&\leq c_1 e^{-n\delta} + \kappa_{(1-2\gamma)/\alpha} n^{-(1-2\gamma)/\alpha} + c_2 n^{-(1-2\gamma)/\alpha}.
\end{aligned}$$

□

4.4 Uniform integrability

For $\delta \in (0,1)$ define $K_\delta = [\delta, \delta^{-1}]$. We have to prove the for any $\varepsilon > 0$, for a.a. $z \in \mathbb{C}$

$$\lim_{\delta \to 0} \sup_n \mathbb{P}\left(\int_{K_\delta^C} |\log(x)| \nu_{X_n - z\mathbb{1}}(dx) > \varepsilon \right) = 0.$$

Take $z \neq 1$. For any and any $t \geq 1$, by Markov inequality

$$\int_t^{+\infty} \log(x)\nu_{X_n - z\mathbb{1}}(dx) \leq \frac{1}{t}\int_1^{+\infty} x\log(x)\nu_{X_n - z\mathbb{1}}(dx) \leq \frac{1}{t}\int_1^{+\infty} x^2 \nu_{X_n - z\mathbb{1}}(dx) \leq \frac{2+2|z|^2}{t},$$

see proof of Lemma 4.1.1. Then the integral over $(\delta^{-1}, +\infty)$ is not an issue. Call now $\sigma_i := s_i(X_n - z\mathbb{1})$, for $i = 1, ..., n$. For the other side, it is sufficient to prove that for any sequence $(\delta_n)_{n\geq 1}$ tending to 0,

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}_{\{\sigma_{n-i} \leq \delta_n\}} \log \sigma_{n-i}^{-2}$$

converges to 0 in probability. By the invertibility lemma we know that $\sigma_n \geq n^{-r}$ for some $r > 0$, if $n \gg 1$, and we can bound some terms in the sum. Let $\gamma \in (0, \alpha/4)$ to be fixed later, for $0 \leq i \leq n^{1-\gamma}$ we use this bound, so that

$$\frac{1}{n}\sum_{i=0}^{n^{1-\gamma}} \log \sigma_{n-i}^{-2} \leq \frac{1}{n} n^{1-\gamma} \log(n^{2r}) \to 0.$$

Call A_n the matrix of the first $n - i/2$ rows of $a_n(X_n - z\mathbb{1})$ and $\vartheta_1 \geq \cdots \geq \vartheta_{n-i/2}$ its singular values. By Cauchy interlacing lemma, see e.g. [16],

$$\sigma_{n-i} \geq \frac{\vartheta_{n-i}}{a_n}$$

and by Tao-Vu negative second moment (lemma 1.2.5),

$$\vartheta_1^{-2} + \cdots + \vartheta_{n-i/2}^{-2} = \mathrm{dist}_1^{-2} + \cdots \mathrm{dist}_{n-i/2}^{-2}$$

where $\mathrm{dist}_j = \mathrm{dist}(R_j, R_{-j})$, as defined in the lemma, is the distance of the j-th row from the subspace spanned by the other rows of A_n.Then

$$\frac{i}{2}\sigma_{n-i}^{-2} \leq a_n^2 \sum_{j=1}^{n-i/2} \mathrm{dist}_j^{-2}.$$

Since the dimension d of the span of the rows is at most $n - i/2$ we can define F_n the event that for all $1 \leq j \leq n - i/2$ $\mathrm{dist}_j \geq n^{(1-2\gamma)/\alpha}$ and use Proposition 4.3.1, to bound $\mathbb{P}\left(F_n^C\right) \leq e^{-n^\delta}$, for a positive δ. Then

$$\mathbb{E}\left[i\sigma_{n-i}^{-2}\mathbb{1}_{F_n} \right] \leq 2a_n^2 n \mathbb{E}\left[\mathrm{dist}_1^{-2} \mathbb{1}_{F_n} \right].$$

Since we are on F_n, $\mathrm{dist}_1 \geq n^{(1-2\gamma)/\alpha}$. By Proposition 4.3.3 there exists an event E independent from all rows $j \neq 1$ such that $\mathbb{P}\left(E^C\right) \leq n^{-(1-2\gamma)/\alpha}$ and for any $W \subset \mathbb{C}$ with dimension $d < n - n^{1-\gamma}$,

$$\mathbb{E}\left[\,\mathrm{dist}(R_1, W)^{-2}\mathbb{1}_E\,\right] \leq c_0(n-d)^{-2/\alpha}.$$

Since d is at most $n - i/2 \leq n - 2n^{1-\gamma}$, then $\mathbb{E}\left[\,\mathrm{dist}_1\mathbb{1}_E\,\right] \leq ci^{-2/\alpha}$. Therefore

$$\begin{aligned}\mathbb{E}\left[\,\mathrm{dist}_1^{-2}\mathbb{1}_{F_n}\,\right] &\leq \mathbb{E}\left[\,\mathrm{dist}_1\mathbb{1}_E\,\right] + \mathbb{P}\left(E^C\right) n^{-2(1-2\gamma)/\alpha} \\ &\leq c_0(i^{-2/\alpha} + n^{-3(1-2\gamma)/\alpha}).\end{aligned}$$

For a suitable γ it holds $3(1-2\gamma)/\alpha > 2/\alpha$, then $n^{-3(1-2\gamma)/\alpha} \leq i^{-2/\alpha}$ and

$$\mathbb{E}\left[\,i\sigma_{n-i}^{-2}\mathbb{1}_{F_n}\,\right] \leq 2a_n^2 n\mathbb{E}\left[\,\mathrm{dist}_1^{-2}\mathbb{1}_{F_n}\,\right] \leq 2c_0 a_n^{-2} n i^{-2/\alpha} = 2c_0(c^{1/\alpha}n^{1/\alpha})^{-2} n i^{-2/\alpha}.$$

Recall that, if we consider variables in $\mathbb{H}_\alpha^*$, a_n has the form $a_n = c^{1/\alpha}n^{1/\alpha}$ and

$$\mathbb{E}\left[\,\sigma_{n-i}^{-2}\mathbb{1}_{F_n}\,\right] \leq \tilde{c}\left(\frac{n}{i}\right)^{1+2/\alpha}$$

Finally by Markov inequality,

$$\mathbb{P}\left(\sigma_{n-i} \leq \delta_n\right) \leq \mathbb{P}\left(F_n^C\right) + c_0\delta_n^2\mathbb{E}\left[\,\sigma_{n-i}^{-2}\mathbb{1}_{F_n}\,\right] \leq e^{-n^\delta} + \tilde{c}\delta_n^2\left(\frac{n}{i}\right)^{1+2/\alpha}.$$

It follows that for any sequence δ_n we can define $\varepsilon_n = \delta_n^{1/(1+2/\alpha)}$ so that

$$\mathbb{P}\left(\sigma_{n-\varepsilon_n n} \leq \delta_n\right) \leq e^{-n^\delta} + \tilde{c}\delta_n^2\left(\frac{n}{\varepsilon_n n}\right)^{1+2/\alpha} = e^{-n^\delta} + \tilde{c}\delta_n \to 0.$$

Then is sufficient to prove

$$\frac{1}{n}\sum_{i=n^{1-\gamma}}^{\varepsilon_n n} \log\sigma_{n-i}^{-2}$$

given F_n, converges to 0 in probability. Thus, for any fixed $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(\frac{1}{n}\sum_{i=n^{1-\gamma}}^{\varepsilon_n n}\log\sigma_{n-i}^{-2} > \varepsilon\,\middle|\,F_n\right) &\leq \varepsilon^{-1}\mathbb{E}\left[\frac{1}{n}\sum_{i=n^{1-\gamma}}^{\varepsilon_n n}\log\sigma_{n-i}^{-2}|F_n\right] \\ &\leq \varepsilon^{-1}\frac{1}{n}\sum_{i=n^{1-\gamma}}^{\varepsilon_n n}\log\mathbb{E}\left[\,\sigma_{n-i}^{-2}|F_n\,\right] \\ &\leq \frac{c_0}{n}\sum_{i=1}^{\varepsilon_n n}\log\left(\frac{n}{i}\right)\end{aligned} \tag{4.4.1}$$

Since

$$\begin{aligned}\sum_{i=1}^{n}\log(i) &= \sum_{i=2}^{n}\log(i) \\ &\geq \sum_{i=2}^{n}\int_{i-1}^{i}\log(x)dx \\ &= \int_1^n \log(x)dx = n\log n - n + 1.\end{aligned}$$

Equation (4.4.1) becomes,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{\varepsilon_n n}\log\left(\frac{n}{i}\right) &= \varepsilon_n\log(n) - \sum_{i=1}^{\varepsilon_n n}\log(i)\\ &\le \varepsilon_n\log(n) - \varepsilon_n\log\varepsilon_n n + \varepsilon_n - \frac{1}{n}\\ &= \varepsilon_n - \varepsilon_n\log\varepsilon_n - \frac{1}{n}.\end{aligned}$$

Which tends to 0 as $n\to\infty$.

4.5 Proof of Theorem 1.6.1

Since we proved that $|\log(x)|$ is uniformly integrable with respect to the measure $\nu_{X_n-z\mathbb{1}}$, for any $z\in\mathbb{C}\setminus\{1\}$, in probability and by theorem1.6.2, $\nu_{X_n-z\mathbb{1}}$ converges almost surely to a measure $\nu_{\alpha,z}$, as a consequence of Girko's Hermitization method, lemma 1.2.3, μ_{X_n} converges in probability to a measure μ_α, that satisfies for any $z\in\mathbb{C}\setminus\{1\}$,

$$U_{\mu_\alpha}(z) = -\int\log(x)\nu_{\alpha,z}(dx).$$

Since $U_{\mu_\alpha}(z)$ is deterministic, we can improve the convergence to almost sure, showing that there exists a deterministic sequence L_n such that, almost surely

$$\lim_{n\to\infty}(U_{\mu_{X_n}}(z) - L_n) = 0.$$

By Lemma 4.1.1 and Lemma 4.2.3, there exists $r>0$ such that, almost surely

$$\operatorname{supp}(\nu_{X_n-z\mathbb{1}}) \subset [s_n(X_n-z\mathbb{1}), s_1(X_n-z\mathbb{1})] \subset [n^{-r}, n^r],$$

if n is large enough. Define $f_n(x) = \mathbb{1}_{x\in[n^{-r},n^r]}\log(x)$, a.s. for $n\gg 1$,

$$U_{\mu_{X_n}}(z) = \int f_n(x)\nu_{X_n-z\mathbb{1}}(dx).$$

Since $\|f_n\|_{TV}\le c\log n$, by theorem 1.1.3,

$$\mathbb{P}\left(\left|\int f_n(x)\nu_{X_n-z\mathbb{1}}(dx) - \mathbb{E}\left[\int f_n(x)\nu_{X_n-z\mathbb{1}}(dx)\right]\right| > \varepsilon\right) \le 2\exp\left\{-2\frac{\varepsilon n}{(c\log n)^2}\right\}.$$

Call $L_n = \mathbb{E}\left[\int f_n(x)\nu_{X_n-z\mathbb{1}}(dx)\right]$, by first Borel-Cantelli lemma, a.s.

$$\lim_{n\to\infty}(U_{\mu_{X_n}}(z) - L_n) = 0.$$

4.6 Non triviality of μ_α

Theorem 1.6.1 gives an existence result for the limiting spectral measure μ_α, in this section we observe simple facts to avoid the possibility of a trivial limiting spectral measure. We indeed will show that μ_α is neither a Dirac's delta in 0, nor concentrated on the boundary of the unitary disc of $\mathbb{C}$.

The first observation is that, since the logarithmic potential of μ_α is not infinite, μ_α is not a delta. By proposition 3.3.3 it is sufficient to check the finiteness of the integral in 0. By almost sure convergence of $U_{\mu_{X_n}}$, we have

$$\int_0^1 |\log(t)|\nu_{X_n-z}(dt) \xrightarrow[n\to+\infty]{} \int_0^1 |\log(t)|\nu_{\alpha,z}(dt),$$

in particular, for $z=0$, we have, almost surely,

$$U_{\mu_\alpha}(0) = \lim_{n\to\infty}\int_0^{+\infty} |\log(t)|\nu_{X_n}(dt).$$

Moreover, by uniform integrability in probability

$$\lim_{y\to\infty}\sup_n \mathbb{P}\left(\int_0^{+\infty} |\log(t)|\nu_{X_n}(dt) > y\right) = 0,$$

so that $\int_0^{+\infty} |\log(t)|\nu_{X_n}(dt)$ can not diverge to infinity as $n\to\infty$ (it would have implied $\mathbb{P}\left(\int_0^{+\infty} |\log(t)|\nu_{X_n}(dt) > y\right) \xrightarrow[n\to+\infty]{} 1$, for any positive y).

The weak convergence of the order statistics of the vector $(X_{1,1},...,X_{1,n})$ to a $PD(\alpha)$ will imply that the limiting spectral distribution does not concentrate on $\{z\in\mathbb{C} : |z| = 1\}$. Indeed, by theorem 1.6.1 and Weyl's equation (1.2.2), we have

$$
\begin{aligned}
\int |z|^2 \mu_\alpha(dz) &= \lim_{n\to\infty} \mathbb{E}\left[\int |z|^2 \mu_{X_n}(dz)\right] \\
&\le \lim_{n\to\infty} \mathbb{E}\left[\int t^2 \nu_{X_n}(dz)\right] \\
&= \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n} s_i(X_n)^2\right] \\
&= \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n} \lambda_i(X_n X_n^*)\right] \\
&= \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}\left[\ \mathrm{Tr}(X_n X_n^*)\right] \\
&= \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}\left[\sum_{i,j=1}^{n} X_{i,j}^2\right] \\
&= \lim_{n\to+\infty} \mathbb{E}\left[\sum_{j=1}^{n} X_{1,j}^2\right] \\
&= \mathbb{E}\left[\sum_{j\ge 1} \zeta_j^2\right] < 1.
\end{aligned}
$$

Where $\{\zeta_j\}_{j\ge 1}$ has Poisson-Dirichlet law of index α.

Bibliography

[1] D. Aldous, *Asymptotics in the random assignment problem*, Probab. Theory Related Fields, 93 (1992), pp. 507–534.

[2] D. Aldous and J. M. Steele, *The objective method: probabilistic combinatorial optimization and local weak convergence*, in Probability on discrete structures, vol. 110 of Encyclopaedia Math. Sci., Springer, Berlin, 2004, pp. 1–72.

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

[4] Z. D. Bai, *Circular law*, Ann. Probab., 25 (1997), pp. 494–529.

[5] S. Belinschi, A. Dembo, and A. Guionnet, *Spectral measure of heavy tailed band and covariance random matrices*, Comm. Math. Phys., 289 (2009), pp. 1023–1055.

[6] G. Ben Arous and A. Guionnet, *The spectrum of heavy tailed random matrices*, Comm. Math. Phys., 278 (2008), pp. 715–751.

[7] C. Bordenave, *Lecture notes: A short course on random matrices*, 2014.

[8] C. Bordenave, P. Caputo, and D. Chafaï, *Spectrum of large random reversible Markov chains: two examples*, ALEA Lat. Am. J. Probab. Math. Stat., 7 (2010), pp. 41–64.

[9] ——, *Spectrum of large random reversible Markov chains: heavy-tailed weights on the complete graph*, Ann. Probab., 39 (2011), pp. 1544–1590.

[10] ——, *Spectrum of non-Hermitian heavy tailed random matrices*, Comm. Math. Phys., 307 (2011), pp. 513–560.

[11] ——, *Circular law theorem for random Markov matrices*, Probab. Theory Related Fields, 152 (2012), pp. 751–779.

[12] ——, *Spectrum of Markov generators on sparse random graphs*, Comm. Pure Appl. Math., 67 (2014), pp. 621–669.

[13] C. Bordenave and D. Chafaï, *Around the circular law*, Probab. Surv., 9 (2012), pp. 1–89.

[14] P. Cizeau and J. P. Bouchaud, *Theory of lévy matrices*, Phys. Rev. E, 50 (1994), pp. 1810–1822.

[15] J. Ginibre, *Statistical ensembles of complex, quaternion, and real matrices*, J. Mathematical Phys., 6 (1965), pp. 440–449.

[16] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[17] R. LePage, M. Woodroofe, and J. Zinn, *Convergence to a stable distribution via order statistics*, Ann. Probab., 9 (1981), pp. 624–632.

[18] C. McDiarmid, *On the method of bounded differences*, in Surveys in combinatorics, 1989 (Norwich, 1989), vol. 141 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 1989, pp. 148–188.

[19] M. L. Mehta, *Random matrices and the statistical theory of energy levels*, Academic Press, New York-London, 1967.

[20] G. Pan and W. Zhou, *Circular law, extreme singular values and potential theory*, J. Multivariate Anal., 101 (2010), pp. 645–656.

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

[22] M. Rudelson and R. Vershynin, *The Littlewood-Offord problem and invertibility of random matrices*, Adv. Math., 218 (2008), pp. 600–633.

[23] E. B. Saff and V. Totik, *Logarithmic potentials with external fields*, vol. 316 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1997. Appendix B by Thomas Bloom.

[24] M. Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*, Inst. Hautes Études Sci. Publ. Math., (1995), pp. 73–205.

[25] T. Tao, *Outliers in the spectrum of iid matrices with bounded rank perturbations*, Probab. Theory Related Fields, 155 (2013), pp. 231–263.

[26] T. Tao and V. Vu, *Random matrices: the circular law*, Commun. Contemp. Math., 10 (2008), pp. 261–307.

[27] ——, *Random matrices: universality of ESDs and the circular law*, Ann. Probab., 38 (2010), pp. 2023–2065. With an appendix by Manjunath Krishnapur.

[28] H. Weyl, *Inequalities between the two kinds of eigenvalues of a linear transformation*, Proc. Nat. Acad. Sci. U. S. A., 35 (1949), pp. 408–411.

[29] E. P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Ann. of Math. (2), 62 (1955), pp. 548–564.

[30] I. Zakharevich, *A generalization of Wigner's law*, Comm. Math. Phys., 268 (2006), pp. 403–414.