Solution of Matrix Dyson Equation for Random Matrices with Fast Correlation Decay

Sofiia Dubova, V.N. Karazin Kharkiv National University

December 14, 2018

Abstract

We consider the solution of Matrix Dyson Equation $-M(z)^{-1} = z + S(M(z))$, where entries of the linear operator $S: \mathbb{C}^{N \times N} \rightarrow \mathbb{C}^{N \times N}$ decay exponentially. We show that $M(z)$ also has exponential off-diagonal decay and can be represented as Laurent series with coefficients determined by entries of $S$. We also prove that for Hermitian random matrices with exponential correlation decay empirical density converges to the deterministic density obtained from $M(z)$. These results have already been proved in [1] with the resolvent method, here we give an alternate proof via conceptually much simpler moment method.

AMS Subject Classification (2010): 60B20.

Keywords: Correlated random matrices, Matrix Dyson equation.

1 Introduction

For Wigner matrices, i.e. Hermitian random matrices with centered i.i.d. entries, it has long been known that the empirical density of eigenvalues converges to the deterministic density $\rho$ called self-consistent density of states. Here $\rho$ is the semicircle density.

By dropping the assumption of identical distribution of entries, one gets a more general class of matrices, called the Wigner-type ensemble. For Wigner-type random matrix $H = (h_{xy})_{x,y=1}^N$ self-consistent density of states can be determined from the solution of vector equation

$$-m_x(z)^{-1} = z + [S_m(z)]_x,$$

where $S = (S_{xy})_{x,y=1}^N$ is the matrix of variances, i.e. $S_{xy} = \mathbb{E}|h_{xy}|^2$. The solution $m(z) = (m_1(z), \ldots, m_N(z))$ is a vector-valued function defined on the upper half plane $\mathbb{H} = \{z \in \mathbb{C} | \text{Im} z > 0\}$ and it satisfies the additional condition $\text{Im} m_x(z) > 0$. The solution is unique analytic function on $\mathbb{H}$. The average of
the entries of vector $m(z)$ is the Stieltjes transform of the self-consistent density of states $\rho$, i.e.

$$\langle m(z) \rangle = \frac{1}{N} \sum_{x=1}^{N} m_x(z) = \int_{\mathbb{R}} \frac{\rho(\tau) d\tau}{\tau - z}.$$  \hspace{1cm} (1.1)

The solution $m(z)$ has been studied in [4] and [6]. It was shown that $m(z)$ can be expressed as Laurent series for large $|z|$. The coefficients are expressed explicitly through the entries of variance matrix $S$.

In this paper we consider a more general class of random matrices. We replace the condition of independence of the entries by the condition on fast correlation decay. This means that the correlation of the entries decreases exponentially with the increase in distance between the entries in the matrix. For such random matrix $H$ of the size $N \times N$ consider the operator $S : \mathbb{C}^{N \times N} \to \mathbb{C}^{N \times N}$ defined by covariances of the entries. More precisely, $S(R) = \mathbb{E} H R H^*$ for each $R \in \mathbb{C}^{N \times N}$. The self-consistent density of states can be obtained from the solution of Matrix Dyson Equation (MDE):

$$- M(z)^{-1} = z + S(M(z)),$$  \hspace{1cm} (1.2)

where $M(z)$ is a matrix-valued function defined on $\mathbb{H}$. Additionally, we are only interested in the solution of MDE with positive imaginary part, i.e. $\text{Im} \ M(z) > 0$. Positivity is meant in the sense of scalar product $\langle R, T \rangle := \frac{1}{N} \text{Tr} R^* T$ in the space of matrices $\mathbb{C}^{N \times N}$.

MDE was studied in [5]. It has unique solution $M(z)$ holomorphic on $\mathbb{H}$. In the first part of the paper we generalize the results from [4] and [6] on the Laurent series representation of the solution of (1.1). We express solution of MDE as the Laurent series with coefficients expressed in terms of the entries of operator $S$. Using this formula, we prove off-diagonal decay of the solution for large $|z|$ (see Theorem 2.1).

In the other part of the paper we prove that the empirical distribution of eigenvalues converges to the self-consistent density of states obtained as the inverse Stieltjes transform of $\frac{1}{N} \text{Tr} M(z)$ using moment method.

Both results in this paper have already been achieved in [4] with the resolvent method and by fairly involved analysis of the matrix Dyson equation. The main goal here is to give shorter and conceptually simpler alternative proofs and demonstrate that the moment method extends to correlated Wigner matrices.

Acknowledgment: This work was done during a summer internship at László Erdős’ research group at IST Austria. The author is very grateful to him for suggesting the problem and for permanent help and guidance during the entire project.

2 Main results

2.1 Frame of the tree

First, we introduce some notation needed for formulation of our result.

2
Let $T_k$ denote the set of the rooted ordered trees $\Gamma = (V(\Gamma), E(\Gamma))$ with $k$ edges. The ordering means the following additional structure on the rooted tree: for every vertex $v \in V(\Gamma)$ the set of children of $v$ is ordered, i.e. children of $v$ are $w_1 < w_2 < \ldots < w_{c(v)}$, where $c(v)$ is the number of children of $v$. For every child $w$ of the vertex $v$ one could define the position of $w$ in the sequence of children of $v$. Let $n(w)$ denote this position. Since any vertex $w \in V(\Gamma) \setminus \{\text{root}\}$ has exactly one parent, $n(w)$ is defined for any such vertex. It is easy to see that the ordering uniquely defines an oriented realization of the tree in the plane, two planar trees being equivalent if they can be deformed into each other by an orientation preserving homeomorphism of the plane.

Every edge $e \in E(\Gamma)$ is incident with two vertices denoted by $e_-, e_+ \in V(\Gamma)$, where the sign indicates child-parent relation, i.e. $e_+$ is a child of $e_-$. For every ordered tree $\Gamma \in T_k$ one could walk around $\Gamma$ starting in the root and each step going from vertex $v$ to the first not yet visited child of $v$ or to the parent of $v$ if all children of $v$ have already been visited. Note that the walk ends at the root of $\Gamma$ and consists of $2k$ steps. This allows us to define the frame of $\Gamma$ denoted by $F(\Gamma)$, which is the path graph with $2k$ edges, fixed direction and association of its vertices with the vertices of $\Gamma$ defined by the following procedure. We start at the first vertex of the path and associate it with the root of $\Gamma$. Then, as we walk around $\Gamma$, each step the next vertex in the path is associated with the vertex of $\Gamma$ where this step arrives. This procedure also gives the association of each edge of the frame with the edge of $\Gamma$, as each step of the walk corresponds to the edge of $\Gamma$.

For $\Gamma \in T_0$ the frame $F(\Gamma)$ is a graph with one vertex and no edges.

Given the oriented planar realization of $\Gamma$, one could draw the frame of $\Gamma$ as the path starting at the vertex near the root of $\Gamma$ and going around the tree in counter-clockwise direction with each vertex drawn near its associated vertex and each edge running parallel with the associated edge of $\Gamma$ (see Figure 2.1).

Denote $[N] = \{1, \ldots, N\}$. It is easy to see that for each vertex $v \in V(\Gamma)$ there are $c(v) + 1$ vertices of $F(\Gamma)$ associated with $v$. For every $v \in V(\Gamma)$ we denote the vertices of $F(\Gamma)$ associated with $v$ by $p^v_0, p^v_1, \ldots, p^v_{c(v)}$ in the order of appearance in the frame. We now label the vertices of $F(\Gamma)$ by assigning a number from $[N]$ to each of them. More precisely, the vertex $p^v_n \in F(\Gamma)$
is labelled with a number \( x^n \in [N] \). A *labelling* of \( \mathcal{F}(\Gamma) \) is a set of arrays \( x = \{x(v) = (x^0_v, x^1_v, \ldots, x^n_v) \mid v \in V(\Gamma)\} \). Let \( \mathcal{X}(\Gamma) \) denote the set of the labellings of \( \mathcal{F}(\Gamma) \). Sets of the labellings with fixed first and last labels are denoted by

\[
\mathcal{X}_{ab}(\Gamma) := \{x \in \mathcal{X}(\Gamma) \mid x^{\text{root}}_0 = a, x^{\text{root}}_{c(\text{root})} = b\}, \ a, b \in [N].
\]

(2.1)

### 2.2 Solution of MDE

In this section we present our result on the form of MDE solution and its off-diagonal decay in large \(|z|\) regime. The proofs are presented in the section 3.

MDE has been studied in [5]. It has unique solution \( M(z) \) with positive imaginary part for \( z \in \mathcal{H} = \{\zeta \in \mathbb{C} \mid \text{Im}\, \zeta > 0\} \). Moreover, the solution holomorphic on \( \mathcal{H} \).

First, we formalize the idea of off-diagonal decay. Define the scalars \( S_{ab,cd} \) for the operator \( \mathcal{S} \) by the identity

\[
[S(R)]_{zt} = \sum_{y,z=1}^N S_{xy,zt} R_{yz}
\]

(2.2)

for any \( R \in \mathbb{C}^{N \times N} \).

We introduce norms on the space of matrices and on operators acting on matrices that reflect exponential decay of the matrix elements as they go farther from the diagonal.

For every matrix \( R = (R_{xy})_{x,y=1}^N \) and \( l > 0 \) we introduce the \( l\)-norm of \( R \):

\[
\|R\|_l = \max_{x,y=1} e^{\frac{l}{2}|x-y|}|R_{xy}|.
\]

(2.3)

Similarly we define the \( l\)-norm of the operator \( \mathcal{S} : \mathbb{C}^{N \times N} \to \mathbb{C}^{N \times N} \):

\[
\|\mathcal{S}\|_l = \max_{x,y,z,t=1} e^{\frac{l}{2}\rho(xy,zt)}|\mathcal{S}_{xy,zt}|,
\]

where

\[
\rho(xy,zt) := \min\{|x-z| + |y-t|, |x-t| + |y-z|\}
\]

(2.4)

is the symmetrized distance between the index pairs \((xy)\) and \((zt)\).

For every tree \( \Gamma \in \mathcal{T}_k \) define the matrix \( \text{val}(\Gamma) = (\text{val}_{ab}(\Gamma))_{a,b=1}^N \), where

\[
\text{val}_{ab}(\Gamma) := \sum_{x \in \mathcal{X}_{ab}(\Gamma)} \prod_{e \in E(\Gamma)} S_{x^{e-}_{n(e_+)-1}x^{e+}_0,x^{e+}_{c(e_+)}x^{e-}_{n(e_+)}},
\]

Note that the edges of the frame connecting the vertices with labels \( x^{e-}_{n(e_+)-1}, x^e_0 \) and the vertices with labels \( x^{e+}_{c(e_+)}, x^{e-}_{n(e_+)} \) are associated with the edge \( e \) of \( \Gamma \).

For \( \Gamma \in \mathcal{T}_0 \) we define \( \text{val}_{ab}(\Gamma) := \delta_{a,b} \).
Theorem 2.1. Suppose for some $l > 0$ and some constant $c > 0$ operator $S$ has finite $l$-norm and $\|S\|_l \leq \frac{c}{N}$. Then for any $\varepsilon > 0$ there is a constant $R(l, \varepsilon)$ such that the solution of the Matrix Dyson Equation for $|z|^2 > R(l, \varepsilon)$ is given by Laurent series

$$M(z) = -\sum_{k=0}^{\infty} C_k z^{-2k-1},$$

(2.5)

where

$$C_k := \sum_{\Gamma \in T_k} \text{val}(\Gamma)$$

(2.6)

for any nonnegative integer $k$.

Moreover, the solution $M(z)$ has finite $(1 + \varepsilon)l$-norm for $|z|^2 > R(l, \varepsilon)$.

2.3 Random matrices with exponentially decaying correlations

For any positive integer $N$ consider $N \times N$ Hermitian random matrix $H = \frac{1}{\sqrt{N}} W$, where $W = (w_{ij})_{i,j=1}^N$. Let $W$ satisfy the following assumptions:

(A) (Finite moments) For all positive integers $k$ there exists a constant $\mu_k$ such that

$$E|w_{xy}|^k \leq \mu_k$$

(2.7)

for any $x, y \in [N]$.

(B) (Decay of the cumulants) The cumulants of the entries of $w$ decay exponentially. More precisely, for some $l > 0$, constant $c$ and any $x, y, z, t \in [N]$

$$|\kappa(w_{xy}, w_{zt})| \leq ce^{-\frac{1}{l} \rho(x, z) \rho(y, t)}$$

(2.8)

and for any $k \geq 3$ and any $x_1, y_1, \ldots, x_k, y_k \in [N]$ the following inequality holds:

$$|\kappa(w_{x_1y_1}, \ldots, w_{x_ky_k})| \leq \prod_{\{i,j\} \in E(T_{\text{min}})} |\kappa(w_{x_iy_i}, w_{x_jy_j})|,$$

(2.9)

where $T_{\text{min}}$ is the minimal spanning tree of the complete graph with vertices $1, \ldots, k$ and edge length $d(\{i, j\}) = \rho(x_iy_i, x_jy_j)$ (see (2.4) for the definition of $\rho$).

These assumptions are the exponential analogue of assumptions made in [3].

Consider the Matrix Dyson Equation with operator $S$ such that $S(R) = EHRH$ for any $R \in \mathbb{C}^{N \times N}$. Then $S_{xy,zt} = \frac{1}{\sqrt{N}} Ew_{xy}w_{zt}$.

Let $\lambda_1, \ldots, \lambda_N$ be eigenvalues of $H$. Define the empirical distribution of the eigenvalues:

$$L^{(N)} = \frac{1}{N} \sum_{j=1}^{N} \delta_{\lambda_j}.$$
Consider the Stieltjes transform of $L^N$:

$$S^{(N)}(z) = \int_{\mathbb{R}} \frac{L^{(N)}(dx)}{x - z}$$

defined on upper half of the complex plane $\mathbb{H}$. 

**Theorem 2.2.** Under the assumptions (A) and (B), the conditions of the Theorem 2.1 hold and $\mathbb{E} S^{(N)}(z) - \frac{1}{N} \text{Tr} M(z)$ converges to 0 for all $z \in \mathbb{H}$.

In the Section 4 we get this result by proving convergence of the moments of empirical distribution to the normalized trace of $C_k$ from (2.6).

Theorem 2.2 implies that the averaged empirical measure $\mathbb{E} L^{(N)}_\cdot$ converges weakly to the self-consistent density of states $\rho$ obtained as inverse Stieltjes transform of $\frac{1}{N} \text{Tr} M(z)$. One could prove that empirical density $L^{(N)}_\cdot$ converges weakly in probability to $\rho$ using similar technique.

## 3 Solution of MDE

### 3.1 Decomposition of the ordered tree

For any two trees $\Gamma_1 \in \mathcal{T}_{n_1}$ and $\Gamma_2 \in \mathcal{T}_{n_2}$ define a new tree $\Gamma = \Gamma_1 \oplus \Gamma_2 \in \mathcal{T}_{n_1 + n_2 + 1}$ with

$$V(\Gamma) = V(\Gamma_1) \cup V(\Gamma_2),$$

$$E(\Gamma) = E(\Gamma_1) \cup E(\Gamma_2) \cup \{\text{root}(\Gamma_1), \text{root}(\Gamma_2)\}.$$ 

The root of $\Gamma_1$ is the root of $\Gamma$ and $\text{root}(\Gamma_2)$ is the last child of $\text{root}(\Gamma)$. Ordering of the children of every other vertex is the same as the ordering of children of the corresponding vertices of $\Gamma_1$ and $\Gamma_2$ (for example, see Figure ...). Every tree $\Gamma \in \mathcal{T}_k$ can be uniquely decomposed as $\Gamma = \Gamma_1 \oplus \Gamma_2$, where $\Gamma_1 \in \mathcal{T}_{n_1}$, $\Gamma_2 \in \mathcal{T}_{n_2}$ and $n_1 + n_2 = k - 1$ (for details see [4]). Note that the operation $\oplus$ is not commutative.

The frame of $\Gamma$ can be naturally split into the frames of $\Gamma_1$, $\Gamma_2$ and two edges associated with the edge $\{\text{root}(\Gamma_1), \text{root}(\Gamma_2)\}$ of $\Gamma$.

### 3.2 Explicit form of the solution of MDE

**Lemma 3.1.** For every $\Gamma \in \mathcal{T}_k$

$$\text{val}(\Gamma) = \text{val}(\Gamma_1) S(\text{val}(\Gamma_2)), \quad (3.1)$$

where $\Gamma = \Gamma_1 \oplus \Gamma_2$.

**Proof.** Let $v$ denote the last child of the root of $\Gamma$, i.e. the root of $\Gamma_2$. For a labelling $\mathbf{x} \in \mathcal{X}_{ab}(\Gamma)$ and fixed $c, d, e \in [N]$, consider the labellings $\mathbf{x}$ with
\(x^{v_{\text{root}}}_{(\text{root})-1} = c, x^{\nu}_0 = d\) and \(x^{v}_{\text{leaf}} = e\). Then \(x = (x_1, x_2)\) is naturally split into two labellings \(x_1 \in \mathcal{X}_{ac}(\Gamma_1)\) and \(x_2 \in \mathcal{X}_{de}(\Gamma_2)\). Therefore, (3.1) is equivalent to

\[
\text{val}_{ab}(\Gamma) = \sum_{c,d,e=1}^N \text{val}_{ac}(\Gamma_1) S_{cd,eb} \text{val}_{de}(\Gamma_2). \tag{3.2}
\]

By definition of the scalars \(S_{cd,eb}\),

\[
\sum_{d,e=1}^N S_{cd,eb} \text{val}_{de}(\Gamma_2) = [S(\text{val}(\Gamma_2))]_{cb}. \tag{3.3}
\]

Plugging in (3.3) into (3.2) we obtain

\[
\text{val}_{ab}(\Gamma) = \sum_{c=1}^N \text{val}_{ac}(\Gamma_1) [S(\text{val}(\Gamma_2))]_{cb} = [\text{val}(\Gamma_1)S(\text{val}(\Gamma_2))]_{ab}. \tag{3.4}
\]

Now we introduce the solution of MDE in the large \(|z|\) regime.

**Proposition 3.2.** Suppose that for some matrix norm \(\|C_k\| \leq CR^k\) for any \(k\) with some constants \(C\) and \(R\) (see (2.6) for the definition of \(C_k\)). Then the Laurent series (2.3) gives the solution of the Matrix Dyson Equation (1.2), defined on the domain of absolute convergence of the series (2.3), in particular for any \(z \in \mathbb{C}\) with \(|z|^2 > R\).

**Proof.** Notice that for \(z \in \mathbb{C}\) with \(|z|^2 > R\) the series (2.5) absolutely converges in the sense of the norm \(\|\|\|\|\).

Multiplying \(MDE\) by \(M(z)\) and introducing \(U(z) := -zM(z)\) we obtain

\[
U(z) = I + z^{-2}U(z)S(U(z)), \tag{3.4}
\]

where \(I\) is identity matrix.

Plugging in the form of the solution \(U(z) = \sum_{m=0}^{\infty} C_m z^{-2k}\) into (3.4) and equating the coefficients, we get

\[
C_k = \sum_{m=0}^{k-1} C_{k-1-m} S(C_m). \tag{3.5}
\]

Since for every \(\Gamma \in \mathcal{T}_k\) there is unique \(m \in \{0, 1, \ldots k-1\}\) and unique trees \(\Gamma_1\) and \(\Gamma_2\) such that \(\Gamma = \Gamma_1 \oplus \Gamma_2\), identity (3.5) is equivalent to

\[
\sum_{\Gamma \in \mathcal{T}_k} \text{val}(\Gamma) = \sum_{\Gamma \in \mathcal{T}_k} \text{val}(\Gamma_1)S(\text{val}(\Gamma_2)),
\]

which follows from Lemma 3.1.

\(\square\)
3.3 Convergence and exponential decay of the solution of MDE

Lemma 3.3. For every positive integer \( k \) and for any \( \varepsilon > 0 \) and \( x \geq 0 \)

\[
\int_{\mathbb{R}^k} e^{-|x-y_1|+|y_1-y_2|+\ldots+|y_k|} dy_1 \ldots dy_k \leq \left( \frac{1 + \varepsilon}{\varepsilon} \right)^{k-1} e^{-\frac{2}{t^2+1}x}. \tag{3.6}
\]

Proof. We introduce the function \( f(x) = e^{-|x|} \). Notice that the LHS of (3.6) is the \( k \)-fold convolution of \( f(x) \), denoted by \( f^*(x) \). Fourier transform of \( f \) is

\[
\mathcal{F}f(t) = \int_{\mathbb{R}} e^{-ixt} f(x) dx = \frac{2}{t^2 + 1}.
\]

Therefore,

\[
(\mathcal{F}f^k)(t) = (\mathcal{F}f(t))^k = \left( \frac{2}{t^2 + 1} \right)^k.
\]

Since \( f^k(x) \) and its Fourier transform are absolutely integrable, we have

\[
f^k(x) = \mathcal{F}^{-1}(\mathcal{F}f^k)(t) = \frac{2^k}{2\pi} \int_{\mathbb{R}} e^{ixt} / (t^2 + 1)^k dt. \tag{3.7}
\]

Consider the function \( g(z) = \frac{e^{ixz}}{(z^2 + 1)^k} \) on the complex plane and a contour \( C_R \) with positive orientation, which is the concatenation of the semicircle \( C^1_R = \{ \text{Re}^{i\theta} \mid \theta \in [0, \pi] \} \) and the segment \( C^2_R = [-R, R] \) with some large \( R \). Integrating \( g(z) \) over \( C_R \) we get

\[
\oint_{C_R} g(z) dz = \int_{C^1_R} g(z) dz + \int_{-R}^R \frac{e^{ixt}}{(t^2 + 1)^k} dt. \tag{3.8}
\]

Notice that \( \left| \frac{1}{(z^2 + 1)^k} \right| \leq \frac{1}{(R-1)^k} \) on the semicircle \( C_R \) with \( R > 1 \). Hence, for \( x > 0 \) the integral \( \int_{C^1_R} g(z) dz \) vanishes as \( R \) goes to infinity by Jordan’s lemma. Similarly, for \( x = 0 \):

\[
\left| \int_{C^1_R} \frac{1}{(z^2 + 1)^k} dz \right| \leq \text{length}(C^1_R) \cdot \left( \frac{1}{(R^2 - 1)^k} \right) = \frac{\pi R}{(R^2 - 1)^k} \to 0.
\]

For \( R > 1 \), the function \( g(z) \) has a pole of order \( k \) at \( z = i \) inside the contour \( C_R \). Hence, by residue theorem,

\[
\oint_{C_R} g(z) dz = 2\pi i \cdot \text{res}_i g(z) = \frac{2\pi i}{(k-1)!} \lim_{z \to i} \frac{d^{k-1}}{dz^{k-1}} [(z - i)^k g(z)]. \tag{3.9}
\]
Plugging in (3.9) into (3.8) and letting $R$ go to infinity we get

$$\int_{\mathbb{R}} \frac{e^{ikt}}{(t^2 + 1)^k} dt = \frac{2\pi i}{(k-1)!} \lim_{z \to i} \frac{d^{k-1}}{dz^{k-1}} [e^{izx} (z+i)^{-k}] =
$$

$$= \frac{2\pi i}{(k-1)!} \sum_{m=0}^{k-1} \binom{k-1}{m} (ix)^{k-1-m} e^{-x}(-k) \cdots (-k-(m-1))(2i)^{-k-m}.$$

(3.10)

Combining (3.10) with (3.7) and taking absolute value of $f^k$ we obtain

$$|f^k(x)| \leq \frac{2^k}{(k-1)!} e^{-x} \sum_{m=0}^{k-1} \binom{k-1}{m} \frac{(k+m-1)!}{(k-1)!} 2^{-k-m}.$$

Notice that

$$\frac{1}{(k-1)!} \binom{k-1}{m} \frac{(k+m-1)!}{(k-1)!} = \binom{k+m-1}{k-1} \frac{1}{(k-1)!}.$$

Estimating the binomial coefficient with $\binom{k+m-1}{k-1} \leq 2^{k+m-1}$ and multiplying powers of 2, we get

$$|f^k(x)| \leq 2^{k-1} e^{-x} \sum_{m=0}^{k-1} \binom{k-1}{m} \frac{1}{(k-1)!} 2^{k-1-m}.$$

Using the inequality $x^{k-1-m} \leq \left( \frac{x}{1+\varepsilon} \right)^{k-1-m}$, for any $\varepsilon > 0$, we see that

$$|f^k(x)| \leq \left( \frac{1+\varepsilon}{\varepsilon} \right)^{k-1} e^{-x} \sum_{m=0}^{k-1} \frac{1}{(k-1)!} \left( \frac{x}{1+\varepsilon} \right)^{k-1-m} \leq \left( \frac{1+\varepsilon}{\varepsilon} \right)^{k-1} e^{-x} \left( \frac{1}{1+\varepsilon} \right)^{k-1}. \quad (3.11)$$

\[ \square \]

**Lemma 3.4.** For fixed $l > 0$ and for any $\varepsilon > 0$ there exists a constant $c(l, \varepsilon) > 1$ such that for any nonnegative integer $k$ and any $a, b \in [N]$ the following inequality holds:

$$Path_{k+1}^{(l)}(a, b) := \sum_{x_1, \ldots, x_k=1}^{N} e^{-\frac{l}{2}(|a-x_1| + |x_1-x_2| + \ldots + |x_k-b|)} \leq c(l, \varepsilon)^{k+1} e^{-\frac{l}{2}(|a-b|)} \quad (3.12)$$

**Remark 1.** We define $Path_0^{(l)}(a, b) = \delta_{a=b}$. Note that (3.12) holds for $k = -1$. 9
We claim that for fixed $x$ Combining (3.15) and (3.16) and integrating over $\Pi$ we get the inequality (3.14).

Replacing the variables $\prod$ where $\Pi = H$ Hence, by Lemma 3.3

**Proof.** For $k = 0$ inequality holds. Assume $k \geq 1$. Without loss of generality assume that $a \geq b$. Adding new summands and the summation indices from $x_i$ to $x_i - b$, we obtain

$$Path_{k+1}^{(l)}(a, b) \leq \sum_{x_1, \ldots, x_k = -N} e^{-\frac{1}{l}((a-b)-x_1+i|x_1-x_2|+\ldots+|x_k|)}.$$ (3.13)

We claim that for fixed $x_1, \ldots, x_k$

$$e^{-\frac{1}{l}((a-b)-x_1+i|x_1-x_2|+\ldots+|x_k|)} \leq e^{\frac{1}{l}k} \int_{\Pi} e^{-\frac{1}{l}((a-b)-y_1+i|y_1-y_2|+\ldots+|y_k|)} dy_1 \ldots dy_k,$$ (3.14)

where $\Pi = \prod_{i=1}^{N} [x_i - \frac{1}{2}, x_i + \frac{1}{2})$.

Indeed, using the inequalities

$$\frac{1}{l}|x_i - x_{i+1}| \leq \frac{1}{l}|y_i - y_{i+1}| + \frac{1}{l}(|x_i - y_i| + |x_{i+1} - y_{i+1}|)$$

and $|x_i - y_i| \leq \frac{1}{2}$, we conclude that

$$e^{-\frac{1}{l}|x_i-x_{i+1}|} \leq e^{\frac{1}{l}} \cdot e^{-\frac{1}{l}|y_i-y_{i+1}|}. \quad (3.15)$$

Similarly,

$$e^{-\frac{1}{l}((a-b)-x_1)} \leq e^{\frac{1}{l}} \cdot e^{-\frac{1}{l}((a-b)-y_1)} \quad \text{and} \quad e^{-\frac{1}{l}|x_k|} \leq e^{\frac{1}{l}} \cdot e^{-\frac{1}{l}|y_k|}. \quad (3.16)$$

Combining (3.15) and (3.16) and integrating over $\Pi$ we get the inequality (3.14).

Plugging in (3.14) into (3.13), we obtain

$$Path_{k+1}^{(l)}(a, b) \leq e^{\frac{1}{l}k} \int_{\prod} e^{-\frac{1}{l}((a-b)-y_1+i|y_1-y_2|+\ldots+|y_k|)} dy_1 \ldots dy_k$$

$$\leq e^{\frac{1}{l}k} \int_{\prod} e^{-\frac{1}{l}((a-b)-y_1+i|y_1-y_2|+\ldots+|y_k|)} dy_1 \ldots dy_k. \quad (3.17)$$

Replacing the variables $y_i$ with $\frac{y_i}{l}$, we see that

$$Path_{k+1}^{(l)}(a, b) \leq l^k e^{\frac{1}{l}k} \int_{\mathbb{R}^k} e^{-\frac{(a-b)-y_1+i|y_1-y_2|+\ldots+|y_k|}} dy_1 \ldots dy_k.$$ 

Hence, by Lemma 3.3

$$Path_{k+1}^{(l)}(a, b) \leq l^k e^{\frac{1}{l}k} \left(2 \frac{1 + \varepsilon}{\varepsilon} \right)^{k-1} e^{-\frac{l}{1+\varepsilon l} (a-b)} \leq c(l, \varepsilon)^k e^{-\frac{l}{1+\varepsilon l} (a-b)},$$

where $c(l, \varepsilon) := le^{\frac{1}{l}} \cdot 2\frac{1 + \varepsilon}{\varepsilon} > 1$. Using the inequality $c(l, \varepsilon)^k \leq c(l, \varepsilon)^{k+1}$ we get the statement of the lemma. \qed
**Corollary 3.5.** For fixed $l > 0$ there exists a constant $c(l)$ such that for any positive integer $k$ the following inequality holds:

$$\text{Cyc}_k^{(l)} := \sum_{x_1,\ldots,x_k=1}^N e^{-\frac{1}{l}(|x_1-x_2|+\ldots+|x_k-x_1|)} \leq Nc(l)^k. \quad (3.18)$$

**Proof.** Choose any $\varepsilon > 0$. We rewrite $\text{Cyc}_k^{(l)}$ as

$$\text{Cyc}_k^{(l)} = \sum_{x_1=1}^N \sum_{x_2,\ldots,x_k=1}^N e^{-\frac{1}{l}(|x_1-x_2|+\ldots+|x_k-x_1|)}.$$  

Applying Lemma 3.4 to the sums with fixed $x_1$, we get

$$\sum_{x_2,\ldots,x_k=1}^N e^{-\frac{1}{l}(|x_1-x_2|+\ldots+|x_k-x_1|)} \leq c(l,\varepsilon)^k. \quad (3.19)$$

Summing (3.19) over $x_1$ and letting $\varepsilon \to \infty$, we obtain (3.18). \qed

For every tree $\Gamma \in \mathcal{T}_k$ define a set of summation graphs $G = (V(G), E(G))$ with the set of vertices $V(G) = V(F(\Gamma))$. Their edges are constructed as follows. Every edge $e \in E(\Gamma)$ gives rise to two edges $(p_{n(e_+)-1}^e, p_0^e)$, $(p_{c(e_+)}^e, p_n(e_+)^e) \in E(F(\Gamma))$ associated with $e$. Two edges of $G$ connect the ends of these two edges of $F(\Gamma)$ (see Fig. 3.1), i.e. either

- $\{p_{n(e_+)-1}^e, p_n^e\}, \{p_0^e, p_{c(e_+)}^e\} \in E(G)$
- $\{p_{n(e_+)}^e, p_{c(e_+)}^e\}, \{p_0^e, p_{n(e_+)-1}^e\} \in E(G)$.

We can make this binary choice for every edge to generate all summation graphs of $\Gamma$. Denote the set of these graphs by $\text{Sum}(\Gamma)$. It is easy to see that $|\text{Sum}(\Gamma)| = 2^k$.

We will draw the edges of the summation graphs as dotted lines. See Figure 3.2 for an example of the summation graph. Note that if vertex $e_+$ does not have children, i.e. $n(e_+) = 0$, the vertices $p_0^e$ and $p_{c(e_+)}^e$ coincide. Hence, either there is a loop $(p_0^e, p_n^e)$ in $E(G)$ or the edges $\{p_{n(e_+)-1}^e, p_n^e\} \in E(G)$ and $\{p_0^e, p_{c(e_+)}^e\} \in E(G)$ are incident.
Lemma 3.6. For any nonnegative integer $k$ and any tree $\Gamma \in T_k$, every summation graph $G \in Sum(\Gamma)$ is a disjoint union of several connected components $G = P \cup \left( \bigcup_{i=1}^{m} C^{(i)} \right)$, where $m$ is nonnegative number, $P$ is a path with the ends $p_0^{root}$ and $p_{c(root)}^{root}$ and $C^{(i)}$ are cycles. Moreover, for every vertex $v \in V(\Gamma)$ the vertices $p_0^v, \ldots, p_{c(v)}^v \in V(G)$ belong to a single component of $G$.

Remark 2. Note that $m \leq k$, since $|V(\Gamma)| = k$.

Proof. We prove the statement by induction on $k$. The base case of $k = 0$ is clear. Fix $k$ and assume that for all $0 \leq n < k$ the statement holds for all $\Gamma \in T_n$. Choose any $\Gamma \in T_k$ and any summation graph $G \in Sum(\Gamma)$. Express $\Gamma$ as $\Gamma = \Gamma_1 \oplus \Gamma_2$, where $\Gamma_1 \in T_{n_1}$, $\Gamma_2 \in T_{n_2}$ and $n_1 + n_2 = k - 1$ (see Fig. 3.3). Denote the last child of the root of $\Gamma$ by $w$. The summation graph $G$ can be naturally split into summation graphs $G_1 \in Sum(\Gamma_1)$ and $G_2 \in Sum(\Gamma_2)$, vertex $p_{c(root)}^{root}$ and two other edges (either $\{p_{c(root)}^{root}, p_{c(root)}^{root} -1\}$ and $\{p_0^w, p_0^w\}$ or $\{p_{c(root)}^{root}, p_{c(root)}^{root} -1\}$ and $\{p_0^w, p_0^w\}$). Here root refers to the root of $\Gamma$. Since $n_1 < k$ and $n_2 < k$, the statement of the lemma holds for the graphs $G_1$ and $G_2$ (see Fig. 3.3).

By induction hypothesis, $G_1 = P_1 \cup \left( \bigcup_{i=1}^{m_1} C_1^{(i)} \right)$ and $G_2 = P_2 \cup \left( \bigcup_{i=1}^{m_2} C_2^{(i)} \right)$. The ends of the path $P_1$ are $p_0^{root}$ and $p_{c(root)}^{root}$, the ends of the path $P_2$ are $p_0^w$ and $p_{c(w)}^w$. 

12
Figure 3.4: Two summation graphs of $\Gamma$ that can be obtained from the summation graphs of $\Gamma_1$ and $\Gamma_2$.

If $\{p_{c(\text{root})}^\text{root}, p_{c(\text{root})}^w\} \in E(G)$ (see Fig. 3.4 on the left), these edges complete $P_1$ and $P_2$ to the path $P$ with the ends $p_0^\text{root}$ and $p_0^w$ and the cycle $C_3$. Hence,

$$G = P \cup C_3 \cup \left( \bigcup_{i=1}^{m_1} C_1^{(i)} \right) \cup \left( \bigcup_{i=1}^{m_2} C_2^{(i)} \right).$$

If $\{p_{c(\text{root})}^\text{root}, p_{c(\text{root})}^w\} \in E(G)$ (see Fig. 3.4 on the right), these edges together with the paths of $G_1$ and $G_2$ form the path $P$ with ends $p_0^\text{root}$ and $p_0^w$. Therefore,

$$G = P \cup \left( \bigcup_{i=1}^{m_1} C_1^{(i)} \right) \cup \left( \bigcup_{i=1}^{m_2} C_2^{(i)} \right).$$

It is easy to see that in any case the vertices associated with a single vertex of $\Gamma$ belong to a single component.

For any tree $\Gamma \in T_k$, any graph $D$ with the set of vertices $V(D) \subset V(\mathcal{F}(\Gamma))$ and any labelling $x \in \mathcal{X}_{ab}(\Gamma)$ define

$$s_D(x) = \sum_{\{p,q\} \in E(D)} |\text{label}(p) - \text{label}(q)|, \quad (3.20)$$

where $\text{label}(p)$ denotes the label $x_n^v$ such that $p = p_n^v$, i.e. $p$ is one of the vertices of $\mathcal{F}(\Gamma)$ associated with $v$ and there are exactly $n$ vertices associated with $v$ before $p$ in the frame $\mathcal{F}(\Gamma)$.

**Lemma 3.7.** Suppose for some $l > 0$ the operator $S$ has finite $l$-norm (see (2.3) for the clarification). Then for any positive integer $k$, any $\Gamma \in T_k$ and any $\epsilon > 0$ the matrix $\text{val}(\Gamma)$ has finite $(1 + \epsilon)l$-norm. Moreover,

$$\|\text{val}(\Gamma)\|_{(1+\epsilon)l} \leq N^k 2^k c(l, \epsilon)^{2k} \|S\|^k_l.$$ 

**Proof.** Recalling the definition of $\text{val}_{\text{ab}}(\Gamma)$ and using the trivial inequality

$$|S_{xy,zt}| \leq e^{-\frac{1}{4}\rho(xy,zt)}\|S\|_l,$$
we get, for any fixed $a$ and $b$,

$$|\text{val}_ab(\Gamma)| \leq \|S\|_l^k \sum_{x \in X_{ab}(\Gamma)} \prod_{e \in E(\Gamma)} e^{-\frac{1}{4} \rho(x_{n(e)}^-,x_{n(e)}^+,x_{n(e)}^+,x_{n(e)}^-)}.$$  \hspace{1cm} (3.21)

Since $\rho(xy,zt) = \min\{|x-t| + |y-z|, |x-z| + |y-t|\}$,

$$e^{-\frac{1}{4} \rho(xy,zt)} \leq e^{-\frac{1}{4} |x-t| + |y-z|} + e^{-\frac{1}{4} |x-z| + |y-t|}.$$  \hspace{1cm}

Applying this inequality to the product in the RHS of (3.21), removing the parentheses and using the notation from (3.20), we obtain

$$\prod_{e \in E(\Gamma)} e^{-\frac{1}{4} \rho(x_{n(e)}^-,x_{n(e)}^+,x_{n(e)}^+,x_{n(e)}^-)} \leq \sum_{G \in \text{Sum}(\Gamma)} e^{-\frac{1}{4} s_G(x)}.$$  \hspace{1cm} (3.22)

Plugging in (3.22) into (3.21) and changing the order of summation, we get

$$|\text{val}_ab(\Gamma)| \leq \|S\|_l^k \sum_{G \in \text{Sum}(\Gamma)} \sum_{x \in X_{ab}(\Gamma)} e^{-\frac{1}{4} s_G(x)}.$$  \hspace{1cm} (3.23)

Choose any $G \in \text{Sum}(\Gamma)$. Let $G = P \cup (\bigcup_{i=1}^m C^{(i)})$ be the representation of $G$ given by Lemma 3.6. Notice that

$$s_G(x) = s_P(x) + \sum_{i=1}^m s_{C^{(i)}}(x)$$

and each label of $x$ is included in exactly one summand. Therefore, recalling the definitions of $\text{Path}_k^{(l)}(a,b)$ and $\text{Cy}_k^{(l)}$

$$\sum_{x \in X_{ab}(\Gamma)} e^{-\frac{1}{4} s_G(x)} = \sum_{x \in X_{ab}(\Gamma)} e^{-\frac{1}{4} s_P(x)} \prod_{i=1}^m e^{-\frac{1}{4} s_{C^{(i)}}(x)} = \text{Path}_k^{(l)}(a,b) \prod_{i=1}^m \text{Cy}_k^{(l)},$$

where $q_0 = |V(P)| - 1$ and $q_i = |V(C^{(i)})|$ for every $i \in \{1, \ldots, m\}$. Hence, $\sum_{i=0}^m q_i = 2k$. Applying Lemma 3.4 and Corollary 3.5 to the RHS of (3.24), we obtain

$$\sum_{x \in X_{ab}(\Gamma)} e^{-\frac{1}{4} s_G(x)} \leq e^{-\frac{1}{4} |a-b|} N_{C^{(l)}}^k l, \varepsilon)^2 \leq e^{-\frac{1}{4} |a-b|} N_{C^{(l)}}^k l, \varepsilon)^2 k. \hspace{1cm} (3.25)$$

Plugging in (3.25) into (3.23) and recalling that $|\text{Sum}(\Gamma)| = 2^k$, we get the statement of the lemma. \hfill \Box

**Corollary 3.8.** Under the conditions of Lemma 3.7

$$\|C_k\|_{(1+\varepsilon)l} \leq N_k^k 8^{k+1} c(l, \varepsilon)^2 k ||S||_l^k,$$

where $C_k$ is defined by (2.7).
Proof. It is sufficient to show that $|\mathcal{T}_k| \leq 4^k$. If this inequality holds, (3.26) follows directly from Lemma 3.7.

As mentioned in Section 2.1, one could walk around ordered tree $\Gamma \in \mathcal{T}_k$ starting in the root and making $2k$ steps of two types: going to the first unvisited child if it exists or going to the parent vertex otherwise. It is easy to see that $\Gamma$ can be recovered from this sequence of steps. Hence, the number of ordered trees with $k$ edges does not exceed the number of binary strings of length $2k$, which equals to $2^{2k}$.

Now we prove the main theorem.

Proof of Theorem 2.1. Choose $R(l, \varepsilon) = 8c(l, \varepsilon)^2c$. Since $\|S\|_l \leq \frac{C}{R}$, the estimate $\|C_k\|_{(1+\varepsilon)^2l} \leq R(l, \varepsilon)^k$ follows from Corollary 3.8. Hence, series (2.25) converges for $|z|^2 > R(l, \varepsilon)$ and $M(z)$ has finite $(1 + \varepsilon)l$-norm. Moreover,

$$\|M(z)\|_{(1+\varepsilon)^2l} \leq \frac{1}{|z|} |1 - R(l, \varepsilon)z^{-2}|$$

for $|z|^2 > R(l, \varepsilon)$. (3.27)

By Proposition 3.2, the series (2.5) gives some solution of the equation (1.2).

Now we show that $\text{Im} M(z)$ is positive for $z \in \mathbb{H}$. Denote the solution of MDE with positive imaginary part by $\tilde{M}(z)$. By Proposition 2.1 in [1], $M(z)$ admits a Stieltjes transform representation

$$\tilde{M}(z) = \int_{\mathbb{R}} \frac{V(d\tau)}{\tau - z}, \quad z \in \mathbb{H}.$$

Moreover, $V(d\tau)$ has compact support, since $\|S\|_{\max} \leq \|S\|_l \leq \frac{C}{R}$ and, hence, $\|S\|$ is finite. Therefore, there is constant $\hat{R}$ such that for any $\tau$ in the support of $V$ we have $\left|\frac{1}{\tau - z}\right| \leq \frac{2}{R}$ for $|z| > \hat{R}$. Thus, there is constant $\hat{C}$ such that

$$\|\tilde{M}(z)\| \leq \frac{\hat{C}}{|z|}$$

for $|z| > \hat{R}$.

It is easy to see that (3.27) implies that there are constants $R$ and $C$ such that $\|M(z)\| \leq \frac{C}{R}$.

Introduce $U(z) = -z M(z)$, $\tilde{U}(z) = -z \tilde{M}(z)$ and consider MDE in the form (3.24). Notice that $\|U(z)\| \leq C$ and $\|\tilde{U}(z)\| \leq C$ for $|z| > R$. Hence, for $|z| > R$ we have

$$\begin{align*}
\|U(z) - \tilde{U}(z)\| &= |z|^{-2} \|U(z)S(U(z)) - \tilde{U}(z)S(\tilde{U}(z))\| \\
&\leq |z|^{-2} \|(U(z) - \tilde{U}(z))S(U(z))\| + |z|^{-2} \|U(z)S(U(z) - \tilde{U}(z))\| \\
&\leq |z|^{-2} \cdot 2C \|S\| \cdot \|U(z) - \tilde{U}(z)\|.
\end{align*}$$

Therefore, there is constant $K$ such that $U(z) = \tilde{U}(z)$ for $|z| > K$. Since both $U(z)$ and $\tilde{U}(z)$ are analytic, $U(z) = \tilde{U}(z)$ in $\mathbb{H}$.
4 Convergence of the moments of empirical distribution of eigenvalues

In this section we consider Hermitian random matrix $H$ from Section 2.3. All results are obtained under the assumptions (A) and (B).

We are interested in the average moments of empirical distribution $m^{(N)}_k := E \int_{\mathbb{R}} x^k dL(x)$ for every nonnegative integer $k$. Notice that

$$m^{(N)}_k = \frac{1}{N} \text{Tr} H^k = N^{-\frac{k}{2}} - 1 \sum_{x_1, \ldots, x_k = 1}^N E w_{x_1 x_2} \cdots w_{x_{k-1} x_k} w_{x_k x_1}. \quad (4.1)$$

The objective of this section is to prove the following proposition. Theorem 2.2 will be obtained as a corollary.

**Proposition 4.1.** For any nonnegative integer $k$:

- If $k$ is odd, $\lim_{N \to \infty} m^{(N)}_k = 0$.
- If $k$ is even, $\lim_{N \to \infty} \left( m^{(N)}_k - \frac{1}{N} \text{Tr} C_k^2 \right) = 0$.

We start by proving several technical lemmas, the proofs of the Proposition 4.1 and Theorem 2.2 will be given at the end of the section.

4.1 Notation

We first introduce some notation. Consider the set $\mathcal{X}^{(k)} = \{ \underline{x} = (x_1, \ldots, x_k) \mid x_j \in [N] \}$. For every $\underline{x} \in \mathcal{X}^{(k)}$ we construct the *semiframe* $\mathfrak{F} = \mathfrak{F}(\underline{x})$ with the set of vertices $V(\mathfrak{F}) = \{1, 2, \ldots, k\}$, edges $E(\mathfrak{F}) = \{\{1, 2\}, \ldots, \{k-1, k\}, \{k, 1\}\}$ and the following additional structure. We define the set of *proximity edges* $Pr(\mathfrak{F}) = \{\{i, j\} \mid |x_i - x_j| \leq (\log N)^2\}$. Let $m = m(\underline{x})$ be such nonnegative integer that the graph with vertices $V(\mathfrak{F})$ and edges $Pr(\mathfrak{F})$ consists of $m + 1$ components called *proximity components*. We denote proximity components by $C_0 = C_0(\underline{x}), \ldots, C_m = C_m(\underline{x})$ and order them so that $1 = \min(C_0) < \min(C_1) < \ldots < \min(C_m)$, where $\min(C)$ is meant as the minimum of the set of integers. Finally, we define the graph $\hat{\Gamma} = \hat{\Gamma}(\underline{x})$ with the set of vertices $V(\hat{\Gamma}) = \{C_0, \ldots, C_m\}$ and $\{C_i, C_j\} \in E(\hat{\Gamma})$ iff there is at least one edge between the proximity components $C_i$ and $C_j$ of the semiframe $\mathfrak{F}$.

We say that edges of the graph $\mathfrak{F}$ connecting components $C_i$ and $C_j$ are associated with the edge $\{C_i, C_j\} \in E(\hat{\Gamma})$.

For $k$ even we call $\underline{x} \in \mathcal{X}^{(k)}$ significant if $m(\underline{x}) = \frac{k}{2}$ and for any edge $\{C_i, C_j\} \in E(\hat{\Gamma})$ there are exactly two associated edges of semiframe $\mathfrak{F}$. In this case $|V(\hat{\Gamma})| = \frac{k}{2} + 1$ and $|E(\hat{\Gamma})| = \frac{k}{2}$, hence $\hat{\Gamma}$ is a tree. Define the ordering
of $\tilde{\Gamma}$ as follows. Let $C_0$ be the root of $\tilde{\Gamma}$. If $i < j$ and $C_i$, $C_j$ have the same parent, they are ordered with respect to the ordering of indices, i.e. $C_i < C_j$. Hence, $\tilde{\Gamma} \in T_2$. Recall the definition of the frame of the ordered tree given in the subsection 2.1. The semi-frame $\mathfrak{g}$ is the frame of $\tilde{\Gamma}$ with $p_0^{\text{root}} = p_c^{\text{root}} = 1$ and vertices of the component $C_i$ are associated with the vertex $C_i \in V(\tilde{\Gamma})$. The labelling $x$ of the frame is defined by the components of the array $x_i$, i.e. for the vertex $p_i = j \in V(\mathfrak{g})$ its label is $x_i = x_j$. Conversely, given the labelling of the frame $x$ with $x_0^{\text{root}} = x_c^{\text{root}}$, one could obtain the array $x$ by ordering the labels of $x$ with respect to the order of appearance of the vertices in the frame.

Now, consider an ordered tree $\Gamma \in T_2$ for some even $k$. The cyclic frame of the tree $\Gamma$ is the frame of $\Gamma$ with $p_0^{\text{root}} = p_c^{\text{root}}$. We denote the cyclic frame by $F_{cyc}(\Gamma)$. The labelling of the cyclic frame is any labelling of the frame $x \in X(\Gamma)$ with $x_0 = x_c^{\text{root}}$. The set of the labellings of the cyclic frame is denoted by $X_{cyc}(\Gamma)$, i.e.

$$X_{cyc}(\Gamma) = \bigcup_{a \in [N]} X_{aa}(\Gamma),$$

where $X_{aa}(\Gamma)$ is defined in 2.1. Notice that, given a tree $\Gamma \in T_2$ and the labelling of the cyclic frame $x \in X_{cyc}(\Gamma)$, one can define the proximity edges and proximity components on the cyclic frame $F_{cyc}(\Gamma)$ analogously to these structures on the semi-frame. We denote the number of proximity components by $m = m(x)$.

We call $x \in X_{cyc}(\Gamma)$ significant if for any $v \in V(\Gamma)$ vertices $p_0^v, \ldots, p_c^v$ form a proximity component. In this case the corresponding $x$ is significant and $\Gamma = \Gamma$. Therefore, for even $k$ we have established the correspondence between the elements of the set $\{x \in X^{(k)} \mid x \text{ is significant}\}$ and the elements of the set

$$\{(\Gamma, x) \mid \Gamma \in T_2, x \in X_{cyc}(\Gamma), x \text{ is significant}\}.$$

Denote the set of partitions of the set $1, \ldots, k$ by $\Pi_k$.

### 4.2 Some estimates on the expectations

**Lemma 4.2.** For any nonnegative integer $k$ and any array of indices $x \in X^{(k)}$ with $m = m(x) \geq \frac{k}{2}$ there is a constant $c$ such that:

- If $x$ is not significant,

$$|\mathbb{E} w_{x_1 x_2} \ldots w_{x_{k-1} x_k} w_{x_k x_1}| \leq |\Pi_k| c^k e^{-\frac{c}{4} (\log N)^2}.$$  \hspace*{1cm} (4.2)

- If $x$ is significant,

$$\left| \mathbb{E} w_{x_1 x_2} \ldots w_{x_{k-1} x_k} w_{x_k x_1} - \prod_{e \in E(\Gamma)} \mathbb{E} w_{x_{\ell_{e}(+)}, x_{\ell_{e}(-)+}} w_{x_{\ell_{e}(+)-}, x_{\ell_{e}(-)-}} \right| \leq |\Pi_k| c^k e^{-\frac{c}{4} (\log N)^2},$$  \hspace*{1cm} (4.3)

where $x$ is corresponding labelling of the frame of $\tilde{\Gamma}$.
Proof. We express the expectation of products of random variables through their cumulants using
the relation

$$\mathbb{E}w_{x_1,x_2} \ldots w_{x_{k-1},x_k} w_{x_k,x_1} = \sum_{\pi \in \Pi_k} \prod_{B \in \pi} \kappa(w_{x_i,x_{i+1}} | i \in B), \quad (4.4)$$

where $x_{k+1} := x_1$ (see Appendix A in [3]).

Consider one summand of this sum with fixed $\pi \in \Pi_k$. If there is $B \in \pi$ such that $|B| = 1$, then $\kappa(w_{x_i,x_{i+1}} | i \in B) = \mathbb{E}w_{x_i,x_{i+1}} = 0$, where $B = \{i\}$. Hence, this summand is 0. Otherwise, by (2.8),

$$|\kappa(w_{x_i,x_{i+1}} | i \in B)| \leq \prod_{(i,j) \in E(T_B)} |\kappa(w_{x_i,x_{i+1}}, w_{x_j,x_{j+1}})|, \quad (4.5)$$

where $T_B$ is the minimal spanning tree of the complete graph with the set of vertices $B$ and edge length $d(\{i, j\}) = \rho(x_i,x_{i+1}, x_j,x_{j+1})$.

Since semiframe $\widehat{x}$ is a cycle, it is connected. Therefore, the graph $\tilde{\Gamma}(\widehat{x})$ is connected and has $m+1$ vertices. Hence, it has at least $m$ edges. Consider the case, where $\widehat{x}$ is not significant. Since $m \geq \frac{1}{\xi}$, there is at least one edge of $\tilde{\Gamma}$ with a single associated edge of semiframe $\widehat{x}$. Denote this edge by $\{i_0, i_0 + 1\}$. Let $B_0 \in \pi$ be such partition set, that $i_0 \in B_0$. Since $|B_0| \geq 2$, there is such $j_0 \in B_0$ that $\{i_0, j_0\}$ is the edge of the minimal spanning tree $T_{B_0}$. The edges $\{i_0, i_0 + 1\}$ and $\{j_0, j_0 + 1\}$ are associated with different edges of $\tilde{\Gamma}(\widehat{x})$. Therefore, either vertices $i_0, j_0$ or vertices $i_0 + 1, j_0 + 1$ belong to different proximity components. Likewise, either vertices $i_0, j_0 + 1$ or vertices $i_0 + 1, j_0$ belong to different proximity components. Hence,

$$\rho(x_{i_0},x_{i_0+1}, x_{j_0},x_{j_0+1}) = \min\{|x_{i_0} - x_{j_0}| + |x_{i_0+1} - x_{j_0+1}|, |x_{i_0} - x_{j_0+1}| + |x_{i_0+1} - x_{j_0}|\} \geq (\log N)^2. \quad (4.6)$$

Using (2.8) and (4.6), we get

$$|\kappa(w_{x_{i_0},x_{i_0+1}}, w_{x_{j_0},x_{j_0+1}})| \leq c e^{-\frac{1}{4}(\log N)^2}. \quad (4.7)$$

For each $B \in \pi$ and each edge $\{i, j\}$ of $T_B$, such that $\{i, j\} \neq \{i_0, j_0\}$, estimate (2.8) implies that

$$|\kappa(w_{x_i,x_{i+1}}, w_{x_j,x_{j+1}})| \leq c. \quad (4.8)$$

Applying (4.7), (4.8) and (4.5) to the summand for $\pi$ in the RHS of (4.4), we obtain

$$\left| \prod_{B \in \pi} \kappa(w_{x_i,x_{i+1}} | i \in B) \right| \leq \left( \prod_{B \in \pi} c^{B|^{-1}} \right) e^{-\frac{1}{4}(\log N)^2} \leq c^k e^{-\frac{1}{4}(\log N)^2}. \quad (4.9)$$

Summing over $\pi \in \Pi_k$, we get the estimate (4.2).

Now, if $\widehat{x}$ is significant, consider the partition $\pi' \in \Pi_k$, such that $i$ and $j$ belong to the same set of $\pi'$ iff $\{i, i + 1\}$ and $\{j, j + 1\}$ are associated with the
same edge of $\tilde{\Gamma}(x)$. Then $|B| = 2$ for any $B \in \pi'$. Hence, any other partition $\pi \in \Pi_k$, $\pi \neq \pi'$ contains either $B \in \pi$ with $|B| = 1$, or $B_0 \in \pi$ such that edges $\{i, i+1\}$ with $i \in B_0$ are not associated with the same edge of $\tilde{\Gamma}(x)$. In the former case, the summand of (4.3) with such $\pi$ is 0. In the latter case, one could find the edge $\{i_0, j_0\}$ of $\tilde{T}_B$, such that $\{i_0, i_0 + 1\}$ and $\{j_0, j_0 + 1\}$ are associated with different edges of $\tilde{\Gamma}(x)$. Hence, the estimate (4.3) holds for such $\pi$.

Since the entries $w_{yz}$ are centered, $\kappa(w_{x_i,x_{i+1}}, w_{x_j,x_{j+1}}) = \mathbb{E}w_{x_i,x_{i+1}}w_{x_j,x_{j+1}}$. Therefore,

$$\prod_{B \in \pi'} \kappa(w_{x_i,x_{i+1}} \mid i \in B) = \prod_{B = \{i_B,j_B\} \in \pi'} \frac{1}{\mathbb{E}w_{i_B,i_{B+1}}w_{j_B,j_{B+1}}} = \prod_{e \in E(\tilde{\Gamma})} \mathbb{E}w^\pm_{x_{n(e_+)},0} w^\pm_{x_{e_+}x_{n(e_+)}} w^\pm_{x_{n(e_+)}x_{e_+}},$$

(4.10)

where the last identity follows from one-to-one correspondence between the partition sets $B \in \pi'$ and edges of $\tilde{\Gamma}$.

Separating $\pi'$ partition in (4.1) and evaluating by (4.10), then applying (4.9) to remaining summands on the RHS of (4.4), we get the estimate (4.3).

\[\square\]

**Lemma 4.3.** For any even nonnegative integer $k$, any $\Gamma \in T_k$ and any labelling of the cyclic frame $x \in X_{\text{cyc}}(\Gamma)$ with $m = m(x) \geq \frac{k}{2}$ such that $x$ is not significant, we have

$$\left| \prod_{e \in E(\Gamma)} \mathbb{E}w^\pm_{x_{n(e_+)},0} w^\pm_{x_{e_+}x_{n(e_+)}} w^\pm_{x_{n(e_+)}x_{e_+}} \right| \leq c e^{-\frac{1}{2}(\log N)^2}. \quad (4.11)$$

**Proof.** If for any edge $e \in E(\Gamma)$ the edges $\{p^e_{n(e_+)-1}, p^e_0\}$ and $\{p^e_{n(e_+)+1}, p^e_{n(e_+)}\}$ are associated with a single edge of $\tilde{\Gamma}$, then the vertices associated with the same vertex of $\Gamma$ belong to the same proximity component (this can be proved similarly to the Lemma 3.9). Hence, either $m < \frac{k}{2}$ or $x$ is significant, which is excluded by the conditions of the lemma.

Thus, there is such edge $\hat{e} \in E(\Gamma)$ that the edges $\{\hat{p}^e_{n(\hat{e}_+)-1}, \hat{p}^e_0\}$ and $\{\hat{p}^e_{n(\hat{e}_+)+1}, \hat{p}^e_{n(\hat{e}_+)}\}$ are associated with different edges of $\tilde{\Gamma}$.

We estimate the expectations using the assumption on the decay of cumulants. More precisely, for the edge $\hat{e}$

$$\left| \mathbb{E}w^\pm_{x_{n(\hat{e}_+)},0} w^\pm_{x_{e_+}x_{n(\hat{e}_+)}} w^\pm_{x_{n(\hat{e}_+)}x_{e_+}} \right| \leq c e^{-\frac{1}{2}(\log N)^2}.$$  

For any other edge $e \in E(\Gamma)$ such that $e \neq \hat{e}$, we have

$$\left| \mathbb{E}w^\pm_{x_{n(e_+)},0} w^\pm_{x_{e_+}x_{n(e_+)}} w^\pm_{x_{n(e_+)}x_{e_+}} \right| \leq c.$$
Hence, the inequality (4.11) holds. □

**Lemma 4.4.** For any nonnegative $k$,

$$\left| \left\{ x \in \mathcal{X}^{(k)} \mid m(x) < \frac{k}{2} \right\} \right| \leq c(k)N^{\frac{k+1}{2}}(\log N)^{2k},$$

(4.12)

where $c(k)$ only depends on $k$.

**Proof.** Denote $l = \left\lfloor \frac{k+1}{2} \right\rfloor$. Let $\tilde{c}(k)$ be the number of partitions of the set $\{1, \ldots, k\}$ into no more than $l$ subsets. Fix some partition $\pi$ and consider the subset $\mathcal{X}_\pi^{(k)}$ of $\mathcal{X}^{(k)}$, consisting of such $x \in \mathcal{X}^{(k)}$ that for each proximity component of $\mathcal{X}(x)$, its vertices form the subsets of partition $\pi$.

Choose a subset $A \subset \{1, \ldots, k\}$ such that $|A| = l$ and $A \cap B$ is not empty for each subset $B$ in $\pi$. There are $N^l$ ways to choose indices $x_i \in [N], i \in A$. If $A \neq \{1, \ldots, k\}$, there is $j \in \{1, \ldots, k\} \setminus A$ such that $j$ is connected by proximity edge to some $i \in A$. Since $i$ and $j$ are connected, $|x_i - x_j| < (\log N)^2$. Hence, there are at most $2(\log N)^2$ allowed values of $x_j$. Repeating this choice until all $x_i$ are fixed for all $i \in \{1, \ldots, k\}$, we get

$$|\mathcal{X}_\pi^{(k)}| \leq N^l(2(\log N)^2)^{k-l} \leq 2^k N^{\frac{k+1}{2}}(\log N)^{2k}.$$ 

Therefore, the inequality (4.12) holds with $c(k) = 2^k \tilde{c}(k)$. □

### 4.3 Proofs of Proposition 4.1 and Theorem 2.2

**Proof of the Proposition 4.1.** We decompose the set $\mathcal{X}^{(k)}$ into the union of three sets $\mathcal{X}^{(k)} = \mathcal{X}^{(k)}_1 \cup \mathcal{X}^{(k)}_2 \cup \mathcal{X}^{(k)}_3$ such that

- $\mathcal{X}^{(k)}_1 = \{ x \in \mathcal{X}^{(k)} \mid m(x) < \frac{k}{2} \}$;
- $\mathcal{X}^{(k)}_2 = \{ x \in \mathcal{X}^{(k)} \mid m(x) \geq \frac{k}{2}, x$ is not significant $\}$;
- $\mathcal{X}^{(k)}_3 = \{ x \in \mathcal{X}^{(k)} \mid x$ is significant $\}.$

Note that this union is disjoint, since $m(x) = \frac{k}{2}$ for significant $x$. Recall the empirical moments from (4.1). We decompose $m_{k}^{(N)}$ into three sums in accordance with the decomposition of $\mathcal{X}^{(k)}$, i.e. $m_{k}^{(N)} = S_1 + S_2 + S_3$, where

$$S_i = N^{-\frac{k}{2}-1} \sum_{x \in \mathcal{X}^{(k)}_i} \mathbb{E} w_{x_1 x_2} \ldots w_{x_{k-1} x_k} w_{x_k x_1}$$

for any $i \in \{1, 2, 3\}$.

In the sum $S_1$ we estimate each summand using Hölder inequality as follows:

$$|\mathbb{E} w_{x_1 x_2} \ldots w_{x_{k-1} x_k} w_{x_k x_1}| \leq \left( \mathbb{E} |w_{x_1 x_2}|^k \cdot \ldots \cdot \mathbb{E} |w_{x_k x_1}|^k \right)^{\frac{1}{k}} \leq \mu_k,$$

for any $i \in \{1, 2, 3\}$. 

20
The sum (4.14) can be decomposed accordingly as

$$|S_1| \leq N^{-\frac{k}{2}}c(k)N^{\frac{k+1}{2}}(\log N)^{2k} \mu_k = c(k)\mu_k N^{-\frac{k}{2}}(\log N)^{2k} \to 0,$$

as $N \to \infty$ for any fixed $k$. Since $x_i \in [N]$ for any $i \in \{1, \ldots, k\}$, $|x_i^{(k)}| \leq N^k$. Applying Lemma 4.2 to each summand of $S_2$, we get

$$|S_2| \leq N^k|\Pi_k|c^k e^{-\frac{k}{2}(\log N)^2} \to 0,$$

as $N \to \infty$, since $l$ is fixed and $N$-independent.

If $k$ is odd, the set $X_3^{(k)}$ is empty. Hence, $S_3 = 0$ and $m_k^{(N)} \to 0$, as $N \to \infty$.

If $k$ is even, from (2.6) we have

$$\frac{1}{N}\text{Tr} C_\frac{k}{2} = \frac{1}{N} \sum_{\Gamma \in T_\frac{k}{2}} \text{Tr} \text{val}(\Gamma). \quad (4.13)$$

Recalling the definition of val($\Gamma$)

$$\text{val}_{ab}(\Gamma) = \sum_{x \in X_{ab}(\Gamma)} \prod_{e \in E(\Gamma)} \frac{1}{N} E_w^* x_{m(x +)}^e x_{m(x -)} = \left\{ E_w^* x_{m(x +)}^e x_{m(x -)} \right\},$$

and plugging it into (4.13), we obtain

$$\frac{1}{N}\text{Tr} C_\frac{k}{2} = N^{-\frac{k}{2}} \sum_{\Gamma \in T_\frac{k}{2}} \sum_{x \in X_{\text{cyc}}(\Gamma)} \prod_{e \in E(\Gamma)} E_w^* x_{m(x +)}^e x_{m(x -)} \cdot (4.14)$$

Consider the set $X^{(k)} = \{(\Gamma, x) \mid \Gamma \in T_\frac{k}{2}, x \in X_{\text{cyc}}(\Gamma)\}$ and decompose it into disjoint union $X^{(k)} = X_1^{(k)} \cup X_2^{(k)} \cup X_3^{(k)}$, where

$$X_1^{(k)} = \{(\Gamma, x) \in X^{(k)} \mid m(x) < \frac{k}{2}\};$$

$$X_2^{(k)} = \{(\Gamma, x) \in X^{(k)} \mid m(x) > \frac{k}{2}, x \text{ is not significant}\};$$

$$X_3^{(k)} = \{(\Gamma, x) \in X^{(k)} \mid x \text{ is significant}\}.$$

The sum (4.14) can be decomposed accordingly as $\frac{1}{N}\text{Tr} C_\frac{k}{2} = \hat{S}_1 + \hat{S}_2 + \hat{S}_3$, where for any $i \in \{1, 2, 3\}$

$$\hat{S}_i = N^{-\frac{k}{2}} \sum_{(\Gamma, x) \in X_i^{(k)} \in X^{(k)}} \prod_{e \in E(\Gamma)} E_w^* x_{m(x +)}^e x_{m(x -)}.$$

Notice that for any fixed $\Gamma \in T_\frac{k}{2}$, the elements of the set $X_{\text{cyc}}(\Gamma)$ are in one-to-one correspondence with the elements of the set $X^{(k)}$. Thus, according to Lemma 4.3

$$|X_1^{(k)}| \leq \left| T_\frac{k}{2} \right| c(k)N^{\frac{k+1}{2}}(\log N)^{2k}. \quad (4.15)$$
By Hölder inequality, for each \((\Gamma, \mathbf{x}) \in \mathcal{X}^{(k)}\) and each \(e \in E(\Gamma)\)
\[
|\mathbb{E} w_{x_{n(e_+)}}^{e_+} x_0^{e_0} w_{x_{n(e_-)}}^{e_-}| \leq \mu_2. \tag{4.16}
\]
Applying (4.13) and (4.16) to the sum \(\hat{S}_1\), we get
\[
|\hat{S}_1| \leq N^{-\frac{K}{2} - 1} |T_{\Gamma}^k| e(k) N \frac{1}{2} \mu_2 \to 0,
\]
as \(N \to \infty\).

Applying Lemma 4.3 to each summand of \(\hat{S}_2\) and using inequality \(|X_2^{(k)}| \leq |T_{\Gamma}^k| N^k\), we obtain
\[
|\hat{S}_2| \leq N^{-\frac{K}{2} - 1} |T_{\Gamma}^k| N^k c^k e^{-\frac{K}{2} (\log N)^2} \to 0,
\]
as \(N \to \infty\).

This proves that
\[
m_k^{(N)} - \frac{1}{N} \text{Tr} C \frac{k}{2} = S_3 - \hat{S}_3 + o(1) \quad \text{as} \quad N \to \infty.
\]

Since there is correspondence between the sets \(\mathcal{X}_3^{(k)}\) and \(\mathcal{X}_3^{(k)}\), the difference \(S_3 - \hat{S}_3\) can be estimated using Lemma 4.2
\[
|S_3 - \hat{S}_3| \leq N^{-\frac{K}{2} - 1} \sum_{x \in \mathcal{X}_3^{(k)}} \left| \mathbb{E} w_{x_1 x_2} \cdots w_{x_{k-1} x_k} w_{x_k x_1} - \prod_{e \in E(\Gamma)} \mathbb{E} w_{x_{n(e_+)}}^{e_+} x_0^{e_0} w_{x_{n(e_-)}}^{e_-} \right|
\leq N^{-\frac{K}{2} - 1} |\mathcal{X}_3^{(k)}| |\Pi_k| c^k e^{-\frac{K}{2} (\log N)^2}.
\]
Applying inequality \(|\mathcal{X}_3^{(k)}| \leq N^k\), we obtain that \(S_3 - \hat{S}_3 \to 0\) as \(N \to \infty\).
Hence, \(m_k^{(N)} - \frac{1}{N} \text{Tr} C \frac{k}{2} \to 0\).

\(\Box\)

**Proof of Theorem 2.2** Armed with Proposition 4.1, this proof is fairly standard. Here we follow the argument in Section 2.1.2 of [2]. From Proposition 2.1 in [1] we know that \(\frac{1}{N} \text{Tr} M(z)\) is Stieltjes transform of some measure \(\mu_N\) for each \(N\) and measures \(\mu_N\) are supported on the set \([-B, B]\) for some constant \(B\). Notice that, since \(\mu_N\) is a probability measure,
\[
\left| \frac{1}{N} \text{Tr} C \right| = \int_{\mathbb{R}} x^{2k} \mu_N(\mathrm{d}x) \leq B^{2k}. \tag{4.17}
\]
Consider the integral
\[
\int_{|x| > (B + 1)^2} x^{2k} \mathbb{E} L^{(N)}(\mathrm{d}x) \leq \frac{1}{(B + 1)^{2k}} \int_{\mathbb{R}} x^{2k} \mathbb{E} L^{(N)}(\mathrm{d}x) = \frac{m_{2k}^{(N)}}{(B + 1)^{2k}}.
\]
Since $m_{2k}^{(N)} = (m_{2k}^{(N)} - \frac{1}{N} \text{Tr} C_k(z)) + \frac{1}{N} \text{Tr} C_k(z)$ and $|m_{2k}^{(N)} - \frac{1}{N} \text{Tr} C_k(z)| \to 0$ as $N \to \infty$ (see Proposition 4.1), we have

$$\limsup_{N \to \infty} \int_{|x|>(B+1)^2} |x|^k E L^{(N)}(dx) = \limsup_{N \to \infty} \frac{1}{N} \text{Tr} C_k(z) \leq \left(\frac{B}{B+1}\right)^{2k}.$$  

Notice that LHS is increasing in $k$ and RHS goes to 0 as $k \to \infty$. Hence,

$$\lim_{N \to \infty} \int_{|x|>(B+1)^2} |x|^k E L^{(N)}(dx) = 0. \quad (4.18)$$

Fix any $z \in \mathbb{H}$ such that $|z| > (B+1)^2$. Denote $(B+1)^2$ by $D$. Since $(x-z)^{-1}$ is a bounded function of $x$, (4.18) for $k = 0$ implies that

$$\left| E S^{(N)}(z) - \int_{-D}^{D} \frac{E L^{(N)}(dx)}{x-z} \right| \to 0 \text{ as } N \to \infty. \quad (4.19)$$

Then $- \sum_{k=0}^{n} x^k z^{-k-1}$ converges to $(x-z)^{-1}$ uniformly in $x$ on $[-D,D]$. Hence,

$$\left| \int_{-D}^{D} \frac{E L^{(N)}(dx)}{x-z} - \int_{-D}^{D} \left( - \sum_{k=0}^{n} x^k z^{-k-1} \right) E L^{(N)}(dx) \right| \to 0, \quad (4.20)$$

uniformly in $N$ as $n \to \infty$.

Notice that

$$\int_{-D}^{D} \left( - \sum_{k=0}^{n} x^k z^{-k-1} \right) E L^{(N)}(dx) = - \sum_{k=0}^{n} \left( \int_{-D}^{D} x^k E L^{(N)}(dx) \right) z^{-k-1}.$$  

Therefore, (4.18) implies that

$$\left| \int_{-D}^{D} \left( - \sum_{k=0}^{n} x^k z^{-k-1} \right) E L^{(N)}(dx) \right| \to 0, \quad (4.21)$$

as $N \to \infty$. By Proposition 4.1

$$\left| \left( - \sum_{k=0}^{n} m_k^{(N)} z^{-k-1} \right) - \left( - \sum_{k=0}^{\infty} \frac{1}{N} \text{Tr} C_k^2 z^{-k-1} \right) \right| \to 0 \text{ as } N \to \infty. \quad (4.22)$$

From (4.17) we have

$$\left| \left( - \sum_{k=0}^{n} \frac{1}{N} \text{Tr} C_k^2 z^{-k-1} \right) - \left( - \sum_{k=0}^{\infty} \frac{1}{N} \text{Tr} C_k^2 z^{-k-1} \right) \right| \to 0, \quad (4.23)$$
uniformly in $N$ as $n \to \infty$.

Combining (4.19), (4.20), (4.21), (4.22), (4.23) and letting $n \to \infty$, $N \to \infty$, we get

$$\left| ES^{(N)}(z) - \frac{1}{N} \text{Tr} M(z) \right| \to 0,$$

(4.24)
as $N \to \infty$ for $|z| > D$.

It is easy to show that (4.24) holds for all $z \in \mathbb{H}$ by applying Montel’s theorem to Stieltjes transforms and functions $\frac{1}{N} \text{Tr} M(z)$.

References

[1] O. H. Ajanki, L. Erdős, and T. Krüger. Stability of the matrix Dyson equation and random matrices with correlations. *Probability Theory and Related Fields*, pages 1–81, 2018.

[2] G. W. Anderson, A. Guionnet, and O. Zeitouni. An introduction to random matrices, volume 118 of Cambridge Studies in Advanced Mathematics, 2010.

[3] L. Erdős, T. Krüger, and D. Schröder. Random matrices with slow correlation decay. *arXiv preprint arXiv:1705.10661*, 2017.

[4] L. Erdős and P. Mühlbacher. Bounds on the norm of Wigner-type random matrices. *arXiv preprint arXiv:1802.05175*, 2018.

[5] J. W. Helton, R. R. Far, and R. Speicher. Operator-valued semicircular elements: solving a quadratic matrix equation with positivity constraints. *International Mathematics Research Notices*, 2007(9), 2007.

[6] M. Ottolini. Spectral norm of random matrices matrices with non-identically distributed entries. Master’s thesis, University Pisa, 2017.