ASPECTS OF RANDOM MATRIX THEORY: CONCENTRATION AND SUBSEQUENCE PROBLEMS

A Thesis
Presented to
The Academic Faculty

by

Hua Xu

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology
December 2008

ASPECTS OF RANDOM MATRIX THEORY: CONCENTRATION AND SUBSEQUENCE PROBLEMS

Approved by:

Professor Christian Houdré, Advisor
School of Mathematics
Georgia Institute of Technology

Professor Yuri Bakhtin
School of Mathematics
Georgia Institute of Technology

Professor Vladimir I Kolchinskii
School of Mathematics
Georgia Institute of Technology

Professor Heinrich Matzinger
School of Mathematics
Georgia Institute of Technology

Professor Ionel Popescu
School of Mathematics
Georgia Institute of Technology

Professor Robert Foley
School of Industrial and System Engineering
Georgia Institute of Technology

Professor Mikhail Lifshits
Math.-Mech. Dept.
St.Petersburg State University

Date Approved: 30 October 2008

To my family.

ACKNOWLEDGEMENTS

I first want to thank my advisor, Dr. Christian Houdré. Not only did he teach me mathematics and inspire my research, he also helped me navigate both life and study issues from the very first day I arrived in Atlanta. I will always be grateful for all the patience he has shown with every aspect of my work.

Next, let me thank Dr. Vladimir Koltchinskii, Dr. Marcus Spruill, Dr. Yang Wang, Dr. Yuri Bakhtin, among others, for teaching me mathematics and for their encouraging discussions. I would also like to extend my sincere gratitude to Dr. Evans Harrell, who extended every effort to help me get used to the new Georgia Tech environment when I first came to begin my study, and to Dr. Luca Dieci, who kindly continued to make everything smooth for me to get work done. I particularly want to thank Ms. Cathy Jacobson, who not only gave the best English training program to make my teaching easy from the beginning, but also enhanced my life here in every possible way. My special thanks goes to Ms. Rena Brakebill, Ms. Klara Grodzinsky, for their gracious support of my teaching, and to Ms. Sharon McDowell, Ms. Annette Rohrs, Ms. Genola Turner and the IT group for their everyday support. Finally, I want to sincerely thank all my thesis defense committee members for their careful reading of my work.

There are many of my fellow graduate students whose friendships have meant so much over the years. I especially want to thank Trevis Litherland for his pleasant company at various conferences as well as our numerous research discussions. I have enjoyed tremendous happiness and shared learning experiences with all my friends, notably Jian Chen, Hao Deng, Ruoting Gong, Alexander Grigo, Xun Huang, Torsten Inkmann, Wen Jiang, David Jimenez, Hwa Kil Kim, Yongfen Li, Nan Lu, Jie Ma, Jinyong Ma, Zixia Song, Ying Wang, Kun Zhao.

I owe so much to my parents who have been educating and supporting me in all of my lifes choices. Finally and most importantly, I owe innumerable thanks to my lovely wife, Jing. Her faith in me over all these years is the ultimate foundation of anything that I have accomplished.

TABLE OF CONTENTS

SUMMARY

The present work studies some aspects of random matrix theory. Its first part is devoted to the asymptotics of random matrices with infinitely divisible, in particular heavy-tailed, entries. Its second part focuses on relations between limiting law in subsequence problems and spectra of random matrices.

In Chapter II, we give concentration inequalities for the spectral measure, with respect to the Wasserstein distance, or for the maximal eigenvalue of random Hermitian matrices with infinitely divisible (not necessarily independent) entries. For such matrices, the classical techniques, which rely on the independence and/or the finite moments properties of the entries, no longer apply. Results for the spectral measure of matrices with stable entries are also obtained; leading to different rates of decay for different ranges of deviation. Finally, concentration results for various functions of Wishart matrices are also derived.

In Chapter III and Chapter IV, interactions between spectra of classical Gaussian ensembles and subsequence problems are studied with the help of the powerful machinery of Young tableaux. For the random word problem, from an ordered finite alphabet, the shape of the associated Young tableaux is shown to converge to the spectrum of the (generalized) traceless GUE. Various properties of the (generalized) traceless GUE are established, such as a law of large numbers for the extreme eigenvalues and the convergence of the spectral measure towards the semicircle law. The limiting shape of the whole tableau is also obtained as a Brownian functional. The Poissonized word problem is finally discussed, and, using Poissonization, the convergence of the whole Poissonized tableaux is derived.

CHAPTER I

INTRODUCTION

Large random matrices have recently attracted a lot of attention in fields such as statistics, mathematical physics or combinatorics (e.g., see Mehta [49], Bai and Silverstein [7], Johnstone [41], Anderson, Guionnet and Zeitouni [3]). For various classes of matrix ensembles, the asymptotic behavior of the, properly centered and normalized, spectral measure or of the largest eigenvalue is understood. Many of these results hold true for matrices with independent entries satisfying some moment conditions (Wigner [66], Tracy and Widom [61], Soshnikov [56], Girko [21], Pastur [51], Bai [6], Götze and Tikhomirov [23]).

There is relatively little work outside the independent or finite second moment assumptions. Let us mention Soshnikov [58] who, using ideas from perturbation theory, studied the distribution of the largest eigenvalue of Wigner matrices with entries having heavy tails. (Recall that a real (or complex) Wigner matrix is a symmetric (or Hermitian) matrix whose entries $\mathbf{M}_{i,i}, 1 \leq i \leq N$, and $\mathbf{M}_{i,j}, 1 \leq i < j \leq N$, form two independent families of iid (complex valued in the Hermitian case) random variables.) In particular, (see [58]), for a properly normalized Wigner matrix with entries belonging to the domain of attraction of an α-stable law, $\lim_{N\to\infty} \mathbb{P}^N(\lambda_{max} \leq x) = \exp(-x^{-\alpha})$ (here λ_{max} is the largest eigenvalue of such a normalized matrix). Further, Soshnikov and Fyodorov [60], using the method of determinants, derived results for the largest singular value of $K \times N$ rectangular matrices with independent Cauchy entries, showing that the largest singular value of such a matrix is of order K^2N^2 (see also the survey article [59], where band and sparse matrices are studied). Very recently, Ben Arous and Guionnet [4] studied the $N \times N$ symmetric matrices whose entries are iid and in the domain of attraction of an α-stable law. They showed, that if the eigenvalues are normalized by a constant of order $N^{1/\alpha}$, the corresponding spectral distribution converges in expectation to a law with heavy-tail. They also explored some properties of the limiting distribution.

On another front, Guionnet and Zeitouni [25], gave concentration results for functionals of the empirical spectral measure of, self-adjoint, random matrices whose entries are independent and either satisfy a logarithmic Sobolev inequality or are compactly supported. They obtained, for such matrices, the subgaussian decay of the tails of the empirical spectral measure when it deviates from its mean. They also noted that their technique could be applied to prove results for the largest eigenvalue or for the spectral radius of such matrices. Alon, Krivelevich and Vu [2] further obtained concentration results for any of the eigenvalues of a Wigner matrix with uniformly bounded entries (see, Ledoux [44] for more developments and references).

The present work, in Chapter II, deals with matrices whose entries form a general infinitely divisible vector, and in particular a stable one (without independence assumption). As well known, unless degenerated, an infinitely divisible random variable cannot be bounded. We obtain concentration results for functionals of the corresponding empirical spectral measure, allowing for any type of light or heavy tails. The methodologies developed here apply as well to the largest eigenvalue or to the spectral radius of such random matrices.

The second part of this thesis studies some of the connections between subsequence problems and classical matrix ensembles. One of the most remarkable achievements in modern random matrix theory is the identification by Tracy and Widom [61] of the distribution which now bears their names. The Tracy-Widom distribution gives the fluctuations of, the properly centered and normalized, largest eigenvalue of a matrix taken from the Gaussian Unitary Ensemble (GUE). Since then, the fluctuations of some apparently disconnected models have also been shown to be governed by the same limiting law. This is, for example, the case for the length of the longest increasing subsequence of a random permutation (Baik, Deift and Johansson [8]) as well as some last-passage time percolation problems (Johansson [38]). We refer the reader to [44] for a survey of such topics. Following this path, the relevance of maximal eigenvalues to longest increasing subsequence problems will be further studied in Chapter III and Chapter IV. Young tableaux are closely related to subsequence problems, last-passage time models and Gaussian ensembles, and so they provide one of the main tools in our approach. We consider the finite alphabet random word problem, both uniform and non-uniform, and prove that the limiting shape of the corresponding Young tableaux is the spectrum of a certain matrix ensemble. Then, letting the size of the alphabet vary, we obtain the Tracy-Widom distribution.

CHAPTER II

CONCENTRATION INEQUALITIES

2.1 Spectral Measure of Random Matrices

The study of concentration of measure phenomenon started in the previous century and it has interesting applications in probability theory (see Ledoux [42]). One of the most powerful example is the Gaussian concentration which states, if a Borel set $B \subset \mathbb{R}^n$ is of canonical Gaussian measure $\gamma(B) \geq 1/2$, for each $r \geq 0$,

$$\gamma(B_r) \geq 1 - e^{-r^2/2},$$

where B_r is the r-th Euclidean neighborhood of B and where

$$\gamma(dx) = (2\pi)^{-n/2} e^{-\sum_{i=1}^n x_i^2/2} dx.$$

This Gaussian concentration property can be expressed equivalently on functions. Let F be a Lipschitz function on $\mathbb{R}^n$ with $\|F\|_{Lip} \leq 1$. The set $B = \{F \leq m(F)\}$ has measure $\gamma(B) \geq 1/2$, where $m(F)$ is any median of F with respect to γ, and moreover for each $r \geq 0$, $\{F - m(F) \leq r\} \subset B_r$. Thus,

$$\gamma\left(F \geq m(F) + r\right) \leq e^{-r^2/2},$$

which when combined with the same inequality for $-F$, namely,

$$\gamma\left(-F \geq m(-F) + r\right) \leq e^{-r^2/2},$$

gives,

$$\gamma\left(\left|F - m(F)\right| \geq r\right) \leq 2e^{-r^2/2}. \tag{2.1.1}$$

A remarkable feature of (2.1.1) is its dimension free character, i.e., its right hand side does not depend on the dimension n. The same inequality holds true when the median $m(F)$ is replaced by the mean $\int F d\gamma$. A natural application of measure concentration to random matrix theory is to provide concentration inequalities for the eigenvalues. For symmetric (or Hermitian) matrices, since all the real eigenvalues are Lipschitz functions of the entries, concentration inequalities can be obtained for various ensembles. This is, for example, the case for the Gaussian Unitary Ensemble (GUE). With this approach, the large deviation bound obtained for the maximal eigenvalue is of the correct order, however, the order of the small deviation is not.

Besides the well known results for Gaussian measures, concentration for infinitely divisible, and in particular for the α-stable measures (Houdré [28], Houdré and Marchal [31], Houdré and Breton [14]) have been recently obtained. These provide the main tools for the following investigations.

Following the lead of Guionnet and Zeitouni [25], let us start by setting our notation and framework.

Let $\mathcal{M}_{N\times N}(\mathbb{C})$ be the set of $N \times N$ Hermitian matrices with complex entries, which is throughout equipped with the Hilbert-Schmidt (or Frobenius or entrywise Euclidean) norm:

$$\|\mathbf{M}\| = \sqrt{tr(\mathbf{M}^*\mathbf{M})} = \sqrt{\sum_{i,j=1}^{N} |\mathbf{M}_{i,j}|^2}.$$

Let f be a real valued function on $\mathbb{R}$. The function f can be viewed as mapping $\mathcal{M}_{N\times N}(\mathbb{C})$ to $\mathcal{M}_{N\times N}(\mathbb{C})$. Indeed, for $\mathbf{M} = (\mathbf{M}_{i,j})_{1\le i,j\le N} \in \mathcal{M}_{N\times N}(\mathbb{C})$, so that $\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{U}^*$, where $\mathbf{D}$ is a diagonal matrix, with real entries $\lambda_1, ..., \lambda_N$, and $\mathbf{U}$ is a unitary matrix, set

$$f(\mathbf{M}) = \mathbf{U} f(\mathbf{D})\mathbf{U}^*, \; f(\mathbf{D}) = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_N) \end{pmatrix}.$$

Let $tr(\mathbf{M}) = \sum_{i=1}^{N} \mathbf{M}_{i,i}$ be the trace operator on $\mathcal{M}_{N\times N}(\mathbb{C})$ and set also

$$tr_N(\mathbf{M}) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{M}_{i,i}.$$

For a $N \times N$ random Hermitian matrix with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_N$, let $F_N(x) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\{\lambda_i \le x\}}$ be the corresponding empirical spectral distribution function. As well known, if $\mathbf{M}$ is a $N \times N$ Hermitian Wigner matrix with $\mathbb{E}[\mathbf{M}_{1,1}] = \mathbb{E}[\mathbf{M}_{1,2}] = 0$, $\mathbb{E}[|\mathbf{M}_{1,2}|^2] = 1$, and $\mathbb{E}[\mathbf{M}_{1,1}^2] < \infty$, the spectral measure of $\mathbf{M}/\sqrt{N}$ converges to the semicircle law: $\sigma(dx) = \sqrt{4-x^2}\mathbf{1}_{\{|x|\le 2\}}dx/2\pi$ ([3]).

We study below the tail behavior of either the spectral measure or the linear statistic of $f(\mathbf{M})$ for classes of matrices $\mathbf{M}$. Still following Guionnet and Zeitouni, we focus on a general random matrix $\mathbf{X_A}$ given as follows:

$$\mathbf{X_A} = ((\mathbf{X_A})_{i,j})_{1\le i,j\le N},\ \mathbf{X_A} = \mathbf{X_A^*},\ (\mathbf{X_A})_{i,j} = \frac{1}{\sqrt{N}} A_{i,j}\omega_{i,j},$$

with $(\omega_{i,j})_{1\le i,j\le N} = (\omega_{i,j}^R + \sqrt{-1}\omega_{i,j}^I)_{1\le i,j\le N}$, $\omega_{i,j} = \overline{\omega_{j,i}}$, and where $\omega_{i,j}$, $1 \le i \le j \le N$, is a complex valued random variable with law $P_{i,j} = P_{i,j}^R + \sqrt{-1}P_{i,j}^I$, $1 \le i \le j \le N$, with $P_{i,i}^I = \delta_0$ (by the Hermite property). Moreover, the matrix $\mathbf{A} = (A_{i,j})_{1\le i,j\le N}$ is Hermitian with, in most cases, non-random complex valued entries uniformly bounded, say, by a. Different choices for the entries of $\mathbf{A}$ allow to cover various types of ensembles. Here are some examples originating in [25].

Example 2.1.1 *If $\omega_{i,j}, 1 \le i < j \le N$, and $\omega_{i,i}, 1 \le i \le N$, are iid $N(0,1)$ random variables, taking $A_{i,i} = \sqrt{2}$ and $A_{i,j} = 1$, for $1 \le i < j \le N$ gives the GOE (Gaussian Orthogonal Ensemble).*

Example 2.1.2 *If $\omega_{i,j}^R, \omega_{i,j}^I, 1 \le i < j \le N$, and $\omega_{i,i}^R, 1 \le i \le N$, are iid $N(0,1)$ random variables, taking $A_{i,i} = 1$ and $A_{i,j} = 1/\sqrt{2}$, for $1 \le i < j \le N$ gives the GUE (Gaussian Unitary Ensemble).*

Example 2.1.3 *If $\omega^R_{i,j}, \omega^I_{i,j}, 1 \leq i < j \leq N$, and $\omega^R_{i,i}, 1 \leq i \leq N$, are two independent families of real valued random variables, taking $A_{i,j} = 0$ for $|i-j|$ large and $A_{i,j} = 1$ otherwise, gives band matrices.*

Proper choices of non-random $A_{i,j}$ also make it possible to cover Wishart matrices, as seen in the later part of this section. In certain instances, A can also be chosen to be random, like in the case of diluted matrices, in which case $A_{i,j}, 1 \leq i \leq j \leq N$, are iid Bernoulli random variables (see [25]).

On $\mathbb{R}^{N^2}$, let $\mathbb{P}^N$ be the joint law of the random vector $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})$, $1 \leq i < j \leq N$, where it is understood that the indices for $\omega^R_{i,i}$ are $1 \leq i \leq N$. Let $\mathbb{E}^N$ be the corresponding expectation. Denote by $\hat{\mu}^N_A$, the empirical spectral measure of the eigenvalues of $\mathbf{X_A}$, and further note that

$$tr_N f(\mathbf{X_A}) = \frac{1}{N} tr(f(\mathbf{X_A})) = \int_{\mathbb{R}} f(x) \hat{\mu}^N_{\mathbf{A}}(dx),$$

for any bounded Borel function f. For a Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$, set

$$\|f\|_{Lip} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|},$$

where throughout $\|\cdot\|$ is the Euclidean norm.

Definition 2.1.1 *For any $c > 0$, $Lip(c) := \{f : \|f\|_{Lip} \leq c\}$.*

Each element $\mathbf{M}$ of $\mathcal{M}_{N\times N}(\mathbb{C})$ has a unique collection of eigenvalues $\lambda = \lambda(\mathbf{M}) = (\lambda_1, \cdots, \lambda_N)$ listed in non increasing order according to multiplicity in the simplex

$$\mathcal{S}^N = \{\lambda_1 \geq \cdots \geq \lambda_N : \lambda_i \in \mathbb{R}, 1 \leq i \leq N\},$$

where throughout $\mathcal{S}^N$ is equipped with the Euclidian norm $\|\lambda\| = \sqrt{\sum_{i=1}^N \lambda_i^2}$. It is a classical result sometimes called Lidskii's theorem ([55]), that the map $\mathcal{M}_{N\times N}(\mathbb{C}) \to \mathcal{S}^N$ which associates to each Hermitian matrix its ordered list of real eigenvalues is 1-Lipschitz ([27], [42]). For a matrix $\mathbf{X_A}$ under consideration with eigenvalues $\lambda(\mathbf{X_A})$, it is then clear that the map $\varphi : (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\le i<j\le N} \mapsto \lambda(\mathbf{X_A})$ is Lipschitz, from $(\mathbb{R}^{N^2}, \|\cdot\|)$ to $(\mathcal{S}^N, \|\cdot\|)$, with Lipschitz constant bounded by $a\sqrt{2/N}$. Moreover, for any real valued Lipschitz function F on $\mathcal{S}^N$ with Lipschitz constant $\|F\|_{Lip}$, the map $F \circ \varphi$ is Lipschitz, from $(\mathbb{R}^{N^2}, \|\cdot\|)$ to $\mathbb{R}$, with Lipschitz constant at most $a\|F\|_{Lip}\sqrt{2/N}$.

Example 2.1.4 *The maximal eigenvalue* $\lambda_{max}(\mathbf{X_A}) = \lambda_1(\mathbf{X_A})$ *and, in fact, any one of the* N *eigenvalues is a Lipschitz function with Lipschitz constant at most* $a\sqrt{2/N}$*.*

Example 2.1.5 *The spectral radius* $\rho(\mathbf{X_A}) = \max_{1\le i\le N}|\lambda_i|$ *is a Lipschitz function with Lipschitz constant at most* $a\sqrt{2/N}$*.*

Example 2.1.6 *For any Lipschitz function* $f : \mathbb{R} \to \mathbb{R}$*,* $tr_N(f(\mathbf{X_A}))$ *is a Lipschitz function with Lipschitz constant at most* $\sqrt{2}a\|f\|_{Lip}/N$*.*

These observations (and our results) are also valid for the real symmetric matrices, with proper modification of the Lipschitz constants.

2.2 Hermitian Random Matrices with Infinitely Divisible Entries

Definition 2.2.1 X *is a* d*-dimensional infinitely divisible random vector without Gaussian component,* $X \sim ID(\beta, 0, \nu)$*, if its characteristic function is given by,*

$$\begin{aligned}\varphi_X(t) &= \mathbb{E}e^{i\langle t,X\rangle} \\ &= \exp\left\{ i\langle t,\beta\rangle + \int_{\mathbb{R}^d} \left(e^{i\langle t,u\rangle} - 1 - i\langle t,u\rangle \mathbf{1}_{\|u\|\le 1}\right)\nu(du)\right\}, \end{aligned} \tag{2.2.1}$$

where $t, \beta \in \mathbb{R}^d$ *and* $\nu \not\equiv 0$ *(the Lévy measure) is a positive measure on* $\mathcal{B}(\mathbb{R}^d)$*, the Borel* σ*-field of* $\mathbb{R}^d$*, without atom at the origin, and such that* $\int_{\mathbb{R}^d}(1 \wedge \|u\|^2)\nu(du) < +\infty$*.*

Example 2.2.1 *Well known examples of the infinitely divisible distributions are Gaussian and the Cauchy distribution on* $\mathbb{R}^d$*, for any* $d \ge 1$*. The Poisson, geometric, negative binomial, exponential and gamma distributions on* $\mathbb{R}$ *are also infinitely divisible.*

The vector X has independent components if and only if its Lévy measure ν is supported on the axes of $\mathbb{R}^d$ and is thus of the form:

$$\nu(dx_1, \dots, dx_d) = \sum_{k=1}^{d} \delta_0(dx_1) \dots \delta_0(dx_{k-1}) \tilde{\nu}_k(dx_k) \delta_0(dx_{k+1}) \dots \delta_0(dx_d), \qquad (2.2.2)$$

for some one-dimensional Lévy measures $\tilde{\nu}_k$. Moreover, the $\tilde{\nu}_k$ are the same for all $k = 1, \dots, d$, if and only if X has identically distributed components.

We start with a proposition, which is a direct consequence of the concentration inequalities obtained in [28] for general Lipschitz function of infinitely divisible random vectors with finite exponential moment.

Proposition 2.2.2 *Let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1 \le i < j \le N}$ be a random vector with joint law $\mathbb{P}^N \sim ID(\beta, 0, \nu)$ such that $\mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty$, for some $t > 0$ and let $T = \sup\{t > 0 : \mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty\}$. Let h^{-1} be the inverse of*

$$h(s) = \int_{\mathbb{R}^{N^2}} \|u\| \left(e^{s\|u\|} - 1\right) \nu(du), \qquad 0 < s < T.$$

(i) For any Lipschitz function f,

$$\mathbb{P}^N\left(tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))] \ge \delta\right) \le \exp\left\{ - \int_0^{\frac{N\delta}{\sqrt{2}a\|f\|_{Lip}}} h^{-1}(s) ds \right\},$$

for all $0 < \delta < \sqrt{2}a\|f\|_{Lip} h\left(T^-\right)/N$.

(ii) Let $\lambda_{max}(\mathbf{X_A})$ be the largest eigenvalue of the matrix $\mathbf{X_A}$. Then,

$$\mathbb{P}^N\left(\lambda_{max}(\mathbf{X_A}) - \mathbb{E}^N[\lambda_{max}(\mathbf{X_A})] \ge \delta\right) \le \exp\left\{ - \int_0^{\frac{\sqrt{N}\delta}{\sqrt{2}a}} h^{-1}(s) ds \right\},$$

for all $0 < \delta < \sqrt{2}a h\left(T^-\right)/\sqrt{N}$.

The following proposition gives an estimate on any median (or the mean, if it exists) of a Lipschitz function of an infinitely divisible vector X. It is used in most of the results presented below. The first part is a consequence of Theorem 1 in [32], while the proof of the second part can be obtained as in [32].

Proposition 2.2.3 *Let* $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\le i<j\le N} \sim ID(\beta, 0, \nu)$ *in* $\mathbb{R}^{N^2}$. *Let* $V^2(x) = \int_{\|u\|\le x} \|u\|^2 \nu(du)$, $\bar{\nu}(x) = \int_{\|u\|>x} \nu(du)$, *and for any* $\gamma > 0$, *let* $p_\gamma = \inf\big\{x > 0 : 0 < V^2(x)/x^2 \le \gamma\big\}$. *Let* $f \in Lip(1)$, *then for any* γ *such that* $\bar{\nu}(p_\gamma) \le 1/4$,

(i) any median $m(f(X))$ *of* $f(X)$ *satisfies*

$$|m(f(X)) - f(0)| \le G_1(\gamma) := p_\gamma\Big(\sqrt{\gamma} + 3k_\gamma(1/4)\Big) + E_\gamma,$$

(ii) the mean $\mathbb{E}^N[f(X)]$ *of* $f(X)$, *if it exists, satisfies*

$$|\mathbb{E}^N[f(X)] - f(0)| \le G_2(\gamma) := p_\gamma\Big(\sqrt{\gamma} + k_\gamma(1/4)\Big) + E_\gamma,$$

where $k_\gamma(x)$, $x > 0$, *is the solution, in* y, *of the equation*

$$y - (y+\gamma)\ln\left(1 + \frac{y}{\gamma}\right) = \ln x,$$

and where

$$E_\gamma = \left(\sum_{k=1}^{N^2}\Big(\langle e_k, \beta\rangle - \int_{p_\gamma<\|y\|\le 1}\langle e_k, y\rangle\nu(dy) + \int_{1<\|y\|\le p_\gamma}\langle e_k, y\rangle\nu(dy)\Big)^2\right)^{1/2}, \quad (2.2.3)$$

with $e_1, e_2, \dots, e_{N^2}$ *being the canonical basis of* $\mathbb{R}^{N^2}$.

Our first result deals with the spectral measure of a Hermitian matrix whose entries on and above the diagonal form an infinitely divisible random vector with finite exponential moments. Below, for any $b > 0$, $c > 0$, let

$$Lip_b(c) = \Big\{f : \mathbb{R} \to \mathbb{R} \ : \|f\|_{Lip} \le c,\ \|f\|_\infty \le b\Big\},$$

while for a fixed compact set $\mathcal{K} \subset \mathbb{R}$, with diameter $|\mathcal{K}| = \sup_{x,y\in\mathcal{K}} |x - y|$, let

$$Lip_{\mathcal{K}}(c) := \{f : \mathbb{R} \to \mathbb{R} \ : \|f\|_{Lip} \le c,\ supp(f) \subset \mathcal{K}\},$$

where $supp(f)$ is the support of f.

Theorem 2.2.4 *Let* $X = (\omega_{i,i}^R, \omega_{i,j}^R, \omega_{i,j}^I)_{1\le i<j\le N}$ *be a random vector with joint law* $\mathbb{P}^N \sim ID(\beta, 0, \nu)$ *such that* $\mathbb{E}^N[e^{t\|X\|}] < +\infty$*, for some* $t > 0$*. Let* $T = \sup\{t \geq 0 : \mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty\}$ *and let* h^{-1} *be the inverse of*

$$h(s) = \int_{\mathbb{R}^{N^2}} \|u\| \big(e^{s\|u\|} - 1\big)\nu(du), \qquad 0 < s < T.$$

(i) For any compact set $\mathcal{K} \subset \mathbb{R}$*,*

$$\mathbb{P}^N\Big(\sup_{f\in Lip_{\mathcal{K}}(1)} |tr_N(f(\mathbf{X_A})) - \mathbb{E}^N\left[tr_N(f(\mathbf{X_A}))\right]| \geq \delta\Big)$$
$$\leq \frac{8|\mathcal{K}|}{\delta} \exp\left\{ -\int_0^{\frac{N\delta^2}{8\sqrt{2}a|\mathcal{K}|}} h^{-1}(s)ds\right\}, \tag{2.2.4}$$

for all $\delta > 0$ *such that* $\delta^2 < 8\sqrt{2}a|\mathcal{K}|h\left(T^-\right)/N$*.*

(ii)

$$\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f(\mathbf{X_A})) - \mathbb{E}^N\left[tr_N(f(\mathbf{X_A}))\right]| \geq \delta\Big)$$
$$\leq \frac{C(\delta,b)}{\delta} \exp\left\{ -\int_0^{\frac{N\delta^2}{\sqrt{2}aC(\delta,b)}} h^{-1}(s)ds\right\}, \tag{2.2.5}$$

for all $\delta > 0$ *such that* $\delta^2 \leq \sqrt{2}aC(\delta,b)h(T^-)/N$*, where*

$$C(\delta,b) = C\left(\frac{\sqrt{2}a}{\sqrt{N}}\Big(G_2(\gamma) + h(t_0)\Big) + b\right),$$

with $G_2(\gamma)$ *as in Proposition 2.2.3,* C *a universal constant, and with* t_0 *the solution, in* t*, of* $th(t) - \int_0^t h(s)ds - \ln(12b/\delta) = 0$*.*

Proof. For part (i), following the proof of Theorem 1.3 of [25], without loss of generality, by shift invariance, assume that $\min\{x : x \in \mathcal{K}\} = 0$. Next, for any $v > 0$, let

$$g_v(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } 0 < x < v \\ v & \text{if } x \geq v. \end{cases} \tag{2.2.6}$$

Clearly $g_v \in Lip(1)$ with $\|g_v\|_\infty = v$. Next for any function $f \in Lip_{\mathcal{K}}(1)$, any $\Delta > 0$, define recursively $f_\Delta(x) = 0$ for $x \le 0$, and for $(j-1)\Delta \le x \le j\Delta$, $j = 1, \ldots, \lceil \frac{x}{\Delta} \rceil$, let

$$f_\Delta(x) = \sum_{j=1}^{\lceil \frac{x}{\Delta} \rceil} g_\Delta^{(j)},$$

where $g_\Delta^{(j)} := (2\mathbf{1}_{\{f(j\Delta) > f_\Delta((j-1)\Delta)\}} - 1) g_\Delta(x - (j-1)\Delta)$. Then $|f - f_\Delta| \le \Delta$ and the 1-Lipschitz function f_Δ is the sum of at most $|\mathcal{K}|/\Delta$ functions $g_\Delta^{(j)} \in Lip(1)$, regardless of the function f. Now, for $\delta > 2\Delta$,

$$\begin{aligned}
&\mathbb{P}^N\left(\sup_{f \in Lip_{\mathcal{K}}(1)} \left| tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))] \right| \ge \delta \right) \\
&\le \mathbb{P}^N\Bigg(\sup_{f \in Lip_{\mathcal{K}}(1)} \Big\{ \left| tr_N(f_\Delta(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\Delta(\mathbf{X_A}))) \right| + \big| tr_N(f(\mathbf{X_A})) \\
&\qquad\qquad - tr_N(f_\Delta(\mathbf{X_A})) \big| + \left| \mathbb{E}^N[tr_N(f(\mathbf{X_A}))] - \mathbb{E}^N[tr_N(f_\Delta(\mathbf{X_A}))] \right| \Big\} \ge \delta \Bigg) \\
&\le \mathbb{P}^N\left(\sup_{f_\Delta} \left| tr_N(f_\Delta(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\Delta(\mathbf{X_A}))) \right| > \delta - 2\Delta \right) \\
&\le \frac{|\mathcal{K}|}{\Delta} \sup_{g_\Delta^{(j)} \in Lip(1)} \mathbb{P}^N\left(\left| tr_N(g_\Delta^{(j)}(\mathbf{X_A})) - \mathbb{E}^N[tr_N(g_\Delta^{(j)}(\mathbf{X_A}))] \right| \ge \frac{\Delta(\delta - 2\Delta)}{|\mathcal{K}|} \right) \\
&\le \frac{8|\mathcal{K}|}{\delta} \exp\left\{ - \int_0^{\frac{N\delta^2}{8\sqrt{2}a|\mathcal{K}|}} h^{-1}(s) ds \right\}, \qquad (2.2.7)
\end{aligned}$$

whenever $0 < \delta < \sqrt{8\sqrt{2}a|\mathcal{K}|h(T^-)/N}$, and where the last inequality follows from part (i) of the previous proposition by taking also $\Delta = \delta/4$.

In order to prove part (ii), for any $f \in Lip_b(1)$, i.e, such that $\|f\|_{Lip} \le 1$, $\|f\|_\infty \le b$, and any $\tau > 0$, let f_τ be given via:

$$f_\tau(x) = \begin{cases} f(x) & \text{if } |x| < \tau \\ f(\tau) - \text{sign}(f(\tau))(x - \tau) & \text{if } \tau \le x < \tau + |f(\tau)| \\ f(-\tau) + \text{sign}(f(-\tau))(x + \tau) & \text{if } -\tau - |f(-\tau)| < x \le -\tau \\ 0 & \text{otherwise.} \end{cases} \qquad (2.2.8)$$

Clearly $f_\tau \in Lip(1)$ and $supp(f_\tau) \subset [-\tau - |f(-\tau)|, \tau + |f(\tau)|]$. Moreover,

$$
\begin{aligned}
\sup_{f\in Lip_b(1)} &\Big| tr_N(f(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f(\mathbf{X_A})))\Big| \\
&\le \sup_{f\in Lip_b(1)} \Big| tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))\Big| \\
&\qquad + \sup_{f\in Lip_b(1)} \Big| tr_N(f(\mathbf{X_A}) - f_\tau(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}) - f_\tau(\mathbf{X_A}))]\Big| \\
&\le \sup_{f\in Lip_b(1)} \Big| tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))\Big| \\
&\qquad + 2tr_N(g_b(|\mathbf{X_A}| - \tau)) + 2\mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))],
\end{aligned}
\tag{2.2.9}
$$

with g_b given as in (2.2.6). Now,

$$
\begin{aligned}
&\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f(\mathbf{X_A})))| \ge \delta \Big) \\
&\le \mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))| \ge \frac{\delta}{3} \Big) \\
&\qquad + \mathbb{P}^N\Big(2tr_N(g_b(|\mathbf{X_A}| - \tau)) + 2\mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))] \ge \frac{2\delta}{3} \Big) \\
&\le \mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))| \ge \frac{\delta}{3} \Big) \\
&\quad + \mathbb{P}^N\Big(tr_N(g_b(|\mathbf{X_A}| - \tau)) - \mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))] \ge \frac{\delta}{3} - 2\mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))] \Big) \\
&\le \mathbb{P}^N\Big(\sup_{f\in Lip_b(1)} |tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))| \ge \frac{\delta}{3} \Big) \\
&\quad + \mathbb{P}^N\Big(tr_N(g_b(|\mathbf{X_A}| - \tau)) - \mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}| - \tau))] \ge \frac{\delta}{3} - 2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}| \ge \tau\}}] \Big).
\end{aligned}
\tag{2.2.10}
$$

Let us first bound the second probability in (2.2.10). Recall that the spectral radius $\rho(\mathbf{X_A}) = \max_{1\le i\le N} |\lambda_i|$ is a Lipschitz function of X with Lipschitz constant at most $a\sqrt{2/N}$. Hence, for any $0 < t \le T$, and $\gamma > 0$ such that $\bar{\nu}(p_\gamma) \le 1/4$,

$$\begin{aligned}
\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})] &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{P}^N\Big(|\lambda_i(\mathbf{X_A})|\geq\tau\Big)\\
&\leq \mathbb{P}^N\big(\rho(\mathbf{X_A})\geq\tau\big)\\
&\leq \mathbb{P}^N\left(\frac{\sqrt{N}}{\sqrt{2}a}\rho(\mathbf{X_A})-\frac{\sqrt{N}}{\sqrt{2}a}\mathbb{E}^N\left[\rho(\mathbf{X_A})\right]\geq\frac{\sqrt{N}}{\sqrt{2}a}\tau-G_2(\gamma)\right)\\
&\leq \exp\left\{H(t)-\left(\frac{\sqrt{N}}{\sqrt{2}a}\tau-G_2(\gamma)\right)t\right\} \qquad (2.2.11)
\end{aligned}$$

where, above, we have used Proposition 2.2.3 in the next to last inequality and where the last inequality follows from Theorem 1 in [28] (p. 1233) with

$$H(t)=\int_0^t h(s)ds=\int_{\mathbb{R}^{N^2}}\big(e^{t\|u\|}-t\|u\|-1\big)\nu(du).$$

We want to choose τ, such that $\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})]\leq\delta/12b$. This can be achieved if

$$\frac{\sqrt{N}}{\sqrt{2}a}\tau-G_2(\gamma)\geq\frac{\ln\frac{12b}{\delta}+H(t)}{t}. \qquad (2.2.12)$$

Since

$$\frac{d}{dt}\left(\frac{\ln\frac{12b}{\delta}+H(t)}{t}\right)=\frac{th(t)-\ln\frac{12b}{\delta}-H(t)}{t^2},$$

and

$$\frac{d^2}{dt^2}\left(\frac{\ln\frac{12b}{\delta}+H(t)}{t}\right)=\frac{t^3H''(t)-2t(th(t)-\ln\frac{12b}{\delta}-H(t))}{t^4},$$

it is clear that the right hand side of (2.2.12) is minimized when $t=t_0$, where t_0 is the solution of

$$th(t)-H(t)-\ln\frac{12b}{\delta}=0,$$

and the minimum is then $h(t_0)$.

Thus, if

$$\tau=C_0(\delta,b):=\frac{\sqrt{2}a}{\sqrt{N}}\Big(G_2(\gamma)+h(t_0)\Big), \qquad (2.2.13)$$

then

$$\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})]\leq\frac{\delta}{12b},$$

and so,

$$\mathbb{P}^N\left(tr_N(g_b(|\mathbf{X_A}|-\tau))-\mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}|-\tau))]\geq\frac{\delta}{3}-2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}}]\right)$$
$$\leq\mathbb{P}^N\left(tr_N(g_b(|\mathbf{X_A}|-\tau))-\mathbb{E}^N[tr_N(g_b(|\mathbf{X_A}|-\tau))]\geq\frac{\delta}{6}\right)$$
$$\leq\exp\left\{-\int_0^{\frac{N\delta}{6\sqrt{2}a}}h^{-1}(s)ds\right\},\qquad(2.2.14)$$

for all $0<\delta<6\sqrt{2}ah\left(T^-\right)/N$, where Proposition 2.2.2 is used in the last inequality.

For τ chosen as in (2.2.13), letting $\mathcal{K}=[-\tau-b,\tau+b]$, it follows that for any $f\in Lip_b(1)$, $f_\tau\in Lip_{\mathcal{K}}(1)$. By part (i), the first term in (2.2.10) is such that

$$\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)}|tr_N(f_\tau(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))|\geq\frac{\delta}{3}\Big)$$
$$\leq\mathbb{P}^N\left(\sup_{f_\tau\in Lip_{\mathcal{K}}(1)}|tr_N(f_\tau(\mathbf{X_A}))-\mathbb{E}^N[tr_N(f_\tau(\mathbf{X_A}))]|\geq\frac{\delta}{3}\right)$$
$$\leq\frac{48(C_0(\delta,b)+b)}{\delta}\exp\left\{-\int_0^{\frac{N\delta^2}{144\sqrt{2}a(C_0(\delta,b)+b)}}h^{-1}(s)ds\right\},\qquad(2.2.15)$$

for all $0<\delta^2\leq144\sqrt{2}a\big(C_0(\delta,b)+b\big)h(T^-)/N$.

Hence, returning to (2.2.10), using (2.2.14) and (2.2.15) and for

$$\delta<\min\left\{6\sqrt{2}ah\left(T^-\right)/N,\sqrt{144\sqrt{2}a\big(C_0(\delta,b)+b\big)h(T^-)/N}\right\},$$

we have

$$\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)}|tr_N(f(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f(\mathbf{X_A})))|\geq\delta\Big)$$
$$\leq2\frac{24(C_0(\delta,b)+b)}{\delta}\exp\left\{-\int_0^{\frac{N\delta}{6\sqrt{2}a}\frac{\delta}{24(C_0(\delta,b)+b)}}h^{-1}(s)ds\right\}+\exp\left\{-\int_0^{\frac{N\delta}{6\sqrt{2}a}}h^{-1}(s)ds\right\}$$
$$\leq\left(2+\frac{1}{12}\right)\frac{24(C_0(\delta,b)+b)}{\delta}\exp\left\{-\int_0^{\frac{N\delta^2}{144\sqrt{2}a(C_0(\delta,b)+b)}}h^{-1}(s)ds\right\},\qquad(2.2.16)$$

since only the case $\delta\leq2b$ presents some interest (otherwise the probability in the statement of the theorem is zero). Part (ii) is then proved.

□

Remark 2.2.5 *(i) The order of $C(\delta, b)$ in part (ii) can be made more specific. Indeed, it will be clear from the proof of this theorem (see (2.2.12)), that for any fixed t^*, $0 < t^* \leq T$,*

$$C(\delta, b) \leq C\left(\frac{\sqrt{2}a}{\sqrt{N}}\left(\frac{\ln \frac{12b}{\delta}}{t^*} + \frac{\int_0^{t^*} h(s)ds}{t^*} + G_2(\gamma)\right)\right).$$

(ii) As seen from the proof (see (2.2.11)), in the statement of the above theorem, $G_2(\gamma)$ can be replaced by $\mathbb{E}^N\left[\|X\|\right]$. Now $\mathbb{E}^N\left[\|X\|\right]$ is of order N, since

$$N \min_{j=1,2,\dots,N^2} \mathbb{E}^N\left[|X_j|\right] \leq \mathbb{E}^N\left[\|X\|\right] \leq N \max_{j=1,2,\dots,N^2} \sqrt{\mathbb{E}^N\left[X_j^2\right]}, \tag{2.2.17}$$

where the X_j, $j = 1, 2, \dots, N^2$ are the components of X. Actually, an estimate more precise than (2.2.17) is given by a result of Marcus and Rosiński [47] which asserts that if $\mathbb{E}[X] = 0$, then

$$\frac{1}{4}x_0 \leq \mathbb{E}\left[\|X\|\right] \leq \frac{17}{8}x_0,$$

where x_0 is the solution of the equation:

$$\frac{V^2(x)}{x^2} + \frac{U(x)}{x} = 1, \tag{2.2.18}$$

with $V^2(x)$ as defined before and $U(x) = \int_{\|u\| \geq x} \|u\| \nu(du)$, $x > 0$.

(iii) When the entries of X are independent, and under a finite exponential moment assumption, the dependency in N of the function h (above and below) can sometimes be improved. We refer the reader to [33] where some of these generic problems are discussed and tackled.

As usual, one can easily pass from the mean $\mathbb{E}^N[tr_N(f)]$ to any median $m(tr_N(f))$ in either (2.2.4) or (2.2.5). Indeed, we have the following proposition.

Proposition 2.2.6 *For any $0 \leq \delta \leq 2b$,*

$$\mathbb{P}^N\left(\sup_{f \in Lip_b(1)} \left|tr_N(f) - m(tr_N(f))\right| \geq \delta\right)$$
$$\leq \mathbb{P}^N\left(\sup_{g \in Lip_b(1)} \left|tr_N(g) - \mathbb{E}^N\left[tr_N(g)\right]\right| \geq \frac{\delta}{4}\right). \tag{2.2.19}$$

Proof. If

$$\sup_{f\in Lip_b(1)} |tr_N(f) - m(tr_N(f))| \geq \delta,$$

then there exists a function $f \in Lip_b(1)$ and a median $m(tr_N(f))$ of $tr_N(f)$, such that either $tr_N(f) - m(tr_N(f)) \geq \delta$ or $tr_N(f) - m(tr_N(f)) \leq -\delta$. Without loss of generality assuming the former, otherwise let the function be $-f$ and then the former inequality holds. Consider the function $g(y) = \min\left(d(y, A), \delta\right)/2,\ y \in \mathbb{R}^{N^2}$, where $A = \{tr_N(f) \leq m(tr_N(f)\}$. Clearly $g \in Lip_b(1)$, $\mathbb{E}^N[tr_N(g)] \leq \delta/4$, and therefore $tr_N(g) - \mathbb{E}^N\left[tr_N(g)\right] \geq \delta/4$, which indicates that

$$\sup_{g\in Lip_b(1)} \Big| tr_N(g) - \mathbb{E}^N\left[tr_N(g)\right]\Big| \geq \frac{\delta}{4}.$$

□

Definition 2.2.7 *The Wasserstein distance between any two probability measures μ_1 and μ_2 on $\mathbb{R}$ is defined by*

$$d_W(\mu_1, \mu_2) = \sup_{f\in Lip_b(1)} \Big| \int_{\mathbb{R}} f d\mu_1 - \int_{\mathbb{R}} f d\mu_2 \Big|. \tag{2.2.20}$$

Theorem 2.2.4 actually gives a concentration result, with respect to the Wasserstein distance, for the empirical spectral measure $\hat{\mu}^N_{\mathbf{A}}$, when it deviates from its mean $\mathbb{E}^N[\hat{\mu}^N_{\mathbf{A}}]$.

As in [25], we can also obtain a concentration result for the distance between any particular probability measure and the empirical spectral measure.

Proposition 2.2.8 *Let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\leq i<j\leq N}$ be a random vector with joint law $\mathbb{P}^N \sim ID(\beta, 0, \nu)$ such that $\mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty$, for some $t > 0$. Let $T = \sup\{t > 0 : \mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty\}$ and let h^{-1} be the inverse of $h(s) = \int_{\mathbb{R}^{N^2}} \|u\|(e^{s\|u\|} - 1)\nu(du),\ 0 < s < T$. Then, for any probability measure μ,*

$$\mathbb{P}^N\left(d_W(\hat{\mu}^N_A, \mu) - \mathbb{E}^N[d_W(\hat{\mu}^N_A, \mu)] \geq \delta\right) \leq \exp\Big\{ -\int_0^{\frac{N\delta}{\sqrt{2}a}} h^{-1}(s) ds \Big\}, \tag{2.2.21}$$

for all $0 < \delta < \sqrt{2}ah\left(T^-\right)/N$.

Proof. As a function of $x \in \mathbb{R}^{N^2}$, $d_W(\hat{\mu}_A^N, \mu)(x)$ is Lipschitz with Lipschitz constant at most $\sqrt{2}a/N$. Indeed, for $x, y \in \mathbb{R}^{N^2}$,

$$
\begin{aligned}
d_W(\hat{\mu}_A^N, \mu)(x) &= \sup_{f \in Lip_b(1)} \left| tr_N\big(f(\mathbf{X_A})(x)\big) - \int_{\mathbb{R}} f d\mu \right| \\
&\leq \sup_{f \in Lip_b(1)} \Big| tr_N(f(\mathbf{X_A})(x)) - tr_N(f(\mathbf{X_A})(y)) \Big| \\
&\qquad\qquad + \sup_{f \in Lip_b(1)} \left| tr_N(f(\mathbf{X_A})(y)) - \int_{\mathbb{R}} f d\mu \right| \\
&\leq \frac{\sqrt{2}a}{N} \|x - y\| + d_W(\hat{\mu}_A^N, \mu)(y). \qquad (2.2.22)
\end{aligned}
$$

Theorem 2.2.8 then follows from Theorem 1 in [28]. □

Of particular importance is the case of an infinitely divisible vector having boundedly supported Lévy measure. We then have:

Corollary 2.2.9 *Let $X = (\omega_{i,i}^R, \omega_{i,j}^R, \omega_{i,j}^I)_{1 \leq i < j \leq N}$ be a random vector with joint law $\mathbb{P}^N \sim ID(\beta, 0, \nu)$ such that ν has bounded support. Let $R = \inf\{r > 0 : \nu(x : \|x\| > r) = 0\}$, let $V^2\big(= V^2(R)\big) = \int_{\mathbb{R}^{N^2}} \|u\|^2 \nu(du)$, and for $x > 0$ let*

$$\ell(x) = (1 + x)\ln(1 + x) - x.$$

(i) For any $\delta > 0$,

$$
\begin{aligned}
&\mathbb{P}^N\Big(\sup_{f \in Lip_b(1)} |tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))]| \geq \delta\Big) \\
&\qquad\qquad \leq \frac{C(\delta, b)}{\delta} \exp\left\{ -\frac{V^2}{R^2} \ell\left(\frac{NR\delta^2}{\sqrt{2}aC(\delta, b)V^2} \right) \right\}, \qquad (2.2.23)
\end{aligned}
$$

where

$$C(\delta, b) = C\left(\frac{\sqrt{2}a}{\sqrt{N}} \Big(G_2(\gamma) + \frac{V^2}{R} \big(e^{t_0 R} - 1\big) \Big) + b \right),$$

with $G_2(\gamma)$ as in Proposition 2.2.3, C a universal constant, and t_0 the solution, in t, of

$$\frac{V^2}{R^2}\Big(tRe^{tR} - e^{tR} + 1\Big) = \ln \frac{12b}{\delta}.$$

(ii) For any probability measure μ on $\mathbb{R}$, and any $\delta > 0$,

$$\begin{aligned}
&\mathbb{P}^N\big(d_W(\hat{\mu}_A^N, \mu) - \mathbb{E}^N[d_W(\hat{\mu}_A^N, \mu)] \geq \delta\big) \\
&\quad \leq \exp\left\{\frac{N\delta}{\sqrt{2}aR} - \left(\frac{N\delta}{\sqrt{2}aR} + \frac{V^2}{R^2}\right)\ln\left(1 + \frac{NR\delta^2}{\sqrt{2}aV^2}\right)\right\}. \qquad (2.2.24)
\end{aligned}$$

Proof. For Lévy measures with bounded support, $\mathbb{E}^N\left[e^{t\|X\|}\right] < +\infty$, for all $t \geq 0$, and moreover

$$h(t) \leq V^2\left(\frac{e^{tR} - 1}{R}\right).$$

Hence

$$H(t) = \int_0^t h(s)ds \leq \frac{V^2}{R^2}\big(s^{tR} - 1 - tR\big),$$

and

$$\exp\left\{-\int_0^x h^{-1}(s)ds\right\} \leq \exp\left\{\frac{x}{R} - \left(\frac{x}{R} + \frac{V^2}{R^2}\right)\ln\left(1 + \frac{Rx}{V^2}\right)\right\}.$$

Thus, one can take

$$C(\delta, b) = C\left(\frac{\sqrt{2}a}{\sqrt{N}}\left(G_2(\gamma) + \frac{V^2}{R}\big(e^{t_0R} - 1\big)\right) + b\right),$$

where t_0 is the solution, in t, of

$$\frac{V^2}{R^2}\left(tRe^{tR} - e^{tR} + 1\right) = \ln\frac{12b}{\delta}.$$

Applying Theorem 2.2.4 (ii) yields the result.

□

Remark 2.2.10 *As in Theorem 2.2.4, the dependence of $C(\delta, b)$ on δ and b can be made more precise. A key step in the proof of (2.2.23) is to choose τ such that*

$$\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}| \geq \tau\}})] \leq \delta/12b,$$

and then $C(\delta, b)$ is determined by τ. Minimizing, in t, the right hand side of (2.2.11), leads to the following estimate

$$\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}| \geq \tau\}})] \leq \exp\left\{-\frac{V^2}{R^2}\ell\left(\frac{R\left(\frac{\sqrt{N}}{\sqrt{2}a}\tau - G_2(\gamma)\right)}{V^2}\right)\right\},$$

where $\ell(x) = (1+x)\ln(1+x) - x$. *For* $x \geq 1$, $2\ell(x) \geq x \ln x$. *Hence one can choose* τ *to be the solution, in* x, *of the equation*

$$\frac{x}{R} \ln \frac{xR}{V^2} = 2 \ln \frac{12b}{\delta}.$$

It then follows that $C(\delta, b)$ *can be taken to be*

$$C\left(\frac{\sqrt{2}a}{\sqrt{N}}\Big(G_2(\gamma) + \tau\Big) + b\right).$$

2.3 Hermitian Random Matrices with Stable Entries

Without the finite exponential moment assumption, an interesting class of random matrices with infinitely divisible entries are the ones with stable entries, which we now analyze.

Definition 2.3.1 *X in $\mathbb{R}^d$ is an α-stable random vector, $(0 < \alpha < 2)$, if its Lévy measure ν is given, for any Borel set $B \in \mathcal{B}(\mathbb{R}^d)$, by*

$$\nu(B) = \int_{S^{d-1}} \sigma(d\xi) \int_0^{+\infty} \mathbf{1}_B(r\xi) \frac{dr}{r^{1+\alpha}}, \tag{2.3.1}$$

where σ, the spherical component of the Lévy measure, is a finite positive measure on S^{d-1}, the unit sphere of $\mathbb{R}^d$.

Example 2.3.1 *The standard Cauchy distribution on $\mathbb{R}$ is stable with $\alpha = 1$. The distribution with density $c(2\pi)^{-1/2} e^{-c^2/(2x)} x^{-3/2} \mathbf{1}_{(0,+\infty)}$ is stable with $\alpha = 1/2$.*

Since the expected value of the spectral measure of a matrix with α-stable entries might fail to exist, we study the deviation from a median. Here is a sample result.

Theorem 2.3.2 *Let $0 < \alpha < 2$, and let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1 \leq i < j \leq N}$ be an α-stable random vector in $\mathbb{R}^{N^2}$, with Lévy measure ν given by (2.3.1).*

(i) Let $f \in Lip(1)$, and let $m(tr_N(f(\mathbf{X_A})))$ be any median of $tr_N(f(\mathbf{X_A}))$. Then,

$$\mathbb{P}^N\big(tr_N(f(\mathbf{X_A})) - m(tr_N(f(\mathbf{X_A}))) \geq \delta\big) \leq C(\alpha)(\sqrt{2}a)^\alpha \frac{\sigma(S^{N^2-1})}{N^\alpha \delta^\alpha}, \quad (2.3.2)$$

whenever $\delta N > \sqrt{2}a\left[2\sigma(S^{N^2-1})C(\alpha)\right]^{1/\alpha}$, and where $C(\alpha) = 4^\alpha(2-\alpha+e\alpha)/\alpha(2-\alpha)$.

(ii) Let $\lambda_{max}(\mathbf{X_A})$ be the largest eigenvalue of $\mathbf{X_A}$, and let $m(\lambda_{max}(\mathbf{X_A}))$ be any median of $\lambda_{max}(\mathbf{X_A})$, then

$$\mathbb{P}^N\big(\lambda_{max}(\mathbf{X_A}) - m(\lambda_{max}(\mathbf{X_A})) \geq \delta\big) \leq C(\alpha)(\sqrt{2}a)^\alpha \frac{\sigma(S^{N^2-1})}{N^{\alpha/2}\delta^\alpha}, \quad (2.3.3)$$

whenever $\delta\sqrt{N} > \sqrt{2}a\left[2\sigma(S^{N^2-1})C(\alpha)\right]^{1/\alpha}$, and where $C(\alpha) = 4^\alpha(2-\alpha+e\alpha)/\alpha(2-\alpha)$.

Remark 2.3.3 *Let $\mathbf{M}$ be a Wigner matrix whose entries $\mathbf{M}_{i,i}, 1\leq i\leq N, \mathbf{M}^R_{i,j}, 1\leq i<j\leq N$, and $\mathbf{M}^I_{i,j}$, $1\leq i<j\leq N$, are iid random variables, such that the distribution of $|\mathbf{M}_{1,1}|$ belongs to the domain of attraction of an α-stable distribution, i.e., for any $\delta > 0$,*

$$\mathbb{P}(|\mathbf{M}_{1,1}| > \delta) = \frac{L(\delta)}{\delta^\alpha},$$

for some slowly varying positive function L such that $\lim_{\delta\to\infty} L(t\delta)/L(\delta) = 1$, for all $t > 0$. Soshnikov [58] showed that, for any $\delta > 0$,

$$\lim_{N\to\infty} \mathbb{P}^N(\lambda_{max}(b_N^{-1}\mathbf{M}) \geq \delta) = 1 - \exp(-\delta^{-\alpha}),$$

where b_N is a normalizing factor such that $N^2 L(b_N)/b_N^\alpha \to 2$ and where $\lambda_{max}(b_N^{-1}\mathbf{M})$ is the largest eigenvalue of $b_N^{-1}\mathbf{M}$. In fact $\lim_{N\to\infty} N^{\frac{2}{\alpha}-\epsilon}/b_N = 0$ and $\lim_{N\to\infty} b_N/N^{\frac{2}{\alpha}+\epsilon} = 0$, for any $\epsilon > 0$. As stated in [31], when the random vector X is in the domain of attraction of an α-stable distribution, concentration inequalities similar to (2.3.2) or (2.3.3) can be obtained for general Lipschitz function. In particular, if the Lévy measure of X is given by

$$\nu(B) = \int_{S^{N^2-1}} \sigma(d\xi) \int_0^{+\infty} \mathbf{1}_B(r\xi)\frac{L(r)dr}{r^{1+\alpha}}, \quad (2.3.4)$$

for some slowly varying function L on $[0,+\infty)$, and if we still choose the normalizing factor b_N such that $\lim_{N\to\infty}\sigma(S^{N^2-1})L(b_N)/b_N^\alpha$ is constant, then,

$$\mathbb{P}^N\big(\lambda_{max}(b_N^{-1}\mathbf{M}) - m(\lambda_{max}(b_N^{-1}\mathbf{M})) \geq \delta\big) \leq \frac{C(\alpha)\sigma(S^{N^2-1})2^{\alpha/2}}{b_N^\alpha}\frac{L\Big(b_N\frac{\delta}{\sqrt{2}}\Big)}{\delta^\alpha}, \qquad (2.3.5)$$

whenever

$$(\delta b_N)^\alpha \geq 2^{1+\alpha/2}C(\alpha)\sigma(S^{N^2-1})L\big(b_N\delta/\sqrt{2}\big).$$

Now, recall that for an N^2 dimensional vector with iid entries, $\sigma(S^{N^2-1}) = N^2(\hat{\sigma}(1)+\hat{\sigma}(-1))$, where $\hat{\sigma}(1)$ is short for $\sigma(1,0,\dots,0)$ and similarly for $\hat{\sigma}(-1)$. Thus, for fixed N, our result gives the correct order of the upper bound for large values of δ, since for $\delta > 1$,

$$\frac{e-1}{e\delta^\alpha} \leq 1 - e^{-\delta^{-\alpha}} \leq \frac{1}{\delta^\alpha}.$$

Moreover, in the stable case, $L(\delta)$ becomes constant, and $b_N = N^{2/\alpha}$. Now, since $\lambda_{max}(N^{-2/\alpha}\mathbf{M})$ is a Lipschitz function of the entries of the matrix $\mathbf{M}$ with Lipschitz constant at most $\sqrt{2}N^{-2/\alpha}$, for any median $m(\lambda_{max}(N^{-2/\alpha}\mathbf{M}))$ of $\lambda_{max}(N^{-2/\alpha}\mathbf{M})$, we have,

$$\mathbb{P}^N\big(\lambda_{max}(N^{-\frac{2}{\alpha}}\mathbf{M}) - m(\lambda_{max}(N^{-\frac{2}{\alpha}}\mathbf{M})) \geq \delta\big) \leq C(\alpha)\frac{\big(\hat{\sigma}(1)+\hat{\sigma}(-1)\big)}{2^{\alpha/2}}\frac{1}{\delta^\alpha}, \qquad (2.3.6)$$

whenever $\delta \geq \big[2C(\alpha)\big(\hat{\sigma}(1)+\hat{\sigma}(-1)\big)\big]^{1/\alpha}$. Furthermore, using Theorem 1 in [32], it is not difficult to see that $m(\lambda_{max}(N^{-2/\alpha}\mathbf{M}))$ can be upper and lower bounded independently of N. Finally, an argument as in Remark 2.4.3 below will give a lower bound on $\lambda_{max}(N^{-2/\alpha}\mathbf{M})$ of the same order as (2.3.6).

The following proposition gives an estimate on any median of a Lipschitz function of X, where X is a stable vector. It is the version of Proposition 2.2.3 for α-stable vectors.

Proposition 2.3.4 *Let $0<\alpha<2$, and let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1\le i<j\le N}$ be an α-stable random vector in $\mathbb{R}^{N^2}$, with Lévy measure ν given by (2.3.1). Let $f \in Lip(1)$, then*

(i) any median $m(f(X))$ of $f(X)$ satisfies

$$
\begin{aligned}
&|m(f(X)) - f(0)| \\
&\le J_1(\alpha) := \left(\frac{\sigma(S^{N^2-1})}{4\alpha}\right)^{1/\alpha}\left(\sqrt{\frac{\alpha}{4(2-\alpha)}} + 3k_{\frac{\alpha}{4(2-\alpha)}}(1/4)\right) + E, \quad (2.3.7)
\end{aligned}
$$

(ii) if $1<\alpha<2$, the mean $\mathbb{E}^N[f(X)]$ of $f(X)$ satisfies

$$
\begin{aligned}
&|\mathbb{E}^N[f(X)] - f(0)| \\
&\le J_2(\alpha) := \left(\frac{\sigma(S^{N^2-1})}{4\alpha}\right)^{1/\alpha}\left(\sqrt{\frac{\alpha}{4(2-\alpha)}} + k_{\frac{\alpha}{4(2-\alpha)}}(1/4)\right) + E, \quad (2.3.8)
\end{aligned}
$$

where $k_{\alpha/4(2-\alpha)}(x)$, $x>0$, is the solution, in y, of the equation

$$
y - \left(y + \frac{\alpha}{4(2-\alpha)}\right)\ln\left(1 + \frac{4(2-\alpha)y}{\alpha}\right) = \ln x,
$$

and where

$$
\begin{aligned}
E = &\left(\sum_{k=1}^{N^2}\left(\langle e_k,\beta\rangle - \int_{\left(\frac{4\sigma(S^{N^2-1})}{\alpha}\right)^{1/\alpha}<\|y\|\le 1}\langle e_k,y\rangle\nu(dy)\right.\right. \\
&\left.\left. + \int_{1<\|y\|\le\left(\frac{4\sigma(S^{N^2-1})}{\alpha}\right)^{1/\alpha}}\langle e_k,y\rangle\nu(dy)\right)^2\right)^{1/2}, \quad (2.3.9)
\end{aligned}
$$

with $e_1, e_2, \ldots, e_{N^2}$ being the canonical basis of $\mathbb{R}^{N^2}$.

Remark 2.3.5 *When the components of X are independent, a direct computation shows that, up to a constant, as $N\to\infty$, E in both $J_1(\alpha)$ and $J_2(\alpha)$ is dominated by $\left(\sigma(S^{N^2-1})/4\alpha\right)^{1/\alpha}$.*

In complete similarity to the finite exponential moments case, we can obtain concentration results for the spectral measure of matrices with α-stable entries.

Theorem 2.3.6 *Let $0 < \alpha < 2$, and let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1 \le i < j \le N}$ be an α-stable random vector in $\mathbb{R}^{N^2}$, with Lévy measure ν given by (2.3.1).*

(i) Then,

$$\mathbb{P}^N \left(\sup_{f \in Lip_b(1)} \left| tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))] \right| \ge \delta \right) \\ \le C(\delta, b, \alpha) \frac{a^\alpha \sigma(S^{N^2-1})}{N^\alpha \delta^\alpha} \wedge 1, \tag{2.3.10}$$

where

$$C(\delta, b, \alpha) = \left(C_1(\alpha) \left(\frac{\sqrt{2}a}{\sqrt{N}} \right)^{1+\alpha} \left(\frac{J_1(\alpha)+1}{\delta} + b \right)^{1+\alpha} + C_2(\alpha) \right),$$

with $C_1(\alpha)$ and $C_2(\alpha)$ constants depending only on α, and with $J_1(\alpha)$ as in Proposition 2.3.4.

(ii) For any probability measure μ,

$$\mathbb{P}^N \left(d_W(\hat{\mu}^N_A, \mu) - m(d_W(\hat{\mu}^N_A, \mu)) \ge \delta \right) \le C(\alpha)(\sqrt{2}a)^\alpha \frac{\sigma(S^{N^2-1})}{N^\alpha \delta^\alpha}, \tag{2.3.11}$$

whenever $\delta N \ge \sqrt{2}a \left[2\sigma(S^{N^2-1}) C(\alpha) \right]^{1/\alpha}$ and where $C(\alpha) = 4^\alpha(2-\alpha+e\alpha)/\alpha(2-\alpha)$.

In order to prove Theorem 2.3.6, we first need the following lemma, whose proof is essentially as the proof of Theorem 1 in [31].

Lemma 2.3.7 *Let $X = (\omega^R_{i,i}, \omega^R_{i,j}, \omega^I_{i,j})_{1 \le i < j \le N}$ be an α-stable vector, $0 < \alpha < 2$, with Lévy measure ν given by (2.3.1). For any $x_0, x_1 > 0$, let $g_{x_0,x_1}(x) = g_{x_1}(x - x_0)$, where $g_{x_1}(x)$ is defined as in (2.2.6). Then,*

$$\mathbb{P}^N \left(\left| tr_N(g_{x_0,x_1}(\mathbf{X_A})) - \mathbb{E}^N[tr_N(g_{x_0,x_1}(\mathbf{X_A}))] \right| \ge \delta \right) \le C(\alpha) \frac{a^\alpha \sigma(S^{N^2-1})}{N^\alpha \delta^\alpha},$$

whenever $\delta^{1+\alpha} > \left(2\sqrt{2}a\right)^{1+\alpha} \sigma(S^{N^2-1}) x_1 / \alpha N^{1+\alpha}$ and where $C(\alpha) = 2^{5\alpha/2}(2e\alpha + 2 - \alpha)/\alpha(2-\alpha)$.

Proof. For any $R > 0$, X could be decomposed in distribution as $X = Y^{(R)} + Z^{(R)}$, where $Y^{(R)}$ and $Z^{(R)}$ are mutually independent infinitely divisible vectors with respective characteristic function $\psi_{Y^{(R)}} = e^{\Psi_Y^{(R)}}$ and $\psi_{Z^{(R)}} = e^{\Psi_Z^{(R)}}$. for $u \in \mathbb{R}^{N^2}$, the exponents are given by

$$\Psi_Z^{(R)}(u) = \int_{\|y\|>R} \big(e^{i\langle u,y\rangle} - 1\big)\nu(dy),$$

$$\Psi_Y^{(R)}(u) = i\langle u, \tilde{\beta}\rangle + \int_{\|y\|\le R} \big(e^{i\langle u,y\rangle} - 1 - i\langle u, y\rangle \mathbf{1}_{\{\|y\|\le 1\}}\big)\nu(dy),$$

with

$$\tilde{\beta} = \beta - \int_{\|y\|>R} y\mathbf{1}_{\{\|y\|\le 1\}}\nu(dy),$$

where the last integral is understood coordinatewise and where ν is the Lévy measure of X.

$Y^{(R)}$ has compactly supported Lévy measure, which gives the following concentration inequality [28]. For any 1-Lipschitz function f, any $x > 0$,

$$\mathbb{P}\Big(f(Y^{(R)}) - \mathbb{E}(f(Y^{(R)})) \ge x\Big) \le e^{\frac{x}{R}} \left(\frac{\sigma(S^{N^2-1})R}{(2-\alpha)R^\alpha x}\right)^{\frac{x}{R}}. \tag{2.3.12}$$

Moreover, $Z^{(R)}$ has a compound Poisson structure and is the same in law as $Z^{(R)} = \sum_{k=1}^{N_R} Z_k$, where $Z_0 = 0$, and Z_k, $k \ge 1$, are iid random vectors with the same law $\nu_{Z^{(R)}}/\nu\{\|u\| > R\}$ and N_R is an independent Poisson random variable with intensity $\nu\{\|u\| > R\}$. It is easy to see that

$$\begin{aligned}
\mathbb{P}(Z^{(R)} \ne 0) &= 1 - \mathbb{P}(Z^{(R)} = 0) \\
&\le 1 - \exp\left(-\int_{\|u\|>R} \nu(du)\right) \\
&= 1 - \exp\left(-\int_{S^{N^2-1}} \sigma(d\xi) \int_{\|r\xi\|>R} \frac{dr}{r^{1+\alpha}}\right) \\
&= 1 - \exp\left(-\frac{\sigma(S^{N^2-1})}{\alpha R^\alpha}\right) \\
&\le \frac{\sigma(S^{N^2-1})}{\alpha R^\alpha},
\end{aligned}$$

and,

$$
\begin{aligned}
&\left| \mathbb{E}^N \Big[tr_N(g_{x_0,x_1}(X)) - \mathbb{E}^N[tr_N(g_{x_0,x_1}(Y^{(R)}))\Big]\right| \\
&\quad \le \mathbb{E}^N \Big[\big| tr_N(g_{x_0,x_1}(X)) - tr_N(g_{x_0,x_1}(Y^{(R)}))\big| \mathbf{1}_{\{\|Z^{(R)}\|>x_1\}}\Big] \\
&\qquad\qquad + \mathbb{E}^N \Big[\big| tr_N(g_{x_0,x_1}(X)) - tr_N(g_{x_0,x_1}(Y^{(R)}))\big| \mathbf{1}_{\{\|Z^{(R)}\|\le x_1\}}\Big] \\
&\quad \le x_1 \mathbb{P}\left(\|Z^{(R)}\| > x_1\right) + \int_0^{x_1} \mathbb{P}(x_1 \ge \|Z^{(R)}\| > x)dx \\
&\quad = \int_0^{x_1} \mathbb{P}(\|Z^{(R)}\| > x)dx \\
&\quad \le x_1 \mathbb{P}(\|Z^{(R)}\| \ne 0) \\
&\quad \le \frac{x_1 \sigma(S^{N^2-1})}{\alpha R^\alpha}.
\end{aligned} \tag{2.3.13}
$$

If $x_1\sigma(S^{N^2-1})/\alpha R^\alpha < \delta/2$,

$$
\begin{aligned}
&\mathbb{P}^N\Big(\big| tr_N(g_{x_0,x_1}(X)) - \mathbb{E}^N[tr_N(g_{x_0,x_1}(X))]\big| > \delta\Big) \\
&\le \mathbb{P}^N\Big(\big| tr_N(g_{x_0,x_1}(Y^{(R)})) - \mathbb{E}^N[tr_N(g_{x_0,x_1}(Y^{(R)}))]\big| > \delta - \big|\mathbb{E}^N[tr_N(g_{x_0,x_1}(X)) \\
&\qquad\qquad - \mathbb{E}^N[tr_N(g_{x_0,x_1}(Y^{(R)}))]\big| + \mathbb{P}^N\Big(\|Z^{(R)}\| \ne 0\Big) \\
&\le \mathbb{P}^N\Big(\big| tr_N(g_{x_0,x_1}(Y^{(R)})) - \mathbb{E}^N[tr_N(g_{x_0,x_1}(Y^{(R)}))]\big| > \frac{\delta}{2}\Big) + \mathbb{P}^N\Big(\|Z^{(R)}\| \ne 0\Big) \\
&\le \frac{2e2^{3\alpha/2}(2a)^\alpha}{2-\alpha}\frac{\sigma(S^{N^2-1})}{N^\alpha\delta^\alpha} + \frac{2^{3\alpha/2}(2a)^\alpha\sigma(S^{N^2-1})}{\alpha N^\alpha \delta^\alpha} \\
&\le C(\alpha)\frac{a^\alpha \sigma(S^{N^2-1})}{N^\alpha\delta^\alpha},
\end{aligned} \tag{2.3.14}
$$

whenever $\delta^{1+\alpha} > \left(2\sqrt{2}a\right)^{1+\alpha}\sigma(S^{N^2-1})x_1/\alpha N^{1+\alpha}$, where $C(\alpha) = 2^{5\alpha/2}(2e\alpha + 2 - \alpha)/\alpha(2-\alpha)$, and we have used (2.3.12) in the second to last inequality with the choice of $N\delta/\sqrt{2}a = 2R$. □

Proof of Theorem 2.3.6. For part (i), first consider $f \in Lip_{\mathcal{K}}(1)$. Using the same approximation as in Theorem 2.2.4, any function $f \in Lip_{\mathcal{K}}(1)$ can be approximated by f_Δ, which is the sum of at most $|\mathcal{K}|/\Delta$ functions $g_\Delta^{(j)} \in Lip(1)$, regardless of the function f. Now, and as before, for $\delta > 2\Delta$,

$$\mathbb{P}^N\left(\sup_{f\in Lip_{\mathcal{K}}(1)}|tr_N(f(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f(\mathbf{X_A})))|\geq\delta\right)$$
$$\leq\frac{|\mathcal{K}|}{\Delta}\sup_{\substack{g_\Delta^{(j)}\in Lip_b(1)\\ j=1,\cdots,\lceil\frac{|\mathcal{K}|}{\Delta}\rceil}}\mathbb{P}^N\left(\left|tr_N(g_\Delta^{(j)}(\mathbf{X_A}))-\mathbb{E}^N[tr_N(g_\Delta^{(j)}(\mathbf{X_A}))]\right|\geq\frac{\Delta(\delta-2\Delta)}{|\mathcal{K}|}\right)$$
$$\leq\frac{4|\mathcal{K}|}{\delta}\frac{8^\alpha a^\alpha C_2(\alpha)\sigma(S^{N^2-1})|\mathcal{K}|^\alpha}{N^\alpha\delta^{2\alpha}},\tag{2.3.15}$$

whenever

$$\frac{\delta^2}{8|\mathcal{K}|}>\frac{2\sqrt{2}a}{N}\Big(\frac{\sigma(S^{N^2-1})\delta}{4\alpha}\Big)^{\frac{1}{1+\alpha}},\tag{2.3.16}$$

and where the last inequality follows from Lemma 2.3.7, taking also $\Delta=\delta/4$.

For any $f\in Lip_b(1)$, and any $\tau>0$, let f_τ be given as in (2.2.8). Then, $f_\tau\in Lip_{\mathcal{K}}(1)$, where $\mathcal{K}=[-\tau-b,\tau+b]$, and moreover,

$$\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)}|tr_N(f(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f(\mathbf{X_A})))|\geq\delta\Big)$$
$$\leq\mathbb{P}^N\Big(tr_N(g_{\tau,b}(|\mathbf{X_A}|))-\mathbb{E}^N[tr_N(g_{\tau,b}(|\mathbf{X_A}|))]\geq\frac{\delta}{3}-2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}}]\Big)$$
$$+\mathbb{P}^N\Big(\sup_{f_\tau\in Lip_{\mathcal{K}}(1)}|tr_N(f_\tau(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))|\geq\frac{\delta}{3}\Big).\tag{2.3.17}$$

The spectral radius $\rho(\mathbf{X_A})$ is a Lipschitz function of X with Lipschitz constant at most $\sqrt{2}a/\sqrt{N}$. Then by Theorem 1 in [31],

$$\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})]=\frac{1}{N}\sum_{i=1}^N\mathbb{P}^N\Big(|\lambda_i(\mathbf{X_A})|\geq\tau\Big)$$
$$\leq\mathbb{P}^N\Big(\rho(\mathbf{X_A})>\tau\Big)$$
$$\leq\mathbb{P}^N\left(\rho(\mathbf{X_A})-m(\rho(\mathbf{X_A}))>\tau-\frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha)\right)$$
$$\leq\frac{C_1(\alpha)2^{\alpha/2}a^\alpha\sigma(S^{N^2-1})}{N^{\alpha/2}\big(\tau-\frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha)\big)^\alpha},\tag{2.3.18}$$

whenever

$$\left(\tau-\frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha)\right)^\alpha\geq\frac{2C_1(\alpha)2^{\alpha/2}a^\alpha\sigma(S^{N^2-1})}{N^{\alpha/2}},\tag{2.3.19}$$

and where $C_1(\alpha) = 4^\alpha(2-\alpha+e\alpha)/\alpha(2-\alpha)$. Now, if τ is chosen such that

$$\frac{C_1(\alpha)2^{\alpha/2}a^\alpha\sigma(S^{N^2-1})}{N^{\alpha/2}\big(\tau - \frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha)\big)^\alpha} \leq \frac{\delta}{12b},$$

that is, if

$$\left(\tau - \frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha)\right)^\alpha \geq \frac{12bC_1(\alpha)2^{\alpha/2}a^\alpha\sigma(S^{N^2-1})}{\delta N^{\alpha/2}}, \tag{2.3.20}$$

it then follows that

$$\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}})] \leq \frac{\delta}{12b}.$$

Since $g_{\tau,b}(|\mathbf{X_A}|)$ is the sum of two functions of the type studied in Lemma 2.3.7 with $x_1 = b$, we have,

$$\begin{aligned}
\mathbb{P}^N\Big(tr_N(g_{\tau,b}(|\mathbf{X_A}|)) - \mathbb{E}^N[tr_N(g_{\tau,b}(|\mathbf{X_A}|))] \geq \frac{\delta}{3} - 2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}}]\Big)& \\
\leq 2\mathbb{P}^N\Big(tr_N(g_{\tau,b}(\mathbf{X_A})) - \mathbb{E}^N[tr_N(g_{\tau,b}(\mathbf{X_A}))] \geq \frac{\delta}{12}\Big)& \\
\leq 2C_2(\alpha)\frac{12^\alpha a^\alpha\sigma(S^{N^2-1})}{N^\alpha\delta^\alpha},&
\end{aligned} \tag{2.3.21}$$

whenever

$$\delta^{1+\alpha} > \Big(\frac{2\sqrt{2}a}{N}\Big)^{1+\alpha}\frac{12^{1+\alpha}\sigma(S^{N^2-1})b}{\alpha}, \tag{2.3.22}$$

and where $C_2(\alpha) = 2^{5\alpha/2}(2e\alpha+2-\alpha)/\alpha(2-\alpha)$. The respective ranges (2.3.20) and (2.3.22) suggest that one can choose, for example,

$$\tau = \frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha) + \frac{\sqrt{2}a}{\sqrt{N}}\delta.$$

Then, there exists $\delta(\alpha, a, N, \nu)$ such that for $\delta > \delta(\alpha, a, N, \nu)$,

$$\begin{aligned}
&\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)}|tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))]| \geq \delta\Big) \\
&\leq \mathbb{P}^N\Big(\sup_{f_\tau Lip_\mathcal{K}(1)}|tr_N(f_\tau(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f_\tau(\mathbf{X_A}))]| \geq \frac{\delta}{3}\Big) \\
&\quad + \mathbb{P}^N\Big(tr_N(g_{\tau,b}(|\mathbf{X_A}|)) - \mathbb{E}^N[tr_N(g_{\tau,b}(|\mathbf{X_A}|))] \geq \frac{\delta}{3} - 2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}}]\Big) \\
&\leq \frac{C_3(\alpha)a^\alpha\sigma(S^{N^2-1})\Big(\frac{\sqrt{2}a}{\sqrt{N}}J_1(\alpha) + b + \frac{\sqrt{2}a}{\sqrt{N}}\delta\Big)^{1+\alpha}}{N^\alpha\delta^{1+2\alpha}} + \frac{C_4(\alpha)a^\alpha\sigma(S^{N^2-1})}{N^\alpha\delta^\alpha},
\end{aligned}$$

where $C_3(\alpha) = 2^{4+2\alpha}12^\alpha C_2(\alpha)$, $C_4(\alpha) = 2(12^\alpha)C_2(\alpha)$ and $\delta(\alpha, a, N, \nu)$ is such that (2.3.16) and (2.3.22) hold. Part (ii) is a direct consequence of Theorem 1 of [31], since $d_W(\hat{\mu}_A^N, \mu) \in Lip(\sqrt{2}a/N)$ as shown in the proof of Proposition 2.2.8. $\square$

It is also possible to obtain concentration results for smaller values of δ. Indeed, the lower and intermediate range for the stable deviation obtained in [14] provide the appropriate tools to achieve such results. We refer to [14] for complete arguments, and only provide below a sample result.

Theorem 2.3.8 *Let $1 < \alpha < 2$, and let $X = (\omega_{i,i}^R, \omega_{i,j}^R, \omega_{i,j}^I)_{1 \le i < j \le N}$ be an α-stable random vector in $\mathbb{R}^{N^2}$, with Lévy measure ν given by (2.3.1). For any $\epsilon > 0$, there exists $\eta(\epsilon)$, and constants $D_1 = D_1\big(\alpha, a, N, \sigma(S^{N^2-1})\big)$ and $D_2 = D_2\big(\alpha, a, N, \sigma(S^{N^2-1})\big)$, such that for all $0 < \delta < \eta(\epsilon)$,*

$$\mathbb{P}^N\Big(\sup_{f \in Lip_b(1)} \big|tr_N(f(\mathbf{X_A})) - \mathbb{E}^N[tr_N(f(\mathbf{X_A}))]\big| \ge \delta\Big) \le (1+\epsilon)\frac{D_1}{\delta^{\frac{\alpha+1}{\alpha}}} \exp\Big(-D_2 \delta^{\frac{2\alpha+1}{\alpha-1}}\Big). \qquad (2.3.23)$$

Proof. For any $f \in Lip(1)$, Theorem 1 in [31] gives a concentration inequality for $f(X)$, when it deviates from one of its medians. For $1 < \alpha < 2$, a completely similar (even simpler) argument gives the following result,

$$\mathbb{P}^N\big(f(X) - \mathbb{E}^N[f(X)] \ge x\big) \le \frac{C(\alpha)\sigma(S^{N^2-1})}{x^\alpha}, \qquad (2.3.24)$$

whenever $x^\alpha \ge K(\alpha)\sigma(S^{N^2-1})$, where $C(\alpha) = 2^\alpha(e\alpha + 2 - \alpha)/(\alpha(2-\alpha))$ and $K(\alpha) = \max\big\{2^\alpha/(\alpha-1), C(\alpha)\big\}$.

Next, following the proof of Theorem 2.2.4, approximate any function $f \in Lip_b(1)$ by $f_\tau \in Lip_{[-\tau-b,\tau+b]}(1)$ defined via (2.2.8). Hence,

$$\mathbb{P}^N\Big(\sup_{f\in Lip_b(1)}|tr_N(f(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f(\mathbf{X_A})))|\geq\delta\Big)$$
$$\leq\mathbb{P}^N\Big(\sup_{f_\tau\in Lip_{\mathcal{K}}(1)}|tr_N(f_\tau(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f_\tau(\mathbf{X_A})))|\geq\frac{\delta}{3}\Big)$$
$$+\mathbb{P}^N\Big(tr_N(g_{\tau,b}(|\mathbf{X_A}|))-\mathbb{E}^N[tr_N(g_{\tau,b}(|\mathbf{X_A}|))]\geq\frac{\delta}{3}-2b\mathbb{E}^N[tr_N(\mathbf{1}_{\{|\mathbf{X_A}|\geq\tau\}}]\Big).\quad(2.3.25)$$

For $\rho(\mathbf{X_A})$ the spectral radius of the matrix $\mathbf{X_A}$, and for any τ, such that $\tau-\mathbb{E}^N[\rho(\mathbf{X_A})]\geq\Big(\frac{\sqrt{2}a}{\sqrt{N}}K(\alpha)\sigma(S^{N^2-1})\Big)^{1/\alpha}$,

$$\mathbb{E}^N\big(tr_N(\mathbf{1}_{\{|\mathbf{X_A}|>\tau\}})\big)\leq\mathbb{P}^N\Big(\rho(\mathbf{X_A})-\mathbb{E}^N[\rho(\mathbf{X_A})]\geq\tau-\mathbb{E}^N[\rho(\mathbf{X_A})]\Big)$$
$$\leq\frac{\Big(\frac{\sqrt{2}a}{\sqrt{N}}\Big)^\alpha C(\alpha)\sigma(S^{N^2-1})}{\big(\tau-\mathbb{E}^N[\rho(\mathbf{X_A})]\big)^\alpha},\quad(2.3.26)$$

where we have used, in the last inequality, (2.3.24) and the fact that $\rho(\mathbf{X_A})\in Lip(\sqrt{2}a/\sqrt{N})$. For $Q>0$, let $\tau=\mathbb{E}^N[\rho(\mathbf{X_A})]+Q\delta^{-1/\alpha}$. With this choice, we then have:

$$\mathbb{E}^N\big(tr_N(\mathbf{1}_{\{|\mathbf{X_A}|>\tau\}})\big)\leq\frac{\Big(\frac{\sqrt{2}a}{\sqrt{N}}\Big)^\alpha C(\alpha)\sigma(S^{N^2-1})}{\big(\tau-\mathbb{E}^N[\rho(\mathbf{X_A})]\big)^\alpha}$$
$$\leq\delta\frac{\Big(\frac{\sqrt{2}a}{\sqrt{N}}\Big)^\alpha C(\alpha)\sigma(S^{N^2-1})}{Q^\alpha}$$
$$\leq\frac{\delta}{12b},\quad(2.3.27)$$

provided $Q^\alpha/\delta>\sqrt{2}aK(\alpha)\sigma(S^{N^2-1})/\sqrt{N}$, and $\Big(\frac{\sqrt{2}a}{\sqrt{N}}\Big)^\alpha C(\alpha)\sigma(S^{N^2-1})/Q^\alpha\leq1/(12b)$. Now, taking $Q=\sqrt{2}a\big(12bC(\alpha)\sigma(S^{N^2-1})\big)^{1/\alpha}/\sqrt{N}$, and recalling, for $1<\alpha<2$, the lower range concentration result for stable vectors (Theorem 1 and Remark 3 in [14]): For any $\epsilon>0$, there exists $\eta_0(\epsilon)$, such that for all $0<\delta<\sqrt{2}a\|f\|_{Lip}\eta_0(\epsilon)/N$,

$$\mathbb{P}^N\big(tr_N(f(\mathbf{X_A}))-\mathbb{E}^N(tr_N(f(\mathbf{X_A})))\geq\delta\big)$$
$$\leq(1+\epsilon)\exp\left\{-\frac{\frac{2-\alpha}{10}\left(\frac{\alpha-1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}}{(\sigma(S^{N^2-1}))^{1/(\alpha-1)}}\left(\frac{N}{\sqrt{2}a\|f\|_{Lip}}\right)^{\frac{\alpha}{\alpha-1}}\delta^{\frac{\alpha}{\alpha-1}}\right\}.\quad(2.3.28)$$

With arguments as in the proof of Theorem 2.2.4, if

$$\delta < \eta(\epsilon) := \left(\frac{72\sqrt{2}a}{N}\left(\frac{\sqrt{2}a}{\sqrt{N}}J_2(\alpha) + b + \left(\frac{\sqrt{2}a}{\sqrt{N}}K(\alpha)\sigma(S^{N^2-1})\right)^{1/\alpha}\right)\eta_0(\epsilon)\right)^{1/2},$$

there exist constants $D_1\big(\alpha, a, N, \sigma(S^{N^2-1})\big)$ and $D_2\big(\alpha, a, N, \sigma(S^{N^2-1})\big)$, such that the first term in (2.3.25) is bounded above by

$$(1+\epsilon)\frac{D_1\big(\alpha, a, N, \sigma(S^{N^2-1})\big)}{\delta^{\frac{\alpha+1}{\alpha}}}\exp\Big(-D_2\big(\alpha, a, N, \sigma(S^{N^2-1})\big)\delta^{\frac{2\alpha+1}{\alpha-1}}\Big). \qquad (2.3.29)$$

Indeed, with the choice of τ above and D^* as in (2.3.30), $2(\tau+b) \leq D^*/\delta^{1/\alpha}$. Moreover, as in obtaining (2.2.7), D_1 can be chosen to be $24D^*$, while D_2 can be chosen to be

$$\frac{\frac{2-\alpha}{10}\left(\frac{\alpha-1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}}{\Big(\sigma(S^{N^2-1})\Big)^{\frac{1}{\alpha-1}}}\left(\frac{N}{\sqrt{2}a}\right)^{\frac{\alpha}{\alpha-1}}\frac{1}{\big(72D^*\big)^{\frac{\alpha}{\alpha-1}}}.$$

Next, as already mentioned, $J_2(\alpha)$ can be replaced by $\mathbb{E}^N[\|X\|]$. In fact, according to (2.2.18) and an estimation in [32], if $\mathbb{E}^N[X] = 0$, then

$$\frac{1}{4(2-\alpha)^{1/\alpha}}\sigma(S^{N^2-1})^{1/\alpha} \leq \mathbb{E}^N[\|X\|] \leq \frac{17}{8\big((2-\alpha)(\alpha-1)\big)^{1/\alpha}}\sigma(S^{N^2-1})^{1/\alpha}.$$

Finally, note that, as in the proof of Theorem 2.2.4 (ii), the second term in (2.3.25) is dominated by the first term. The theorem is then proved, with the constant $D_1\big(a, N, \sigma(S^{N^2-1})\big)$ magnified by 2. □

Remark 2.3.9 *(i) In (2.3.2), (2.3.3) or (2.3.11), the constant $C(\alpha)$ is not of the right order as $\alpha \to 2$. It is, however, a simple matter to adapt Theorem 2 of [31] to obtain, at the price of worsening the range of validity of the concentration inequalities, the right order in the constants as $\alpha \to 2$.*

(ii) Let us now provide some estimation of D_1 and D_2, which are needed for comparison with the GUE results of [25] (see (iii) below). Let $C(\alpha) = 2^\alpha(e\alpha + 2 -$

$\alpha)/(2(2-\alpha))$, $K(\alpha) = \max\left\{2^\alpha/(\alpha-1), C(\alpha)\right\}$, $L(\alpha) = \left((\alpha-1)/\alpha\right)^{\alpha/(\alpha-1)}(2-\alpha)/10$ *and let*

$$D^* = 2\left(\frac{\sqrt{2}a}{\sqrt{N}}\right)^{\frac{2\alpha-1}{\alpha}}\left(12\frac{C(\alpha)}{K(\alpha)}\right)^{\frac{1}{\alpha}}J_2(\alpha)b^{\frac{1}{\alpha}} + 2\left(\frac{\sqrt{2}a}{\sqrt{N}}\right)^{\frac{\alpha-1}{\alpha}}\left(12\frac{C(\alpha)}{K(\alpha)}\right)^{\frac{1}{\alpha}}b^{\frac{\alpha+1}{\alpha}} + \frac{2\sqrt{2}a}{\sqrt{N}}\left(12C(\alpha)\sigma(S^{N^2-1})\right)^{\frac{1}{\alpha}}b^{\frac{1}{\alpha}}. \quad (2.3.30)$$

As shown in the proof of the theorem, $D_1 = 24D^*$, *while*

$$D_2 = \frac{L(\alpha)}{\left(\sigma(S^{N^2-1})\right)^{\frac{1}{\alpha-1}}}\left(\frac{N}{\sqrt{2}a}\right)^{\frac{\alpha}{\alpha-1}}\frac{1}{(72D^*)^{\frac{\alpha}{\alpha-1}}}.$$

Thus, as $N \to +\infty$, D_1 *is of order* $N^{-1/2}\left(\sigma(S^{N^2-1})\right)^{1/\alpha}$, *while* D_2 *is of order* $N^{3\alpha/(2\alpha-2)}\left(\sigma(S^{N^2-1})\right)^{2/(1-\alpha)}$.

(iii) Guionnet and Zeitouni [25], obtained concentration results for the spectral measure of matrices with independent entries, which are either compactly supported or satisfy a logarithmic Sobolev inequality. In particular for the elements of the GUE, their upper bound of concentration for the spectral measure is

$$\frac{C_1 + b^{3/2}}{\delta^{3/2}}\exp\left\{-\frac{C_2}{8ca^2}N^2\frac{\delta^5}{(C_1+b^{3/2})^2}\right\}, \quad (2.3.31)$$

where C_1 *and* C_2 *are universal constants. In Theorem 2.3.8, the order, in* b, *of* D_1 *is at most* $b^{\alpha+1/\alpha}$, *while that of* D_2 *is at least* $b^{-(\alpha+1)/(\alpha-1)}$. *For* α *close to 2, this order is thus consistent with the one in (2.3.31). Taking into account part (ii) above, the order of the constants in (2.3.23) are correct when* $\alpha \to 2$. *Following [14] (see also Remark 4 in [46]), we can recover a suboptimal Gaussian result by considering a particular stable random vector* $X^{(\alpha)}$ *and letting* $\alpha \to 2$. *Toward this end, let* $X^{(\alpha)}$ *be the stable random vector whose Lévy measure has for spherical component* σ, *the uniform measure with total mass* $\sigma(S^{N^2-1}) = N^2(2-\alpha)$. *As* α *converges to 2,* $X^{(\alpha)}$ *converges in distribution to a standard* *normal random vector. Also, as $\alpha \to 2$, the range of δ in Theorem 2.3.8 becomes $(0, +\infty)$ while the constants in the concentration bound do converge. Thus, the right hand side of (2.3.23) becomes*

$$\frac{D_1}{\delta^{3/2}} \exp\left\{ - D_2 \delta^5 \right\},$$

which is of the same order, in δ, as (2.3.31). However our order in N is suboptimal.

(iv) *In the proof of Theorem 2.3.8, the desired estimate in (2.3.27) is achieved through a truncation of order $\delta^{-1/\alpha}$, which, when $\alpha \to 2$, is of the same order as the one used in obtaining (2.3.31). However, for the GUE result, using Gaussian concentration, a truncation of order $\sqrt{\ln(12b/\delta)}$ gives a slightly better bound, namely,*

$$\frac{C_1 \sqrt{\ln \frac{12b}{\delta}}}{\delta} \exp\left\{ - \frac{C_2 N^2 \delta^4}{8ca^2 \ln \frac{12b}{\delta}} \right\},$$

where C_1 and C_2 are absolute constants (different from those of (2.3.31)).

2.4 Wishart Matrices

Wishart matrices are of interest in many contexts, in particular as sample covariance matrices in statistics.

Definition 2.4.1 $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$ *is a complex Wishart matrix if $\mathbf{Y}$ is a $K \times N$ matrix, $K > N$, with entries $\mathbf{Y}_{i,j} = \mathbf{Y}^R_{i,j} + \sqrt{-1}\mathbf{Y}^I_{i,j}$ (a real Wishart matrix is defined similarly with $\mathbf{Y}^I_{i,j} = \delta_0$ and $\mathbf{M} = \mathbf{Y}^t\mathbf{Y}$).*

Recall that if the entries of $\mathbf{Y}$ are iid centered random variables with finite variance σ^2, the empirical distribution of the eigenvalues of $\mathbf{Y}^*\mathbf{Y}/N$ converges as $K \to \infty$, $N \to \infty$, and $K/N \to \gamma \in (0, +\infty)$ to the Marčenko-Pastur law ([7], [48]) with density

$$p_\gamma(x) = \frac{1}{2\pi x \gamma \sigma^2} \sqrt{(c_2 - x)(x - c_1)}, \quad c_1 \le x \le c_2,$$

where $c_1 = \sigma^2(1-\gamma^{-1/2})^2$ and $c_2 = \sigma^2(1+\gamma^{-1/2})^2$. When the entries of $\mathbf{Y}$ are iid normal, Johansson [36] and Johnstone [40] showed, in the complex and real case respectively, that the properly normalized largest eigenvalue converges in distribution to the Tracy-Widom law ([61], [62]). Soshnikov [57] extended the results to Wishart matrices with symmetric subgaussian entries under the condition that $K - N = O(N^{1/3})$. This last condition on $K - N$ was, very recently, removed by Péché [52]. Ben Arous and Péché [5] also proved the same universal fluctuation for the largest eigenvalue when the entries of $\mathbf{Y}$ are iid with a distribution which they called Gaussian divisible and for $K/N \to 1$. The Gaussian divisible distributions, they refer to, are those probability distributions μ on $\mathbb{C}$, which can be written as a convolution of P_μ and G_μ for some probability measure P_μ such that $\int xP_\mu(dx) = 0$ and $\int |x|^2 P_\mu(dx) = 1$, and a complex-centered Gaussian law G_μ with positive finite variance. Soshnikov and Fyodorov [60] studied the distribution of the largest eigenvalue of the Wishart matrix $\mathbf{Y}^*\mathbf{Y}$, when the entries of $\mathbf{Y}$ are iid Cauchy random variables. We are interested here in concentration for the linear statistics of the spectral measure and for the largest eigenvalue of the Wishart matrix $\mathbf{Y}^*\mathbf{Y}$, where the entries of $\mathbf{Y}$ form an infinitely divisible vector and, in particular, a stable one. We restrict our work to the complex framework, the real framework being essentially the same.

It is not difficult to see that if $\mathbf{Y}$ has iid Gaussian entries, $\mathbf{Y}^*\mathbf{Y}$ has infinitely divisible entries, each with a Lévy measure without a known explicit form. However the dependence structure among the entries of $\mathbf{Y}^*\mathbf{Y}$ prevents the vector of entries to be, itself, infinitely divisible (this is a well known fact originating with Lévy, see [53]). Thus the methodology, we used till this point, cannot be directly applied to deal with functions of eigenvalues of $\mathbf{Y}^*\mathbf{Y}$. However, concentration results can be obtained when we consider the following facts, due to Guionnet and Zeitouni [25] and already used for the same purpose in their paper.

Let

$$A_{i,j} = \begin{cases} 0 & \text{for } 1 \leq i \leq K, 1 \leq j \leq K \\ 0 & \text{for } N+1 \leq i \leq K+N, K+1 \leq j \leq K+N \\ 1 & \text{for } 1 \leq i \leq K, K+1 \leq j \leq K+N \\ 1 & \text{for } N+1 \leq i \leq K+N, 1 \leq j \leq K, \end{cases} \tag{2.4.1}$$

and

$$\omega_{i,j} = \begin{cases} 0 & \text{for } 1 \leq i \leq K, 1 \leq j \leq K \\ 0 & \text{for } N+1 \leq i \leq K+N, K+1 \leq j \leq K+N \\ \bar{Y}_{i,j} & \text{for } 1 \leq i \leq K, K+1 \leq j \leq K+N \\ Y_{i,j} & \text{for } N+1 \leq i \leq K+N, 1 \leq j \leq K, \end{cases} \tag{2.4.2}$$

then $\mathbf{X_A} = \begin{pmatrix} \mathbf{0} & \mathbf{Y}^* \\ \mathbf{Y} & \mathbf{0} \end{pmatrix} \in \mathcal{M}_{(K+N)\times(K+N)}(\mathbb{C})$, and

$$\mathbf{X_A^2} = \begin{pmatrix} \mathbf{Y}^*\mathbf{Y} & \mathbf{0} \\ \mathbf{0} & \mathbf{YY}^* \end{pmatrix}.$$

Moreover, since the spectrum of $\mathbf{Y}^*\mathbf{Y}$ differs from that of $\mathbf{YY}^*$ only by the multiplicity of the zero eigenvalue, for any function f, one has

$$tr(f(\mathbf{X_A^2})) = 2tr(f(\mathbf{Y}^*\mathbf{Y})) + (K-N)f(0),$$

and

$$\lambda_{max}(\mathbf{M}^{1/2}) = \max_{1\leq i\leq N} |\lambda_i(\mathbf{X_A})|,$$

where $\mathbf{M}^{1/2}$ is the unique positive semi-definite square root of $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$.

Next, let $\mathbb{P}^{K,N}$ be the joint law of $(\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})_{1\leq i\leq K, 1\leq j\leq N}$ on $\mathbb{R}^{2KN}$, and let $\mathbb{E}^{K,N}$ be the corresponding expectation. We present below, in the infinitely divisible case, a concentration result for the largest eigenvalue $\lambda_{max}(\mathbf{M})$, of the Wishart matrices $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$. The concentration for the linear statistic $tr_N(f(\mathbf{M}))$ could also be obtained using the above observations.

Corollary 2.4.2 *Let* $\mathbf{M} = \mathbf{Y}^*\mathbf{Y}$*, with* $\mathbf{Y}_{i,j} = \mathbf{Y}^R_{i,j} + \sqrt{-1}\mathbf{Y}^I_{i,j}$.

(i) Let $X = (\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})_{1\le i\le N, 1\le j\le K}$ *be a random vector with joint law* $\mathbb{P}^{K,N} \sim ID(\beta, 0, \nu)$ *such that* $\mathbb{E}^{K,N}[e^{t\|X\|}] < +\infty$*, for some* $t > 0$*. Let* $T = \sup\{t > 0 : \mathbb{E}^{K,N}[e^{t\|X\|}] < +\infty\}$ *and let* h^{-1} *be the inverse of*

$$h(s) = \int_{\mathbb{R}^{2KN}} \|u\|(e^{s\|u\|} - 1)\nu(du), \qquad 0 < s < T.$$

Then,

$$\mathbb{P}^{K,N}\left(\lambda_{max}(\mathbf{M}^{1/2}) - \mathbb{E}^{K,N}[\lambda_{max}(\mathbf{M}^{1/2})] \ge \delta\right) \le e^{-\int_0^{\delta/\sqrt{2}} h^{-1}(s)ds}, \qquad (2.4.3)$$

for all $0 < \delta < h(T^-)$.

(ii) Let $X = (\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})_{1\le i\le K, 1\le j\le N}$ *be an* α*-stable random vector with Lévy measure* ν *given by* $\nu(B) = \int_{S^{2KN-1}} \sigma(d\xi) \int_0^{+\infty} \mathbf{1}_B(r\xi)dr/r^{1+\alpha}$*. Then,*

$$\mathbb{P}^{K,N}\left(\lambda_{max}(\mathbf{M}^{1/2}) - m(\lambda_{max}(\mathbf{M}^{1/2})) \ge \delta\right) \le C(\alpha)(\sqrt{2})^\alpha \frac{\sigma(S^{2KN-1})}{\delta^\alpha},$$

whenever $\delta > \sqrt{2}a\left[2\sigma(S^{2KN-1})C(\alpha)\right]^{1/\alpha}$ *and where* $C(\alpha) = 4^\alpha(2-\alpha+e\alpha)/\alpha(2-\alpha)$.

Proof. As a function of $(\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})_{1\le i\le K, 1\le j\le N}$, with the choice of A made in (2.4.1), $\lambda_{max}(\mathbf{X_A}) \in Lip(\sqrt{2})$. Hence part(i) of the corollary is a direct application of Theorem 1 in [28], while part(ii) can be obtained by applying Theorem 2.3.2. □

Remark 2.4.3 *(i) As already mentioned, Soshnikov and Fyodorov ([60]) obtained the asymptotic behavior of the largest eigenvalue of the Wishart matrix* $\mathbf{Y}^*\mathbf{Y}$ *when the entries of, the* $K\times N$ *matrix,* $\mathbf{Y}$ *are iid Cauchy random variables. They further argue that although the typical eigenvalues of* $\mathbf{Y}^*\mathbf{Y}$ *are of order* KN*, the correct order for the largest one is* K^2N^2*. The above corollary combined with Remark 2.3.5 and the estimate* (2.3.7)*, shows that when the entries of* $\mathbf{Y}$

form an α-stable random vector, the largest eigenvalue of $\mathbf{Y}^*\mathbf{Y}$ is of order at most $\sigma(S^{2KN-1})^{2/\alpha}$. There is also a lower concentration result, described next, which leads to a lower bound on the order of this largest eigenvalue. Thus, from these two estimates, if the entries of $\mathbf{Y}$ are iid α-stable, the largest eigenvalue of $\mathbf{Y}^*\mathbf{Y}$ is of order $K^{2/\alpha}N^{2/\alpha}$.

(ii) Let $X \sim ID(\beta, 0, \nu)$ in $\mathbb{R}^d$, then (see Lemma 5.4 in [16]) for any $x > 0$, and any norm $\|\cdot\|_{\mathcal{N}}$ on $\mathbb{R}^d$,

$$\mathbb{P}\big(\|X\|_{\mathcal{N}} \geq x\big) \geq \frac{1}{4}\Big(1 - \exp\Big\{ -\nu\big(\{u \in \mathbb{R}^d : \|u\|_{\mathcal{N}} \geq 2x\}\big)\Big\}\Big).$$

But, $\lambda_{max}(\mathbf{M}^{1/2})$ is a norm of the vector $X = (\mathbf{Y}^R_{i,j}, \mathbf{Y}^I_{i,j})$, which we denote by $\|X\|_\lambda$, if X is a stable vector in $\mathbb{R}^{2KN}$.

$$\begin{aligned}
&\mathbb{P}^{K,N}\Big(\lambda_{max}(\mathbf{M}^{1/2}) - m(\lambda_{max}(\mathbf{M}^{1/2})) \geq \delta\Big)\\
&= \mathbb{P}^{K,N}\Big(\lambda_{max}(\mathbf{M}^{1/2}) \geq \delta + m(\lambda_{max}(\mathbf{M}^{1/2}))\Big)\\
&\geq \frac{1}{4}\Big(1 - \exp\Big\{ -\nu\big(\{\lambda_{max}(\mathbf{M}^{1/2}) \geq 2\big(\delta + m(\lambda_{max}(\mathbf{M}^{1/2}))\big)\}\big)\Big\}\Big)\\
&\geq \frac{1}{4}\Big(1 - \exp\Big\{ -\nu\big(\{\|X\|_\lambda \geq 2\big(\delta + m(\lambda_{max}(\mathbf{M}^{1/2}))\big)\}\big)\Big\}\Big)\\
&= \frac{1}{4}\Bigg(1 - \exp\Bigg\{ -\frac{\tilde{\sigma}\big(S^{2KN-1}_{\|\cdot\|_\lambda}\big)}{\alpha\big(\delta + m(\lambda_{max}(\mathbf{M}^{1/2}))\big)^\alpha}\Bigg\}\Bigg),
\end{aligned} \tag{2.4.4}$$

where $S^{2KN-1}_{\|\cdot\|_\lambda}$ is the unit sphere relative to the norm $\|\cdot\|_\lambda$ and where $\tilde{\sigma}$ is the spherical part of the Lévy measure corresponding to this norm. Moreover, if the components of X are independent, in which case the Lévy measure is supported on the axes of $\mathbb{R}^{2KN}$, $\tilde{\sigma}\big(S^{2KN-1}_{\|\cdot\|_\lambda}\big)$ is of order KN, and so, as above, the largest eigenvalue of $\mathbf{M}^{1/2}$ is of order $K^{1/\alpha}N^{1/\alpha}$.

(iii) For any function f such that $g(x) = f(x^2)$ is Lipschitz with Lipschitz constant $\|g\|_{Lip} := |||f|||_{\mathcal{L}}$, $tr(g(\mathbf{X_A})) = tr(f(\mathbf{X}^2_\mathbf{A}))$ is a Lipschitz function of the entries

of $\mathbf{Y}$ *with Lipschitz constant at most* $\sqrt{2}|||f|||_{\mathcal{L}}\sqrt{K+N}$. *Hence, under the assumptions of part (i) of Corollary 2.4.2,*

$$\mathbb{P}^{K,N}\Big(tr_N(f(\mathbf{M})) - \mathbb{E}^{K,N}[tr_N(f(\mathbf{M}))] \geq \delta\frac{K+N}{N}\Big)$$
$$\leq \exp\Big\{-\int_0^{\sqrt{2(K+N)}\delta/|||f|||_{\mathcal{L}}} h^{-1}(s)ds\Big\}, \tag{2.4.5}$$

for all $0 < \delta < |||f|||_{\mathcal{L}} h\left(T^-\right)/\sqrt{2(K+N)}$.

(iv) Under the assumptions of part (ii) of Corollary 2.4.2, for any function f *such that* $g(x) = f(x^2)$ *is Lipschitz with* $\|g\|_{Lip} = |||f|||_{\mathcal{L}}$, *and any median* $m(tr_N(f(\mathbf{M})))$ *of* $tr_N(f(\mathbf{M}))$ *we have:*

$$\mathbb{P}^{K,N}\Big(tr_N(f(\mathbf{M})) - m(tr_N(f(\mathbf{M}))) \geq \delta\frac{K+N}{N}\Big)$$
$$\leq C(\alpha)\frac{|||f|||_{\mathcal{L}}^{\alpha}}{\sqrt{2^{\alpha}(K+N)^{\alpha}}}\frac{\sigma(S^{2KN-1})}{\delta^{\alpha}}, \tag{2.4.6}$$

whenever $\delta > |||f|||_{\mathcal{L}}\left[2\sigma(S^{2KN-1})C(\alpha)\right]^{1/\alpha}/\sqrt{2(K+N)}$, *and where* $C(\alpha) = 4^{\alpha}(2-\alpha+e\alpha)/\alpha(2-\alpha)$.

Remark 2.4.4 *In the absence of finite exponential moments, the methods described in this chapter extend beyond the heavy tail case and apply to any random matrix whose entries on and above the main diagonal form an infinitely divisible vector* X. *However, to obtain explicit concentration estimates, we do need explicit bounds on* V^2 *and on* $\bar{\nu}$. *Such bounds are not always available when further knowledge on the Lévy measure of* X *is lacking.*

CHAPTER III

LONGEST INCREASING SUBSEQUENCE FOR UNIFORM FINITE ALPHABETS

3.1 Traceless GUE

For any integer $M \geq 2$, an element of the $M \times M$ GUE is a complex $M \times M$ Hermitian matrix whose entries on and above the main diagonal are independent random variables; moreover, the diagonal entries are real $N(0,1)$ random variables while both the real and imaginary parts of each off diagonal entry are independent $N(0,1/2)$ random variables.

Equivalently, the random matrix $\mathbf{X}$ is an element of the $M \times M$ GUE if it is distributed according to the probability distribution

$$\mathbb{P}(d\mathbf{X}) = \frac{1}{C_M} e^{-tr\left(\mathbf{X}^2\right)/2} d\mathbf{X}, \tag{3.1.1}$$

on the space $\mathcal{H}_M$ of $M \times M$ Hermitian matrices where $C_M = 2^{M/2}\pi^{M^2/2}$ and where

$$d\mathbf{X} = \prod_{1 \leq i \leq M} d\mathbf{X}_{i,i} \prod_{1 \leq i < j \leq M} d\mathrm{Re}\left(\mathbf{X}_{i,j}\right) d\mathrm{Im}\left(\mathbf{X}_{i,j}\right),$$

is the Lebesgue measure on $\mathcal{H}_M$.

The $M \times M$ traceless GUE is the set of all the matrices of the form $\mathbf{X} - tr(\mathbf{X})\mathbf{I}_M/M$, where $\mathbf{X}$ is an element of the $M \times M$ GUE whose trace is denoted by $tr(\mathbf{X})$. Equivalently, we also have:

Proposition 3.1.1 *A random matrix $\mathbf{X}^0$ is an element of the $M \times M$ traceless GUE if and only if $\mathbf{X}^0$ is distributed according to the probability distribution*

$$\mathbb{P}\left(d\mathbf{X}^0\right) = \pi^{-\frac{M(M-1)}{2}} e^{-\sum_{1 \leq i < j \leq M} \left|\mathbf{X}^0_{i,j}\right|^2} \gamma\left(d\mathbf{X}^0_{1,1}, ..., d\mathbf{X}^0_{M,M}\right) \prod_{1 \leq i < j \leq M} dRe\left(\mathbf{X}^0_{i,j}\right) dIm\left(\mathbf{X}^0_{i,j}\right), \tag{3.1.2}$$

on the space of $M \times M$ trace zero Hermitian matrices, where $\gamma\left(d\mathbf{X}^0_{1,1}, ..., d\mathbf{X}^0_{M,M}\right)$ is the distribution of an M-dimensional centered (degenerate) multivariate normal with covariance matrix

$$\mathbf{\Sigma}_0 := (\sigma_{i,j})_{1\le i,j\le M} = \frac{1}{M}\begin{pmatrix} M-1 & -1 & \cdots & -1 \\ -1 & M-1 & \cdots & -1 \\ \vdots & \ddots & \ddots & \vdots \\ -1 & \cdots & -1 & M-1 \end{pmatrix}.$$

Proof. Let $\mathbf{X}$ be an element of the $M \times M$ GUE, and let $\mathbf{X}^0 = \mathbf{X} - tr(\mathbf{X})\mathbf{I}_M/M$. Then,

$$\left(\mathbf{X}^0_{1,1}, \mathbf{X}^0_{2,2}, \cdots, \mathbf{X}^0_{M,M}\right)' = \mathbf{\Sigma}_0\left(\mathbf{X}_{1,1}, \mathbf{X}_{2,2}, \cdots, \mathbf{X}_{M,M}\right)'.$$

Since $(\mathbf{X}_{1,1}, ..., \mathbf{X}_{M,M}) \sim \mathbf{N}(\mathbf{0}, \mathbf{I}_M)$, and $\mathbf{\Sigma}^0\mathbf{\Sigma}^0 = \mathbf{\Sigma}^0$, it follows that $\left(\mathbf{X}^0_{1,1}, ..., \mathbf{X}^0_{M,M}\right) \sim \mathbf{N}(\mathbf{0}, \mathbf{\Sigma}^0)$. The random variables $\mathbf{X}_{1,1} - tr\left(\mathbf{X}\right)/M, ..., \mathbf{X}_{M,M} - tr\left(\mathbf{X}\right)/M$ are independent of the off diagonal entries of $\mathbf{X}$, thus the distribution of $\mathbf{X}^0$ is given by (3.1.2). On the other hand, suppose the matrix $\mathbf{X}^0$ is distributed according to the probability distribution (3.1.2). Clearly, the diagonal entries $\left\{\mathbf{X}^0_{i,i}\right\}_{1\le i\le M}$ are independent of the off diagonal ones. Moreover,

$$\mathbb{E}\left[\left(tr\mathbf{X}^0\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^{M}\left(\mathbf{X}^0_{i,i}\right)^2 + 2\sum_{1\le i<j\le M}\mathbf{X}^0_{i,i}\mathbf{X}^0_{j,j}\right] = \sum_{i,j=1}^{M}\sigma_{i,j} = 0. \tag{3.1.3}$$

Therefore, $tr\left(\mathbf{X}^0\right) = 0$. Since $\left(\mathbf{X}^0_{1,1}, ..., \mathbf{X}^0_{M,M}\right) \sim \mathbf{N}(\mathbf{0}, \mathbf{\Sigma}^0)$, and $\mathbf{\Sigma}^0\mathbf{\Sigma}^0 = \mathbf{\Sigma}^0$, there exists a vector $(Z_1, ..., Z_M) \sim \mathbf{N}(\mathbf{0}, \mathbf{I}_M)$ such that $\left(\mathbf{X}^0_{1,1}, ..., \mathbf{X}^0_{M,M}\right)' = \mathbf{\Sigma}^0\left(Z_1, ..., Z_M\right)'$, and, moreover, the vector $(Z_1, ..., Z_M)$ can be also chosen to be independent of the off diagonal entries of $\mathbf{X}^0$. Let $\mathbf{X}$ be the matrix $\mathbf{X}^0$ with the diagonal entries $\mathbf{X}^0_{1,1}, ..., \mathbf{X}^0_{M,M}$ replaced by $Z_1, ..., Z_M$. Then $\mathbf{X}$ is an element from the $M \times M$ GUE and moreover $\mathbf{X}^0 = \mathbf{X} - tr(\mathbf{X})\mathbf{I}_M/M$. □

Let $\lambda^1_{GUE,M} \ge \lambda^2_{GUE,M} \ge ... \ge \lambda^M_{GUE,M}$ be the eigenvalues, written in non-increasing order, of an element of the $M \times M$ GUE. It is well known that the empirical distribution of the eigenvalues $\left(\lambda^i_{GUE,M}/\sqrt{M}\right)_{1\le i\le M}$ converges almost surely to the semicircle law ν with density $\sqrt{4-x^2}/2\pi$, $-2\le x\le 2$. Let now $\lambda^{1,0}_{GUE,M}\ge\lambda^{2,0}_{GUE,M}\ge ...\ge\lambda^{M,0}_{GUE,M}$ be the eigenvalues, written in non-increasing order, of an element of the $M\times M$ traceless GUE. Equivalently, the semicircle law is also the almost sure limit of the empirical spectral measure for the traceless GUE.

Proposition 3.1.2 *The empirical distribution of the eigenvalues* $\left(\lambda^{i,0}_{GUE,M}/\sqrt{M}\right)_{1\le i\le M}$ *converges almost surely to the semicircle law* ν *with density* $\sqrt{4-x^2}/2\pi$, $-2\le x\le 2$.

Proof. We provide two proofs of this result. By Wigner's theorem [49], the spectral measure of an element of the $M\times M$ GUE, normalized by $\sqrt{M}$, converges weakly to the semicircle law ν almost surely, i.e., for any bounded continuous function $f:\mathbb{R}\to\mathbb{R}$,

$$\frac{1}{M}\sum_{i=1}^M f\left(\lambda^i_{GUE,M}/\sqrt{M}\right)\longrightarrow\int f(x)d\nu(x),$$

almost surely. Now, if $\mathbf{G}$ is an element of the $M\times M$ GUE with eigenvalues $\lambda^1_{GUE,M}\ge ...\ge\lambda^M_{GUE,M}$, by the very definition of the traceless GUE, for each $i=1,...,M$,

$$\lambda^{i,0}_{GUE,M}=\lambda^i_{GUE,M}-\frac{tr\left(\mathbf{G}\right)}{M}.$$

Moreover, by the strong law of large numbers,

$$\frac{tr\left(\mathbf{G}\right)}{M}=\frac{1}{M}\sum_{i=1}^M\mathbf{G}_{i,i}\xrightarrow{a.s.}0.$$

Next, for any bounded Lipschitz function f, almost surely,

$$\begin{aligned}&\left|\frac{1}{M}\sum_{i=1}^M f\left(\frac{\lambda^{i,0}_{GUE,M}}{\sqrt{M}}\right)-\int f(x)d\nu(x)\right|\\&\le\left|\frac{1}{M}\sum_{i=1}^M f\left(\frac{\lambda^{i,0}_{GUE,M}}{\sqrt{M}}\right)-\frac{1}{M}\sum_{i=1}^M f\left(\frac{\lambda^{i}_{GUE,M}}{\sqrt{M}}\right)\right|+\left|\frac{1}{M}\sum_{i=1}^M f\left(\frac{\lambda^{i}_{GUE,M}}{\sqrt{M}}\right)-\int f(x)d\nu(x)\right|.\end{aligned}\tag{3.1.4}$$

Now,

$$\left|\frac{1}{M}\sum_{i=1}^M f\left(\frac{\lambda^{i,0}_{GUE,M}}{\sqrt{M}}\right)-\frac{1}{M}\sum_{i=1}^M f\left(\frac{\lambda^{i}_{GUE,M}}{\sqrt{M}}\right)\right|\le\frac{\|f\|_{Lip}}{\sqrt{M}}\left|\frac{tr\left(\mathbf{G}\right)}{M}\right|\longrightarrow 0,\tag{3.1.5}$$

and the proposition is proved, since bounded Lipschitz functions form a determining class (see section 9.3 of [19]).

Here is an alternative proof. Recall a concentration argument from [44] for proving the almost sure convergence of the empirical spectral measure of the GUE to the semicircle law. Namely, let μ^M denote the mean spectral measure, i.e., for any bounded continuous function $g : \mathbb{R} \to \mathbb{R}$,

$$\int g d\mu^M = \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M} g\left(\frac{\lambda^i_{GUE,M}}{\sqrt{M}}\right)\right].$$

Then, by the Gaussian concentration inequality, for any bounded 1-Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ and for any $x > 0$,

$$\mathbb{P}\left(\left|\frac{1}{M}\sum_{i=1}^{M} f\left(\frac{\lambda^i_{GUE,M}}{\sqrt{M}}\right) - \int f d\mu^M\right| \geq x\right) \leq 2e^{-Mx^2/4}. \tag{3.1.6}$$

Therefore,

$$\frac{1}{M}\sum_{i=1}^{M} f\left(\frac{\lambda^i_{GUE,M}}{\sqrt{M}}\right) - \int f d\mu^M \longrightarrow 0,$$

almost surely for every bounded 1-Lipschitz function $f : \mathbb{R} \to \mathbb{R}$. Since, as $M \to \infty$, the mean spectral measure converges to the semicircle law by the (weak) Wigner's theorem [49], and again since bounded Lipschitz functions form a determining class for the weak convergence, it follows that, almost surely,

$$\frac{1}{M}\sum_{i=1}^{M} \delta_{\lambda^i_{GUE,M}/\sqrt{M}} \longrightarrow \nu,$$

weakly as probability measures on $\mathbb{R}$. Now, if $\mathbf{G}$ is an element of the $M \times M$ GUE, by the strong law of large numbers,

$$\frac{tr\left(\mathbf{G}\right)}{M} = \frac{1}{M}\sum_{i=1}^{M} \mathbf{G}_{i,i} \overset{a.s.}{\longrightarrow} 0.$$

For any 1-Lipschitz function $f : \mathbb{R} \to \mathbb{R}$,

$$\left|\frac{1}{M}\sum_{i=1}^{M} f\left(\frac{\lambda^{i,0}_{GUE,M}}{\sqrt{M}}\right) - \frac{1}{M}\sum_{i=1}^{M} f\left(\frac{\lambda^i_{GUE,M}}{\sqrt{M}}\right)\right| \leq \frac{1}{\sqrt{M}}\left|\frac{tr\left(\mathbf{G}\right)}{M}\right|.$$

Again, via concentration and the Borel-Cantelli Lemma, we have,

$$\frac{1}{M}\sum_{i=1}^{M} f\left(\frac{\lambda^{i,0}_{GUE,M}}{\sqrt{M}}\right) - \int f d\mu^M \longrightarrow 0,$$

almost surely for every bounded Lipschitz function $f : \mathbb{R} \to \mathbb{R}$. Thus, almost surely,

$$\frac{1}{M}\sum_{i=1}^{M} \delta_{\lambda^{i,0}_{GUE,M}/\sqrt{M}} \longrightarrow \nu,$$

weakly as probability measures on $\mathbb{R}$. □

The joint probability density of the eigenvalues of a matrix from the GUE is given, for any $x = (x_1, ..., x_M) \in \mathbb{R}^M$, by

$$\phi_{GUE,M}(x) := c_M e^{-\frac{\sum_{j=1}^{M} x_j^2}{2}} \prod_{1\le i<j\le M} (x_i - x_j)^2, \tag{3.1.7}$$

where $c_M = (2\pi)^{-M/2}\prod_{j=1}^{M} 1/j!$ (see Mehta [49]). Let $\lambda^{max}_{GUE,M}$ denote the largest eigenvalue of an element of the $M \times M$ GUE, its probability distribution function is then,

$$\mathbb{P}\left(\lambda^{max}_{GUE,M} \le s\right) = \int_{-\infty}^{s} \cdots \int_{-\infty}^{s} \phi_{GUE,M}(x) dx_1 \cdots dx_M,$$

for any $s \in \mathbb{R}$.

For any $s_1, ..., s_M$, let

$$\mathcal{L}_{(s_1,...,s_M)} := \Bigg\{ x = (x_1, ..., x_M) \in \mathbb{R}^M : \sum_{j=1}^{M} x_j = 0,\ x_1 > \cdots > x_M, \\ \text{and } x_j < s_j, \text{for each } j{=}1, ..., M \Bigg\}. \tag{3.1.8}$$

The following proposition gives the joint distribution function of the eigenvalues of the traceless GUE.

Proposition 3.1.3 *The joint distribution function of the ordered eigenvalues of an element of the $M \times M$ traceless GUE is given by*

$$\mathbb{P}\Big(\lambda^{1,0}_{GUE,M} \le s_1, \lambda^{2,0}_{GUE,M} \le s_2, ..., \lambda^{M,0}_{GUE,M} \le s_M\Big) \\ = \sqrt{2\pi M} M! \int_{\mathcal{L}_{(s_1,...,s_M)}} \phi_{GUE,M}(x) dx_1 \cdots dx_{M-1}, \tag{3.1.9}$$

for any $s_1, ..., s_M$.

Proof. Let $\mathbf{X}$ be an element of the $M \times M$ GUE with eigenvalues $\lambda^1_{GUE,M} \geq \cdots \geq \lambda^M_{GUE,M}$. Let $\mathbf{X}^0 = \mathbf{X} - tr(\mathbf{X})\mathbf{I}_M/M$, then for any $i = 1, ..., M$,

$$\lambda^{i,0}_{GUE,M} = \lambda^i_{GUE,M} - \frac{\sum_{j=1}^M \lambda^j_{GUE,M}}{M}.$$

Clearly,

$$\sum_{j=1}^M \lambda^{j,0}_{GUE,M} = 0.$$

Let us now compute the joint density of $\lambda^{1,0}_{GUE,M}, \cdots, \lambda^{M-1,0}_{GUE,M}$. Recall that the joint density of $\left(\lambda^1_{GUE,M}, \cdots, \lambda^M_{GUE,M}\right)$ is $M!\phi_{GUE,M}(x)$, for any $x \in \mathbb{R}^M$, where the factor $M!$ comes from the fact that the eigenvalues are ordered. Consider the change of variables from $\left(\lambda^1_{GUE,M}, \cdots, \lambda^{M-1}_{GUE,M}, \lambda^M_{GUE,M}\right)$ to $\left(\lambda^{1,0}_{GUE,M}, \cdots, \lambda^{M-1,0}_{GUE,M}, Y\right)$, where $Y = \sum_{j=1}^M \lambda^j_{GUE,M}/M$. The determinant of the Jacobian J of this transformation is given by:

$$\frac{1}{\det(J)} = \det \begin{pmatrix} \frac{M-1}{M} & -\frac{1}{M} & \cdots & \cdots & -\frac{1}{M} \\ -\frac{1}{M} & \frac{M-1}{M} & \ddots & & -\frac{1}{M} \\ \vdots & & \ddots & \ddots & \vdots \\ -\frac{1}{M} & \cdots & -\frac{1}{M} & \frac{M-1}{M} & -\frac{1}{M} \\ \frac{1}{M} & \cdots & \cdots & \frac{1}{M} & \frac{1}{M} \end{pmatrix} \tag{3.1.10}$$

$$= \det \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ \frac{1}{M} & \cdots & \cdots & \frac{1}{M} & \frac{1}{M} \end{pmatrix}$$

$$= \frac{1}{M}. \tag{3.1.11}$$

The joint density of $\left(\lambda^{1,0}_{GUE,M},\cdots,\lambda^{M-1,0}_{GUE,M},Y\right)$ is then given by:

$$\begin{aligned}
& f_{\lambda^{1,0}_{GUE,M},\cdots,\lambda^{M-1,0}_{GUE,M},Y}\left(x^0_1,...,x^0_{M-1},y\right)\\
&\qquad = M!Mc_M e^{-\frac{My^2+\sum_{j=1}^M x_j^{0\,2}}{2}} \prod_{1\le i<j\le M}(x_i^0-x_j^0)^2\\
&\qquad = \sqrt{2\pi M}M!c_M\frac{\sqrt{M}}{\sqrt{2\pi}}e^{-\frac{My^2}{2}}e^{-\frac{\sum_{j=1}^M x_j^{0\,2}}{2}} \prod_{1\le i<j\le M}(x_i^0-x_j^0)^2\\
&\qquad = \sqrt{2\pi M}\frac{\sqrt{M}}{\sqrt{2\pi}}e^{-\frac{My^2}{2}}M!\phi_{GUE,M}(x^0), \qquad (3.1.12)
\end{aligned}$$

where $c_M = (2\pi)^{-M/2}\prod_{j=1}^M 1/j!$, $x^0 = (x_1^0,...,x_M^0)$ and $x_M^0 = -\sum_{j=1}^{M-1}x_j^0$. Integrating y from $-\infty$ to ∞, the joint density of $\left(\lambda^{1,0}_{GUE,M},\cdots,\lambda^{M-1,0}_{GUE,M}\right)$ is thus $\sqrt{2\pi M}M!\phi_{GUE,M}(x^0)$. □

For any $s>0$, set

$$\mathcal{L}_s := \Big\{ x=(x_1,...,x_M)\in\ \mathbb{R}^M : \sum_{j=1}^M x_j = 0,\ x_1>\cdots>x_M, \text{ and } x_j<s, \text{for each } j=1,...,M \Big\}, \qquad (3.1.13)$$

If $\lambda^{max,0}_{GUE,M}$ denotes the largest eigenvalue of an element of the $M\times M$ traceless GUE, then, from the previous result,

$$\mathbb{P}\left(\lambda^{max,0}_{GUE,M}\le s\right) = \sqrt{2\pi M}M!\int_{\mathcal{L}_s}\phi_{GUE,M}(x)dx_1\cdots dx_{M-1}. \qquad (3.1.14)$$

Extending a computation from [64], the next proposition gives a relation in law between the whole spectrum of an element from the $M\times M$ GUE and the spectrum of an element of the $M\times M$ traceless GUE.

Proposition 3.1.4 *For any $M\ge 2$, let $\lambda^1_{GUE,M}\ge\cdots\ge\lambda^M_{GUE,M}$ be the eigenvalues of an element of the $M\times M$ GUE, and let $\lambda^{1,0}_{GUE,M}\ge\cdots\ge\lambda^{M,0}_{GUE,M}$ be the eigenvalues of an element of the $M\times M$ traceless GUE. Then*

$$\left(\lambda^1_{GUE,M},\cdots,\lambda^M_{GUE,M}\right) \overset{d}{=} \left(\lambda^{1,0}_{GUE,M},\cdots,\lambda^{M,0}_{GUE,M}\right)+\left(Z^1_M,\cdots,Z^M_M\right),$$

where $\left(Z_M^1, \cdots, Z_M^M\right)$ *is a centered normal vector with covariance matrix*

$$\frac{1}{M}\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}. \tag{3.1.15}$$

Moreover, $\left(\lambda_{GUE,M}^{1,0}, \cdots, \lambda_{GUE,M}^{M,0}\right)$ *and* $\left(Z_M^1, \cdots, Z_M^M\right)$ *are independent.*

Proof. Recall that, for any $s_1, ..., s_M$,

$$\begin{aligned} &\mathbb{P}\Big(\lambda_{GUE,M}^1 \le s_1, ..., \lambda_{GUE,M}^M \le s_M\Big) \\ &\qquad = c_M \int_{-\infty}^{s_1} \cdots \int_{-\infty}^{s_M} e^{-\frac{\sum_{j=1}^M x_j^2}{2}} \prod_{1\le i<j\le M} (x_i - x_j)^2 dx_1 \cdots dx_M, \end{aligned} \tag{3.1.16}$$

where $c_M = (2\pi)^{-M/2} \prod_{j=1}^M 1/j!$. Consider the change of variables

$$y = \sum_{j=1}^M x_j,\ x_j = x_j^0 + \frac{y}{M}, \qquad j = 1, ..., M. \tag{3.1.17}$$

Clearly, $\sum_{j=1}^M x_j^0 = 0$, and changing the variables x_j^0 back to x_j, we get

$$\begin{aligned} &\mathbb{P}\Big(\lambda_{GUE,M}^1 \le s_1, ..., \lambda_{GUE,M}^M \le s_M\Big) \\ &= c_M \int_{-\infty}^{\infty} dy \int_{\mathcal{L}_{\left(s_1 - \frac{y}{M}, ..., s_M - \frac{y}{M}\right)}} e^{-\frac{\frac{y^2}{M} + \sum_{j=1}^M {x_j^0}^2}{2}} \prod_{1\le i<j\le M} (x_i^0 - x_j^0)^2 dx^0 \\ &= M c_M \int_{-\infty}^{\infty} e^{-\frac{My^2}{2}} dy \int_{\mathcal{L}_{(s_1 - y, ..., s_M - y)}} e^{-\frac{\sum_{j=1}^M {x_j}^2}{2}} \prod_{1\le i<j\le M} (x_i - x_j)^2 dx^0 \\ &= \int_{-\infty}^{\infty} \frac{\sqrt{M}}{\sqrt{2\pi}} e^{-\frac{My^2}{2}} \mathbb{P}\left(\lambda_{GUE,M}^{1,0} < s_1 - y, \cdots, \lambda_{GUE,M}^{M,0} < s_M - y\right) dy \\ &= \mathbb{E}\left[\mathbb{P}\left(\lambda_{GUE,M}^{1,0} < s_1 - Y, \cdots, \lambda_{GUE,M}^{M,0} < s_M - Y \Big| Y\right)\right], \end{aligned} \tag{3.1.18}$$

where dx^0 is the Lebesgue measure on $\left\{x = (x_1, ..., x_M) \in \mathbb{R}^M : \sum_{l=1}^M x_l = 0\right\}$ and $Y \sim N(0, 1/M)$. The right hand side of (3.1.18) is the distribution function of the sum of the independent random vectors $\left(\lambda_{GUE,M}^{1,0}, \cdots, \lambda_{GUE,M}^{M,0}\right)$ and $\left(Z_M^1, \cdots, Z_M^M\right)$, where $\left(Z_M^1, \cdots, Z_M^M\right) \stackrel{d}{=} \left(1/\sqrt{M}, ..., 1/\sqrt{M}\right) Z$ with $Z \sim N(0,1)$, i.e., the vector $\left(Z_M^1, \cdots, Z_M^M\right)$ is a centered normal vector with covariance matrix given by (3.1.15).

□

3.2 Homogeneous Random Words

Let us now turn to subsequences.

Definition 3.2.1 *Let $X_1, X_2, ..., X_N$ be a finite sequence of random variables taking values in an ordered set. The length of the longest (weakly) increasing subsequence of $X_1, X_2, ..., X_N$, denoted by LI_N, is the maximal $k \leq N$ such that there exists an increasing sequence of integers $1 \leq i_1 < i_2 < \cdots < i_k \leq N$ with $X_{i_1} \leq X_{i_2} \leq \cdots \leq X_{i_k}$, i.e.,*

$$LI_N = \max \{k : \exists\ 1 \leq i_1 < i_2 < \cdots < i_k \leq N, \ \textit{with}\ X_{i_1} \leq X_{i_2} \leq \cdots \leq X_{i_k}\}.$$

A very efficient method to study longest increasing subsequences is through Young tableaux ([1], [18], [39]). To further study asymptotics related to the spectrum of random matrices and the shapes of Young tableaux, let us recall a few definitions and facts.

Definition 3.2.2 *A Young diagram of size n is a collection of n boxes arranged in left-justified rows, with a weakly decreasing number of boxes from row to row.*

Definition 3.2.3 *A (semi-standard) Young tableau is a Young diagram, with a filling of a positive integer in each box, in such a way that the integers are weakly increasing along the rows and strictly increasing down the columns. A standard Young tableau of size n is a Young tableau in which the fillings are the integers from 1 to n. The shape of a Young tableau is the vector $\lambda = (\lambda_1, \lambda_2, ..., \lambda_k)$, where $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$ and for each i, λ_i is the number of boxes in the i-th row while k is the total number of rows of the tableau (and so $\lambda_1 + \cdots + \lambda_k = n$).*

When $X_1, X_2, ..., X_N$ is a random permutation of $1, 2, ..., N$, Baik, Deift and Johansson [8] showed that,

$$\frac{LI_N - 2\sqrt{N}}{N^{1/6}} \Longrightarrow F_{TW},$$

where $\Longrightarrow$ indicates convergence in distribution and where F_{TW} is the Tracy-Widom distribution. When each X_i takes values independently and uniformly in an M-letter ordered alphabet, through a careful analysis of the exponential generating function of LI_N, Tracy and Widom [64] gave the limiting distribution of LI_N (properly centered and normalized) as that of the largest eigenvalue of a matrix drawn from the $M \times M$ traceless GUE.

Johansson [39], using discrete orthogonal polynomial methods, proved that the limiting shape of the Young tableaux, associated to the iid uniform m-letter framework through the RSK correspondence (described below), is given by the joint distribution of the eigenvalues of the traceless $M \times M$ GUE. Since LI_N is equal to the length of the first row of the associate Young tableaux, the results of [64] were recovered. The permutation case is also obtained by Johansson [39] and, independently, by Okounkov [50] as well as Borodin, Okounkov and Olshanki [13]. Its, Tracy and Widom ([34], [35]) further analyzed the independent but no longer uniform framework. They showed that the corresponding limiting distribution of LI_N can be written in terms of the distribution function of the eigenvalues of the direct sum of mutually independent GUEs with an overall trace constraint. More recently, in [29] and [30], via simple probabilistic tools, the limiting law of the longest increasing subsequence for finite and countable alphabets is given as a Brownian functional. To recall their results, let

$$\tilde{H}_M = \sqrt{\frac{M-1}{M}} \sup_{0=t_0 \le t_1 \le \cdots \le t_{M-1} \le t_M = 1} \sum_{i=1}^{M} \left(\tilde{B}^i(t_i) - \tilde{B}^i(t_{i-1}) \right),$$

where $\left(\tilde{B}^1(t), \tilde{B}^2(t), ..., \tilde{B}^M(t)\right)$ is an M-dimensional Brownian motion having covariance matrix

$$t \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}, \tag{3.2.1}$$

with $\rho = -1/(M-1)$. Then,

Theorem 3.2.4 *Let $X_1, X_2, ..., X_N, ...$ be an iid sequence of random variables with values taken uniformly in an M-letter ordered alphabet. Then,*

$$\frac{LI_N - N/M}{\sqrt{N/M}} \overset{N\to\infty}{\Longrightarrow} \tilde{H}_M.$$

We investigate below the connections between random matrix models and longest increasing subsequence problems for finite alphabets through this Brownian functional. In studying a queueing problem, Glynn and Whitt [22] had previously introduced a related Brownian functional

$$D_M = \sup_{0=t_0\le t_1\le\cdots\le t_{M-1}\le t_M=1} \sum_{i=1}^{M} \big(B^i(t_i) - B^i(t_{i-1})\big), \tag{3.2.2}$$

where now $\big(B^1(t), B^2(t), ..., B^M(t)\big)$ is a standard M-dimensional Brownian motion. Moreover, analyzing a one-dimensional discrete space and time growth model, Gravner, Tracy and Widom [24] proved that,

$$D_M \overset{d}{=} \lambda^{max}_{GUE,M}, \tag{3.2.3}$$

where $\overset{d}{=}$ denotes equality in distribution. Independently, Baryshnikov [10], studying monotonous paths on the integer lattice, showed that the process $(D_M)_{M\ge 1}$ has the same law as the process of the largest eigenvalues of the principal minors of an infinite random matrix drawn from the GUE. Doumerc [18], using a matrix central limit theorem making the GUE appear as a limit of the Laguerre Unitary Ensemble, gave a probabilistic proof of (3.2.3) (also involving Young tableaux and the RSK correspondence) with further extensions to the whole spectrum as well as a process version of this equality.

We provide bellow another proof of the fact that $\tilde{H}_M$ and $\lambda^{max,0}_{GUE,M}$ are equal in distribution, with an extension to the whole spectrum. The proof only involves very basic probabilistic arguments and some combinatorics tools, namely Young tableaux and the Robinson-Schensted-Knuth (RSK) correspondence [20], which we now recall.

Definition 3.2.5 *Let $\{1, 2, ..., M\}$ be an M-letter alphabet. A word of length N is a mapping W from $\{1, 2, ..., N\}$ to $\{1, 2, ..., M\}$. Let $[M]^N$ denote the set of words of length N with letters taken from the alphabet $\{1, 2, ..., M\}$. A word is a permutation if $M = N$, and W is onto.*

(The Robinson-Schensted correspondence) The Robinson-Schensted correspondence is a bijection between the set of words $[M]^N$ and the set of pairs of Young tableaux $\{(P, Q)\}$, where P is a semi-standard Young tableau with entries from $\{1, 2, ..., M\}$, Q is a standard Young tableau with entries from $\{1, 2, ..., N\}$. Moreover P and Q share the same shape which is a partition of N. From now on, we do not distinguish between shape and partition. If the word is a permutation, then P is also a standard Young tableau.

A word W in $[M]^N$ can be represented uniquely as an $M \times N$ matrix $\mathbf{X}_W$ with entries

$$(\mathbf{X}_W)_{i,j} = \mathbf{1}_{W(j)=i}. \tag{3.2.4}$$

The Robinson-Schensted correspondence actually gives a one to one correspondence between the set of pairs of Young tableaux and the set of matrices whose entries are either 0 or 1 and with exactly a unique 1 in each column. This was generalized by Knuth to the set of $M \times N$ matrices with nonnegative integer entries.

(The Knuth generalizations) Let $\mathcal{M}(M, N)$ be the set of $M \times N$ matrices with nonnegative integer entries. Let $\mathcal{P}(P, Q)$ be the set of pairs of semi-standard Young tableaux (P, Q) sharing the same shape and whose size is the sum of all the entries, where P has elements in $\{1, ..., M\}$ and Q has elements in $\{1, ..., N\}$. The Robinson-Schensted-Knuth (RSK) correspondence is a one to one mapping between $\mathcal{M}(M, N)$ and $\mathcal{P}(P, Q)$. If the matrix corresponds to a word in $[M]^N$, Q is a standard tableau. If the matrix corresponds to a permutation, both P and Q are standard tableaux. For $\mathbf{X}_W \in \mathcal{M}(M, N)$, let $\lambda(RSK(\mathbf{X}_W))$ be the common shape, i.e., $\lambda(RSK(\mathbf{X}_W)) =$

$(\lambda_1, ..., \lambda_M)$ where, for each $k = 1, ..., M$, λ_k is the length of the k-th row, of the corresponding Young tableaux of $\mathbf{X}_W$ through the RSK correspondence.

The following theorem shows that the spectrum of the traceless $M \times M$ GUE appears to be the limiting law of the, properly normalized, shape of the Young tableaux associated to a random word with letters independently and uniformly drawn from an M-letter alphabets. The proof heavily depends on the RSK correspondence and on the asymptotics of the number of Young tableaux of a given shape, as studied, for example, by Johansson in [39].

Theorem 3.2.6 *Let $\mathbf{X}_W$ be the matrix corresponding to a random word of length N as in (3.2.4), with each letter independently and uniformly draw from an M-letter alphabet. Let $\lambda(RSK(\mathbf{X}_W)) = (\lambda_1, ..., \lambda_M)$ be the common shape of the associated Young tableaux through the RSK correspondence. Then, as $N \to \infty$,*

$$\left(\frac{\lambda_1 - N/M}{\sqrt{N/M}}, ..., \frac{\lambda_M - N/M}{\sqrt{N/M}}\right) \Longrightarrow \left(\lambda^{1,0}_{GUE,M}, ..., \lambda^{M,0}_{GUE,M}\right). \tag{3.2.5}$$

Proof. For any partition $\lambda^0 = (\lambda^0_1, \lambda^0_2, ..., \lambda^0_M)$ of N,

$$\mathbb{P}(\lambda(RSK(\mathbf{X}_W)) = \lambda^0) = \frac{1}{M^N} L(\lambda^0, M, N),$$

where $L(\lambda^0, M, N)$ is the number of matrices $\mathbf{X}_W$ such that $\lambda(RSK(\mathbf{X}_W)) = \lambda^0$. Johansson [39] has studied the large N asymptotics of $L(\lambda^0, M, N)$, and this is described next. Since the RSK correspondence is one to one,

$$L(\lambda^0, M, N) = d_{\lambda^0}(M) f^{\lambda^0},$$

where $d_{\lambda^0}(M)$ is the number of Young tableaux of shape λ^0 with elements in $\{1, ..., M\}$, and f^{λ^0} is the number of Young tableaux of shape λ^0 with elements in $\{1, ..., N\}$. Expressions for these quantities are well known in combinatorics (see [20] Section 4.3, for example), namely,

$$d_{\lambda^0}(M) = \left(\prod_{j=1}^{M-1} \frac{1}{j!}\right) \prod_{1 \leq i < j \leq M} (\lambda^0_i - \lambda^0_j + j - i),$$

and

$$f^{\lambda^0} = N! \prod_{1\le i<j\le M} (\lambda_i^0 - \lambda_j^0 + j - i) \prod_{j=1}^{M} \frac{1}{(\lambda_j^0 + M - j)!}.$$

Hence,

$$\mathbb{P}(\lambda(RSK(\mathbf{X}_W)) = \lambda^0) = \frac{N!}{M^N} \left(\prod_{j=1}^{M-1} \frac{1}{j!} \right) \prod_{1\le i<j\le M} (\lambda_i^0 - \lambda_j^0 + j - i)^2 \prod_{j=1}^{M} \frac{1}{(\lambda_j^0 + M - j)!}. \tag{3.2.6}$$

The change of variables

$$x_j = \frac{\lambda_j^0 - N/M}{\sqrt{N/M}},$$

for $j = 1, ..., M$, and Stirling's formula show that, as $N \to \infty$,

$$\begin{aligned}(\lambda_j^0 + M - j)! &= \left(\sqrt{\frac{N}{M}} x_j + \frac{N}{M} + M - j \right)! \\ &\sim \sqrt{\frac{2\pi N}{M}} \left(\frac{N}{M} \right)^{N/M + M - j} \exp\left(\frac{x_j^2}{2} - \frac{N}{M} \right).\end{aligned}$$

Hence, as $N \to \infty$,

$$\prod_{j=1}^{M} \frac{1}{(\lambda_j^0 + M - j)!} \sim \left(\frac{2\pi N}{M} \right)^{-M/2} \left(\frac{M}{N} \right)^{N + M(M-1)/2} e^N \prod_{j=1}^{M} e^{-x_j^2/2}.$$

Therefore,

$$\begin{aligned}&\mathbb{P}(\lambda(RSK(\mathbf{X}_W)) = \lambda^0) \\ &\qquad \xrightarrow{N\to\infty} \sqrt{2\pi M} (2\pi)^{-M/2} 2^{M/2} \left(\prod_{j=1}^{M-1} \frac{1}{j!} \right) \prod_{1\le i<j\le M} (x_i - x_j)^2 \prod_{j=1}^{M} e^{-x_j^2/2} \\ &\qquad = \sqrt{2\pi M} M! \phi_{GUE,M}(x),\end{aligned}$$

where $\phi_{GUE,M}(x)$ is the joint density of the eigenvalues of the $M \times M$ GUE ([49], [3]).

Now, for any $s_1, ..., s_M \in \mathbb{R}$,

$$\begin{aligned}&\mathbb{P}\left(\frac{\lambda^1 - N/M}{\sqrt{N/M}} \le s_1, ..., \frac{\lambda^M - N/M}{\sqrt{N/M}} \le s_M \right) \\ &\qquad = \mathbb{E}\left[\mathbf{1}_{\frac{\lambda^1 - N/M}{\sqrt{N/M}} \le s_1, ..., \frac{\lambda^M - N/M}{\sqrt{N/M}} \le s_M} \right] \\ &\qquad = \sum_{\frac{\lambda_1^0 - N/M}{\sqrt{N/M}} \le s_1, ..., \frac{\lambda_M^0 - N/M}{\sqrt{N/M}} \le s_M} \mathbb{P}(\lambda(RSK(\mathbf{X}_W)) = \lambda_0).\end{aligned} \tag{3.2.7}$$

The right hand side of (3.2.7), as $N \to \infty$, is approximately a Riemann sum for $\sqrt{2\pi M} M! \int\limits_{\mathcal{L}_{(s_1,\ldots,s_M)}} \phi_{GUE,M}(x) dx_1 \cdots dx_{M-1}$. Hence,

$$\begin{aligned}
\mathbb{P}\left(\frac{\lambda^1 - N/M}{\sqrt{N/M}} \le s_1, ..., \frac{\lambda^M - N/M}{\sqrt{N/M}} \le s_M \right) \\
&\xrightarrow{N\to\infty} \sqrt{2\pi M} M! \int\limits_{\mathcal{L}_{(s_1,\ldots,s_M)}} \phi_{GUE,M}(x) dx_1 \cdots dx_{M-1} \\
&= \mathbb{P}\left(\lambda^{1,0}_{GUE,M} \le s_1, ..., \lambda^{M,0}_{GUE,M} \le s_M \right). \qquad (3.2.8)
\end{aligned}$$

□

Still let $\left(\tilde{B}^1(t), \tilde{B}^2(t), ..., \tilde{B}^M(t)\right)$ be the M-dimensional Brownian motion having covariance matrix (3.2.1). For $k = 1, ..., M$, as in [30] (see also [18]), let

$$\tilde{H}^k_M = \sqrt{\frac{M-1}{M}} \sup \sum_{i=1}^{M} \sum_{p=1}^{k} \left(\tilde{B}^i(t^p_{i-p+1}) - \tilde{B}^i(t^p_{i-p})\right),$$

where the sup is taken over all the subdivisions (t^p_i) of $[0,1]$ of the form:

$$t^p_i \in [0,1], \ t^{p+1}_i \le t^p_i, \ t^p_i = 0 \ for \ i \le 0 \ and \ t^p_i = 1 \ for \ i \ge M-k+1. \qquad (3.2.9)$$

We are now ready to prove the following Brownian functional representations of the spectrum of the traceless GUE.

Theorem 3.2.7 *For each $M \ge 2$,*

$$\left(\tilde{H}^1_M, \tilde{H}^2_M - \tilde{H}^1_M, ..., \tilde{H}^M_M - \tilde{H}^{M-1}_M\right) \stackrel{d}{=} \left(\lambda^{1,0}_{GUE,M}, \lambda^{2,0}_{GUE,M}, ..., \lambda^{M,0}_{GUE,M}\right). \qquad (3.2.10)$$

Proof. Consider the random vector $\mathbf{V} = (V_1, V_2, ..., V_M) \in \{0,1\}^M$, where the random variables $V_1, V_2, ..., V_M$ have the joint probability mass function

$$\mathbb{P}\left(V_p = 1; V_q = 0, \ for \ all \ q \ne p\right) = \frac{1}{M}, \quad p = 1, 2, ..., M,$$

i.e., if $(\mathbf{e_k})_{k=1,...,M}$ is the canonical basis of $\mathbb{R}^M$,

$$\mathbb{P}\left(\mathbf{V} = \mathbf{e_k}\right) = \frac{1}{M}, \quad k = 1, ..., M.$$

Clearly, for each $1 \le p \le M$,

$$\mathbb{E}\left(V_p\right) = \frac{1}{M},\ Var(V_p) = \frac{M-1}{M^2},$$

and for $p_1 \neq p_2$, $Cov(V_{p_1}, V_{p_2}) = -1/M^2$. Hence the covariance matrix of $\mathbf{V}$ is

$$\mathbf{\Sigma} = \frac{M-1}{M^2}\begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}, \tag{3.2.11}$$

with $\rho = -1/(M-1)$. Let $\mathbf{V_1}, \mathbf{V_2}, ..., \mathbf{V_N}$ be independent copies of $\mathbf{V}$, where $\mathbf{V_i} = (V_{i,1}, V_{i,2}, ..., V_{i,M})$. Let

$$G(M,N) = \max\left\{\sum_{(i,j)\in\pi} V_{i,j}\ ; \pi \in \mathcal{P}(M,N)\right\}, \tag{3.2.12}$$

where $\mathcal{P}(M,N)$ is the set of all paths π taking only unit steps up or to the right in the rectangle $\{1,...,N\} \times \{1,...,M\}$.

Each path π is uniquely determined by a weakly increasing sequence of its $M-1$ jumps, namely $0 = t_0 \le t_1 \le \cdots \le t_{M-1} \le 1$, such that π is horizontal on $[\lfloor t_{j-1}N\rfloor, \lfloor t_j N\rfloor] \times \{j\}$ and vertical on $\{\lfloor t_j N\rfloor\} \times [j, j+1]$. Hence

$$G(M,N) = \sup_{0=t_0\le t_1\le\cdots\le t_{M-1}\le t_M=1} \sum_{j=1}^{M} \sum_{i=\lfloor t_{j-1}N\rfloor}^{\lfloor t_j N\rfloor} V_{i,j}.$$

Now,

$$\frac{G(M,N) - N/M}{\sqrt{N/M}} = \sup \sum_{j=1}^{M} \frac{\sum_{i=\lfloor t_{j-1}N\rfloor}^{\lfloor t_j N\rfloor} V_{i,j} - (t_j - t_{j-1})N/M}{\sqrt{N/M}}. \tag{3.2.13}$$

Notice that the random vectors $\mathbf{V_1}, \mathbf{V_2}, ..., \mathbf{V_N}$ are independent. The probability measures $\mathbb{P}_N$ generated by $\left(\left(\sum_{i=1}^{\lfloor tN\rfloor} V_{i,j} - tN/M\right)/\sqrt{N}\right)_{1\le j\le M}$ satisfy $\mathbb{P}_N(A) \to \mathbb{P}_\infty(A)$, for all Borel subsets A of the space of continuous functions $C([0,1]^M)$ for which $\mathbb{P}_\infty(\partial A) = 0$, where $\mathbb{P}_\infty$ is the M-dimensional Wiener measure. With the uniform metric $d^M(x,y) = \sup_{t\in[0,1]^M} |x(t)-y(t)|$, $C([0,1]^M)$ is complete and separable. A sequence of stochastic processes converges weakly if and only if the finite dimensional distribution converges and the sequence is tight. We claim that, as $N \to \infty$, for any $t > 0$,

$$\left(\frac{\sum_{i=1}^{\lfloor tN \rfloor} V_{i,j} - tN\frac{1}{M}}{\sqrt{N}}\right)_{1\le j\le M} \Longrightarrow \left(\tilde{B}^j(t)\right)_{1\le j\le M},$$

where $\left(\tilde{B}^j(t)\right)_{1\le j\le M}$ is an M-dimensional Brownian motion with covariance matrix $t\mathbf{\Sigma}$. Indeed, for any $t > 0$, since $\mathbf{V_1}, \mathbf{V_2}, ...$ are independent, each with mean vector $\mathbf{1}_M/M$, where $\mathbf{1}_M = (1,1,...,1)$, and covariance matrix $\mathbf{\Sigma}$,

$$\frac{\sum_{i=1}^{\lfloor tN \rfloor} \mathbf{V_i} - tN\frac{\mathbf{1}_M}{M}}{\sqrt{N}} \Longrightarrow \left(\tilde{B}^j(t)\right)_{1\le j\le M},$$

by the central limit theorem for iid random vectors and Slutsky's lemma. Next, for any $t > s > 0$,

$$\left(\frac{\sum_{i=1}^{\lfloor (t-s)N \rfloor} \mathbf{V_i} - \lfloor (t-s)N\rfloor\frac{\mathbf{1}_M}{M}}{\sqrt{N}}, \frac{\sum_{i=1}^{\lfloor sN \rfloor} \mathbf{V_i} - \lfloor sN\rfloor\frac{\mathbf{1}_M}{M}}{\sqrt{N}}\right)$$
$$\Longrightarrow \left(\left(\tilde{B}^j(t-s)\right)_{1\le j\le M}, \left(\tilde{B}^j(s)\right)_{1\le j\le M}\right). \quad (3.2.14)$$

The continuous mapping theorem and Slutsky's lemma allow us to conclude that

$$\left(\frac{\sum_{i=1}^{\lfloor tN \rfloor} \mathbf{V_i} - tN\frac{\mathbf{1}_M}{M}}{\sqrt{N}}, \frac{\sum_{i=1}^{\lfloor sN \rfloor} \mathbf{V_i} - sN\frac{\mathbf{1}_M}{M}}{\sqrt{N}}\right)$$
$$\Longrightarrow \left(\left(\tilde{B}^j(t)\right)_{1\le j\le M}, \left(\tilde{B}^j(s)\right)_{1\le j\le M}\right). \quad (3.2.15)$$

The convergence for the time points $t_1 > t_2 > \cdots > t_n > 0$ can be treated in a similar fashion. Thus the finite dimensional distribution converges to that of $\left(\tilde{B}^j(t)\right)_{1\le j\le M}$. Since tightness in $\left(C([0,1]^M), d^M(\cdot,\cdot)\right)$ is as in the proof of Donsker's invariance principle (e.g., see[11]), we are just left with identifying the covariance structure of the limiting Brownian motion $\left(\tilde{B}^j(t)\right)_{1\le j\le M}$. For each N,

$$\mathbb{E}\left(\left|\frac{\sum_{i=1}^{\lfloor tN \rfloor} V_{i,j} - \lfloor tN\rfloor\frac{1}{M}}{\sqrt{N}}\right|^2\right) = \frac{\lfloor tN\rfloor \mathbb{E}\left(\left|V_{1,j} - \frac{1}{M}\right|^2\right)}{N} \le t\mathbb{E}\left(\left|V_{1,j} - \frac{1}{M}\right|^2\right). \quad (3.2.16)$$

Therefore, $\sup\limits_{N\geq 1}\mathbb{E}\left(\left|\sum_{i=1}^{\lfloor tN\rfloor}\left(V_{i,j}-\lfloor tN\rfloor\frac{1}{M}\right)/\sqrt{N}\right|^2\right)<\infty$. As $N\to\infty$, for each $1\leq j\leq M$,

$$Var\left(\tilde{B}^j(t)\right)=\lim_{N\to\infty}Var\left(\sum_{i=1}^{\lfloor tN\rfloor}\frac{V_{i,j}}{\sqrt{N}}\right)=\lim_{N\to\infty}\frac{\lfloor tN\rfloor}{N}Var\left(V_{1,j}\right)=t\frac{M-1}{M^2}.$$

Moreover, for any $j_1\neq j_2$, by the continuous mapping theorem,

$$\left(\frac{\sum_{i=1}^{\lfloor tN\rfloor}V_{i,j_1}-\lfloor tN\rfloor\frac{1}{M}}{\sqrt{N}}\right)\left(\frac{\sum_{i=1}^{\lfloor tN\rfloor}V_{i,j_2}-\lfloor tN\rfloor\frac{1}{M}}{\sqrt{N}}\right)\overset{N\to\infty}{\Longrightarrow}\tilde{B}^{j_1}(t)\tilde{B}^{j_2}(t).$$

Since

$$\begin{aligned}
\sup_{N\geq 1}&\left(\mathbb{E}\left(\frac{\sum_{i=1}^{\lfloor tN\rfloor}V_{i,j_1}-\lfloor tN\rfloor\frac{1}{M}}{\sqrt{N}}\frac{\sum_{i=1}^{\lfloor tN\rfloor}V_{i,j_2}-\lfloor tN\rfloor\frac{1}{M}}{\sqrt{N}}\right)\right)^2\\
&\leq\sup_{N\geq 1}\left(\mathbb{E}\left(\frac{\sum_{i=1}^{\lfloor tN\rfloor}V_{i,j_1}-\lfloor tN\rfloor\frac{1}{M}}{\sqrt{N}}\right)^2\right)\sup_{N\geq 1}\left(\mathbb{E}\left(\frac{\sum_{i=1}^{\lfloor\lfloor tN\rfloor\rfloor}V_{i,j_2}-\lfloor tN\rfloor\frac{1}{M}}{\sqrt{N}}\right)^2\right)\\
&<\infty,
\end{aligned}\tag{3.2.17}$$

therefore,

$$\begin{aligned}
Cov\left(\tilde{B}^{j_1}(t),\tilde{B}^{j_2}(t)\right)&=\lim_{N\to\infty}Cov\left(\frac{\sum_{i=1}^{\lfloor tN\rfloor}V_{i,j_1}}{\sqrt{N}},\frac{\sum_{i=1}^{\lfloor tN\rfloor}V_{i,j_2}}{\sqrt{N}}\right)\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{\lfloor tN\rfloor}Cov\left(V_{1,j_1},V_{1,j_2}\right)\\
&=tCov\left(V_{1,j_1},V_{1,j_2}\right).
\end{aligned}\tag{3.2.18}$$

Hence the M-dimensional Brownian motion $\left(\tilde{B}^j(t)\right)_{1\leq j\leq M}$ has covariance matrix $t\mathbf{\Sigma}$ with $\mathbf{\Sigma}$ given in (3.2.11), and from the continuous mapping theorem, we conclude that

$$\frac{G(M,N)-N/M}{\sqrt{N/M}}\overset{d}{\Longrightarrow}\tilde{H}_M.\tag{3.2.19}$$

Next, let $\mathbf{X}_W$ be the matrix formed by all the $V_{i,j}$ on the lattice $\{1,...,N\}\times\{1,...,M\}$. Each realization of the matrix $\mathbf{X}_W$ is then an element of $\mathcal{M}(M,N)$, and moreover it corresponds to a random word of length N. Again, denote by

$\lambda(RSK(\mathbf{X}_W))$ the common shape of the corresponding pair of Young tableaux of $\mathbf{X}_W$ through the RSK correspondence. It is a well known combinatorial fact (see Lemma 1 of Section 3.2 in [20]) that, for all $1 \le k \le M$,

$$\lambda_1 + \cdots + \lambda_k = G^k(M,N) := \sup\Big\{ \sum_{(i,j)\in\pi_1\cup\cdots\cup\pi_k} V_{i,j} : \pi_1, ..., \pi_k \in \mathcal{P}(M,N),$$
$$\textit{and } \pi_1, ..., \pi_k \textit{ are all disjoint}\Big\}, \tag{3.2.20}$$

where, by disjoint, it is meant that any two paths do not share a common point (i, j) in the rectangle $\{1, ..., N\} \times \{1, ..., M\}$. Through a consideration of the particular k disjoint paths on the lattice, as in [30] and [18], where the sup in (3.2.20) is actually attained, it is more generally true that

$$\frac{G^k(M,N) - kN/M}{\sqrt{N/M}} \overset{N\to\infty}{\Longrightarrow} \tilde{H}^k_M. \tag{3.2.21}$$

Now, for all $1 \le L \le M$, any linear combination of $\frac{\lambda_1 - N/M}{\sqrt{N/M}}, ..., \frac{\sum_{k=1}^L \lambda_k - LN/M}{\sqrt{N/M}}$ is still a continuous functional of the process $\left(\left(\sum_{i=1}^{\lfloor tN\rfloor} V_{i,j} - tN/M\right)/\sqrt{N}\right)_{1\le j\le M}$ which converges weakly to the M-dimensional Brownian motion with covariance matrix $t\mathbf{\Sigma}$ as $N \to \infty$. Hence any linear combination converges in distribution to the corresponding linear combination of $\tilde{H}^1_M, \tilde{H}^2_M, ..., \tilde{H}^L_M$ by the continuous mapping theorem. With the help of the Cramér-Wold theorem, we conclude that, for any $1 \le L \le M$, as $N \to \infty$,

$$\left(\frac{\lambda_1 - N/M}{\sqrt{N/M}}, \frac{\sum_{k=1}^2 \lambda_k - 2N/M}{\sqrt{N/M}}, ..., \frac{\sum_{k=1}^L \lambda_k - LN/M}{\sqrt{N/M}}\right)$$
$$\Longrightarrow \left(\tilde{H}^1_M, \tilde{H}^2_M, ..., \tilde{H}^L_M\right), \tag{3.2.22}$$

therefore, for any $1 \le L \le M$, as $N \to \infty$,

$$\left(\frac{\lambda_1 - N/M}{\sqrt{N/M}}, \frac{\lambda_2 - N/M}{\sqrt{N/M}}, ..., \frac{\lambda_L - N/M}{\sqrt{N/M}}\right)$$
$$\Longrightarrow \left(\tilde{H}^1_M, \tilde{H}^2_M - \tilde{H}^1_M, ..., \tilde{H}^L_M - \tilde{H}^{L-1}_M\right). \tag{3.2.23}$$

On the other hand, by Theorem 3.2.6, as $N \to \infty$,

$$\left(\frac{\lambda_1 - N/M}{\sqrt{N/M}}, ..., \frac{\lambda_M - N/M}{\sqrt{N/M}}\right) \Longrightarrow \left(\lambda^{1,0}_{GUE,M}, ..., \lambda^{M,0}_{GUE,M}\right). \tag{3.2.24}$$

Combining (3.2.23) and (3.2.24), the theorem is proved. $\square$

Remark 3.2.8 *Let $\left(B^1(t), B^2(t), ..., B^M(t)\right)$ be a standard M-dimensional Brownian motion. For $k = 1, ..., M$, let*

$$D_M^k = \sup \sum_{i=1}^{M} \sum_{p=1}^{k} \left(B^i(t_{i-p+1}^p) - B^i(t_{i-p}^p)\right),$$

where the sup *is taken over all the subdivisions (t_i^p) of $[0,1]$ described in (3.2.9). The very approach to prove Theorem 4.2.1 can be used to obtain a Brownian functional representation of the spectrum of the $M \times M$ GUE, namely,*

$$\left(D_M^1, D_M^2 - D_M^1, ..., D_M^M - D_M^{M-1}\right) \stackrel{d}{=} \left(\lambda_{GUE,M}^1, \lambda_{GUE,M}^2, ..., \lambda_{GUE,M}^M\right). \qquad (3.2.25)$$

From the observation that the supremum in the definition of $G^k(M,N)$ is attained on a particular set of k disjoint northeast paths for each $k = 1, ..., M$, Doumerc ([18]) found Brownian functional representations for $\sum_{i=1}^k \lambda_{GUE,M}^i$. These functionals are similar to the D_M^k except that the supremum is taken over a different set of subdivisions of $[0,1]$. In fact, we believe that the subdivisions given in (3.2.9) should be the ones present in [18]. With a similar consideration of k disjoint increasing subsequences, a specific expression for the sum of the first k rows of the Young tableau associated to a random Markovian word is obtained, in [30], in terms of the number of occurrences of the letters among the sequence. The multidimensional convergence of the whole tableau to a corresponding multidimensional Brownian functional is also obtained in [30].

In contrast to the approach in [18], our potential proof of (3.2.25) does not require passing through the matrix central limit theorem. To briefly describe the approach in [18], let the $V_{i,j}$ in (4.2.7) be iid geometric random variables, i.e., for $r = 0, 1, ...,$ let $\mathbb{P}\left(V_{i,j} = r\right) = q(1-q)^r$. With such $\{V_{i,j}\}$, the probability of a given matrix realization only depend on the sum of the matrix entries, which is also the sum of the entries in the shape of the associate Young tableaux. The joint probability mass function of *the shape of the associate Young tableaux through the RSK correspondence can then be expressed through the well known number of Young tableaux sharing this given shape. Next, by setting* $q = 1 - L^{-1}$*, and letting* $L \to \infty$*, the random variables on the lattice converge to iid exponential random variables with parameter one, while the corresponding shape of the associated Young tableaux converges to the spectrum of the* $M \times N$ *Laguerre Unitary Ensemble. As* $N \to \infty$*, for any* $k = 1, ..., M$*, the corresponding* $G^k(M, N)$*, properly normalized, converge in distribution to* D_M^k*. With the same normalization, it is proved in [18] that the spectrum of the* $M \times N$ *Laguerre Unitary Ensemble converges to the spectrum of the* $M \times M$ *GUE. Hence, the continuous mapping theorem, gives* $\sum_{j=1}^{k} \lambda_{GUE,M}^{j} \stackrel{d}{=} D_M^k$*. Via the large* N *asymptotics of the corresponding numbers of Young tableaux, we are able to directly show that the limiting joint probability mass function of the shape of the tableaux converges to the joint probability density function of the eigenvalues of an element of the GUE. Thus,* $\sum_{j=1}^{k} \lambda_{GUE,M}^{j} \stackrel{d}{=} D_M^k$*, and (3.2.25) follows from the Cramér-Wold theorem. Similar ideas are already developed by Johansson (Theorem 1.1 in [39]) to prove that the Poissonized Plancherel measure can be obtained as a limit of the Meixner measure. Johansson also proves the convergence of the whole tableau corresponding to a random word for uniform alphabets, and obtains the joint density of the limiting law.*

As a consequence of Theorem 3.2.7, $\lambda_{GUE,M}^{max,0} \stackrel{d}{=} \tilde{H}_M^1$. Together with the fact that $\lambda_{GUE,M}^{max} \stackrel{d}{=} D_M$, Proposition 3.1.4 indicates that the difference between D_M and $\tilde{H}_M^1$ is a centered normal random variable with variance $1/M$. Now, it is well known that as $M \to \infty$, $D_M/\sqrt{M} \to 2$ $a.s.$ and in L^1, and that $(D_M - 2\sqrt{M})M^{1/6} \Rightarrow F_{TW}$. The same asymptotics also hold for the Brownian functional $\tilde{H}_M^1$.

Below, we give a simple proof of the fact that $\tilde{H}_M^1/\sqrt{M} \to 2$ almost surely. This proof is based on a "tridiagonalization" technique originating in Trotter [65] (see also Silverstein [54] where similar ideas are used). Our first result is the well known Householder representation of Hermitian matrices.

Lemma 3.2.9 *Let $\mathbf{G} = (G_{i,j})_{1\le i,j\le M}$ be a matrix from the GUE. Then, there exists a unitary matrix $\mathbf{U}$, such that*

$$\mathbf{UGU}^* = \begin{pmatrix} A_{1,1} & \chi_{M-1} & 0 & \cdots & 0 \\ \chi_{M-1} & A_{2,2} & \chi_{M-2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \chi_2 & A_{M-1,M-1} & \chi_1 \\ 0 & \cdots & 0 & \chi_1 & A_{M,M} \end{pmatrix}, \tag{3.2.26}$$

where $A_{1,1}, ..., A_{M,M}$ are independent $N(0,1)$ random variables, and for each $1 \le k \le M-1$, χ_{M-k} has a chi distribution, with $M-k$ degrees of freedom. Moreover, for each $k = 1, ..., M-1$, $A_{k,k}$ is independent of $\chi_{M-k}, ..., \chi_1$.

Proof. We will construct the unitary matrix $\mathbf{U}$ as a product of $M-1$ unitary matrices $\mathbf{U_{M-1}} \cdots \mathbf{U_2 U_1}$. The vectors $\left(\sqrt{\sum_{j=2}^M |G_{1,j}|^2}, 0, ..., 0\right)$ and $\left(G_{1,2}, G_{1,3}, ..., G_{1,M}\right)$ have the same length. Let $\mathbf{U^{(1)}}$ be the $(M-1)\times(M-1)$ unitary matrix such that

$$\left(G_{1,2}, G_{1,3}, ..., G_{1,M}\right)\mathbf{U^{(1)}} = \left(\sqrt{\sum_{j=2}^M |G_{1,j}|^2}, 0, ..., 0\right).$$

Let

$$\mathbf{U_1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{U^{(1)}} & \\ 0 & & & \end{pmatrix}. \tag{3.2.27}$$

Then

$$\mathbf{U_1 G U_1^*} = \begin{pmatrix} G_{1,1} & \sqrt{\sum_{j=2}^M |G_{1,j}|^2} & 0 & \cdots & 0 \\ \sqrt{\sum_{i=2}^M |G_{i,1}|^2} & & & & \\ 0 & & & & \\ \vdots & & & \mathbf{G^{(1)}} & \\ 0 & & & & \end{pmatrix}, \tag{3.2.28}$$

where the $(M-1)\times(M-1)$ matrix

$$\mathbf{G^{(1)}} = \mathbf{U^{(1)}} \begin{pmatrix} G_{2,2} & \cdots & G_{2,M} \\ \vdots & \ddots & \vdots \\ G_{M,2} & \cdots & G_{M,M} \end{pmatrix} \mathbf{U^{(1)}}^*. \tag{3.2.29}$$

Let $A_{1,1} = G_{1,1}$ and $\chi_{M-1} = \sqrt{\sum_{j=2}^{M} |G_{1,j}|^2}$. Clearly, $A_{1,1}$ and χ_{M-1} are independent. Actually $A_{1,1}$ is independent of the whole matrix $\mathbf{G^{(1)}}$. The crucial fact here is that the matrix $\mathbf{G^{(1)}}$ is still an element of the $(M-1)\times(M-1)$ GUE. The same procedure can then be applied to $\mathbf{G^{(1)}}$, namely, there is a unitary matrix

$$\mathbf{U_2} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \mathbf{U^{(2)}} & \\ 0 & 0 & & & \end{pmatrix}, \tag{3.2.30}$$

such that

$$\mathbf{U_2 U_1 G U_1^* U_2^*} = \begin{pmatrix} G_{1,1} & \chi_{M-1} & 0 & 0 & \cdots & 0 \\ \chi_{M-1} & G^{(1)}_{1,1} & \sqrt{\sum_{j=2}^{M-1} |G^{(1)}_{1,j}|^2} & 0 & \cdots & 0 \\ 0 & \sqrt{\sum_{i=2}^{M-1} |G^{(1)}_{i,1}|^2} & & & & \\ 0 & 0 & & & & \\ \vdots & \vdots & & \mathbf{G^{(2)}} & & \\ 0 & 0 & & & & \end{pmatrix}, \tag{3.2.31}$$

where $\mathbf{G^{(2)}}$ is an element of $(M-2)\times(M-2)$ GUE. Let $A_{2,2} = G^{(1)}_{1,1}$ and let $\chi_{M-2} = \sqrt{\sum_{j=2}^{M-1} |G^{(1)}_{1,j}|^2}$. The matrices $\mathbf{U_3}$,...,$\mathbf{U_{M-1}}$ are chosen in a completely similar fashion, and the lemma follows. □

Theorem 3.2.10 *As $M \to \infty$,*

$$\frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}} \to 2, \qquad \textit{almost surely},$$

equivalently,

$$\frac{\lambda_{GUE,M}^{max}}{\sqrt{M}} \to 2, \qquad \textit{almost surely.}$$

Proof. Let $\mathbf{G} = (G_{i,j})_{1\le i,j\le M}$ be a matrix from the $M \times M$ GUE with maximal eigenvalue $\lambda_{GUE,M}^{max}$. Clearly,

$$\lambda_{GUE,M}^{max,0} = \lambda_{GUE,M}^{max} - \frac{tr\,(\mathbf{G})}{M}.$$

By the strong law of large numbers, $tr\,(\mathbf{G})\,/M \to 0$, and it suffices to prove that, almost surely,

$$\frac{\lambda_{GUE,M}^{max}}{\sqrt{M}} \to 2.$$

By Lemma 3.2.9, there exists a unitary matrix $\mathbf{U}$, such that

$$\mathbf{T} := \mathbf{U}\mathbf{G}\mathbf{U}^* = \begin{pmatrix} A_{1,1} & \chi_{M-1} & 0 & \cdots & 0 \\ \chi_{M-1} & A_{2,2} & \chi_{M-2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \chi_2 & A_{M-1,M-1} & \chi_1 \\ 0 & \cdots & 0 & \chi_1 & A_{M,M} \end{pmatrix}, \tag{3.2.32}$$

where $A_{1,1}, ..., A_{k,k}$ are independent $N(0,1)$ random variable, and for each $1 \le k \le M-1$, χ_k has a chi distribution with k degrees of freedom. Clearly $\mathbf{G}$ and $\mathbf{T}$ share the same eigenvalues.

By the Gerŝgorin circle theorem (see [27]), for any eigenvalue λ_i of $\mathbf{G}$, letting also $\chi_0 = \chi_M = 0$,

$$\lambda_i \in \bigcup_{k=1,\ldots,M} \left[A_{k,k} - \chi_{M-k+1} - \chi_{M-k}, A_{k,k} + \chi_{M-k+1} + \chi_{M-k}\right].$$

Hence

$$\begin{aligned} \frac{\lambda_{GUE,M}^{max}}{\sqrt{M}} &\le \max_{k=1,\ldots,M}\left(\frac{A_{k,k}}{\sqrt{M}} + \frac{\chi_{M-k+1}}{\sqrt{M}} + \frac{\chi_{M-k}}{\sqrt{M}}\right) \\ &\le \max_{k=1,\ldots,M}\frac{A_{k,k}}{\sqrt{M}} + \max_{k=1,\ldots,M}\frac{\chi_{M-k+1}}{\sqrt{M}} + \max_{k=1,\ldots,M}\frac{\chi_{M-k}}{\sqrt{M}}. \end{aligned} \tag{3.2.33}$$

For each $k = 1, ..., M$, $A_{k,k} \sim N\big(0,1\big)$, thus, for any fixed $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(\left|\max_{k=1,...,M}\frac{A_{k,k}}{\sqrt{M}}\right| > \varepsilon\right) &\leq \sum_{k=1}^{M}\mathbb{P}\left(\frac{|A_{k,k}|}{\sqrt{M}} > \varepsilon\right)\\ &= \frac{2M}{\sqrt{2\pi}}\int_{\varepsilon\sqrt{M}}^{\infty}e^{-x^2/2}dx\\ &\leq \frac{2M}{\varepsilon\sqrt{2\pi M}}\int_{\varepsilon\sqrt{M}}^{\infty}xe^{-x^2/2}dx\\ &= \frac{\sqrt{2M}}{\varepsilon\sqrt{\pi}}e^{-\varepsilon^2 M/2}. \end{aligned} \tag{3.2.34}$$

Therefore,

$$\sum_{M=1}^{\infty}\mathbb{P}\left(\left|\max_{k=1,...,M}\frac{A_{k,k}}{\sqrt{M}}\right| > \varepsilon\right) \leq \sum_{M=1}^{\infty}\frac{\sqrt{2M}}{\varepsilon\sqrt{\pi}}e^{-\varepsilon^2 M/2} < \infty,$$

and so, by the Borel-Cantelli lemma,

$$\max_{k=1,...,M}\frac{A_{k,k}}{\sqrt{M}} \xrightarrow{a.s.} 0. \tag{3.2.35}$$

Next, for any fixed $\varepsilon > 0$,

$$\begin{aligned}&\mathbb{P}\left(\left|\max_{k=1,...,M}\frac{\chi^2_{M-k+1}}{M} - 1\right| > \varepsilon\right)\\ &= \mathbb{P}\left(\left|\max_{k=1,...,M}\frac{\chi^2_{k}}{M} - 1\right| > \varepsilon\right)\\ &= \mathbb{P}\left(\max_{k=1,...,M}\chi^2_k < M(1-\varepsilon)\right) + \mathbb{P}\left(\max_{k=1,...,M}\chi^2_k > M(1+\varepsilon)\right)\\ &\leq \mathbb{P}\left(\chi^2_M < M(1-\varepsilon)\right) + \sum_{k=1}^{M}\mathbb{P}\left(\chi^2_k > M(1+\varepsilon)\right)\\ &\leq \mathbb{P}\left(\chi^2_M < M(1-\varepsilon)\right) + M\mathbb{P}\left(\chi^2_M > M(1+\varepsilon)\right). \end{aligned} \tag{3.2.36}$$

Now, for any $y > 0$,

$$\begin{aligned}&\int_y^{\infty}u^{\frac{M}{2}-1}e^{-\frac{u}{2}}du\\ &= 2e^{-\frac{y}{2}}y^{\frac{M}{2}}\left(\sum_{k=1}^{\lfloor M/2\rfloor}\prod_{l=1}^{k}\frac{M-2l}{y}\right) + 2(M-2)!!\int_{\sqrt{y}}^{\infty}x^{M \bmod (2)}e^{-\frac{x^2}{2}}dx,\end{aligned} \tag{3.2.37}$$

where

$$(M-2)!! = \begin{cases} (M-2)(M-4)\cdots(3)(1), & \text{if M is odd, } M \geq 3; \\ (M-2)(M-4)\cdots(4)(2), & \text{if M is even, } M \geq 4; \\ 1, & \text{if } M = 2. \end{cases} \tag{3.2.38}$$

For $y = M(1+\varepsilon)$,

$$\begin{aligned} &\mathbb{P}\left(\chi_M^2 > M(1+\varepsilon)\right) \\ &= \frac{1}{\Gamma\left(\frac{M}{2}\right) 2^{\frac{M}{2}}} \int_{M(1+\varepsilon)}^{\infty} u^{\frac{M}{2}-1} e^{-\frac{u}{2}} du \\ &\leq \frac{2e^{-\frac{M(1+\varepsilon)}{2}} \left(M(1+\varepsilon)\right)^{\frac{M}{2}}}{\Gamma\left(\frac{M}{2}\right)} \left(\sum_{k=1}^{\lfloor M/2 \rfloor} \prod_{l=1}^{k} \frac{M-2l}{M(1+\varepsilon)} \right) + \frac{2e^{-\frac{M(1+\varepsilon)}{2}} (M-2)!!}{\Gamma\left(\frac{M}{2}\right)} \\ &\leq \frac{Me^{-\frac{M(1+\varepsilon)}{2}} \left(M(1+\varepsilon)\right)^{\frac{M}{2}}}{\Gamma\left(\frac{M}{2}\right)} + \frac{2e^{-\frac{M(1+\varepsilon)}{2}} (M-2)!!}{\Gamma\left(\frac{M}{2}\right)}. \end{aligned} \tag{3.2.39}$$

But, as $M \to \infty$,

$$\frac{Me^{-\frac{M(1+\varepsilon)}{2}} \left(M(1+\varepsilon)\right)^{\frac{M}{2}}}{\Gamma(\frac{M}{2}) 2^{\frac{M}{2}}} \sim \frac{\sqrt{M} e^{-\frac{M\varepsilon}{2}-1} (1+\varepsilon)^{\frac{M}{2}}}{\sqrt{\pi}},$$

and

$$(M-2)!! \sim \Gamma\left(\frac{M}{2}\right) 2^{\frac{M}{2}},$$

thus $\sum_{M=1}^{\infty} M\mathbb{P}\left(\chi_M^2 > M(1+\varepsilon)\right) < \infty$. Next, since $f(u) = u^{\frac{M}{2}-\varepsilon} e^{-\frac{u}{2}}$ is increasing for $u \leq M - 2\varepsilon$, we have for $M \geq 2$:

$$\begin{aligned} \mathbb{P}\left(\chi_M^2 < M(1-\varepsilon)\right) &= \frac{1}{\Gamma\left(\frac{M}{2}\right) 2^{\frac{M}{2}}} \int_0^{M(1-\varepsilon)} u^{\frac{M}{2}-1} e^{-\frac{u}{2}} du \\ &= \frac{1}{\Gamma\left(\frac{M}{2}\right) 2^{\frac{M}{2}}} \int_0^{M(1-\varepsilon)} u^{\frac{M}{2}-\varepsilon} e^{-\frac{u}{2}} \frac{1}{u^{1-\varepsilon}} du \\ &\leq \frac{1}{\Gamma\left(\frac{M}{2}\right) 2^{\frac{M}{2}}} \left(M(1-\varepsilon)\right)^{\frac{M}{2}-\varepsilon} e^{-\frac{M(1-\varepsilon)}{2}} \frac{\left(M(1-\varepsilon)\right)^{\varepsilon}}{\varepsilon} \\ &= \frac{\left(M(1-\varepsilon)\right)^{\frac{M}{2}} e^{-\frac{M}{2}+\frac{M\varepsilon}{2}}}{\varepsilon\Gamma\left(\frac{M}{2}\right) 2^{\frac{M}{2}}}. \end{aligned} \tag{3.2.40}$$

By Stirling's Formula, as $M \to \infty$,

$$\Gamma\left(\frac{M}{2}\right) \sim \sqrt{\pi M} \left(\frac{M}{2}\right)^{\frac{M}{2}-1} e^{-\frac{M}{2}}.$$

So, for $0 < \varepsilon < 1$,

$$\frac{(M(1-\varepsilon))^{\frac{M}{2}}\, e^{-\frac{M}{2}+\frac{M\varepsilon}{2}}}{\varepsilon\Gamma\left(\frac{M}{2}\right)2^{\frac{M}{2}}} \sim \frac{\sqrt{M}}{2\varepsilon\sqrt{\pi}} e^{\frac{M}{2}(\varepsilon+\log(1-\varepsilon))}$$
$$\leq \frac{\sqrt{M}}{2\varepsilon\sqrt{\pi}} e^{-\frac{M\varepsilon^2}{4}}. \tag{3.2.41}$$

But

$$\sum_{M=2}^{\infty} \frac{\sqrt{M}}{2\varepsilon\sqrt{\pi}} e^{-\frac{M\varepsilon^2}{4}} < \infty,$$

and thus,

$$\sum_{M=2}^{\infty} \mathbb{P}\left(\chi_M^2 < M(1-\varepsilon)\right) < \infty.$$

Therefore,

$$\sum_{M=1}^{\infty} \mathbb{P}\left(\left|\max_{k=1,\dots,M} \frac{\chi_{M-k+1}^2}{M} - 1\right| > \varepsilon\right)$$
$$\leq \sum_{M=1}^{\infty} \mathbb{P}\left(\chi_M^2 < M(1-\varepsilon)\right) + \sum_{M=1}^{\infty} M\mathbb{P}\left(\chi_M^2 > M(1+\varepsilon)\right)$$
$$< \infty. \tag{3.2.42}$$

Hence, $\max_{k=1,\dots,M} \chi_{M-k+1}^2/M \xrightarrow{a.s.} 1$, and thus

$$\max_{k=1,\dots,M} \frac{\chi_{M-k+1}}{\sqrt{M}} \xrightarrow{a.s.} 1. \tag{3.2.43}$$

With a completely similar argument,

$$\max_{k=1,\dots,M} \frac{\chi_{M-k}}{\sqrt{M}} \xrightarrow{a.s.} 1. \tag{3.2.44}$$

Thus, almost surely,

$$\limsup_{M\to\infty} \frac{\lambda_{GUE,M}^{max}}{\sqrt{M}} \leq 2. \tag{3.2.45}$$

Next, since the empirical distribution of the eigenvalues $\left(\lambda_{GUE,M}^{i}/\sqrt{M}\right)_{1\leq i\leq M}$ converges almost surely to the semicircle law ν with density $\sqrt{4-x^2}/2\pi$. We claim that, for any $\varepsilon > 0$,

$$\mathbb{P}\left(\liminf_{M\to\infty} \frac{\lambda_{GUE,M}^{max}}{\sqrt{M}} > 2-\varepsilon\right) = 1, \tag{3.2.46}$$

If not, i.e., if

$$\mathbb{P}\left(\liminf_{M\to\infty} \lambda_{GUE,M}^{max}/\sqrt{M} > 2-\varepsilon\right) < 1,$$

then,

$$\mathbb{P}\Big(\liminf_{M\to\infty} \lambda_{GUE,M}^{max}/\sqrt{M} \le 2-\varepsilon\Big) > 0,$$

and thus,

$$\mathbb{P}\Big(\liminf_{M\to\infty} \lambda_{GUE,M}^{i}/\sqrt{M} \le 2-\varepsilon, i=1,...,M\Big) > 0.$$

Let the event $A_\varepsilon := \left\{\liminf_{M\to\infty} \lambda_{GUE,M}^{i}/\sqrt{M} \le 2-\varepsilon, i=1,...,M\right\}$ and consider the bounded continuous function

$$f_\varepsilon(x) = \begin{cases} 1, & x < 2-\varepsilon; \\ \frac{2-x}{\varepsilon}, & 2-\varepsilon \le x \le 2; \\ 0, & x > 2. \end{cases} \tag{3.2.47}$$

Then, on A_ε, $\sum_{i=1}^{M} f_\varepsilon\left(\lambda_{GUE,M}^{i}/\sqrt{M}\right)/M = 1$ and $\int f_\varepsilon d\nu < 1$, and so

$$\mathbb{P}\Big(\liminf_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} f_\varepsilon\left(\lambda_{GUE,M}^{i}/\sqrt{M}\right) \ne \int f_\varepsilon d\nu\Big) \ge \mathbb{P}\left(A_\varepsilon\right) > 0,$$

which is clearly a contradiction. Letting $\varepsilon \to 0$ in (3.2.46) yields,

$$\liminf_{M\to\infty} \frac{\lambda_{GUE,M}^{max}}{\sqrt{M}} \ge 2. \quad a.s. \tag{3.2.48}$$

Therefore, combining (3.2.45) and (3.2.48),

$$\frac{\lambda_{GUE,M}^{max}}{\sqrt{M}} \to 2. \quad a.s.$$

□

To prove our next convergence result, we first need a well known lemma (e.g., see [45]) and then a less known one.

Lemma 3.2.11 *For any $M \ge 2$, let $A_1, ..., A_M$ be centered Gaussian random variables, such that $\mathbb{E}A_i^2 \le 1$, for all $i = 1, ..., M$, then*

$$0 \le \mathbb{E}\left(\max_{k=1,...,M} A_k\right) \le \sqrt{2\ln M}. \tag{3.2.49}$$

Proof. First,

$$\mathbb{E}\left(\max_{k=1,\ldots,M} A_k\right) \geq \max_{k=1,\ldots,M} \mathbb{E}\left(A_k\right) = 0.$$

Next, since the logarithm is a concave function, for any $t > 0$,

$$\begin{aligned}
\mathbb{E}\left(\max_{k=1,\ldots,M} tA_k\right) &\leq \ln\left(\sum_{k=1}^{M} \mathbb{E}e^{tA_k}\right) \\
&\leq \ln M + \ln\left(e^{t^2/2}\right) \\
&= \ln M + t^2/2. \qquad (3.2.50)
\end{aligned}$$

The lemma is proved by taking $t = \sqrt{2\ln M}$. □

Lemma 3.2.12 *For each $k = 1, \ldots$, let χ_k^2 be a chi-square random variable with k degrees of freedom. Then,*

$$\lim_{M\to\infty} \mathbb{E}\left(\frac{\max_{k=1,\ldots,M} \chi_k^2}{M}\right) = 1. \qquad (3.2.51)$$

Proof. First,

$$\mathbb{E}\left(\max_{k=1,\ldots,M} \chi_k^2\right) \geq \mathbb{E}\left(\chi_M^2\right) = M.$$

Next, by the concavity of the logarithm, for any $0 < t < 1/2$,

$$\begin{aligned}
t\mathbb{E}\left(\frac{\max_{k=1,\ldots,M} \chi_k^2}{M}\right) &\leq \frac{1}{M}\ln\left(\sum_{k=1}^{M} \mathbb{E}e^{t\chi_k^2}\right) \\
&\leq \frac{1}{M}\ln\left(M\frac{1}{(1-2t)^{M/2}}\right) \\
&= \frac{\ln M}{M} - \frac{1}{2}\ln\left(1-2t\right). \qquad (3.2.52)
\end{aligned}$$

Hence for any $0 < t < 1/2$,

$$t\limsup_{M\to\infty} \mathbb{E}\left(\frac{\max_{k=1,\ldots,M} \chi_k^2}{M}\right) \leq -\frac{1}{2}\ln\left(1-2t\right).$$

By letting $t \to 0$,

$$\limsup_{M\to\infty} \mathbb{E}\left(\frac{\max_{k=1,\ldots,M} \chi_k^2}{M}\right) \leq \lim_{t\to 0} -\frac{\ln\left(1-2t\right)}{2t} = 1.$$

$\Big($Since $-\ln(1-2t) \le 2t + 4t^2$, for $0 \le t \le 1/3$, taking $t = \sqrt{\ln M/2M}$ in (3.2.52), will give $\mathbb{E}\left(\max_{k=1,\dots,M} \chi_k^2/M\right) \le 1 + 2\sqrt{2\ln M/M}$, for $M > 10$. $\Big)$ □

Theorem 3.2.13 *As $M \to \infty$,*

$$\frac{\lambda_{GUE,M}^{max,0}}{\sqrt{M}} \to 2, \quad in\ L^1.$$

Equivalently,

$$\frac{\lambda_{GUE,M}^{max}}{\sqrt{M}} \to 2, \quad in\ L^1.$$

Equivalently,

$$\frac{\tilde{H}_M^1}{\sqrt{M}} \to 2, \quad in\ L^1.$$

Proof. Note that when $p_1 = \cdots = p_M = 1/M$, $\mathcal{L}_{(s_1,\dots,s_M)}$, given by (3.1.8), is the empty set when $s_1 < 0$, and so $\lambda_{GUE,M}^{max,0}$ is nonnegative (this is actually clear from the traceless requirement). By Theorem 3.2.7, $\tilde{H}_M^1$ and $\lambda_{GUE,M}^{max,0}$ are equal in distribution, and so it suffices to prove that, as $M \to \infty$,

$$\frac{\mathbb{E}\left(\lambda_{GUE,M}^{max,0}\right)}{\sqrt{M}} \to 2. \tag{3.2.53}$$

Next, by Proposition 3.1.4, $\mathbb{E}\left(\lambda_{GUE,M}^{max,0}\right) = \mathbb{E}\left(\lambda_{GUE,M}^{max}\right)$. Moreover, taking expectations on both sides of (3.2.33) gives:

$$\mathbb{E}\left(\lambda_{GUE,M}^{max}\right) \le \mathbb{E}\left(\max_{k=1,\dots,M} A_{k,k}\right) + \mathbb{E}\left(\max_{k=1,\dots,M} \chi_{M-k+1}\right) + \mathbb{E}\left(\max_{k=1,\dots,M} \chi_{M-k}\right).$$

By Lemma 3.2.11,

$$\limsup_{M\to\infty} \mathbb{E}\left(\max_{k=1,\dots,M} \frac{A_{k,k}}{\sqrt{M}}\right) \le \limsup_{M\to\infty} \frac{\sqrt{2\ln M}}{\sqrt{M}} = 0,$$

while, by Lemma 3.2.12,

$$\limsup_{M\to\infty} \mathbb{E}\left(\max_{k=1,\dots,M} \frac{\chi_k}{\sqrt{M}}\right) = 1,$$

leading to

$$\limsup_{M\to\infty} \mathbb{E}\left(\frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}}\right) \leq 2.$$

Now, $\lambda^{max,0}_{GUE,M}$ is nonnegative and by Theorem 3.2.10, almost surely,

$$\frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}} \to 2.$$

Thus, by Fatou's Lemma,

$$\liminf_{M\to\infty} \mathbb{E}\left(\frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}}\right) \geq \mathbb{E}\left(\liminf_{M\to\infty} \frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}}\right) = 2,$$

and so,

$$\lim_{M\to\infty} \mathbb{E}\left(\frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}}\right) = 2.$$

Using once more the fact that $\lambda^{max,0}_{GUE,M} \geq 0$, we conclude that

$$\lim_{M\to\infty} \mathbb{E}\left|\frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}} - 2\right| = 0,$$

and by the weak law of large number, $\lim_{M\to\infty} \mathbb{E}\left|\frac{\lambda^{max}_{GUE,M}}{\sqrt{M}} - 2\right| = 0.$ □

Remark 3.2.14 *Following Davidson and Szarek [17] (see also [44]), an alternative proof of Theorem 3.2.13 is presented next. Let* $\mathbf{G}$ *be an element of the* $M \times M$ *GUE, let*

$$\lambda^{max}_{GUE,M} = \sup_{|u|=1} u\mathbf{G}^T u^*, \tag{3.2.54}$$

and consider the real-valued Gaussian process

$$G_u = u\mathbf{G}^T u^* = \sum_{i,j=1}^{M} \mathbf{G}_{i,j} u_i \bar{u}_j, \quad |u| = 1, \quad u = (u_1, ..., u_M) \in \mathbb{C}^M.$$

For any $u, v \in \mathbb{C}^M$, such that $|u| = |v| = 1$,

$$\begin{aligned}
\mathbb{E}\left(|G_u - G_v|^2\right) &= \mathbb{E}\left(\left|\sum_{i,j=1}^{M} \mathbf{G}_{i,j}\left(u_i\bar{u}_j - v_i\bar{v}_j\right)\right|^2\right) \\
&= \mathbb{E}\left(\sum_{i,j=1}^{M} \mathbf{G}_{i,j}\left(u_i\bar{u}_j - v_i\bar{v}_j\right)\mathbf{G}_{j,i}\left(u_j\bar{u}_i - v_j\bar{v}_i\right)\right) \\
&= \mathbb{E}\left(\sum_{i,j=1}^{M} |\mathbf{G}_{i,j}|^2 \left|u_i\bar{u}_j - v_i\bar{v}_j\right|^2\right) \\
&= \sum_{i,j=1}^{M} \left|u_i\bar{u}_j - v_i\bar{v}_j\right|^2 \\
&= \sum_{i,j=1}^{M} \left(u_i\bar{u}_j - v_i\bar{v}_j\right)\left(\bar{u}_i u_j - \bar{v}_i v_j\right) \\
&= \sum_{i=1}^{M} |u_i|^2 + \sum_{i=1}^{M} |v_i|^2 - \sum_{i,j=1}^{M} u_i\bar{u}_j\bar{v}_i v_j - \sum_{i,j=1}^{M} \bar{u}_i u_j v_i\bar{v}_j \\
&= 2 - 2\left|\sum_{i=1}^{M} u_i\bar{v}_i\right|^2. \qquad (3.2.55)
\end{aligned}$$

Now, define the Gaussian process indexed by $u \in \mathbb{C}^M$, $|u| = 1$,

$$K_u = \sum_{i=1}^{M} g_i Re(u_i) + \sum_{j=1}^{M} h_j Im(u_j), \qquad (3.2.56)$$

where $g_1, ..., g_M, h_1, ..., h_M$ are iid $N(0,1)$ random variables. Then for any $u, v \in \mathbb{C}^M$, such that $|u| = |v| = 1$,

$$\begin{aligned}
&\mathbb{E}\left(|K_u - K_v|^2\right) \\
&= \mathbb{E}\left(\left|\sum_{i=1}^{M} g_i\left(Re(u_i) - Re(v_i)\right) + \sum_{j=1}^{M} h_j\left(Im(u_j) - Im(v_j)\right)\right|^2\right) \\
&= \sum_{i=1}^{M} \left(Re(u_i) - Re(v_i)\right)^2 + \sum_{j=1}^{M} \left(Im(u_j) - Im(v_j)\right)^2 \\
&= \sum_{i=1}^{M} \left(Re(u_i - v_i)\right)^2 + \sum_{i=1}^{M} \left(Im(u_i - v_i)\right)^2
\end{aligned}$$

(3.2.57)

$$
\begin{aligned}
&= \sum_{i=1}^{M} |u_i - v_i|^2 \\
&= \sum_{i=1}^{M} (u_i - v_i)(\bar{u}_i - \bar{v}_i) \\
&= \sum_{i=1}^{M} |u_i|^2 + \sum_{i=1}^{M} |v_i|^2 - \sum_{i=1}^{M} u_i \bar{v}_i - \sum_{i=1}^{M} \bar{u}_i v_i \\
&= 2 - 2Re\left(\sum_{i=1}^{M} u_i \bar{v}_i\right). \qquad (3.2.58)
\end{aligned}
$$

Now

$$
\mathbb{E}\left(|G_u - G_v|^2\right) \le 2\mathbb{E}\left(|K_u - K_v|^2\right),
$$

since

$$
\begin{aligned}
&2\mathbb{E}\left(|K_u - K_v|^2\right) - \mathbb{E}\left(|G_u - G_v|^2\right) \\
&\quad = \left(4 - 4Re\left(\sum_{i=1}^{M} u_i \bar{v}_i\right)\right) - \left(2 - 2\left|\sum_{i=1}^{M} u_i \bar{v}_i\right|^2\right) \\
&\quad = 2\left(\left|\sum_{i=1}^{M} u_i \bar{v}_i\right|^2 - 2Re\left(\sum_{i=1}^{M} u_i \bar{v}_i\right) + 1\right) \\
&\quad = 2\left(Re\left(\sum_{i=1}^{M} u_i \bar{v}_i\right) - 1\right)^2 + 2Im\left(\sum_{i=1}^{M} u_i \bar{v}_i\right)^2 \\
&\quad \ge 0. \qquad (3.2.59)
\end{aligned}
$$

Then, by the Slepian-Fernique Lemma [45],

$$
\mathbb{E}\left(\sup_{|u|=1} G_u\right) \le \sqrt{2}\mathbb{E}\left(\sup_{|u|=1} K_u\right).
$$

Finally, for $i = 1, ..., 2M$, *let*

$$
a_i = \begin{cases} g_i, & \text{i=1,...,M;} \\ h_{i-M}, & \text{i=M+1,...,2M,} \end{cases}
$$

and

$$
b_i = \begin{cases} Re(u_i), & \text{i=1,...,M;} \\ Im(u_{i-M}), & \text{i=M+1,...,2M.} \end{cases}
$$

Then,

$$
\begin{aligned}
\mathbb{E}\left(\lambda_{GUE,M}^{max}\right) &= \mathbb{E}\left(\sup_{|u|=1} G_u\right) \\
&\leq \sqrt{2}\mathbb{E}\left(\sup_{|u|=1} K_u\right) \\
&\leq \sqrt{2}\mathbb{E}\left(\sup_{|u|=1}\left(\sum_{i=1}^{M} g_i Re(u_i) + \sum_{j=1}^{M} h_j Im(u_j)\right)\right) \\
&= \sqrt{2}\mathbb{E}\left(\sup_{|u|=1}\left(\sum_{i=1}^{2M} a_i b_i\right)\right) \\
&\leq \sqrt{2}\mathbb{E}\left(\sup_{|u|=1}\left(\sqrt{\sum_{i=1}^{2M} a_i^2}\sqrt{\sum_{i=1}^{2M} b_i^2}\right)\right) \\
&= \sqrt{2}\mathbb{E}\left(\sup_{|u|=1}\left(\sqrt{\sum_{i=1}^{M} (g_i^2 + h_i^2)}\sqrt{\sum_{i=1}^{M} Re(u_i)^2 + Im(u_j)^2}\right)\right) \\
&\leq 2\mathbb{E}\left(\sqrt{\sum_{i=1}^{M} g_i^2}\right) \\
&\leq 2\sqrt{M}. \qquad (3.2.60)
\end{aligned}
$$

Therefore,

$$
\frac{\mathbb{E}\left(\lambda_{GUE,M}^{max}\right)}{\sqrt{M}} \leq 2. \qquad (3.2.61)
$$

We have established in Theorem 3.2.10 that $\lambda_{GUE,M}^{max}/\sqrt{M} \xrightarrow{a.s.} 2$, *as* $M \to \infty$. *Hence* $\left(\lambda_{GUE,M}^{max}/\sqrt{M} - 2\right)^{-} \xrightarrow{a.s.} 0$, *and since* $0 \leq \left(\lambda_{GUE,M}^{max}\right)^{-}/\sqrt{M} \leq \left(\lambda_{GUE,M}^{max}/\sqrt{M} - 2\right)^{-}$, *we have*

$$
\frac{\left(\lambda_{GUE,M}^{max}\right)^{-}}{\sqrt{M}} \xrightarrow{a.s.} 0,
$$

and therefore

$$
\frac{\left(\lambda_{GUE,M}^{max}\right)^{+}}{\sqrt{M}} \xrightarrow{a.s.} 2.
$$

Now,

$$
0 \leq \mathbb{E}\left(\frac{\lambda_{GUE,M}^{max,0}}{\sqrt{M}}\right) = \mathbb{E}\left(\frac{\lambda_{GUE,M}^{max}}{\sqrt{M}}\right) \leq 2,
$$

hence,

$$0 \le \mathbb{E}\left(2 - \frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}}\right) \le \mathbb{E}\left(2 - \frac{\lambda^{max}_{GUE,M}}{\sqrt{M}}\right)^+ \le 2.$$

Since $0 \le \left(2 - \lambda^{max}_{GUE,M}/\sqrt{M}\right)^+ \le 2$, *and since* $\lambda^{max,0}_{GUE,M}/\sqrt{M} \xrightarrow{a.s.} 2$, *dominated convergence ensures that*

$$\mathbb{E}\frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}} \longrightarrow 2,$$

from which it follows that

$$\mathbb{E}\left|\frac{\lambda^{max,0}_{GUE,M}}{\sqrt{M}} - 2\right| \longrightarrow 0,$$

since $\lambda^{max,0}_{GUE,M} \ge 0$.

Here is another almost sure convergence result, for $\tilde{H}^1_M$.

Theorem 3.2.15 *As* $M \to \infty$, *almost surely,*

$$\frac{\tilde{H}^1_M}{\sqrt{M}} \longrightarrow 2.$$

Proof. By Theorem 3.2.13, for any $\varepsilon > 0$, when M is large enough,

$$\left|\mathbb{E}\left[\tilde{H}^1_M/\sqrt{M}\right] - 2\right| < \varepsilon/2.$$

Thus

$$\begin{aligned}\mathbb{P}\left(\left|\frac{\tilde{H}^1_M}{\sqrt{M}} - 2\right| > \varepsilon\right) &\le \mathbb{P}\left(\left|\frac{\tilde{H}^1_M}{\sqrt{M}} - \mathbb{E}\left(\frac{\tilde{H}^1_M}{\sqrt{M}}\right)\right| > \varepsilon - \left|2 - \mathbb{E}\left(\frac{\tilde{H}^1_M}{\sqrt{M}}\right)\right|\right) \\ &\le \mathbb{P}\left(\left|\tilde{H}^1_M - \mathbb{E}\left(\tilde{H}^1_M\right)\right| > \frac{\sqrt{M}\varepsilon}{2}\right) \\ &\le 2e^{-\frac{M\varepsilon^2}{8\frac{M-1}{M}}} \\ &< 2e^{-\frac{M\varepsilon^2}{8}}, \end{aligned} \tag{3.2.62}$$

where, in the second to last inequality, we have used Gaussian concentration for $\tilde{H}^1_M$ since it is a 1-Lipschitz function of the Brownian entries each one having variance

$(M-1)/M$. Hence, for any ε,

$$\sum_{M=1}^{\infty} \mathbb{P}\left(\left|\frac{\tilde{H}_M^1}{\sqrt{M}} - 2\right| > \varepsilon\right) < \infty,$$

and the almost sure convergence of $\tilde{H}_M^1/\sqrt{M}$ to 2 follows from the Borel-Cantelli lemma. □

CHAPTER IV

LONGEST INCREASING SUBSEQUENCE FOR NON-UNIFORM FINITE ALPHABETS

4.1 Generalized Traceless GUE

The arguments used in the previous chapter can be modified to obtain the limiting law of the shape of the Young tableaux associated to a random word whose letters are independently, but no longer uniformly, drawn from an M-letter alphabet. This limiting law is closely related to the spectrum of certain random matrix ensembles and it will be described in this section.

Definition 4.1.1 *For $M \geq 1$, $K = 1, ..., M$ and $d_1, ..., d_K$ such that $\sum_{k=1}^{K} d_k = M$, let $\mathcal{G}_M(d_1, ..., d_K)$ be the set of random matrices $\mathbf{X}$ which are direct sums of mutually independent elements of the $d_k \times d_k$ GUE, $k = 1, ..., K$ (i.e., $\mathbf{X}$ is an $M \times M$ block diagonal matrix whose K blocks are mutually independent elements of the $d_k \times d_k$ GUE, $k = 1, ..., K$).*

Example 4.1.1 *An element of $\mathcal{G}_5(2, 1, 2)$ is*

$$\begin{bmatrix} X_{1,1} & X_{1,2} & 0 & 0 & 0 \\ X_{2,1} & X_{2,2} & 0 & 0 & 0 \\ 0 & 0 & X_{3,3} & 0 & 0 \\ 0 & 0 & 0 & X_{4,4} & X_{4,5} \\ 0 & 0 & 0 & X_{5,4} & X_{5,5} \end{bmatrix},$$

where $\begin{bmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{bmatrix}$ and $\begin{bmatrix} X_{4,4} & X_{4,5} \\ X_{5,4} & X_{5,5} \end{bmatrix}$ are elements of the 2×2 GUE and where $X_{3,3} \sim N(0, 1)$; these three matrices being mutually independent.

Definition 4.1.2 *Let $p_1, \cdots, p_M > 0$, $\sum_{j=1}^M p_j = 1$, be such that the multiplicities of the K distinct probabilities $p^{(1)}, ..., p^{(K)}$ are respectively $d_1, ..., d_K$, i.e., let $m_1 = 0$ and for any $k = 2, ..., K$, let $m_k = \sum_{j=1}^{k-1} d_j$, and $p_{m_k+1} = \cdots = p_{m_k+d_k} = p^{(k)}$, $k = 1, ..., K$. The generalized $M \times M$ traceless GUE associated to the probabilities $p_1, ..., p_M$ is the set, denoted by $\mathcal{G}^0(p_1, ..., p_M)$, of $M \times M$ matrices $\mathbf{X}^0$, of the form*

$$\mathbf{X}^0_{i,j} = \begin{cases} \mathbf{X}_{i,i} - \sqrt{p_i} \sum_{l=1}^M \sqrt{p_l} \mathbf{X}_{l,l}, & \text{if } i = j; \\ \mathbf{X}_{i,j}, & \text{if } i \neq j, \end{cases} \tag{4.1.1}$$

where $\mathbf{X} \in \mathcal{G}_M(d_1, ..., d_K)$.

Clearly, from (4.1.1), $\sum_{i=1}^M \sqrt{p_i} \mathbf{X}^0_{i,i} = 0$. Note also that the case $K = 1$ (for which $d_1 = M$), recovers the previously defined traceless GUE.

Here is an equivalent way of defining the generalized traceless GUE: let $\mathbf{X}^{(k)}$ be the $M \times M$ diagonal matrix such that

$$\mathbf{X}^{(k)}_{i,i} = \begin{cases} \sqrt{p^{(k)}} \sum_{l=1}^M \sqrt{p_l} \mathbf{X}_{l,l}, & \text{if } m_k < i \leq m_k + d_k; \\ 0, & \text{otherwise}, \end{cases} \tag{4.1.2}$$

and let $\mathbf{X} \in \mathcal{G}_M(d_1, ..., d_K)$. Then, $\mathbf{X}^0 := \mathbf{X} - \sum_{k=1}^K \mathbf{X}^{(k)} \in \mathcal{G}^0(p_1, ..., p_M)$.

Example 4.1.2 *An element of* $\mathcal{G}^0(0.1, 0.1, 0.4, 0.2, 0.2)$ *is*

$$\begin{bmatrix} X_{1,1} - \sqrt{0.1}Y & X_{1,2} & 0 & 0 & 0 \\ X_{2,1} & X_{2,2} - \sqrt{0.1}Y & 0 & 0 & 0 \\ 0 & 0 & X_{3,3} - \sqrt{0.4}Y & 0 & 0 \\ 0 & 0 & 0 & X_{4,4} - \sqrt{0.2}Y & X_{4,5} \\ 0 & 0 & 0 & X_{5,4} & X_{5,5} - \sqrt{0.2}Y \end{bmatrix},$$

where

$$\begin{bmatrix} X_{1,1} & X_{1,2} & 0 & 0 & 0 \\ X_{2,1} & X_{2,2} & 0 & 0 & 0 \\ 0 & 0 & X_{3,3} & 0 & 0 \\ 0 & 0 & 0 & X_{4,4} & X_{4,5} \\ 0 & 0 & 0 & X_{5,4} & X_{5,5} \end{bmatrix},$$

is an element of $\mathcal{G}_5(2,1,2)$ *and* $Y = \sqrt{0.1}X_{1,1} + \sqrt{0.1}X_{2,2} + \sqrt{0.4}X_{3,3} + \sqrt{0.2}X_{4,4} + \sqrt{0.2}X_{5,5}$.

Equivalently, there is an "ensemble" description of $\mathcal{G}^0(p_1, ..., p_M)$.

Proposition 4.1.3 $\mathbf{X}^0 \in \mathcal{G}^0(p_1, ..., p_M)$ *if and only if* $\mathbf{X}^0$ *is distributed according to the probability distribution*

$$\mathbb{P}\left(d\mathbf{X}^0\right) = C\gamma\left(d\mathbf{X}^0_{1,1}, ..., d\mathbf{X}^0_{M,M}\right)\prod_{k=1}^{K}\left(e^{-\sum\limits_{m_k<i<j\le m_k+d_k}\left|\mathbf{X}^0_{i,j}\right|^2}\prod_{m_k<i<j\le m_k+d_k} dRe\left(\mathbf{X}^0_{i,j}\right)dIm\left(\mathbf{X}^0_{i,j}\right)\right), \tag{4.1.3}$$

on the space of $M \times M$ *Hermitian matrices, which are direct sum of* $d_k \times d_k$ *Hermitian matrices,* $k = 1, ..., K$, $\sum_{k=1}^{K} d_k = M$, *and where* $m_1 = 0$, $m_k = \sum_{j=1}^{k-1} d_j$, *for any* $k = 2, ..., K$. *Above,* $C = \pi^{-\sum_{k=1}^{K} d_k(d_k-1)/2}$ *and* $\gamma\left(d\mathbf{X}^0_{1,1}, ..., d\mathbf{X}^0_{M,M}\right)$ *is the distribution of an* M*-dimensional centered (degenerate) multivariate Gaussian law with covariance matrix*

$$\mathbf{\Sigma}_0 = \begin{pmatrix} 1-p_1 & -\sqrt{p_1p_2} & \cdots & -\sqrt{p_1p_M} \\ -\sqrt{p_2p_1} & 1-p_2 & \cdots & -\sqrt{p_2p_M} \\ \vdots & \ddots & \ddots & \vdots \\ -\sqrt{p_Mp_1} & \cdots & -\sqrt{p_Mp_{M-1}} & 1-p_M \end{pmatrix}.$$

Proof. Let $\mathbf{X} \in \mathcal{G}_M(d_1, ..., d_K)$, and let $\mathbf{X}^0$ be given as in (4.1.1). Then,

$$\left(\mathbf{X}^0_{1,1}, \mathbf{X}^0_{2,2}, ..., \mathbf{X}^0_{M,M}\right)' = \mathbf{\Sigma}_0\left(\mathbf{X}_{1,1}, \mathbf{X}_{2,2}, ..., \mathbf{X}_{M,M}\right)'.$$

Since $(\mathbf{X}_{1,1}, ..., \mathbf{X}_{M,M}) \sim \mathbf{N}(\mathbf{0}, \mathbf{I}_M)$, and $\mathbf{\Sigma}^0\mathbf{\Sigma}^0 = \mathbf{\Sigma}^0$, it follows that $\left(\mathbf{X}^0_{1,1}, ..., \mathbf{X}^0_{M,M}\right) \sim \mathbf{N}(\mathbf{0}, \mathbf{\Sigma}^0)$. Now, the off diagonal entries of $\mathbf{X}$ are independent of the random variables $\mathbf{X}_{1,1} - \sqrt{p_1}\sum_{l=1}^M \sqrt{p_l}\mathbf{X}_{l,l}, ..., \mathbf{X}_{M,M} - \sqrt{p_M}\sum_{l=1}^M \sqrt{p_l}\mathbf{X}_{l,l}$, and thus the distribution of $\mathbf{X}^0$ is given by (4.1.3). On the other hand, suppose the matrix $\mathbf{X}^0$ is distributed according to the probability distribution (4.1.3). Clearly, the diagonal entries $\mathbf{X}^0_{1,1}, ..., \mathbf{X}^0_{M,M}$ are independent of the off diagonal ones. Moreover,

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{i=1}^M \sqrt{p_i}\mathbf{X}^0_{i,i}\right)^2\right] &= \mathbb{E}\left[\sum_{i,j=1}^M \sqrt{p_i}\sqrt{p_j}\mathbf{X}^0_{i,i}\mathbf{X}^0_{j,j}\right] \\
&= \left(\sqrt{p_1}, ..., \sqrt{p_M}\right)\mathbf{\Sigma}^0\left(\sqrt{p_1}, ..., \sqrt{p_M}\right)' \\
&= \sum_{i=1}^M p_i\left(1-p_i\right) - \sum_{i=1}^M \sum_{j\neq i} p_i p_j \\
&= \sum_{i=1}^M p_i - \sum_{i=1}^M p_i \\
&= 0. \qquad (4.1.4)
\end{aligned}$$

Therefore, $\sum_{i=1}^M \sqrt{p_i}\mathbf{X}^0_{i,i} = 0$. Since $\left(\mathbf{X}^0_{1,1}, ..., \mathbf{X}^0_{M,M}\right) \sim \mathbf{N}(\mathbf{0}, \mathbf{\Sigma}^0)$ and $\mathbf{\Sigma}^0\mathbf{\Sigma}^0 = \mathbf{\Sigma}^0$, there exists a vector $(Z_1, ..., Z_M) \sim \mathbf{N}(\mathbf{0}, \mathbf{I}_M)$ such that

$$\left(\mathbf{X}^0_{1,1}, ..., \mathbf{X}^0_{M,M}\right)' = \mathbf{\Sigma}^0\left(Z_1, ..., Z_M\right)',$$

and, moreover, the vector $(Z_1, ..., Z_M)$ can also be chosen to be independent of the off diagonal entries of $\mathbf{X}^0$. Let $\mathbf{X}$ be the matrix $\mathbf{X}^0$ with the diagonal entries $\mathbf{X}^0_{1,1}, ..., \mathbf{X}^0_{M,M}$ replaced by $Z_1, ..., Z_M$. Then $\mathbf{X} \in \mathcal{G}_M\left(d_1, ..., d_K\right)$ and $\mathbf{X}^0$ is given as in (4.1.1). □

Next, we provide a relation in law between the spectrum of $\mathbf{X}$ and that of $\mathbf{X}^0$.

Proposition 4.1.4 *Let $\mathbf{X} \in \mathcal{G}_M\left(d_1, ..., d_K\right)$, and let $\mathbf{X}^0 \in \mathcal{G}^0\left(p_1, ..., p_M\right)$. Let $\lambda^1, \cdots, \lambda^M$ be the eigenvalues of $\mathbf{X}$, where for each $k = 1, ..., K$, $\lambda^{m_k+1}, \cdots, \lambda^{m_k+d_k}$ are the eigenvalues of the kth diagonal block (an element of the $d_k \times d_k$ GUE). Then, the* *eigenvalues of* $\mathbf{X}^0$ *are given by:*

$$\lambda_0^i = \lambda^i - \sqrt{p_i} \sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l} = \lambda^i - \sqrt{p_i} \sum_{l=1}^{M} \sqrt{p_l} \lambda^l, \quad i = 1, ..., M.$$

Proof. It is clear that $\lambda_0^i = \lambda^i - \sqrt{p_i} \sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l}$, $i = 1, ..., M$, are the eigenvalues of $\mathbf{X}^0$. Next, to prove that $\sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l} = \sum_{l=1}^{M} \sqrt{p_l} \lambda^l$, let $\mathbf{X}_p$ be the $M \times M$ matrix obtained by multiplying the kth diagonal block of $\mathbf{X}$ by $\sqrt{p^{(k)}}$. For each $i = 1, ..., M$, there exists a unique $1 \leq k \leq K$, such that $m_k < i \leq m_k + d_k$, and λ^i is an eigenvalue of the kth diagonal block of $\mathbf{X}$. Moreover, $p_i = p^{(k)}$, thus $\sqrt{p_i}\lambda^i$ is an eigenvalue of the kth diagonal block of $\mathbf{X}_p$, which is an eigenvalue of $\mathbf{X}_p$ as well. Then, $\sqrt{p_1}\lambda^1, ..., \sqrt{p_M}\lambda^M$ are the eigenvalues of $\mathbf{X}_p$. The proposition is proved since both $\sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l}$ and $\sum_{l=1}^{M} \sqrt{p_l} \lambda^l$ are equal to $tr\mathbf{X}_p$. □

The semicircle law is again the almost sure limit of the empirical spectral measure for the kth block of the generalized traceless GUE, provided $d_k \to \infty$, $k = 1, ..., K$. This is, for example, the case of the uniform alphabet, where again $K = 1$, $d_1 = M$ and $p^{(1)} = 1/M$.

Proposition 4.1.5 *Let* $\lambda_0^1, \lambda_0^2, ..., \lambda_0^M$ *be the eigenvalues of an element of the* $M \times M$ *generalized traceless GUE, such that* $\lambda_0^{m_k+1}, \cdots, \lambda_0^{m_k+d_k}$ *are the eigenvalues of the* k*th diagonal block,* $k = 1, ..., K$*. For any* $k = 1, ..., K$*, the empirical distribution of the eigenvalues* $\left(\lambda_0^i/\sqrt{d_k}\right)_{m_k<i\leq m_k+d_k}$ *converges almost surely to the semicircle law* ν *with density* $\sqrt{4-x^2}/2\pi$*,* $-2 \leq x \leq 2$*, whenever* $d_k \to \infty$*.*

Proof. By Proposition 4.1.4, there exists an $\mathbf{X} \in \mathcal{G}_M(d_1, ..., d_K)$, whose eigenvalues $\lambda^1, \cdots, \lambda^M$ satisfy,

$$\lambda_0^i = \lambda^i - \sqrt{p_i} \sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l} = \lambda^i - \sqrt{p_i} \sum_{l=1}^{M} \sqrt{p_l} \lambda^l, \quad i = 1, ..., M.$$

If $d_k \to \infty$ for some $k = 1, ..., K$, by Wigner's theorem [49], the spectral measure of $\left(\lambda^i/\sqrt{d_k}\right)_{m_k<i\leq m_k+d_k}$ converges weakly to the semicircle law ν almost surely, i.e., for

any bounded continuous function $f : \mathbb{R} \to \mathbb{R}$,

$$\frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda^i}{\sqrt{d_k}}\right) \longrightarrow \int f d\nu,$$

almost surely. Now, $\sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l} \sim N(0,1)$, and $p^{(k)}/d_k \to 0$ as $d_k \to \infty$, hence,

$$\frac{\sqrt{p^{(k)}} \sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l}}{\sqrt{d_k}} \xrightarrow{a.s.} 0.$$

Next, for any bounded Lipschitz function f, almost surely,

$$\begin{aligned} &\left| \frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda_0^i}{\sqrt{d_k}}\right) - \int f d\nu \right| \\ &\leq \left| \frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda_0^i}{\sqrt{d_k}}\right) - \frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda^i}{\sqrt{d_k}}\right) \right| + \left| \frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda^i}{\sqrt{d_k}}\right) - \int f d\nu \right|. \end{aligned} \tag{4.1.5}$$

Now,

$$\left| \frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda_0^i}{\sqrt{d_k}}\right) - \frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda^i}{\sqrt{d_k}}\right) \right| \leq \|f\|_{Lip} \left| \frac{\sqrt{p^{(k)}} \sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l}}{\sqrt{d_k}} \right| \longrightarrow 0, \tag{4.1.6}$$

and the proposition is proved, since the bounded Lipschitz functions form a determining class for weak convergence ([19, Section 9.3]).

An alternative argument, similar to the one given in the proof of Proposition 3.1.2, can also be applied to the non-uniform case. Indeed, if $d_k \to \infty$ for some $k = 1, ..., K$,

$$\frac{\sqrt{p^{(k)}} \sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l}}{\sqrt{d_k}} \xrightarrow{a.s.} 0.$$

Next, for any 1-Lipschitz function $f : \mathbb{R} \to \mathbb{R}$,

$$\left| \frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda_0^i}{\sqrt{d_k}}\right) - \frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda^i}{\sqrt{d_k}}\right) \right| \leq \left| \frac{\sqrt{p^{(k)}} \sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l}}{\sqrt{d_k}} \right|.$$

Again, via concentration and the Borel-Cantelli Lemma, we have,

$$\frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} f\left(\frac{\lambda_0^i}{\sqrt{d_k}}\right) - \int f d\mu^M \longrightarrow 0,$$

almost surely for every bounded Lipschitz function $f : \mathbb{R} \to \mathbb{R}$. Thus, almost surely,

$$\frac{1}{d_k} \sum_{i=m_k+1}^{m_k+d_k} \delta\left(\frac{\lambda_0^i}{\sqrt{d_k}}\right) \longrightarrow \nu,$$

weakly as probability measures on $\mathbb{R}$. □

Now for $p_1, ..., p_M$ considered, so far, i.e., such that the multiplicities of the K distinct probabilities $p^{(1)}, ..., p^{(K)}$ are respectively $d_1, ..., d_K$ and $p_{m_k+1} = \cdots = p_{m_k+d_k} = p^{(k)}$, $k = 1, ..., K$, let

$$\mathcal{L}^{p_1,...,p_M} := \Big\{ x = (x_1, ..., x_M) \in \mathbb{R}^M : \ x_{m_k+1} \geq \cdots \geq x_{m_k+d_k}, \ k = 1, ..., K; \\ \sum_{j=1}^{M} \sqrt{p_j} x_j = 0 \Big\}. \tag{4.1.7}$$

In other words, $\mathcal{L}^{p_1,...,p_M}$ is a subset of the hyperplane $\sum_{j=1}^{M} \sqrt{p_j} x_j = 0$, where within each block of size d_k, $k = 1, ..., K$, the coordinates $x_{m_k+1}, ..., x_{m_k+d_k}$, are ordered. For any $s_1, ..., s_M \in \mathbb{R}$, let also

$$\mathcal{L}^{p_1,...,p_M}_{(s_1,...,s_M)} := \mathcal{L}^{p_1,...,p_M} \cap \Big\{ (x_1, ..., x_M) \in \mathbb{R}^M : x_i \leq s_i, \ i = 1, ..., M \Big\}. \tag{4.1.8}$$

The distribution function of the eigenvalues, written in non-increasing order within each $d_k \times d_k$ GUE, of an element of $\mathcal{G}^0(p_1, ..., p_M)$ is obtained now.

Proposition 4.1.6 *The joint distribution function of the eigenvalues, written in non-increasing order within each $d_k \times d_k$ GUE, of an element of $\mathcal{G}^0(p_1, ..., p_M)$ is given, for any $s_1, ..., s_M \in \mathbb{R}$, by*

$$\mathbb{P}\Big(\lambda_0^1 \leq s_1, \lambda_0^2 \leq s_2, ..., \lambda_0^M \leq s_M\Big) = \int_{\mathcal{L}^{p_1,...,p_M}_{(s_1,...,s_M)}} f(x) dx_1 \cdots dx_{M-1}, \tag{4.1.9}$$

where for $x = (x_1, ..., x_M) \in \mathbb{R}^M$,

$$f(x) := c_M \prod_{k=1}^{K} \Delta_k(x)^2 e^{-\sum_{i=1}^{M} x_i^2/2} \mathbf{1}_{\mathcal{L}^{p_1,...,p_M}}(x), \tag{4.1.10}$$

with $c_M = (2\pi)^{-(M-1)/2} \prod_{k=1}^{K} (0!1!\cdots(d_k-1)!)^{-1}$ *and where* $\Delta_k(x)$ *is the Vandermonde determinant associated to those* x_i *for which* $p_i = p^{(k)}$*, i.e.,*

$$\Delta_k(x) = \prod_{m_k \le i < j \le m_k + d_k} (x_i - x_j).$$

Remark 4.1.7 *When the eigenvalues are not ordered within each* $d_k \times d_k$ *GUE, the identity (4.1.9) remains valid, multiplying* c_M*, above, by* $\prod_{k=1}^{K} (d_k!)^{-1}$*, and also by omitting the ordering constraint* $x_{m_k+1} \ge \cdots \ge x_{m_k+d_k}$*,* $k = 1, ..., K$*, in the definition of* $\mathcal{L}^{p_1,...,p_M}$*.*

Proof. From Proposition 4.1.4,

$$\lambda_0^i = \lambda^i - \sqrt{p_i} \sum_{l=1}^{M} \sqrt{p_l} \lambda^l, \quad i = 1, ..., M,$$

where $\lambda^1, \cdots, \lambda^M$ are the eigenvalues of an element of $\mathcal{G}_M(d_1, ..., d_K)$, and where $\lambda^{m_k+1} \ge \cdots \ge \lambda^{m_k+d_k}$ are the eigenvalues of the kth diagonal block (an element of the $d_k \times d_k$ GUE), for each $k = 1, ..., K$. Clearly, $\sum_{l=1}^{M} \sqrt{p_l} \lambda_0^l = 0$. Let us now compute the joint density of $\lambda_0^1, \cdots, \lambda_0^{M-1}$. Recall that the joint density of $(\lambda^1, \cdots, \lambda^M)$ is, for any $x \in \mathbb{R}^M$, given by

$$\frac{1}{\sqrt{2\pi}} c_M \prod_{k=1}^{K} \Delta_k(x)^2 e^{-\sum_{i=1}^{M} x_i^2/2}, \tag{4.1.11}$$

where $c_M = (2\pi)^{-(M-1)/2} \prod_{k=1}^{K} (0!1!\cdots(d_k-1)!)^{-1}$, and where

$$\Delta_k(x) = \prod_{m_k \le i < j \le m_k + d_k} (x_i - x_j),$$

with again $m_1 = 0$ and $m_k = \sum_{j=1}^{k-1} d_j$, for $k = 2, ..., K$. Consider the change of variables from $(\lambda^1, \cdots, \lambda^{M-1}, \lambda^M)$ to $(\lambda_0^1, \cdots, \lambda_0^{M-1}, Y)$, where $Y = \sqrt{p_M} \sum_{l=1}^{M} \sqrt{p_l} \lambda^l$. Then,

$$\frac{1}{\det(J)} = \det \begin{pmatrix} 1 - p_1 & -\sqrt{p_1}\sqrt{p_2} & \cdots & -\sqrt{p_1}\sqrt{p_M} \\ \vdots & \ddots & \ddots & \vdots \\ -\sqrt{p_{M-1}}\sqrt{p_1} & \cdots & 1 - p_{M-1} & -\sqrt{p_{M-1}}\sqrt{p_M} \\ \sqrt{p_1} & \cdots & \sqrt{p_{M-1}} & \sqrt{p_M} \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ \sqrt{p_1} & \cdots & \cdots & \sqrt{p_{M-1}} & \sqrt{p_M} \end{pmatrix}$$

$$= \sqrt{p_M}. \tag{4.1.12}$$

where J is the Jacobian of this transformation. The joint density of $\left(\lambda_0^1, \cdots, \lambda_0^{M-1}, Y\right)$ is thus given by:

$$\begin{aligned} & f_{\lambda_0^1, \cdots, \lambda_0^{M-1}, Y}\left(x_1^0, ..., x_{M-1}^0, y\right) \\ &= \frac{1}{\sqrt{2\pi p_M}} c_M e^{-\frac{\frac{y^2}{p_M} + \sum_{j=1}^M x_j^{0\,2}}{2}} \prod_{1 \le i < j \le M} (x_i^0 - x_j^0)^2 \mathbf{1}_{\mathcal{L}^{p_1, \ldots, p_M}}\left(x_1^0, ..., x_{M-1}^0, y\right) \\ &= \frac{1}{\sqrt{2\pi p_M}} e^{-\frac{y^2}{2p_M}} f(x^0), \end{aligned} \tag{4.1.13}$$

where $x^0 = (x_1^0, ..., x_M^0)$ and $x_M^0 = -\sum_{j=1}^{M-1} x_j^0$. Integrating y from $-\infty$ to ∞, shows that the joint density of $\left(\lambda_0^1, \cdots, \lambda_0^{M-1}\right)$ is $f(x^0)$. $\square$

For $p_1, ..., p_M$, as in Definition 4.1.2, the forthcoming proposition gives a relation in law between the spectra of elements of $\mathcal{G}_M(d_1, ..., d_K)$ and of $\mathcal{G}^0(p_1, ..., p_M)$.

Proposition 4.1.8 *For any $M \ge 2$, let $p_1, ..., p_M$, $\mathbf{X}$ and $\mathbf{X}^0$ be as in Definition 4.1.2. Let $\lambda^1, \cdots, \lambda^M$ be the eigenvalues of $\mathbf{X}$, and let $\lambda_0^1, \cdots, \lambda_0^M$ be the eigenvalues of $\mathbf{X}^0$ as given in Proposition 4.1.4. Then*

$$\left(\lambda^1, \cdots, \lambda^M\right) \stackrel{d}{=} \left(\lambda_0^1, \cdots, \lambda_0^M\right) + (Z_1, \cdots, Z_M),$$

where $(Z_1, \cdots, Z_M)$ is a centered (degenerate) multivariate Gaussian vector with covariance matrix $\left(\sqrt{p_i p_j}\right)_{1 \le i,j \le M}$. Moreover, $\left(\lambda_0^1, \cdots, \lambda_0^M\right)$ and $(Z_1, \cdots, Z_M)$ are independent.

Proof. For any $s_1, ..., s_M \in \mathbb{R}$,

$$\mathbb{P}\Big(\lambda^1 \leq s_1, ..., \lambda^M \leq s_M\Big)$$
$$= \frac{c_M}{\sqrt{2\pi}} \int_{-\infty}^{s_1} \cdots \int_{-\infty}^{s_M} \prod_{k=1}^{K} \Delta_k(x)^2 e^{-\sum_{i=1}^{M} x_i^2/2} dx_1 \cdots dx_M, \tag{4.1.14}$$

where $c_M = (2\pi)^{-(M-1)/2} \prod_{k=1}^{K} (0!1! \cdots (d_k - 1)!)^{-1}$. Consider the change of variables

$$y = \sum_{j=1}^{M} \sqrt{p_j} x_j, \;\; x_j = x_j^0 + \sqrt{p_j} y, \qquad j = 1, ..., M. \tag{4.1.15}$$

Clearly, $\sum_{j=1}^{M} \sqrt{p_j} x_j^0 = 0$. With $\mathcal{L}_{(s_1,\ldots,s_M)}^{p_1,\ldots,p_M}$ as in (4.1.8), we have

$$\mathbb{P}\Big(\lambda^1 \leq s_1, ..., \lambda^M \leq s_M\Big)$$
$$= \frac{c_M}{\sqrt{2\pi}} \int_{-\infty}^{s_1} \cdots \int_{-\infty}^{s_M} \prod_{k=1}^{K} \left(\Delta_k(x)^2 e^{-\sum_{i=m_k}^{m_k+d_k} x_i^2/2}\right) dx_1 \cdots dx_M$$
$$= \frac{c_M}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dy \int_{\mathcal{L}_{(s_1-\sqrt{p_1}y,\ldots,s_M-\sqrt{p_M}y)}^{p_1,\ldots,p_M}} \prod_{k=1}^{K} \left(\Delta_k(x)^2 e^{-\sum_{i=m_k}^{m_k+d_k} x_i^2/2}\right) e^{-\frac{y^2}{2}} dx^0$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \mathbb{P}\left(\lambda_0^1 < s_1 - \sqrt{p_1}y, \cdots, \lambda_0^M < s_M - \sqrt{p_M}y\right) dy$$
$$= \mathbb{E}\left[\mathbb{P}\left(\lambda_0^1 < s_1 - \sqrt{p_1}Y, \cdots, \lambda_0^M < s_M - \sqrt{p_M}Y \Big| Y\right)\right], \tag{4.1.16}$$

where dx^0 is the Lebesgue measure on $\Big\{x = (x_1, ..., x_M) \in \mathbb{R}^M : \sum_{l=1}^{M} \sqrt{p_l} x_l = 0\Big\}$. The right hand side of (4.1.16) is the distribution function of the sum of the mutually independent random vectors $\left(\lambda_0^1, \cdots, \lambda_0^M\right)$ and $(Z_1, \cdots, Z_M)$, where $(Z_1, \cdots, Z_M) \stackrel{d}{=} \left(\sqrt{p_1}, ..., \sqrt{p_M}\right) Z$ with $Z \sim N(0, 1)$. □

With the conclusions of Theorem 3.2.10 and Theorem 3.2.13, the asymptotic behavior of the maximal eigenvalues, within each block, of $\mathbf{X}^0 \in \mathcal{G}^0(p_1, ..., p_M)$ is well understood.

Corollary 4.1.9 *For $k = 1, ..., K$, let $\max\limits_{m_k < i \leq m_k + d_k} \lambda_0^i$ be the largest eigenvalue of the $d_k \times d_k$ block of $\mathbf{X}^0 \in \mathcal{G}^0(p_1, ..., p_M)$, then*

$$\lim_{d_k \to \infty} \frac{\max\limits_{m_k < i \leq m_k + d_k} \lambda_0^i}{\sqrt{d_k}} = 2,$$

with probability one, or in the mean.

Proof. By Proposition 4.1.4,

$$\max_{m_k < i \le m_k + d_k} \lambda_0^i = \max_{m_k < i \le m_k + d_k} \lambda^i - \sqrt{p^{(k)}} \sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l}.$$

Since $\max\limits_{m_k < i \le m_k + d_k} \lambda^i$ is the maximal eigenvalue of an element of the $d_k \times d_k$ GUE, with probability one or in the mean,

$$\lim_{d_k \to \infty} \frac{\max\limits_{m_k < i \le m_k + d_k} \lambda^i}{\sqrt{d_k}} = 2.$$

Moreover, $\sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l}$ is a centered Gaussian random variable with variance

$$Var\left(\sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l} \right) = \sum_{l=1}^{M} p_l = 1.$$

Hence, with probability one or in the mean,

$$\lim_{d_k \to \infty} \frac{\sqrt{p^{(k)}} \sum_{l=1}^{M} \sqrt{p_l} \mathbf{X}_{l,l}}{\sqrt{d_k}} = 0.$$

□

4.2 Inhomogeneous Random Words

Let $\mathcal{A}_M = \{\alpha_1, ..., \alpha_M\}$, $\alpha_1 < \cdots < \alpha_M$, be an M-letter ordered alphabet and let $X_1 X_2 \cdots X_N$ be a random word, where $X_1, X_2, \cdots, X_N$ are iid random variables with $\mathbb{P}(X_1 = \alpha_j) = p_j$, with $p_j > 0$, and $\sum_{j=1}^{M} p_j = 1$. Assume also there are $K = 1, ..., M$, distinct probabilities in $\{p_1, p_2, ..., p_M\}$, and reorder them as $p^{(1)} > \cdots > p^{(K)}$ in such a way that the multiplicity of each $p^{(k)}$ is d_k, for each $k = 1, ..., K$. In our notation, $K = 1$ corresponds to the uniform alphabet case, where $d_1 = M$. Let $m_1 = 0$ and for any $k = 2, ..., K$, let $m_k = \sum_{j=1}^{k-1} d_j$. Finally, let τ be a permutation of $\{1, ..., M\}$ corresponding to a non-increasing ordering of $p_1, p_2, ..., p_M$, i.e., $p_{\tau(1)} \ge \cdots \ge p_{\tau(M)}$.

Its, Tracy and Widom ([34], [35]) have obtained the limiting law of the length of the longest increasing subsequence of such a random word. To recall their result, let $(\lambda_1, ..., \lambda_M)$ be the eigenvalues of an element of $\mathcal{G}^0\left(p_{\tau(1)}, ..., p_{\tau(M)}\right)$, written in such a way that $(\lambda_1, ..., \lambda_M) = \left(\lambda_1^{d_1}, ..., \lambda_{d_1}^{d_1}, ..., \lambda_1^{d_K}, ..., \lambda_{d_K}^{d_K}\right)$, i.e., $\lambda_1^{d_k}, ..., \lambda_{d_k}^{d_k}$ are the eigenvalues of the kth block, $k = 1, ..., K$. Then (see [35]), the limiting law of the length of the longest increasing subsequence, properly centered and normalized, is the law of $\max_{1\le i\le d_1} \lambda_i^{d_1}$. A representation of this limiting law, as a Brownian functional is given in [29]. A multidimensional Brownian functional representation of the whole tableaux associated to a Markovian random word is further given in [30]. Below, we recover the convergence of the whole tableau, in the iid nonuniform case, through the approach we have used till now, which is related to the work of Baryshnikov [10], Gravner, Tracy and Widom [24], Doumerc [18] or Houdré and Litherland [30].

Let $\left(\hat{B}^1(t), \hat{B}^2(t), ..., \hat{B}^M(t)\right)$ be the M-dimensional Brownian motion having covariance matrix

$$\mathbf{\Sigma}_t := t \begin{pmatrix} p_{\tau(1)}\left(1-p_{\tau(1)}\right) & -p_{\tau(1)}p_{\tau(2)} & \cdots & -p_{\tau(1)}p_{\tau(M)} \\ -p_{\tau(2)}p_{\tau(1)} & p_{\tau(2)}\left(1-p_{\tau(2)}\right) & \cdots & -p_{\tau(2)}p_{\tau(M)} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{\tau(M)}p_{\tau(1)} & -p_{\tau(M)}p_{\tau(2)} & \cdots & p_{\tau(M)}\left(1-p_{\tau(M)}\right) \end{pmatrix}. \tag{4.2.1}$$

For each $l = 1, ..., M$, there is a unique $1 \le k \le K$ such that $p_{\tau(l)} = p^{(k)}$, and let

$$\hat{L}_M^l = \sum_{j=1}^{m_k} \hat{B}^{\tau(j)}(1) + \sup_{J(l-m_k,d_k)} \sum_{j=m_k+1}^{m_k+d_k} \sum_{i=1}^{k-m_k} \left(\hat{B}^{\tau(j)}(t_{j-i+1}^i) - \hat{B}^{\tau(j)}(t_{j-i}^i)\right), \tag{4.2.2}$$

where the set $J(l-m_k, d_k)$ consists of all the subdivisions (t_j^i) of $[0,1]$, $1 \le j \le l-m_k$, $0 \le i \le d_k$, of the form:

$$t_j^i \in [0,1];\ t_j^{i+1} \le t_j^i;\ t_j^i = 0\ for\ j \le 0;\ t_j^i = 1\ for\ j \ge M-k+1. \tag{4.2.3}$$

With these preliminaries, we have:

Theorem 4.2.1 *Let $\mathbf{X}_W$ be the matrix corresponding to a random word W of length N as in (3.2.4), with each letter independently drawn from an M-letter alphabet*

$\{\alpha_1 < ... < \alpha_M\}$ with $\mathbb{P}(X_i = \alpha_j) = p_j$, for each $i = 1, ..., N$, where $p_j > 0$ and $\sum_{j=1}^{M} p_j = 1$. Let τ be a permutation of $\{1, ..., M\}$ corresponding to a non-increasing ordering of $p_1, p_2, ..., p_M$. Assume that there are K distinct probabilities in $\{p_1, p_2, ..., p_M\}$, and reorder them as $p^{(1)} > \cdots > p^{(K)}$ in such a way that the multiplicity of each $p^{(k)}$ is d_k, for each $k = 1, ..., K$. Let $m_1 = 0$ and for any $k = 2, ..., K$, let $m_k = \sum_{j=1}^{k-1} d_j$, and so the multiplicity of each $p_{\tau(j)}$ is d_k if $m_k < \tau(j) \leq m_k + d_k$, $j = 1, ..., M$. Let $\lambda(RSK(\mathbf{X}_W)) = (\lambda_1, ..., \lambda_M)$ be the common shape of the associated Young tableaux through the RSK correspondence. Let $(\xi_1, ..., \xi_M)$ be the vector of eigenvalues of an element of $\mathcal{G}^0\left(p_{\tau(1)}, ..., p_{\tau(M)}\right)$, written in such a way that $\xi_{m_k+1} \geq \cdots \geq \xi_{m_k+d_k}$ for $k = 1, ..., K$. Then:

(i) As $N \to \infty$,

$$\left(\frac{\lambda_1 - Np_{\tau(1)}}{\sqrt{Np_{\tau(1)}}}, ..., \frac{\lambda_M - Np_{\tau(M)}}{\sqrt{Np_{\tau(M)}}}\right) \Longrightarrow \left(\hat{L}_M^1, \hat{L}_M^2 - \hat{L}_M^1, ..., \hat{L}_M^M - \hat{L}_M^{M-1}\right). \tag{4.2.4}$$

(ii)

$$\left(\hat{L}_M^1, \hat{L}_M^2 - \hat{L}_M^1, ..., \hat{L}_M^M - \hat{L}_M^{M-1}\right) \stackrel{d}{=} (\xi_1, ..., \xi_M). \tag{4.2.5}$$

Proof. To prove (i), let $(\mathbf{e_j})_{j=1,...,M}$ be the canonical basis of $\mathbb{R}^M$, and let $\mathbf{V} = (V_1, ..., V_M)$ be the random vector such that

$$\mathbb{P}(\mathbf{V} = \mathbf{e_j}) = p_j, \quad j = 1, ..., M.$$

Clearly, for each $1 \leq p \leq M$,

$$\mathbb{E}(V_j) = p_j, \ Var(V_j) = p_j(1 - p_j),$$

and for $j_1 \neq j_2$, $Cov(V_{j_1}, V_{j_2}) = -p_{j_1}p_{j_2}$. Hence the covariance matrix of $\mathbf{V}$ is

$$\mathbf{\Sigma} = \begin{pmatrix} p_j(1-p_j) & -p_1p_2 & \cdots & -p_1p_M \\ -p_2p_1 & p_2(1-p_2) & \cdots & -p_2p_M \\ \vdots & \vdots & \ddots & \vdots \\ -p_Mp_1 & -p_Mp_2 & \cdots & p_M(1-p_M) \end{pmatrix}. \tag{4.2.6}$$

Let $\mathbf{V_1}, \mathbf{V_2}, ..., \mathbf{V_N}$ be independent copies of $\mathbf{V}$, where $\mathbf{V_i} = (V_{i,1}, V_{i,2}, ..., V_{i,M})$, $i = 1, ..., N$. Then $\mathbf{X}_W$ has the same law as the matrix formed by all the $V_{i,j}$ on the lattice $\{1, ..., N\} \times \{1, ..., M\}$. As in the proof of Theorem 3.2.7, for all $1 \leq l \leq M$,

$$\lambda_1 + \cdots + \lambda_l = G^l(M,N) := \sup \left\{ \sum_{(i,j) \in \pi_1 \cup \cdots \cup \pi_l} V_{i,j} : \pi_1, ..., \pi_l \in \mathcal{P}(M, N), \right.$$

$$\left. \textit{and } \pi_1, ..., \pi_l \textit{ are all disjoint} \right\}, \tag{4.2.7}$$

where, by disjoint, it is meant that any two paths do not share a common point (i, j) in the rectangle $\{1, ..., N\} \times \{1, ..., M\}$. We prove next that, for any $l = 1, ..., M$,

$$\frac{G^l(M, N) - N s_l}{\sqrt{N}} \overset{N \to \infty}{\Longrightarrow} \hat{L}^l_M, \tag{4.2.8}$$

where $s_l = \sum_{j=1}^{l} p_{\tau(j)}$. For $l = 1$,

$$G^1(M, N) = \max \left\{ \sum_{(i,j) \in \pi} V_{i,j} \ ; \pi \in \mathcal{P}(M, N) \right\}. \tag{4.2.9}$$

Moreover, each path π is uniquely determined by the weakly increasing sequence of its $M - 1$ jumps, namely $0 = t_0 \leq t_1 \leq \cdots \leq t_{M-1} \leq 1$, such that π is horizontal on $[\lfloor t_{j-1} N \rfloor, \lfloor t_j N \rfloor] \times \{j\}$ and vertical on $\{\lfloor t_j N \rfloor\} \times [j, j+1]$. Hence

$$G^1(M, N) = \sup_{0 = t_0 \leq t_1 \leq \cdots \leq t_{M-1} \leq t_M = 1} \sum_{j=1}^{M} \sum_{i = \lfloor t_{j-1} N \rfloor}^{\lfloor t_j N \rfloor} V_{i,j}.$$

Let $p_{max} = \max_{1 \leq j \leq M} p_j$, $J(M) = \{j : p_j = p_{max}\} \subset \{1, ..., M\}$ and so $d_1 = card\,(J(M))$ (the α_j, where $j \in J(M)$, correspond to the most probable letters). As shown in [30, Section 5], the distribution of $G^1(M, N)$ is very close, for large N, to that of a very similar expression which involves only those $V_{i,j}$ for which $j \in J(M)$. To recall this result, if

$$\hat{G}^1(M, N) = \sup_{\substack{0 = t_0 \leq t_1 \leq \cdots \leq t_{M-1} \leq t_M = 1 \\ t_{j-1} = t_j \ for \ j \notin J(M)}} \sum_{j=1}^{M} \sum_{i = \lfloor t_{j-1} N \rfloor}^{\lfloor t_j N \rfloor} V_{i,j},$$

then, as $N \to \infty$,

$$\frac{G^1(M, N)}{\sqrt{N}} - \frac{\hat{G}^1(M, N)}{\sqrt{N}} \xrightarrow{\mathbb{P}} 0, \tag{4.2.10}$$

i.e., as $N \to \infty$, the distribution of the maximum (over all the northeast paths) in (4.2.9) is approximately the distribution of the maximum over the northeast paths going eastbound only along the rows corresponding to the most probable letters. Now,

$$\frac{\hat{G}^1(M,N) - Np_{max}}{\sqrt{N}} = \sup_{\substack{0=t_0 \le t_1 \le \cdots \\ \le t_{M-1} \le t_M = 1 \\ t_{j-1}=t_j \ for \ j \notin J(M)}} \sum_{j=1}^{M} \frac{\sum_{i=\lfloor t_{j-1}N \rfloor}^{\lfloor t_j N \rfloor} V_{i,j} - (t_j - t_{j-1})Np_{max}}{\sqrt{N}}, \tag{4.2.11}$$

and notice that the random vectors $\mathbf{V_1}, \mathbf{V_2}, ..., \mathbf{V_N}$ are independent. As in the proof of Theorem 3.2.7, we have, as $N \to \infty$, for any $t > 0$,

$$\left(\frac{\sum_{i=1}^{\lfloor tN \rfloor} V_{i,j} - tNp_{max}}{\sqrt{N}} \right)_{1 \le j \le M, \ j \in J(M)} \Longrightarrow \left(\hat{B}^j(t) \right)_{1 \le j \le M, \ j \in J(M)},$$

where $\left(\hat{B}^j(t) \right)_{1 \le j \le M, \ j \in J(M)}$ is a d_1-dimensional Brownian motion with $d_1 \times d_1$ covariance matrix

$$t \begin{pmatrix} p_{max}(1-p_{max}) & -p_{max}^2 & \cdots & -p_{max}^2 \\ -p_{max}^2 & p_{max}(1-p_{max}) & \cdots & -p_{max}^2 \\ \vdots & \vdots & \ddots & \vdots \\ -p_{max}^2 & -p_{max}^2 & \cdots & p_{max}(1-p_{max}) \end{pmatrix}. \tag{4.2.12}$$

By the continuous mapping theorem,

$$\frac{\hat{G}^1(M,N) - Np_{max}}{\sqrt{N}} \overset{N\to\infty}{\Longrightarrow} \sup_{J(1,d_1)} \sum_{j=1}^{d_1} \left(\hat{B}^{\tau(j)}(t_j) - \hat{B}^{\tau(j)}(t_{j-1}) \right), \tag{4.2.13}$$

and the right hand side of (4.2.13) is exactly $\hat{L}_M^1$, then (4.2.10), leads to

$$\frac{G^1(M,N) - Np_{max}}{\sqrt{N}} \overset{N\to\infty}{\Longrightarrow} \hat{L}_M^1. \tag{4.2.14}$$

Now, for $l \ge 2$, $G^l(M,N)$ is the maximum, of the sums of the $V_{i,j}$, over l disjoint paths. Still by the argument in [30], as $N \to \infty$, $\left(G^l(M,N) - \hat{G}^l(M,N) \right)/\sqrt{N} \xrightarrow{\mathbb{P}} 0$, where $\hat{G}^l(M,N)$ is the maximal sums of the $V_{i,j}$ over l disjoint paths we now describe. Let $1 \le k \le K$ be the unique integer such that $p_{\tau(l)} = p^{(k)}$. Denote by $\alpha_{j(1)}, ..., \alpha_{j(m_k)}$ the letters corresponding to the m_k probabilities that are strictly larger than $p_{\tau(l)}$. For each $1 \le s \le m_k$, the horizontal path from $(1, j(s))$ to $(N, j(s))$ is included, and thus so are these m_k paths. The remaining $l - m_k$ disjoint paths only go eastbound along the rows corresponding to the d_k letters having probability $p_{\tau(l)}$. The set of these $l - m_k$ paths is in a one to one correspondence with the set of subdivisions of $[0, 1]$ given in (4.2.3). Therefore

$$\hat{G}^l(M, N) = \sum_{j=1}^{m_k} \sum_{i=1}^{N} V_{i,\tau(j)} + \sup_{J(l-m_k,d_k)} \sum_{j=m_k+1}^{m_k+d_k} \sum_{i=1}^{k-m_k} \sum_{\lfloor t^i_{j-i} N \rfloor}^{\lfloor t^i_{j-i+1} N \rfloor} \left(V_{\lfloor t^i_{j-i+1} N \rfloor, \tau(j)} - V_{\lfloor t^i_{j-i} N \rfloor, \tau(j)} \right). \quad (4.2.15)$$

Now,

$$\frac{\hat{G}^l(M, N) - N s_l}{\sqrt{N}} = \sum_{j=1}^{m_k} \frac{\sum_{i=1}^{N} V_{i,\tau(j)} - N p_{\tau(j)}}{\sqrt{N}} + \sup_{J(l-m_k,d_k)} \sum_{j=m_k+1}^{m_k+d_k} \sum_{i=1}^{k-m_k} \frac{\sum_{\lfloor t^i_{j-i} N \rfloor}^{\lfloor t^i_{j-i+1} N \rfloor} \left(V_{\lfloor t^i_{j-i+1} N \rfloor, \tau(j)} - V_{\lfloor t^i_{j-i} N \rfloor, \tau(j)} \right) - N p^{(k)}}{\sqrt{N}}. \quad (4.2.16)$$

Since the column vectors $\mathbf{V_1'}, \mathbf{V_2'}, ..., \mathbf{V_N'}$ are iid, a multivariate Donsker invariance principle argument, as in the proof of Theorem 3.2.7, shows that, as $N \to \infty$, for any $t > 0$,

$$\left(\frac{\sum_{i=1}^{\lfloor tN \rfloor} V_{i,\tau(j)} - tN p_{\tau(j)}}{\sqrt{N}} \right)_{1 \le j \le M} \Longrightarrow \left(\hat{B}^j(t) \right)_{1 \le j \le M},$$

where $\left(\hat{B}^j(t) \right)_{1 \le j \le M}$ is an M-dimensional Brownian motion with covariance matrix given by (4.2.1). The continuous mapping theorem, Slutsky's lemma, together with (4.2.16), allow us to conclude that,

$$\frac{G^l(M, N) - N s_l}{\sqrt{N}} \overset{N \to \infty}{\Longrightarrow} \hat{L}^l_M.$$

Finally, by the Cramér-Wold theorem, as $N \to \infty$,

$$\left(\frac{\lambda_1 - Ns_1}{\sqrt{N}}, \frac{\sum_{j=1}^{2} \lambda_j - Ns_2}{\sqrt{N}}, ..., \frac{\sum_{j=1}^{M} \lambda_j - Ns_M}{\sqrt{N}} \right) \Longrightarrow \left(\hat{L}_M^1, \hat{L}_M^2, ..., \hat{L}_M^M \right), \quad (4.2.17)$$

therefore, as $N \to \infty$, by the continuous mapping theorem,

$$\begin{aligned} &\left(\frac{\lambda_1 - Np_{\tau(1)}}{\sqrt{N}}, \frac{\lambda_2 - Np_{\tau(2)}}{\sqrt{N}}, ..., \frac{\lambda_M - Np_{\tau(M)}}{\sqrt{N}} \right) \\ &= \left(\frac{G^1 - Ns_1}{\sqrt{N}}, \frac{(G^2 - Ns_2) - (G^1 - Ns_1)}{\sqrt{N}}, ..., \frac{\left(G^M - Ns_M\right) - \left(G^{M-1} - Ns_{M-1}\right)}{\sqrt{N}} \right) \\ &\Longrightarrow \left(\hat{L}_M^1, \hat{L}_M^2 - \hat{L}_M^1, ..., \hat{L}_M^M - \hat{L}_M^{M-1} \right). \end{aligned} \quad (4.2.18)$$

Part (i) is thus proved.

To prove (ii), recall that in [34] Its, Tracy and Widom have obtained the induced probability measure on the shape of the associated Young tableaux and also the limiting asymptotics for the non-uniform finite alphabets. Namely, for any partition $\lambda^0 = (\lambda_1^0, \lambda_2^0, ..., \lambda_M^0)$ of N,

$$\mathbb{P}(\lambda(RSK(\mathbf{X}_W)) = \lambda^0) = s_{\lambda^0}(p) f^{\lambda^0},$$

where f^{λ^0} is again the number of Young tableaux of shape λ^0 with elements in $\{1, ..., N\}$:

$$f^{\lambda^0} = N! \prod_{1 \le i < j \le M} (\lambda_i^0 - \lambda_j^0 + j - i) \prod_{j=1}^{M} \frac{1}{(\lambda_j^0 + M - j)!},$$

and where $s_{\lambda^0}(p)$ is the Schur function of shape λ^0 in the variable $p = \left(p_{\tau(1)}, ..., p_{\tau(M)}\right)$ which we describe next. Let $\mathcal{A}_1, ..., \mathcal{A}_K$ be the decomposition of $\{1, ..., M\}$ such that $p_{\tau(i)} = p_{\tau(j)} = p^{(k)}$ if and only if $i, j \in \mathcal{A}_k$, for some $1 \le k \le K$. Clearly, $d_k = card\,(\mathcal{A}_k)$. Then,

$$s_{\lambda^0}(p) = \frac{\sum_{\sigma \in \mathcal{S}_M} (-1)^\sigma \prod_{k=1}^{K} \prod_{i \in \mathcal{A}_k} \left(p_{\tau(i)}^{M-\sigma(i)-m_k-d_k+\tau(i)} h_{\sigma(i)}^{m_k+d_k-\tau(i)} \right)}{\prod_{k=1}^{K} (0!1! \cdots (d_k-1)!) \prod_{k<l} \left(p^{(k)} - p^{(l)}\right)^{d_k d_l}}, \quad (4.2.19)$$

where $\mathcal{S}_M$ is the set of all the permutations of $\{1, ..., M\}$ and where $h_i = \lambda_i^0 + M - i$ for $i = 1, ..., M$. Next, for $i = 1, ..., M$, let

$$x_i = \frac{\lambda_i^0 - Np_{\tau(i)}}{\sqrt{Np_{\tau(i)}}},$$

then, as $N \to \infty$,

$$N! \prod_{j=1}^{M} \frac{1}{(\lambda_j^0 + M - j)!}$$
$$\sim (2\pi)^{-(M-1)/2} N^{-M(M-1)/2} \left(\prod_{i=1}^{M-1} p_{\tau(i)}^{\tau(i)-M} \right) e^{-\sum_{i=1}^{M} x_i^2/2} \mathbf{1}_{\left\{ \sum_{i=1}^{M} \sqrt{p_{\sigma^{-1}(i)}} x_i = 0 \right\}}, \quad (4.2.20)$$

and

$$\prod_{1 \le i < j \le M} (\lambda_i^0 - \lambda_j^0 + j - i)$$
$$\sim N^{M(M-1)/2 - \sum_{k=1}^{K} d_k(d_k-1)/4} \prod_{k=1}^{K} \left(\left(p^{(k)}\right)^{d_k(d_k-1)/4} \Delta_k(x) \right) \prod_{k<l} \left(p^{(k)} - p^{(l)}\right)^{d_k d_l}. \quad (4.2.21)$$

Together with

$$\sum_{\sigma \in \mathcal{S}_M} (-1)^\sigma \prod_{k=1}^{K} \prod_{i \in \mathcal{A}_k} \left(p_{\tau(i)}^{M-\sigma(i)-m_k-d_k+\tau(i)} h_{\sigma(i)}^{m_k+d_k-\tau(i)} \right)$$
$$\sim \prod_{i=1}^{M} p_{\tau(i)}^{M-\tau(i)} \prod_{k=1}^{K} \left(p^{(k)}\right)^{-d_k(d_k-1)/2} N^{\sum_{k=1}^{K} d_k(d_k-1)/4} \prod_{k=1}^{K} \left(\left(p^{(k)}\right)^{d_k(d_k-1)/4} \Delta_k(x) \right), \quad (4.2.22)$$

the limiting density of $\left(\left(\lambda_1 - Np_{\tau(1)}\right)/\sqrt{Np_{\tau(1)}}, ..., \left(\lambda_M - Np_{\tau(M)}\right)/\sqrt{Np_{\tau(M)}} \right)$, as $N \to \infty$, is f given in (4.1.10). This is the joint density of the eigenvalues of an element of $\mathcal{G}^0\left(p_{\tau(1)}, ..., p_{\tau(M)}\right)$. Then, as in the proof of Theorem 3.2.6, for any $s_1, ..., s_M \in \mathbb{R}$,

$$\mathbb{P}\left(\frac{\lambda_1 - Np_{\tau(1)}}{\sqrt{Np_{\tau(1)}}} \le s_1, ..., \frac{\lambda_M - Np_{\tau(M)}}{\sqrt{Np_{\tau(M)}}} \le s_M \right)$$
$$= \mathbb{E}\left[\mathbf{1}_{\frac{\lambda_1 - Np_{\tau(1)}}{\sqrt{Np_{\tau(1)}}} \le s_1, ..., \frac{\lambda_M - Np_{\tau(M)}}{\sqrt{Np_{\tau(M)}}} \le s_M} \right]$$
$$= \sum_{\frac{\lambda_1^0 - Np_{\tau(1)}}{\sqrt{Np_{\tau(1)}}} \le s_1, ..., \frac{\lambda_M^0 - Np_{\tau(M)}}{\sqrt{Np_{\tau(M)}}} \le s_M} \mathbb{P}(\lambda(RSK(\mathbf{X}_W)) = \left(\lambda_1^0, ..., \lambda_M^0\right)) \quad (4.2.23)$$

$$\overset{N\to\infty}{\longrightarrow} \int\limits_{\mathcal{L}^p_{(s_1,\dots,s_M)}} f(x)dx_1\cdots dx_{M-1}$$

$$= \mathbb{P}\left(\xi_1 \le s_1, ..., \xi_M \le s_M\right). \quad (4.2.24)$$

Part (ii) of the theorem then follows from part (i). □

4.3 Poissonized Word Problem

"Poissonization" is another useful tool in dealing with length asymptotics for longest increasing subsequence problems. It was introduced by Hammersley in [26] in order to show the existence of $\lim_{N\to\infty} \mathbb{E}\left(LI_N\right)/\sqrt{N}$, for a random permutation of $\{1, 2, ..., N\}$. Since then, this technique has been widely used, and we intend, below, to use it in connection with the inhomogeneous word problem.

Let $\mathbf{X}_W$ be the matrix corresponding to a random word W of length N from an M letter alphabet as in (3.2.4), and let $\lambda(RSK(\mathbf{X}_W)) = (\lambda_1, ..., \lambda_M)$ be the common shape of the associated Young tableaux through the RSK correspondence. Johansson [39] studied the Poissonized measure on the set of shapes of Young tableaux associated to the homogeneous random word, i.e., each letter is independently and uniformly drawn from the alphabet. Moreover, Its Tracy and Widom [35] analyzed the Poissonization of LI_N for an inhomogeneous random word, and showed that the Poissonized distribution of the length of the longest increasing subsequence, as a function of $p_1, ..., p_M$, can be identified as the solution of a certain integrable system of nonlinear PDEs. Below, we show that the Poissonized distribution of the shape of the whole Young tableaux associated to an inhomogeneous random word converges to the spectrum of the corresponding direct sum of GUEs. Next, using this result, together with "de-Poissonization", we obtain the asymptotic behavior of the shape of the tableaux.

Let $W = X_1 X_2 \cdots X_N$ be a random word of length N, with each letter independently drawn from $\mathcal{A}_M = \{\alpha_1 < ... < \alpha_M\}$ with $\mathbb{P}_M (X_i = \alpha_j) = p_j$, $i = 1, ..., N$, where $p_j > 0$ and $\sum_{j=1}^{M} p_j = 1$, i.e., the random word is distributed according to $\mathbb{P}_{W,M,N} = \mathbb{P}_M \times \cdots \times \mathbb{P}_M$ on the set of words $[M]^N$. Using the terminology of [39], with $\mathbb{N} = \{0, 1, 2, \cdots\}$, let

$$\mathcal{P}_M^{(N)} := \left\{ \lambda = (\lambda_1, ..., \lambda_M) \in \mathbb{N}^M : \lambda_1 \geq \cdots \geq \lambda_M, \ \sum_{i=1}^{M} \lambda_i = N \right\},$$

denote the set of partitions of N, of length at most M. The RSK correspondence defines a bijection from $[M]^N$ to the set of pairs of Young tableaux (P, Q) of common shape $\lambda \in \mathcal{P}_M^{(N)}$, where P is semi-standard with elements in $\{1, ..., M\}$ and Q is standard with elements in $\{1, ..., N\}$.

Definition 4.3.1 *Let T be a function from a probability space $(A, \mathbb{P}_A)$ to a set B, the image of $\mathbb{P}_A$ by T (also known as the push forward of $\mathbb{P}_A$ by T) is the probability distribution on $T(A) \subset B$ given by $\mathbb{P}_A (T^{-1}(C))$, for any $C \subset T(A)$.*

For any $W \in [M]^N$, let $S(W)$ be the common shape of the Young tableaux associated to W by the RSK correspondence. Then S is a mapping from $[M]^N$ to $\mathcal{P}_M^{(N)}$, which, moreover, is a surjection. The push-forward of $\mathbb{P}_{W,M,N}$ by S is the measure $\mathbb{P}_{M,N}$ given, for any $\lambda_0 \in \mathcal{P}_M^{(N)}$, by

$$\mathbb{P}_{M,N} (\lambda_0) := \mathbb{P}_{W,M,N} (\lambda (RSK(\mathbf{X}_W)) = \lambda_0).$$

Next, let

$$\mathcal{P}_M := \left\{ \lambda = (\lambda_1, ..., \lambda_M) \in \mathbb{N}^M : \lambda_1 \geq \cdots \geq \lambda_M \right\},$$

be the set of partitions, of elements of $\mathbb{N}$, of length at most M. The set $\mathcal{P}_M$ consists of the shapes of the Young tableaux associated to the random words of any finite length made up from the M letter alphabet $\mathcal{A}_M$.

For $\alpha > 0$, the Poissonized measure of $\mathbb{P}_{M,N}$ on the set $\mathcal{P}_M$ is then defined as

$$\mathbb{P}_M^{\alpha} (\lambda_0) := e^{-\alpha} \sum_{N=0}^{\infty} \mathbb{P}_{M,N} (\lambda_0) \frac{\alpha^N}{N!}. \tag{4.3.1}$$

The Poissonized measure $\mathbb{P}_M^\alpha$ coincides with the distribution of the shape of the Young tableaux associated to a random word, taking its values in the alphabet $\mathcal{A}_M$ and, whose length is a Poisson random variable with mean α. Such a random word is called Poissonized, and LI_α denote the length of its longest increasing subsequence.

The Charlier ensemble is closely related to the Poissonized word problem. It is used by Johansson [39] to investigate the asymptotics of LI_N for finite uniform alphabets. For the non-uniform alphabets we consider, let us define the generalized Charlier ensemble on $\mathcal{P}_M$.

Definition 4.3.2 *For any $\alpha > 0$, the generalized Charlier ensemble on $\mathcal{P}_M$ is*

$$\mathbb{P}_{Ch,M}^\alpha\left(\lambda^0\right) = \prod_{1\le i<j\le M} (\lambda_i^0 - \lambda_j^0 + j - i) \prod_{j=1}^{M} \frac{1}{(\lambda_j^0 + M - j)!} s_{\lambda^0}(p) e^{-\alpha} \prod_{i=1}^{M} \alpha^{\lambda_i^0}, \quad (4.3.2)$$

for all $\lambda^0 = (\lambda_1^0, \lambda_2^0, ..., \lambda_M^0) \in \mathcal{P}_M$, and where $s_{\lambda^0}(p)$ is the Schur function of shape λ^0 given in (4.2.19).

The next theorem gives, for inhomogeneous random words, both $\mathbb{P}_{M,N}(\lambda_0)$ and the distribution of LI_α. The first statement is due to Its, Tracy and Widom ([34], [35]), while the second follows directly from the fact that the length of the longest increasing subsequence is equal to the length of the first row of the corresponding Young tableau.

Theorem 4.3.3 *(i) The image of $\mathbb{P}_{W,M,N}$ on the set $[M]^N$ by the mapping $S : [M]^N \to \mathcal{P}_M^{(N)}$ is, for any $\lambda^0 = (\lambda_1^0, \lambda_2^0, ..., \lambda_M^0) \in \mathcal{P}_M^{(N)}$, given by*

$$\mathbb{P}_{M,N}(\lambda^0) = s_{\lambda^0}(p) f^{\lambda^0}, \quad (4.3.3)$$

where

$$f^{\lambda^0} = N! \prod_{1\le i<j\le M} (\lambda_i^0 - \lambda_j^0 + j - i) \prod_{j=1}^{M} \frac{1}{(\lambda_j^0 + M - j)!},$$

and where $s_{\lambda^0}(p)$ *is the Schur function of shape* λ^0 *in the variable* $p = \big(p_{\tau(1)}, ..., p_{\tau(M)}\big)$ *given in (4.2.19), with* τ *a permutation of* $\{1, ..., M\}$ *corresponding to a non-increasing ordering of* $p_1, p_2, ..., p_M$*.*

(ii) The Poissonization of $\mathbb{P}_{M,N}$ *is the generalized Charlier ensemble* $\mathbb{P}^{\alpha}_{Ch,M}$ *defined in (4.3.2). In particular, for the Poissonized word problem,*

$$\mathbb{P}^{\alpha}_{W,M}\left(LI_\alpha \le t\right) := e^{-\alpha} \sum_{N=0}^{\infty} \mathbb{P}_{M,N}\left(\lambda_1 \le t\right) \frac{\alpha^N}{N!} = \mathbb{P}^{\alpha}_{Ch,M}\left(\lambda_1 \le t\right). \tag{4.3.4}$$

For uniform alphabet, Johansson [39] showed the convergence, as $\alpha \to \infty$, of the Poissonized measure on $\mathcal{P}_M$ to the joint law of the ordered eigenvalues of the GUE. Next, following his lead and techniques, we generalize this result to the nonuniform case, where the convergence is towards the joint law of the eigenvalues $(\xi_1, ..., \xi_M)$, ordered within each block, of an element of $\mathcal{G}_M(d_1, ..., d_K)$. The density of $(\xi_1, ..., \xi_M)$ is, for any $x \in \mathbb{R}^M$, given by

$$f_{\xi_1,...,\xi_M}(x) = \frac{1}{\sqrt{2\pi}} c_M \prod_{k=1}^{K} \Delta_k(x)^2 e^{-\sum_{i=1}^{M} x_i^2/2}, \tag{4.3.5}$$

where $c_M = (2\pi)^{-(M-1)/2} \prod_{k=1}^{K} \left(0!1!\cdots(d_k-1)!\right)^{-1}$, and where

$$\Delta_k(x) = \prod_{m_k \le i < j \le m_k + d_k} (x_i - x_j).$$

Theorem 4.3.4 *Let* $\mathbf{X}_W$ *be the matrix corresponding to a random word* W *of length* N *as in (3.2.4), with each letter independently drawn from an* M*-letter alphabet* $\mathcal{A}_M = \{\alpha_1 < ... < \alpha_M\}$*, with* $\mathbb{P}(X_i = \alpha_j) = p_j$*,* $i = 1, ..., N$*, where* $p_j > 0$ *and* $\sum_{j=1}^{M} p_j = 1$*. Let* τ *be a permutation of* $\{1, ..., M\}$ *corresponding to a non-increasing ordering of* $p_1, p_2, ..., p_M$*. Assume that there are* K *distinct probabilities in* $\{p_1, p_2, ..., p_M\}$*, and reorder them as* $p^{(1)} > \cdots > p^{(K)}$ *in such a way that the multiplicity of each* $p^{(k)}$ *is* d_k*, for each* $k = 1, ..., K$*. For any* $k = 2, ..., K$*, let* $m_k = \sum_{j=1}^{k-1} d_j$*,* $m_1 = 0$*, and* *so the multiplicity of each* $p_{\tau(j)}$ *is* d_k *if* $m_k < \tau(j) \le m_k + d_k$, $j = 1, ..., M$. *Let* $\lambda(RSK(\mathbf{X}_W)) = (\lambda_1, ..., \lambda_M)$ *be the common shape of the Young tableaux associated through the RSK correspondence. Let* $(\xi_1, ..., \xi_M)$ *be the eigenvalues of an element of* $\mathcal{G}_M(d_1, ..., d_K)$*, written in such a way that* $\xi_{m_k+1} \ge \cdots \ge \xi_{m_k+d_k}$ *for* $k = 1, ..., K$*, and let* $f_{\xi_1,...,\xi_M}$ *be its density given by (4.3.5). Then, for any continuous function* g *on* $\mathbb{R}^M$*,*

$$\lim_{\alpha\to\infty} \mathbb{E}_M^\alpha \left(g\left(\frac{\lambda_1 - \alpha p_{\tau(1)}}{\sqrt{\alpha p_{\tau(1)}}}, ..., \frac{\lambda_M - \alpha p_{\tau(M)}}{\sqrt{\alpha p_{\tau(M)}}} \right)\right) = \int_{\mathbb{R}^M} g(x) f_{\xi_1,...,\xi_M}(x)dx. \qquad (4.3.6)$$

Proof. By Theorem 4.3.3, for any partition $\lambda^0 = (\lambda_1^0, \lambda_2^0, ..., \lambda_M^0)$ of $N \in \mathbb{N}$,

$$\mathbb{P}_{M,N}(\lambda(RSK(\mathbf{X}_W)) = \lambda^0) = s_{\lambda^0}(p) f^{\lambda^0},$$

where

$$f^{\lambda^0} = N! \prod_{1\le i<j\le M} (\lambda_i^0 - \lambda_j^0 + j - i) \prod_{j=1}^M \frac{1}{(\lambda_j^0 + M - j)!},$$

and where $s_{\lambda^0}(p)$ is the Schur function of shape λ^0 in the variable $p = (p_{\tau(1)}, ..., p_{\tau(M)})$ as given in (4.2.19). Hence the Poissonized measure is

$$\mathbb{P}_M^\alpha(\lambda^0) = e^{-\alpha} \sum_{N=0}^\infty N! \prod_{1\le i<j\le M} (\lambda_i^0 - \lambda_j^0 + j - i) \prod_{j=1}^M \frac{1}{(\lambda_j^0 + M - j)!} s_{\lambda^0}(p) \frac{\alpha^N}{N!}.$$

Next, for $i = 1, ..., M$, let

$$x_i = \frac{\lambda_i^0 - \alpha p_{\tau(i)}}{\sqrt{\alpha p_{\tau(i)}}},$$

then, as $\alpha \to \infty$,

$$\prod_{j=1}^M \frac{1}{(\lambda_j^0 + M - j)!} \sim (2\pi)^{-M/2} \frac{e^\alpha}{\alpha^N} \alpha^{-M(M-1)/2} \left(\prod_{i=1}^M p_{\tau(i)}^{\tau(i)-M} \right) e^{-\sum_{i=1}^M x_i^2/2}, \qquad (4.3.7)$$

and

$$\prod_{1\le i<j\le M} (\lambda_i^0 - \lambda_j^0 + j - i)$$
$$\sim \alpha^{M(M-1)/2 - \sum_{k=1}^K d_k(d_k-1)/4} \prod_{k=1}^K \left(\left(p^{(k)}\right)^{d_k(d_k-1)/4} \Delta_k(x) \right) \prod_{k<l} \left(p^{(k)} - p^{(l)}\right)^{d_k d_l}. \quad (4.3.8)$$

Together with

$$\sum_{\sigma\in\mathcal{S}_M}(-1)^\sigma\prod_{k=1}^K\prod_{i\in\mathcal{A}_k}\left(p_{\tau(i)}^{M-\sigma(i)-m_k-d_k+\tau(i)}h_{\sigma(i)}^{m_k+d_k-\tau(i)}\right)$$
$$\sim\prod_{i=1}^M p_{\tau(i)}^{M-\tau(i)}\prod_{k=1}^K\left(p^{(k)}\right)^{-d_k(d_k-1)/2}\alpha^{\sum_{k=1}^K d_k(d_k-1)/4}\prod_{k=1}^K\left(\left(p^{(k)}\right)^{d_k(d_k-1)/4}\Delta_k(x)\right),$$
(4.3.9)

as $\alpha\to\infty$, the limiting density of $\left(\frac{\lambda_1-\alpha p_{\tau(1)}}{\sqrt{\alpha p_{\tau(1)}}},...,\frac{\lambda_M-\alpha p_{\tau(M)}}{\sqrt{\alpha p_{\tau(M)}}}\right)$ is

$$\sqrt{2\pi}c_M\prod_{k=1}^K\Delta_k(x)^2e^{-\sum_{i=1}^M x_i^2/2},\quad x=(x_1,...,x_M)\in\mathbb{R}^M,$$

which is just the joint density of the eigenvalues, ordered within each block, of an element of $\mathcal{G}_M(d_1,...,d_K)$. The statement then follows from a Riemann integral approximation argument completely analogous to the one used at the end of the proof of Theorem 3.2.6. □

The next result is concerned with "de-Poissonization", and again is the nonuniform version of a result of Johansson.

Proposition 4.3.5 *Let* $\alpha_N=N+3\sqrt{N\ln N}$ *and* $\beta_N=N-3\sqrt{N\ln N}$. *Then there is a constant* C *such that, for sufficiently large* N, *and for any* $0\le n_i\le N$, $i=1,...,M$,

$$\mathbb{P}_M^{\alpha_N}(\lambda_1\le n_1,...,\lambda_M\le n_M)-\frac{C}{N^2}\le\mathbb{P}_{M,N}(\lambda_1\le n_1,...,\lambda_M\le n_M)$$
$$\le\mathbb{P}_M^{\beta_N}(\lambda_1\le n_1,...,\lambda_M\le n_M)+\frac{C}{N^2}.\quad(4.3.10)$$

Proof. The proof is analogous to the proof of the corresponding uniform alphabet result, given in [39] (see also Lemma 4.7 in [13]). Following Johansson, let us first prove that $\mathbb{P}_{M,N}(\lambda_1\le n_1,...,\lambda_M\le n_M)$ is decreasing in N. Denote a random word in $[M]^N$ by $W^{(N)}=X_1X_2\cdots X_N$ and let $[M]^{N+1}(j)$ be the set of the random words $W^{(N+1)}=X_1X_2\cdots X_NX_{N+1}$ such that $X_{N+1}=j$, $j=1,...,M$. Each word $W^{(N+1)}$ in $[M]^{(N+1)}(j)$ is mapped into a word $F_j\left(W^{(N+1)}\right)\in[M]^N$, by deleting the last letter

X_{N+1}. Clearly, F_j is a bijection from $[M]^{N+1}(j)$ to $[M]^N$. Moreover adding the letter X_{N+1} can only increase a row. Therefore, for all $i = 1, ..., M$,

$$\lambda_i \left(F_j \left(W^{(N+1)}\right)\right) \leq \lambda_i \left(W^{(N+1)}\right).$$

Now, let $g\left(W^{(N)}\right) = 1$, if $\lambda_i \left(W^{(N)}\right) \leq n_i$, for all $i = 1, ..., M$, and let $g\left(W^{(N)}\right) = 0$, otherwise. We have

$$g\left(F_j \left(W^{((N+1))}\right)\right) \geq g\left(W^{(N+1)}\right).$$

Hence,

$$\begin{aligned}
\mathbb{P}_{M,N+1}(\lambda_1 \leq n_1, ..., \lambda_M \leq n_M) &= \sum_{W^{(N+1)} \in [M]^{N+1}} g\left(W^{(N+1)}\right) \mathbb{P}\left(W^{(N+1)}\right) \\
&= \sum_{j=1}^{M} \left(p_k \sum_{W^{(N+1)} \in [M]^{N+1}(j)} g\left(W^{(N+1)}\right) \mathbb{P}\left(W^{(N+1)}\right) \right) \\
&\leq \sum_{j=1}^{M} \left(p_k \sum_{W^{(N+1)} \in [M]^{N+1}(j)} g\left(F_j\left(W^{(N+1)}\right)\right) \mathbb{P}\left(W^{(N+1)}\right) \right) \\
&= \sum_{W^{(N)} \in [M]^N} g\left(W^{(N)}\right) \mathbb{P}\left(W^{(N)}\right) \sum_{j=1}^{M} p_j \\
&= \mathbb{P}_{M,N}\left(\lambda_1 \leq n_1, ..., \lambda_M \leq n_M\right). \qquad (4.3.11)
\end{aligned}$$

Next, let $h_N(\alpha) = \alpha^N e^{-\alpha}/N!$, then

$$\mathbb{P}_M^{\alpha}\left(\lambda_1 \leq n_1, ..., \lambda_M \leq n_M\right) = \sum_{N=0}^{\infty} h_N(\alpha) \mathbb{P}_{M,N}\left(\lambda_1 \leq n_1, ..., \lambda_M \leq n_M\right).$$

By Stirling's formula, for large N,

$$h_N(\alpha) \sim \frac{1}{\sqrt{2\pi N}} \exp\left\{-\alpha\varphi\left(\frac{N}{\alpha}\right)\right\} \leq C \exp\left\{-\alpha\varphi\left(\frac{N}{\alpha}\right)\right\}, \qquad (4.3.12)$$

where $\varphi(x) = x \ln x + 1 - x$. It is easy to check that

$$\varphi(x) \geq \begin{cases} (x-1)^2/4, & \text{if } 0 \leq x \leq 2; \\ x/10, & \text{if } x \geq 2. \end{cases} \qquad (4.3.13)$$

Hence,

$$h_N(\alpha) \leq \begin{cases} C\exp\left\{-\frac{(N-\alpha)^2}{4\alpha}\right\}, & \text{if } 0 \leq N \leq 2\alpha; \\ Ce^{-N/10}, & \text{if } N \geq 2\alpha, \end{cases} \tag{4.3.14}$$

for some absolute constant C, which might differ from line to line. Therefore,

$$\begin{aligned} \sum_{N\geq 2\alpha} h_N(\alpha) &\leq C \sum_{N\geq 2\alpha} e^{-N/10} \\ &= Ce^{-\alpha/5} \sum_{N\geq 0} e^{-N/10} \\ &\leq Ce^{-\alpha/5}, \end{aligned} \tag{4.3.15}$$

and

$$\begin{aligned} \sum_{N\leq \alpha-\sqrt{8\alpha\ln\alpha}} h_N(\alpha) &\leq C \sum_{N\leq \alpha-\sqrt{8\alpha\ln\alpha}} \exp\left\{-\frac{(N-\alpha)^2}{4\alpha}\right\} \\ &\leq \frac{C}{\alpha^2} \sum_{N\geq 0} \exp\left\{-\frac{N^2}{4\alpha} - \frac{\sqrt{2\ln\alpha}N}{\sqrt{\alpha}}\right\} \\ &\leq \frac{C}{\alpha^2}, \end{aligned} \tag{4.3.16}$$

with also

$$\begin{aligned} \sum_{\alpha+\sqrt{8\alpha\ln\alpha}\leq N\leq 2\alpha} h_N(\alpha) &\leq C \sum_{\alpha+\sqrt{8\alpha\ln\alpha}\leq N\leq 2\alpha} \exp\left\{-\frac{(N-\alpha)^2}{4\alpha}\right\} \\ &\leq \frac{C}{\alpha^2} \sum_{N\geq 0} \exp\left\{-\frac{N^2}{4\alpha} - \frac{\sqrt{2\ln\alpha}N}{\sqrt{\alpha}}\right\} \\ &\leq \frac{C}{\alpha^2}. \end{aligned} \tag{4.3.17}$$

Together with the fact that $0 \leq \mathbb{P}_{M,N}(\lambda_1 \leq n_1, ..., \lambda_M \leq n_M) \leq 1$, we have

$$\begin{aligned} \left|\mathbb{P}_M^\alpha(\lambda_1 \leq n_1, ..., \lambda_M \leq n_M) - \sum_{|N-\alpha|\leq\sqrt{8\alpha\ln\alpha}} h_N(\alpha)\mathbb{P}_{M,N}(\lambda_1 \leq n_1, ..., \lambda_M \leq n_M)\right| \\ \leq \frac{C}{\alpha^2}, \end{aligned} \tag{4.3.18}$$

for sufficiently large α and any $0 \le n_i \le N$, $i = 1, ..., M$. Next, and still following Johansson [39], let $\alpha_N = N + 3\sqrt{N \ln N}$, then

$$\alpha_N - \sqrt{8\alpha_N \ln \alpha_N} = N + 3\sqrt{N \ln N} - \sqrt{8\left(N + 3\sqrt{N \ln N}\right) \ln \left(N + 3\sqrt{N \ln N}\right)} \ge N,$$

for sufficiently large N. Also since $\mathbb{P}_{M,N}(\lambda_1 \le n_1, ..., \lambda_M \le n_M)$ is non-increasing in N, we have

$$\begin{aligned}
&\mathbb{P}_{M,N}(\lambda_1 \le n_1, ..., \lambda_M \le n_M) \\
&\qquad \ge \mathbb{P}_{M,\alpha_N - \sqrt{8\alpha_N \ln \alpha_N}}(\lambda_1 \le n_1, ..., \lambda_M \le n_M) \\
&\qquad \ge \sum_{|N-\alpha_N| \le \sqrt{8\alpha_N \ln \alpha_N}} h_N(\alpha_N) \mathbb{P}_{M,N}(\lambda_1 \le n_1, ..., \lambda_M \le n_M) \\
&\qquad \ge \mathbb{P}_M^{\alpha_N}(\lambda_1 \le n_1, ..., \lambda_M \le n_M) - \frac{C}{\alpha_N^2} \\
&\qquad \ge \mathbb{P}_M^{\alpha_N}(\lambda_1 \le n_1, ..., \lambda_M \le n_M) - \frac{C}{N^2}.
\end{aligned} \tag{4.3.19}$$

This gives the left inequality in (4.3.10). For the right inequality there, let $\beta_N = N - 3\sqrt{N \ln N}$, then

$$\beta_N + \sqrt{8\beta_N \ln \beta_N} = N - 3\sqrt{N \ln N} + \sqrt{8\left(N - 3\sqrt{N \ln N}\right) \ln \left(N - 3\sqrt{N \ln N}\right)} \le N,$$

for sufficiently large N. Again since $\mathbb{P}_{M,N}(\lambda_1 \le n_1, ..., \lambda_M \le n_M)$ is non-increasing in N, we have

$$\begin{aligned}
&\mathbb{P}_{M,N}\Big(\lambda_1 \le n_1, ..., \lambda_M \le n_M\Big) \\
&\qquad \le \mathbb{P}_{M,\beta_N + \sqrt{8\beta_N \ln \beta_N}}(\lambda_1 \le n_1, ..., \lambda_M \le n_M) \\
&\qquad \le \sum_{|N-\beta_N| \le \sqrt{8\beta_N \ln \beta_N}} h_N(\beta_N) \mathbb{P}_{M,N}(\lambda_1 \le n_1, ..., \lambda_M \le n_M) \\
&\qquad\qquad + \sum_{|N-\beta_N| \ge \sqrt{8\beta_N \ln \beta_N}} h_N(\beta_N) \\
&\qquad \le \mathbb{P}_M^{\beta_N}(\lambda_1 \le n_1, ..., \lambda_M \le n_M) + \frac{C}{\beta_N^2} \\
&\qquad \le \mathbb{P}_M^{\beta_N}(\lambda_1 \le n_1, ..., \lambda_M \le n_M) + \frac{C}{N^2},
\end{aligned} \tag{4.3.20}$$

where in the last inequality, we have used the fact that $\beta_N = N - 3\sqrt{N \ln N} \geq N/2$, for sufficiently large N. $\square$

We are now ready to obtain asymptotics for the shape of the Young tableaux associated to a random word $W \in [M]^N$, when M and N go to infinity. Before stating our result, let us recall the well known, large M, asymptotic behavior of the spectrum of the $M \times M$ GUE ([61], [63], [39]). Let $\mathbb{P}_{GUE,M}$ denote the distribution of the $M \times M$ GUE given in (3.1.1).

Theorem 4.3.6 (Tracy-Widom) *Let $\lambda^j_{GUE,M}$ be the jth largest eigenvalue of an element of the $M \times M$ GUE. For each $r \geq 1$, there is a distribution function F_r on $\mathbb{R}^r$, such that,*

$$\lim_{M\to\infty} \mathbb{P}_{GUE,M}\left(\lambda^j_{GUE,M} \leq 2\sqrt{M} + t_j/M^{1/6}, j = 1, ..., r\right) = F_r(t_1, ..., t_r),$$

for $(t_1, ..., t_r) \in \mathbb{R}^r$.

Remark 4.3.7 *The multivariate distribution function F_r originates in [61] and [63], another expression for it is also given in [39] (see (3.48) there). For each $r = 1, 2, \cdots$, the first marginal of F_r is the Tracy-Widom distribution F_{TW}.*

Again, our next theorem is already present, for uniform alphabets, in Johansson [39].

Theorem 4.3.8 *For each $r \geq 1$, let $F_r(t_1, ..., t_r)$ on $\mathbb{R}^r$ be the distribution function obtained in Theorem 4.3.6. Assume that $d_1 \to +\infty$, as $M \to +\infty$. Then, for all $(t_1, ..., t_r) \in \mathbb{R}^r$,*

$$\lim_{M\to\infty}\lim_{\alpha\to\infty} \mathbb{P}^\alpha_M\left(\lambda_j \leq \alpha p_{max} + 2\sqrt{d_1 \alpha p_{max}} + t_j d_1^{-1/6}\sqrt{\alpha p_{max}}, j = 1, ..., r\right)$$
$$= F_r(t_1, ..., t_r), \qquad (4.3.21)$$

and,

$$\lim_{M\to\infty}\lim_{N\to\infty}\mathbb{P}_{M,N}\Big(\lambda_j \le Np_{max} + 2\sqrt{d_1 Np_{max}} + t_j d_1^{-1/6}\sqrt{Np_{max}}, j=1,...,r\Big)$$
$$= F_r(t_1,...,t_r). \tag{4.3.22}$$

In particular, for any $t \in \mathbb{R}$,

$$\lim_{M\to\infty}\lim_{N\to\infty}\mathbb{P}_{W,M,N}\Big(LI_N \le Np_{max} + 2\sqrt{d_1 Np_{max}} + td_1^{-1/6}\sqrt{Np_{max}}\Big) = F_{TW}(t). \tag{4.3.23}$$

Proof. By Theorem 4.3.4, for each $r \ge 1$, and for all $(s_1,...,s_r) \in \mathbb{R}^r$,

$$\lim_{\alpha\to\infty}\mathbb{P}^{\alpha}_{W,M}\left(\frac{\lambda_j - \alpha p_{max}}{\sqrt{\alpha p_{max}}} \le s_j,\ j=1,...,r\right) = \mathbb{P}_{GUE,d_1}\left(\xi_j \le s_j,\ j=1,...,r\right), \tag{4.3.24}$$

where ξ_j is the jth largest eigenvalue of the $d_1 \times d_1$ GUE. Hence, for any $(t_1,...,t_r) \in \mathbb{R}^r$,

$$\begin{aligned}
&\lim_{\alpha\to\infty}\mathbb{P}^{\alpha}_{M}\left(\lambda_j \le \alpha p_{max} + 2\sqrt{d_1\alpha p_{max}} + t_j d_1^{-1/6}\sqrt{\alpha p_{max}}, j=1,...,r\right)\\
&= \lim_{\alpha\to\infty}\mathbb{P}^{\alpha}_{M}\left(\frac{\lambda_j - \alpha p_{max}}{\sqrt{\alpha p_{max}}} \le 2\sqrt{d_1} + t_j d_1^{-1/6}, j=1,...,r\right)\\
&= \mathbb{P}\left(\xi_j \le 2\sqrt{d_1} + t_j d_1^{-1/6}, j=1,...,r\right).
\end{aligned} \tag{4.3.25}$$

As $d_1 \to \infty$, Theorem 4.3.6 gives the first conclusion, proving (4.3.21). Next, by Proposition 4.3.5, with $\alpha_N = N + 3\sqrt{N\ln N}$ and $\beta_N = N - 3\sqrt{N\ln N}$, there is a constant C such that, for sufficiently large N, and for any $0 \le s_j \le N$, $j=1,...,r$,

$$\begin{aligned}
\mathbb{P}^{\alpha_N}_{M}\left(\lambda_j \le s_j, j=1,...,r\right) - \frac{C}{N^2} &\le \mathbb{P}_{M,N}\left(\lambda_j \le s_j, j=1,...,r\right)\\
&\le \mathbb{P}^{\beta_N}_{M}\left(\lambda_j \le s_j, j=1,...,r\right) + \frac{C}{N^2}.
\end{aligned} \tag{4.3.26}$$

Next, $N = (1-\varepsilon_\alpha)\alpha_N$, with $\varepsilon_\alpha = 3\sqrt{N\log N}/\left(N - 3\sqrt{N\log N}\right)$, whereas $N = (1+\varepsilon_\beta)\beta_N$ with $\varepsilon_\beta = 3\sqrt{N\log N}/\left(N + 3\sqrt{N\log N}\right)$. Since $\varepsilon_\alpha, \varepsilon_\beta \to 0$, as $N\to\infty$, it follows from (4.3.26), by setting $s_j = p_{max}N + 2\sqrt{d_1 p_{max} N} + t_j d_1^{-1/6}\sqrt{p_{max}N}$, that

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{P}_M^{\alpha_N}\left(\lambda_j \le \alpha_N p_{max} + 2\sqrt{d_1 \alpha_N p_{max}} + t_j d_1^{-1/6}\sqrt{\alpha_N p_{max}}, j = 1, ..., r\right) \\
&\le \lim_{N\to\infty} \mathbb{P}_{M,N}\left(\lambda_j \le N p_{max} + 2\sqrt{d_1 N p_{max}} + t_j d_1^{-1/6}\sqrt{N p_{max}}, j = 1, ..., r\right) \\
&\le \lim_{N\to\infty} \mathbb{P}_M^{\beta_N}\left(\lambda_j \le \beta_N p_{max} + 2\sqrt{d_1 \beta_N p_{max}} + t_j d_1^{-1/6}\sqrt{\beta_N p_{max}}, j = 1, ..., r\right).
\end{aligned} \tag{4.3.27}$$

Now, (4.3.25) holds true with α replaced by α_N or β_N. Finally, (4.3.22) follows from (4.3.27) by letting $d_1 \to \infty$. □

Remark 4.3.9 *The convergence results in Theorem 4.3.8 are obtained by taking successive limits, e.g., first in N and then in M. For uniform finite alphabet, in which case $d_1 = M$, Johansson [39] had previously obtained the convergence of the whole shape of the associated Young tableaux towards F_r, via a careful analysis of corresponding Kernel and methods of orthogonal polynomials. His results, which are for simultaneous limits, require $(\ln N)^{3/2}/M \to 0$ and $\sqrt{N}/M \to \infty$. Also in the uniform case, under the assumption $M = o\left(N^{3/7}(\ln N)^{-6/7}\right)$, the convergence result (4.3.22) is obtained in [15] for simultaneous limits, via Gaussian approximation and a method originating in Baik and Suidan [9] and Bodineau and Martin [12]. Non-uniform results are also given in [15].*

CHAPTER V

CONCLUSION

We have tackled two sets of problems involving random matrices. The first set is concerned with concentration inequalities for various statistics of random matrices with infinitely divisible entries. The second set analyzes how subsequence problems relate to spectra of certain matrix ensembles. We comment, below, on our approaches and draw some possible future directions of studies.

For random matrices with infinitely divisible entries, we gave concentration inequalities for the spectral measure and for the maximal eigenvalue in Chapter II. Our results stem from the fact that both the linear statistic of the empirical spectral measure and the maximal eigenvalue are Lipschitz functions of the matrix entries. However, in such settings, investigating properties other than the Lipschitz one might lead to shaper probabilistic bounds in certain ranges. Moving beyond concentration inequalities, it will be of great interest to identify the limiting law of the spectral measure of random matrices with infinitely divisible entries, in particular, in our framework where no independence assumption is required.

In Chapter III and Chapter IV, the limiting shape of the Young tableaux associated to random words, with letters independently drawn from a finite ordered alphabets, is shown to be the spectrum of the (generalized) traceless GUE. This limit is also identified as the law of a multidimensional Brownian functional. This equality in law is established through an argument involving a last passage percolation model with appropriate choices of Bernoulli random variables on the lattice. The dimension of this percolation model is determined by the size of the alphabet and the length of the random word. We suspect that these arguments could be applied in obtaining the limiting law for the shape of the Young tableaux associated to random words drawn from a countable infinite alphabet.

Johansson [39] used the Charlier ensemble to study the Poissonized word problem in the uniform setting. He showed the Poissonized Plancherel measure can be obtained as a limit of the Charlier ensemble as the size of the alphabet grows without bound. In Chapter IV, we introduced the generalized Charlier ensemble to deal with Poissonized nonuniform random word. When the size of the alphabet goes to infinity, there should be a corresponding limiting measure for the generalized Charlier ensemble, which would recover the Poissonized Plancherel measure in the uniform case. This convergence depends on the alphabet's growth, and the interplay between the distinct probabilities and their multiplicities.

We focused our research on the shape of the Young tableaux associated to random words, because the size of the first row is equal to the length of the longest increasing subsequence. Other models such as the longest alternating subsequence or the longest unimodel subsequence are also of great interest. There might be ways to connect these problems to different ones such as last passage percolation or random matrix problems, and our approaches might or might not be efficient in studying these subsequences. They do, however, illustrate how longest increasing subsequence, last passage percolation and random matrices are linked, while a priori seemingly rather disconnected. This is the main contribution of the present thesis.

REFERENCES

[1] D. Aldous, P. Diaconis, Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem. *Bull. Amer. Math. Soc. (N.S.)*, **36**, (1999), no. 4, 413-432.

[2] N. Alon, M. Krivelevich, V. Vu, On the concentration of eigenvalues of random symmetric matrices. *Isr. J. Math.*, **131**, (2002), 259-267.

[3] G.W. Anderson, A. Guionnet, O. Zeitouni, Lecture notes on random matrices. *SAMSI*. September, (2006).

[4] G. Ben Arous, A. Guionnet, The spectrum of heavy tailed random matrices. Available at Math arXiv:0707.2159 (2007).

[5] G. Ben Arous, S. Péché, Universality of local statistics for some sample covariance matrices. *Comm. Pure Appl. Math.*, **58**, no. 10, (2005), 1316-1357.

[6] Z.D. Bai, Circular law. *Ann. Probab.*, **25**, (1997), 494-529.

[7] Z.D. Bai, J.W. Silverstein, Spectral analysis of large dimensional random matrices. Science Press, Beijing, (2006).

[8] J. Baik, P.A. Deift, K. Johansson, On the distribution of the length of the longest increasing subsequence in a random permutation. *J. Amer. Math. Soc.*, **12**, (1999), 1119-1178.

[9] J. Baik, T. Suidan. A GUE central limit theorem and universality of directed first and last passage percolation site. *Int. Math. Res. Not.*, **6**, (2005), 325-337.

[10] Y. Baryshnikov, GUEs and queues. *Probab. Theor. Rel. Fields.*, **119**, (2001), 256-274.

[11] P. Billingsley, Convergence of probability measures, 2nd ed.. John Wiley and Sons, Inc., (1999).

[12] T. Bodineau, J. Martin. A universality property for last-passage percolation paths close to the axis. *Elect. Comm. Probab.* **10**, (2005), 105-112.

[13] A. Borodin, A. Okounkov, G. Olshanki. Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.*, **13** (2000), 481-515.

[14] J.-C. Breton, C. Houdré, On finite range stable-type concentration. To appear *Theory Probab. Appl.* (2008).

[15] J.-C. Breton, C. Houdré, Simultaneous asymptotics for the shape of random Young tableaux with a growing alphabet. Preprint.

[16] J.-C. Breton, C. Houdré, N. Privault, Dimension free and infinite variance tail estimates on Poisson space. *Acta Appli. Math.*, **95**, (2007), 151-203.

[17] K. Davidson, S. Szarek. Local operator theory, random matrices and Banach spaces. *Handbook of the Geometry of Banach Spaces*, vol. I, North Holland, (2001), 317-366.

[18] Y. Doumerc, A note on representations of classical Gaussian matrices. *Sémimaire de Probabilités XXXVII.*, *Lecture Notes in Math., No. 1832*, Springer, Berlin, (2003), 370-384.

[19] R.M. Dudley, Real analysis and probability. Cambridge University Press, (2002).

[20] W. Fulton, Young tableaux: with applications to representation theory and geometry. Cambridge University Press, (1997).

[21] V.L. Girko, Spectral theory of random matrices. Moscow, Nauka, (1988).

[22] P.W. Glynn, W, Whitt, Departures from many queues in series. *Ann. Appl. Probab.* **1** (1991), 546-572.

[23] F. Götze, A. Tikhomirov, On the circular law. Preprint. Avalaible at Math arXiv 0702386, (2007).

[24] J. Gravner, J. Tracy, H. Widom, Limit theorems for height fluctuations in a class of discrete space and time growth models. *J. Stat. Phys.*, **102**, Nos. 5-6, (2001), 1085-1132.

[25] A. Guionnet, O. Zeitouni, Concentration of spectral measure for large matrices. *Electron. Comm. Probab.* **5** (2000), 119-136.

[26] J.M. Hammersley, A few seedings of research. *Proc. Sixth Berkeley Symp. Math. Statist. and probability, vol 1*, University of california Press, (1972), 345-394.

[27] R.A. Horn, C. Johnson, Topics in matrix analysis. Cambridge Univ. Press, (1991).

[28] C. Houdré, Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.* **30** (2002), 1223-1237.

[29] C. Houdré, T. Litherland, On the longest increasing subsequence for finite and countable alphabets. Preprint. Available at Math arXiv 0612364, (2006).

[30] C. Houdré, T. Litherland, On the limiting shape of Markovian random Young tableaux. Preprint. Avalaible at Math arXiv:0810.2982, (2008).

[31] C. Houdré, P. Marchal, On the concentration of measure phenomenon for stable and related random vectors. *Ann. Probab.* **32** (2004), 1496-1508.

[32] C. Houdré, P. Marchal, Median, concentration and fluctuations for Lévy processes. *Stochastic Processes and their Applications* **118/5** (2008), 852-863.

[33] C. Houdré, P. Marchal, P. Reynaud-Bouret, Concentration for norms of infinitely divisible vectors with independent components. *Bernoulli* (2008).

[34] A.R. Its, C. Tracy, H. Widom, Random words, Toeplitz determinants, and integrable systems. I. Random matrx models and their applications, *Math. Sci. Res. Inst. Publ.*, **40** Cambridge Univ. Press, Cambridge, (2001), 245-258.

[35] A.R. Its, C. Tracy, H. Widom, Random words, Toeplitz determinants, and integrable systems. II. Advances in nonlinear mathematics and science, *Phys. D.*, vol. 152-153 (2001), 199-224.

[36] K. Johansson, On fluctuations of eigenvalues of random Hermitian matrices. *Duke math. J.* **91** (1998), 151-204.

[37] K. Johansson, The longest increasing subsequence in a random permutation and a unitary random matrix model. *Math. Res. Lett.* **5** (1998), 63-82.

[38] K. Johansson, Shape fluctuation and random matrices. *Comm. Math. Phys.* **209** (2000), 437-476.

[39] K. Johansson, Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. Math.* **153** (2001), 199-224.

[40] I.M. Johnstone, On the distribution of the largest principal component. *Ann. Statist.*, **29**, (2001), 295-327.

[41] I.M. Johnstone, High dimensional statistical inference and random matrices. International Congress of Mathematicians. Vol. I, Eur. Math. Soc., Zürich, (2007), 307–333.

[42] M. Ledoux, The concentration of measure phenomenon. *Math. Surveys Monogr.*, **89** American Mathematical Society, Providence, RI, (2001).

[43] M. Ledoux, A remark on hypercontractivity and tail inequalities for the largest eigenvalues of random matrices. *Séminaire de Probabilités XXXVII.*, *Lecture Notes in Math., No. 1832*, Springer, Berlin, (2003), 360-369.

[44] M. Ledoux, Deviation inequalities on largest eigenvalues. *Geometric Aspects of Functional Analysis, Israel Seminar 2004-2005*, Lecture Notes in Math. **1910**, 167-219. Springer (2007).

[45] M. Ledoux, M. Talagrand, Probability in Banach spaces (Isoperimetry and processes). Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag, Berlin, (1991).

[46] P. Marchal, Measure concentration for stable laws with index close to 2. *Electron. Comm. Probab.* **10** (2005), 29-35.

[47] M.B. Marcus, J. Rosiński, L^1-norms of infinitely divisible random vectors and certain stochastic integrals. *Electron. Comm. Probab.* **6** (2001), 15-29.

[48] V.A. Marčenko, L.A. Pastur. Distributions of eigenvalues of some sets of random matrices. *Math. USSR-Sb.*, **1** (1967), 507-536.

[49] M.L. Mehta, Random matrices, 2nd ed. Academic Press, San Diego, (1991).

[50] A. Okounkov, Random matrices and random permutations. *Internat. Math. Res. Notice* **2000**, no. 20, (2000), 1043-1095.

[51] L.A. Pastur, On the spectrum of random matrices. *Teor. Mat. Fiz.* **10** (1972), 102-112.

[52] S. Péché, Universality results for largest eigenvalues of some sample covariance matrix ensembles. Preprint, Available at Math arXiv:0705.1701 (2007).

[53] K-I. Sato, Lévy processes and infinitely divisible distributions. Cambridge University Press, Cambridge, (1999).

[54] J. Silverstein, The smallest eigenvalue of a large-dimensional Wishart matrix. *Ann. Probab.* **13**, no. 4, (1985), 1364-1368.

[55] B. Simon, Trace ideals and their applications. Cambridge University Press, (1979).

[56] A. Soshnikov, Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.* **207** (1999), 697-733.

[57] A. Soshnikov, A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays. *J. Stat. Phys.*, **108, Nos 5/6**, (2002), 1033-1056.

[58] A. Soshnikov, Poisson statistics for the largest eigenvalues of Wigner random matrices with heavy tails. *Elec. Commun. Probab.* **9** (2004), 82-91.

[59] A. Soshnikov, Poisson statistics for the largest eigenvalues in random matrix ensembles. *Mathematical physics of quantum mechanics*, 351–364, Lecture Notes in Phys., **690**, Springer, Berlin, (2006).

[60] A. Soshnikov, Y. V. Fyodorov, On the largest singular values of random matrices with independent Cauchy entries. *J. Math. Phys.* **46** (2005), 033302.

[61] C. Tracy, H. Widom, Level-spacing distribution and the Airy kernel. *Comm. Math. Phys.* **159** (1994), 151-174.

[62] C. Tracy, H. Widom, On orthogonal and symplectic random matrix ensembles. *Comm. Math. Phys.* **177** (1996), 724-754.

[63] C. Tracy, H. Widom, Correlation functions, cluster functions, and spacing distributions for random matrices. *J. Statist. Phys.* **92**, no. 5-6, (1998), 809-835.

[64] C. Tracy, H. Widom, On the distribution of the lengths of the longest inceasing monotone subsequences in random words. *Probab. Theor. Rel. Fields.* **119** (2001), 350-380.

[65] H.F. Trotter, Eigenvalue distribution of large Hermitian matrices; Wigner's semicircle law and a theorem of Kac, Murdock, and Szegö. *Adv. Math.* **54** (1984), 67-82.

[66] E.P. Wigner, On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)* **67** (1958), 325-327.

VITA

Hua Xu was born in Wuhan, China on September 26, 1979. He went to the University of Science and Technology of China in Hefei in 1998 and graduated with a Bachelors of Science in mathematics in June of 2003. In July of that year, he moved to Atlanta and jointed the Georgia Institute of Technology to work in Probability Theory with Professor Christian Houdré in the School of Mathematics. During his stay at the Georgia Institute of Technology, he also received a dual Masters degree of Science in Statistics from the School of Industrial and System Engineering.

On July 23, 2003, in Wuhan, Hua married Jing Xu, who later joined the Georgia Institute of Technology and graduated with a Masters degree in City and Regional Planning in 2008.