# Eigenvalue variance bounds for Wigner and covariance random matrices

Sandrine Dallaporta

# Eigenvalue variance bounds for Wigner and covariance random matrices

S. Dallaporta

University of Toulouse, France

**Abstract.** *This work is concerned with finite range bounds on the variance of individual eigenvalues of Wigner random matrices, in the bulk and at the edge of the spectrum, as well as for some intermediate eigenvalues. Relying on the GUE example, which needs to be investigated first, the main bounds are extended to families of Hermitian Wigner matrices by means of the Tao and Vu Four Moment Theorem and recent localization results by Erdös, Yau and Yin. The case of real Wigner matrices is obtained from interlacing formulas. As an application, bounds on the expected 2-Wasserstein distance between the empirical spectral measure and the semicircle law are derived. Similar results are available for random covariance matrices.*

Two different models of random Hermitian matrices were introduced by Wishart in the twenties and by Wigner in the fifties. Wishart was interested in modeling tables of random data in multivariate analysis and worked on random covariance matrices. In this paper, the results for covariance matrices are very close to those for Wigner matrices. Therefore, it deals mainly with Wigner matrices. Definitions and results regarding covariance matrices are available in the last section.

Random Wigner matrices were first introduced by Wigner to study eigenvalues of infinite-dimensional operators in statistical physics (see [17]) and then propagated to various fields of mathematics involved in the study of spectra of random matrices. Under suitable symmetry assumptions, the asymptotic properties of the eigenvalues of a random matrix were soon conjectured to be universal, in the sense they do not depend on the individual distribution of the matrix entries. This opened the way to numerous developments on the asymptotics of various statistics of the eigenvalues of random matrices, such as for example the global behavior of the spectrum, the spacings between the eigenvalues in the bulk of the spectrum or the behavior of the extreme eigenvalues. Two main models have been considered, invariant matrices and Wigner matrices. In the invariant matrix models, the matrix law is unitary invariant and the eigenvalue joint distribution can be written explicitly in terms of a given potential. In the Wigner models, the matrix entries are independent (up to symmetry conditions). The case where the entries are Gaussian is the only model belonging to both types. In the latter case, the joint distribution of the eigenvalues is thus explicitly known and the previous statistics have been completely studied.

One main focus of random matrix theory in the past decades was to prove that these asymptotic behaviors were the same for non-Gaussian matrices (see for instance [1], [3] and [19]).

However, in several fields such as computer science or statistics for example, asymptotic statements are often not enough, and more quantitative finite range results are required. Several recent developments have thus been concerned with non-asymptotic random matrix theory towards quantitative bounds (for instance on the probability for a certain event to occur) which are valid for all $N$, where $N$ is the size of the given matrix. See for example [30] for an introduction to some problems considered in non-asymptotic random matrix theory. In this paper, we investigate in this respect variance bounds on the eigenvalues of families of Wigner random matrices.

Wigner matrices are Hermitian or real symmetric matrices $M_N$ such that, if $M_N$ is complex, for $i < j$, the real and imaginary parts of $(M_N)_{ij}$ are independent and identically distributed (iid) with mean 0 and variance $\frac{1}{2}$, $(M_N)_{ii}$ are iid, with mean 0 and variance 1. In the real case, $(M_n)_{ij}$ are iid, with mean 0 and variance 1 and $(M_N)_{ii}$ are iid, with mean 0 and variance 2. In both cases, set $W_N = \frac{1}{\sqrt{N}} M_N$. An important example of Wigner matrices is the case where the entries are Gaussian. If $M_N$ is complex, then it belongs to the so-called Gaussian Unitary Ensemble (GUE). If it is real, it belongs to the Gaussian Orthogonal Ensemble (GOE). The matrix $W_N$ has $N$ real eigenvalues $\lambda_1 \leqslant \cdots \leqslant \lambda_N$. In the Gaussian case, the joint law of the eigenvalues is known, allowing for complete descriptions of their limiting behavior both in the global and local regimes (see for example [1], [3] and [19]).

Among universality results, at the global level, the classical Wigner's Theorem states that the empirical distribution $L_N = \frac{1}{N} \sum_{j=1}^{N} \delta_{\lambda_j}$ on the eigenvalues of $W_N$ converges weakly almost surely to the semicircle law $d\rho_{sc}(x) = \frac{1}{2\pi} \sqrt{4 - x^2} 1_{[-2,2]}(x) dx$ (see for example [3] for a proof in the more general setting). This gives the global asymptotic behavior of the spectrum. However, this theorem is not enough to deduce some information on individual eigenvalues. Define, for all $1 \leqslant j \leqslant N$, the theoretical location of the $j^{\text{th}}$ eigenvalue $\gamma_j$ by $\int_{-2}^{\gamma_j} d\rho_{sc}(x) = \frac{j}{N}$. Bai and Yin proved in [2] with assumptions on higher moments that almost surely the smallest and the largest eigenvalues converge to their theoretical locations, which means that $\lambda_N \to 2$ and $\lambda_1 \to -2$ almost surely (see also [3] for a proof of this theorem). From this almost sure convergence of the extreme eigenvalues and Wigner's Theorem, it is possible to deduce information on individual eigenvalues in the bulk of the spectrum. Indeed, according to the Glivenko-Cantelli Theorem (see for example [7]), the normalized eigenvalue function $\frac{1}{N} \mathcal{N}_x$, where $\mathcal{N}_x$ is the number of eigenvalues which are in $(-\infty, x]$, converges uniformly on $\mathbb{R}$ almost surely to the distribution function of the semicircle law $G$ (with no more assumptions on the matrix entries). Then, using that $\frac{1}{N} \mathcal{N}_{\lambda_j} = \frac{j}{N} = G(\gamma_j)$ together with crude bounds on the semicircle density function and the fact that $\lambda_j$ is almost surely between $-2 - \varepsilon$ and $2 + \varepsilon$ shows that almost surely $\lambda_j - \gamma_j \to 0$ uniformly for $\eta N \leqslant j \leqslant (1 - \eta)N$ (for any fixed $\eta > 0$).

At the fluctuation level, eigenvalues inside the bulk and at the edge of the spectrum do not have the same behavior. Tracy and Widom showed in [28] that the largest eigenvalue fluctuates around 2 according to the so-called Tracy-Widom law

$F_2$. Namely,

$$N^{2/3}(\lambda_N - 2) \to F_2$$

in distribution as $N$ goes to infinity. They proved this result for Gaussian matrices of the GUE, later extended to families of non-Gaussian Wigner matrices by Soshnikov (see [21]). Recent results by Tao and Vu (see [27]) and by Erdös, Yau and Yin (see [8]) provide alternate proofs of this fact for larger families of Wigner matrices. According to this asymptotic property, the variance of the largest eigenvalue $\lambda_N$ is thus of the order of $N^{-4/3}$. In the bulk, Gustavsson proved in [14] again for the GUE, that, for any fixed $\eta > 0$ and all $\eta N \leqslant j \leqslant (1 - \eta)N$,

$$\frac{\lambda_j - \gamma_j}{\sqrt{\frac{2 \log N}{(4 - \gamma_j^2)N^2}}} \to \mathcal{N}(0, 1) \tag{1}$$

in distribution as $N$ goes to infinity. This result was extended by Tao and Vu in [26] to large families of non-Gaussian Wigner matrices. The variance of an eigenvalue $\lambda_j$ in the bulk is thus of the order of $\frac{\log N}{N^2}$. Right-side intermediate eigenvalues consist in the $\lambda_j$'s with $\frac{N}{2} \leqslant j \leqslant N$ such that $\frac{j}{N} \to 1$ but $N - j \to \infty$ as $N$ goes to infinity (the left-side can be deduced by symmetry). Gustavsson proved a Central Limit Theorem for these eigenvalues (see [14]) from which their variance is guessed to be of the order of $\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}$. This result was again extended to large classes of Wigner matrices by Tao and Vu in [27].

The previous results are however asymptotic. As announced, the purpose of this work is to provide quantitative bounds on the variance of the eigenvalues of the correct order, in the bulk and at the edge of the spectrum, as well as for some intermediate eigenvalues. In the statements below, $M_N$ is a complex (respectively real) Wigner matrix satisfying condition $(C0)$. This condition, which will be detailed in Section 2.1, provides an exponential decay of the matrix entries. Assume furthermore that the entries of $M_N$ have the same first four moments as the entries of a GUE matrix (respectively GOE). Set $W_N = \frac{1}{\sqrt{N}}M_N$.

**Theorem 1** (in the bulk). *For all $0 < \eta \leqslant \frac{1}{2}$, there exists a constant $C(\eta) > 0$ such that, for all $N \geqslant 2$ and for all $\eta N \leqslant j \leqslant (1 - \eta)N$,*

$$\mathrm{Var}(\lambda_j) \leqslant C(\eta)\frac{\log N}{N^2}. \tag{2}$$

**Theorem 2** (at the edge). *There exists a universal constant $C > 0$ such that, for all $N \geqslant 1$,*

$$\mathrm{Var}(\lambda_1) \leqslant CN^{-4/3} \quad and \quad \mathrm{Var}(\lambda_N) \leqslant CN^{-4/3}. \tag{3}$$

**Theorem 3** (between the bulk and the edge). *For all $0 < \eta \leqslant \frac{1}{2}$ and for all $K > 20\sqrt{2}$, there exists a constant $C(\eta, K) > 0$ such that, for all $N \geqslant 2$, for all $K \log N \leqslant j \leqslant \eta N$ and $(1 - \eta)N \leqslant j \leqslant N - K \log N$,*

$$\mathrm{Var}(\lambda_j) \leqslant C(\eta, K)\frac{\log\big(\min(j, N - j)\big)}{N^{4/3}\big(\min(j, N - j)\big)^{2/3}}. \tag{4}$$

3

It should be mentioned that Theorem 1 does not seem to be known even for Gaussian matrices. The first step is thus to prove it for the GUE. This will be achieved via the analysis of the eigenvalue counting function, which due to the particular determinantal structure in this case, has the same distribution as a sum of independent Bernoulli variables. Sharp standard deviation inequalities are thus available in this case. These may then be transferred to the eigenvalues in the bulk together with Gustavsson's bounds on the variance of the eigenvalue counting function. As a result, we actually establish that

$$E\big[|\lambda_j - \gamma_j|^2\big] \leqslant C(\eta)\frac{\log N}{N^2}$$

leading thus to Theorem 1 in this case. Similarly, Theorem 3 does not seem to be known for Gaussian matrices. The proof follows exactly the same scheme and we establish that

$$E\big[|\lambda_j - \gamma_j|^2\big] \leqslant C(\eta, K)\frac{\log\big(\min(j, N-j)\big)}{N^{4/3}\big(\min(j, N-j)\big)^{2/3}}.$$

On the other hand, Theorem 2 for the GUE and GOE has been known for some time (see [15]). On the basis of these results for the GUE (and GOE), Theorems 1, 2 and 3 are then extended to families of Wigner matrices by a suitable combination of Tao and Vu's Four Moment Theorem (see [26], [27]) and Erdös, Yau and Yin's Localization Theorem (see [8]). The basic idea is that while the localization properties almost yield the correct order, the Four Moment Theorem may be used to reach the optimal bounds by comparison with the Gaussian models. Theorems 1, 2 and 3 are established first in the complex case. The real case is deduced by means of interlacing formulas. Furthermore, analogous results are established for covariance matrices, for which non-asymptotic quantitative results are needed, as they are useful in several fields, such as compressed sensing (see [30]), wireless communication and quantitative finance (see [3]).

The method developed here do not seem powerful enough to strengthen the variance bounds into exponential tail inequalities. In the recent contribution [24], Tao and Vu proved such exponential tail inequalities by a further refinement of the replacement method leading to the Four Moment Theorem. While much more powerful than variance bounds, they do not seem to yield at this moment the correct order of the variance bounds of Theorems 1, 2 and 3.

As a corollary of the latter results, a bound on the rate of convergence of the empirical spectral measure $L_N$ can be achieved. This bound is expressed in terms of the so-called 2-Wasserstein distance $W_2$ between $L_N$ and the semicircle law $\rho_{sc}$. For $p \in [1, \infty)$, the $p$-Wasserstein distance $W_p(\mu, \nu)$ between two probability measures $\mu$ and $\nu$ on $\mathbb{R}$ is defined by

$$W_p(\mu, \nu) = \inf\left(\int_{\mathbb{R}^2}|x-y|^p\, d\pi(x,y)\right)^{1/p}$$

where the infimum is taken over all probability measure $\pi$ on $\mathbb{R}^2$ such that its first marginal is $\mu$ and its second marginal is $\nu$. Note for further purposes that the rate

of convergence of this empirical distribution has been investigated in various forms. For example, the Kolmogorov distance between $L_N$ and $\rho_{sc}$ has been considered in this respect in several papers (see for example [11], [5] and [10]). It is given by

$$d_K(L_N, \rho_{sc}) = \sup_{x \in \mathbb{R}} \left| \tfrac{1}{N} \mathcal{N}_x - G(x) \right|,$$

where $\mathcal{N}_x$ is the eigenvalue counting function and $G$ is the distribution function of the semicircle law. More and more precise bounds were established. For example, under the hypothesis of an exponential decay of the matrix entries, Götze and Tikhomirov recently showed that, with high probability,

$$d_K(L_N, \rho_{sc}) \leqslant \frac{(\log N)^c}{N}$$

for some universal constant $c > 0$ (see [10]). The rate of convergence in terms of $W_1$, also called the Kantorovich-Rubinstein distance, was studied by Guionnet and Zeitouni in [13] and recently by Meckes and Meckes in [16]. In [16], the authors proved that $\mathrm{E}[W_1(L_N, \mathrm{E}[L_N])] \leqslant C N^{-2/3}$. The following is concerned with the distance between $L_N$ and $\rho_{sc}$ and strengthens the preceding conclusion on $W_1$.

**Corollary 4.** *There exists a numerical constant $C > 0$ such that, for all $N \geqslant 2$,*

$$\mathrm{E}\left[W_2^2(L_N, \rho_{sc})\right] \leqslant C \frac{\log N}{N^2}. \tag{5}$$

The proof of this corollary relies on the fact that $\mathrm{E}\left[W_2^2(L_N, \rho_{sc})\right]$ is bounded above, up to a constant, by the sum of the expectations $\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right]$. The previously established bounds then easily yield the result.

Turning to the organization of the paper, Section 1 describes Theorems 1, 2 and 3 in the GUE case. Section 2 emphasizes to start with the Four Moment Theorem of Tao and Vu (see [26] and [27]) and the Localization Theorem of Erdös, Yau and Yin (see [8]). On the basis of these results and the GUE case, the main Theorems 1, 2 and 3 are then established for families of Wigner matrices. Section 3 is devoted to the corresponding statements for real matrices, while Section 5 describes the analogous results for covariance matrices. Section 4 deals with Corollary 4 and the rate of convergence of $L_N$ towards $\rho_{sc}$ in terms of 2-Wasserstein distance.

Throughout this paper, $C(\cdot)$ and $c(\cdot)$ will denote positive constants which depend on the specified quantities and may differ from one line to another.

# 1 Deviation inequalities and variance bounds in the GUE case

The results and proofs developed in this section for the GUE model heavily rely on its determinantal structure which allows for a complete description of the joint law of the eigenvalues (see [1]). In particular the eigenvalue counting function is known to have the same distribution as a sum of independent Bernoulli variables (see [4], [1]), whose mean and variance were computed by Gustavsson (see [14]). Deviation inequalities for individual eigenvalues can thus be established, leading to the announced bounds on the variance.

## 1.1 Eigenvalues in the bulk of the spectrum

The aim of this section is to establish the following result.

**Theorem 5.** *Let $M_N$ be a GUE matrix. Set $W_N = \frac{1}{\sqrt{N}} M_N$. For any $0 < \eta \leqslant \frac{1}{2}$, there exists a constant $C(\eta) > 0$ such that for all $N \geqslant 2$ and all $\eta N \leqslant j \leqslant (1 - \eta)N$,*

$$\mathrm{E}\left[|\lambda_j - \gamma_j|^2\right] \leqslant C(\eta) \frac{\log N}{N^2}. \tag{6}$$

*In particular,*

$$\mathrm{Var}(\lambda_j) \leqslant C(\eta) \frac{\log N}{N^2}. \tag{7}$$

As announced, the proof is based on the connection between the distribution of eigenvalues and the eigenvalue counting function. For every $t \in \mathbb{R}$, let $\mathcal{N}_t = \sum_{i=1}^{N} \mathbf{1}_{\lambda_i \leqslant t}$ be the eigenvalue counting function. Due to the determinantal structure of the GUE, it is known (see [4]) that $\mathcal{N}_t$ has the same distribution as a sum of independent Bernoulli random variables. Bernstein's inequality for example (although other, even sharper, inequalities may be used) may then be applied to get that for every $u \geqslant 0$,

$$\mathrm{P}\left(|\mathcal{N}_t - \mathrm{E}[\mathcal{N}_t]| \geqslant u\right) \leqslant 2 \exp\left(-\frac{u^2}{2\sigma_t^2 + u}\right), \tag{8}$$

where $\sigma_t^2$ is the variance of $\mathcal{N}_t$ (see for example [29]). Note that the upper-bound is non-increasing in $u$ while non-decreasing in the variance. Set for simplicity $\rho_t = \rho_{sc}\left((-\infty, t]\right)$, $t \in \mathbb{R}$. It has been shown in [12] that for some numerical constant $C_1 > 0$,

$$\sup_{t \in \mathbb{R}} \left|\mathrm{E}[\mathcal{N}_t] - N\rho_t\right| \leqslant C_1.$$

In particular thus, together with (8), for every $u \geqslant 0$,

$$\mathrm{P}\left(|\mathcal{N}_t - N\rho_t| \geqslant u + C_1\right) \leqslant 2 \exp\left(-\frac{u^2}{2\sigma_t^2 + u}\right). \tag{9}$$

As a main conclusion of the work by Gustavsson [14], for every $\delta \in (0,2)$, there exists $c_\delta > 0$ such that

$$\sup_{t \in I_\delta} \sigma_t^2 \leqslant c_\delta \log N, \tag{10}$$

where $I_\delta = [-2+\delta, 2-\delta]$. On the basis of inequalities (9) and (10), it is then possible to derive a deviation inequality for eigenvalues $\lambda_j$ in the bulk from their theoretical locations $\gamma_j \in [-2, 2]$, $1 \leqslant j \leqslant N$, defined by $\rho_{\gamma_j} = \frac{j}{N}$.

**Proposition 6** (Deviation inequality for $\lambda_j$). *Let $\eta \in (0, \frac{1}{2}]$ and $\eta N \leqslant j \leqslant (1-\eta)N$. There exist $C > 0$, $c > 0$, $c' > 0$ and $\delta \in (0,2)$ (all depending on $\eta$) such that, for all $c \leqslant u \leqslant c'N$,*

$$\mathrm{P}\left(|\lambda_j - \gamma_j| \geqslant \frac{u}{N}\right) \leqslant 4 \exp\left(-\frac{C^2 u^2}{2c_\delta \log N + Cu}\right). \tag{11}$$

*Proof.* Let $\eta \in (0, \frac{1}{2}]$. To start with, evaluate, for $\frac{N}{2} \leqslant j \leqslant (1 - \eta)N$ and $u \geqslant 0$, the probability $\mathrm{P}(|\lambda_j - \gamma_j| > \frac{u}{N})$. We have

$$
\begin{aligned}
\mathrm{P}\left(\lambda_j > \gamma_j + \frac{u}{N}\right) &= \mathrm{P}\left(\sum_{i=1}^{N} \mathbf{1}_{\lambda_i \leqslant \gamma_j + \frac{u}{N}} < j\right) \\
&= \mathrm{P}\left(\mathcal{N}_{\gamma_j + \frac{u}{N}} < j\right) \\
&= \mathrm{P}\left(N\rho_{\gamma_j + \frac{u}{N}} - \mathcal{N}_{\gamma_j + \frac{u}{N}} > N\rho_{\gamma_j + \frac{u}{N}} - j\right) \\
&= \mathrm{P}\left(N\rho_{\gamma_j + \frac{u}{N}} - \mathcal{N}_{\gamma_j + \frac{u}{N}} > N(\rho_{\gamma_j + \frac{u}{N}} - \rho_{\gamma_j})\right)
\end{aligned}
$$

where it has been used that $\rho_{\gamma_j} = \frac{j}{N}$. Then

$$
\mathrm{P}\left(\lambda_j > \gamma_j + \frac{u}{N}\right) \leqslant \mathrm{P}\left(|\mathcal{N}_{\gamma_j + \frac{u}{N}} - N\rho_{\gamma_j + \frac{u}{N}}| > N(\rho_{\gamma_j + \frac{u}{N}} - \rho_{\gamma_j})\right).
$$

But

$$
\begin{aligned}
\rho_{\gamma_j + \frac{u}{N}} - \rho_{\gamma_j} &= \int_{\gamma_j}^{\gamma_j + \frac{u}{N}} \frac{1}{2\pi}\sqrt{4 - x^2}\, dx \\
&\geqslant \frac{1}{\sqrt{2}\pi} \int_{\gamma_j}^{\gamma_j + \frac{u}{N}} \sqrt{2 - x}\, dx \\
&\geqslant \frac{\sqrt{2}}{3\pi}(2 - \gamma_j)^{3/2}\left(1 - \left(1 - \frac{\frac{u}{N}}{2 - \gamma_j}\right)^{3/2}\right) \\
&\geqslant \frac{\sqrt{2}}{3\pi}(2 - \gamma_j)^{1/2}\frac{u}{N},
\end{aligned}
$$

if $u \leqslant (2 - \gamma_j)N$. By definition, $1 - \frac{j}{N} = \int_{\gamma_j}^{2} \frac{1}{2\pi}\sqrt{4 - x^2}\, dx$. Then

$$
\frac{1}{\sqrt{2}\pi} \int_{\gamma_j}^{2} \sqrt{2 - x}\, dx \leqslant 1 - \frac{j}{N} \leqslant \frac{1}{\pi} \int_{\gamma_j}^{2} \sqrt{2 - x}\, dx.
$$

Computing the preceding integrals yields

$$
\left(\frac{3\pi}{2}\frac{N - j}{N}\right)^{2/3} \leqslant 2 - \gamma_j \leqslant \left(\frac{3\pi}{\sqrt{2}}\frac{N - j}{N}\right)^{2/3}. \tag{12}
$$

Therefore, $u \leqslant (2 - \gamma_j)N$ if $u \leqslant c'N$ with $c' \leqslant (\frac{3\pi}{2})^{2/3}\eta^{2/3}$. In this case, (12) yields

$$
\rho_{\gamma_j + \frac{u}{N}} - \rho_{\gamma_j} \geqslant 2C\frac{u}{N}
$$

where $C > 0$. Therefore

$$
\begin{aligned}
\mathrm{P}\left(\lambda_j > \gamma_j + \frac{u}{N}\right) &\leqslant \mathrm{P}\left(|\mathcal{N}_{\gamma_j + \frac{u}{N}} - N\rho_{\gamma_j + \frac{u}{N}}| > 2Cu\right) \\
&\leqslant \mathrm{P}\left(|\mathcal{N}_{\gamma_j + \frac{u}{N}} - N\rho_{\gamma_j + \frac{u}{N}}| > Cu + C_1\right)
\end{aligned}
$$

when $u \geqslant \frac{C_1}{C} = c$. Then, applying (9) leads to

$$
\mathrm{P}\left(\lambda_j > \gamma_j + \frac{u}{N}\right) \leqslant 2\exp\left(-\frac{C^2 u^2}{2\sigma_{\gamma_j + \frac{u}{N}}^2 + Cu}\right).
$$

As $\gamma_j$ is in the bulk, there exist $\delta$ and $c' < (\frac{3\pi}{2})^{2/3}\eta^{2/3}$ such that $\gamma_j + \frac{u}{N} \in I_\delta$, for all $\eta N \leqslant j \leqslant (1-\eta)N$ and for all $c \leqslant u \leqslant c'N$, for all $N \geqslant 1$ (both $\delta$ and $c'$ depend on $\eta$). Then

$$\mathrm{P}\left(\lambda_j > \gamma_j + \frac{u}{N}\right) \leqslant 2\exp\left(-\frac{C^2 u^2}{2c_\delta \log N + Cu}\right).$$

Repeating the argument leads to the same bound on $\mathrm{P}\left(\lambda_j < \gamma_j - \frac{u}{N}\right)$. Therefore,

$$\mathrm{P}\left(|\lambda_j - \gamma_j| \geqslant \frac{u}{N}\right) \leqslant 4\exp\left(-\frac{C^2 u^2}{2c_\delta \log N + Cu}\right).$$

The proposition is thus established.

□

*Proof of Theorem 5.* Note first that, for every $j$,

$$\mathrm{E}\left[\lambda_j^4\right] \leqslant \sum_{i=1}^{N} \mathrm{E}\left[\lambda_i^4\right] = \mathrm{E}\left[\mathrm{Tr}(W_N^4)\right].$$

The mean of this trace can be easily computed and is equal to $2N + \frac{1}{N}$. Consequently, for all $N \geqslant 1$,

$$\mathrm{E}\left[\lambda_j^4\right] \leqslant 3N. \tag{13}$$

Choose next $M = M(\eta) > 0$ large enough such that $\frac{C^2 M^2}{2c_\delta + CM} > 5$. Setting $Z = N|\lambda_j - \gamma_j|$,

$$\begin{aligned}
\mathrm{E}[Z^2] &= \int_0^\infty \mathrm{P}(Z \geqslant v)2v\,dv \\
&= \int_0^c \mathrm{P}(Z \geqslant v)2v\,dv + \int_c^{M\log N} \mathrm{P}(Z \geqslant v)2v\,dv + \int_{M\log N}^\infty \mathrm{P}(Z \geqslant v)2v\,dv \\
&\leqslant c^2 + I_1 + I_2.
\end{aligned}$$

The two latter integrals are handled in different ways. The first one $I_1$ is bounded using (11) while $I_2$ is controlled using the Cauchy-Schwarz inequality and (13). Starting thus with $I_2$,

$$\begin{aligned}
I_2 &= \int_{M\log N}^{+\infty} \mathrm{P}(Z \geqslant v)2v\,dv \\
&\leqslant \mathrm{E}\left[Z^2 \mathbf{1}_{Z \geqslant M\log N}\right] \\
&\leqslant \sqrt{\mathrm{E}\left[Z^4\right]}\sqrt{\mathrm{P}\left(Z \geqslant M\log N\right)} \\
&\leqslant 2A\exp\left(\frac{1}{2}\left(5 - \frac{C^2 M^2}{2c_\delta + CM}\right)\log N\right),
\end{aligned}$$

where $A > 0$ is a numerical constant. As $\exp\left(\frac{1}{2}\left(5 - \frac{C^2 M^2}{2c_\delta + CM}\right)\log N\right) \xrightarrow{N\to\infty} 0$, there exists a constant $C(\eta) > 0$ such that

$$I_2 \leqslant C(\eta).$$

Turning to $I_1$, recall that Proposition 6 gives, for $c \leqslant v \leqslant c'N$,

$$P(Z \geqslant v) = \mathrm{P}\Big(|\lambda_j - \gamma_j| \geqslant \frac{v}{N}\Big) \leqslant 4\exp\Big(-\frac{C^2 v^2}{2c_\delta \log N + Cv}\Big).$$

Hence in the range $v \leqslant M\log N$,

$$P(Z \geqslant v) \leqslant 4\exp\Big(-\frac{B}{\log N}v^2\Big),$$

where $B = B(\eta) = \frac{C^2}{2c_\delta + CM}$. There exists thus a constant $C(\eta) > 0$ such that

$$I_1 \leqslant C(\eta)\log N.$$

Summarizing the previous steps, $\mathrm{E}\big[Z^2\big] \leqslant C(\eta)\log N$. Therefore

$$\mathrm{E}\big[|\lambda_j - \gamma_j|^2\big] \leqslant C(\eta)\frac{\log N}{N^2},$$

which is the claim. The proof of Theorem 5 is complete. $\qquad\square$

It may be shown similarly that, under the assumptions of Theorem 5,

$$\mathrm{E}\big[|\lambda_j - \gamma_j|^p\big] \leqslant C(p,\eta)\frac{(\log N)^{p/2}}{N^p}.$$

## 1.2 Eigenvalues at the edge of the spectrum

In [15], Ledoux and Rider gave unified proofs of precise small deviation inequalities for the extreme eigenvalues of $\beta$-ensembles. The results hold in particular for GUE matrices ($\beta = 2$) and for GOE matrices ($\beta = 1$). The following theorem summarizes some of the relevant inequalities for the GUE.

**Theorem 7.** *There exists a universal constant $C > 0$ such that the following holds. Let $M_N$ be a GUE matrix. Set $W_N = \frac{1}{\sqrt{N}}M_N$ and denote by $\lambda_N$ the maximal eigenvalue of $W_N$. Then, for all $N \in \mathbb{N}$ and all $0 < \varepsilon \leqslant 1$,*

$$\mathrm{P}\big(\lambda_N \geqslant 2(1+\varepsilon)\big) \leqslant C\exp\Big(-\frac{2N}{C}\varepsilon^{3/2}\Big), \tag{14}$$

*and*

$$\mathrm{P}\big(\lambda_N \leqslant 2(1-\varepsilon)\big) \leqslant C^2\exp\Big(-\frac{2N^2}{C}\varepsilon^3\Big). \tag{15}$$

*There exists also a right-tail large deviation inequality for $\varepsilon = O(1)$ of the form*

$$\mathrm{P}\big(\lambda_N \geqslant 2(1+\varepsilon)\big) \leqslant C\exp\Big(-\frac{2N}{C}\varepsilon^2\Big), \tag{16}$$

*where $C > 0$ is a universal constant.* Similar inequalities hold for the smallest eigenvalue $\lambda_1$. As stated in [15], bounds on the variance straightly follow from these deviation inequalities.

**Corollary 8.** *Let $M_N$ be a GUE matrix. Set $W_N = \frac{1}{\sqrt{N}}M_N$. Then there exists a universal constant $C > 0$ such that for all $N \geqslant 1$,*

$$\mathrm{Var}(\lambda_N) \leqslant \mathrm{E}\big[(\lambda_N - 2)^2\big] \leqslant CN^{-4/3}.$$

Similar results are probably true for the $k^{\mathrm{th}}$ smallest or largest eigenvalue (for $k \in \mathbb{N}$ fixed).

## 1.3 Eigenvalues between the bulk and the edge of the spectrum

The aim of this section is to establish a bound on the variance for some intermediate eigenvalues. The proof is very similar to what was done for eigenvalues in the bulk. It relies on the fact that the eigenvalue counting function has the same distribution as a sum of independent Bernoulli variables due to the determinantal properties of the eigenvalues. A deviation inequality for individual eigenvalues can then be derived and the bound on the variance straightly follows. Parts of the proof which are identical to the proof for eigenvalues in the bulk will be omitted. In what follows we only consider the right-side of the spectrum. Results and proofs for the left-side can be deduced by replacing $N - j$ by $j$. The precise statement is the following.

**Theorem 9.** *Let $M_N$ be a GUE matrix. Set $W_N = \frac{1}{\sqrt{N}} M_N$. For all $K \geqslant 20\sqrt{2}$ and for all $\eta \in (0, \frac{1}{2}]$, there exists a constant $C(\eta, K) > 0$ such that for all $N \geqslant 2$ and all $(1 - \eta)N \leqslant j \leqslant N - K \log N$,*

$$\mathrm{E}\left[|\lambda_j - \gamma_j|^2\right] \leqslant C(\eta, K) \frac{\log(N - j)}{N^{4/3}(N - j)^{2/3}}. \tag{17}$$

*In particular,*

$$\mathrm{Var}(\lambda_j) \leqslant C(\eta, K) \frac{\log(N - j)}{N^{4/3}(N - j)^{2/3}}. \tag{18}$$

The preceding theorem does not concern all intermediate eigenvalues since $N - j$ has to be at least of the order of $20\sqrt{2} \log N$ for the method used here to yield the correct order on the variance. This restriction seems however to be technical since Gustavsson [14] proved that a Central Limit Theorem holds for all eigenvalues $\lambda_j$ such that $N - j \to \infty$ but $\frac{N-j}{N} \to 0$. From this CLT, the variance of such an individual eigenvalue $\lambda_j$ is guessed to be similarly of the order of $\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}$ for this range.

As for eigenvalues in the bulk, the proof relies on the deviation inequality for the eigenvalue counting function (9). From this and a bound on the variance of this counting function (10), it was then possible to derive a deviation inequality for eigenvalues in the bulk. The work of Gustavsson [14] suggests that for all $0 < \tilde{\eta} < 4$ and for all $\tilde{K} > 0$, there exists a constant $c_{\tilde{\eta}, \tilde{K}} > 0$ such that the following holds. For every sequence $(t_N)_{N \in \mathbb{N}}$ such that, for all $N$, $0 < 2 - t_N \leqslant \tilde{\eta}$ and $N(2 - t_N)^{3/2} \geqslant \tilde{K} \log N$,

$$\sigma_{t_N}^2 \leqslant c_{\tilde{\eta}, \tilde{K}} \log\left(N(2 - t_N)^{3/2}\right). \tag{19}$$

Similarly to the bulk case, the following proposition can be established.

**Proposition 10** (Deviation inequality for intermediate $\lambda_j$)**.** *There exist universal positive constants $C$ and $c$ such that the following holds. Let $K > 20\sqrt{2}$ and $\eta \in (0, \frac{1}{2}]$. Set $(1 - \eta)N \leqslant j \leqslant N - K \log N$. There exists $C' > 0$ and $c' > 0$ depending on $K$ and $\eta$ such that for all $c \leqslant u \leqslant c'(N - j)$,*

$$\mathrm{P}\left(|\lambda_j - \gamma_j| \geqslant \frac{u}{N^{2/3}(N - j)^{1/3}}\right) \leqslant 4 \exp\left(-\frac{C^2 u^2}{C' \log(N - j) + Cu}\right). \tag{20}$$

*Proof.* Set $C = 2^{-5/6}(3\pi)^{-2/3}$. Let $K > 20\sqrt{2}$ and $\eta \in (0, \frac{1}{2}]$. Take $\alpha \in \left(\frac{20\sqrt{2}}{K}, 1\right)$ and set $c' = \alpha\left(\frac{3\pi}{2}\right)^{2/3}$. For $(1-\eta)N \leqslant j \leqslant N - K \log N$ and $u \geqslant 0$, set $u_{N,j} = \frac{u}{N^{2/3}(N-j)^{1/3}}$. As in the proof of Proposition 6, evaluating the probability $\mathrm{P}(|\lambda_j - \gamma_j| > u_{N,j})$ yields

$$\mathrm{P}\left(\lambda_j > \gamma_j + u_{N,j}\right) \leqslant \mathrm{P}\left(|\mathcal{N}_{\gamma_j + u_{N,j}} - N\rho_{\gamma_j + u_{N,j}}| > N(\rho_{\gamma_j + u_{N,j}} - \rho_{\gamma_j})\right).$$

But, if $u \leqslant N^{2/3}(N-j)^{1/3}(2 - \gamma_j)$,

$$\rho_{\gamma_j + u_{N,j}} - \rho_{\gamma_j} \geqslant \frac{\sqrt{2}}{3\pi}(2 - \gamma_j)^{1/2} u_{N,j}.$$

Similarly to the proof of Proposition 6, this condition holds if $u \leqslant c'(N-j)$. In this case,

$$\mathrm{P}\left(\lambda_j > \gamma_j + u_{N,j}\right) \leqslant \mathrm{P}\left(|\mathcal{N}_{\gamma_j + u_{N,j}} - N\rho_{\gamma_j + u_{N,j}}| > Cu + C_1\right),$$

when $u \geqslant c = \frac{C_1}{C}$. Then, applying (9) leads to

$$\mathrm{P}\left(\lambda_j > \gamma_j + u_{N,j}\right) \leqslant 2\exp\left(-\frac{C^2 u^2}{2\sigma_{\gamma_j + u_{N,j}}^2 + Cu}\right)$$

for all $c \leqslant u \leqslant c'(N-j)$. Gustavsson's result (19) gives a bound on $\sigma_{\gamma_j + u_{N,j}}^2$. Set $t_N = \gamma_j + u_{N,j}$. As $j \geqslant (1-\eta)N$ and $u \geqslant 0$, $0 \leqslant 2 - t_N \leqslant 2 - \gamma_j \leqslant \left(\frac{3\pi}{2}\eta\right)^{2/3} = \tilde{\eta}$ for all $N$. Moreover,

$$
\begin{aligned}
N(2 - t_N)^{3/2} &= N(2 - \gamma_j - u_{N,j})^{3/2} \\
&= N(2 - \gamma_j)^{3/2}\left(1 - \frac{u_{N,j}}{2 - \gamma_j}\right)^{3/2} \\
&\geqslant \frac{3\pi}{2}(N-j)\left(1 - \frac{c'(N-j)^{2/3}}{N^{2/3}(2 - \gamma_j)}\right)^{3/2} \\
&\geqslant \frac{3\pi}{2}(1 - \alpha)^{3/2} K \log N \\
&\geqslant \tilde{K}\log N,
\end{aligned}
$$

where $\tilde{K} = \frac{3\pi}{2}(1 - \alpha)^{3/2} K > 0$. From (19), for all $c \leqslant u \leqslant c'(N-j)$,

$$\mathrm{Var}(\mathcal{N}_{\gamma_j + u_{N,j}}) \leqslant c_{\tilde{\eta}, \tilde{K}} \log\left(N(2 - t_N)^{3/2}\right).$$

But

$$
\begin{aligned}
N(2 - t_N)^{3/2} &= N(2 - \gamma_j - u_{N,j})^{3/2} \\
&\leqslant N(2 - \gamma_j)^{3/2} \\
&\leqslant \frac{3\pi}{\sqrt{2}}(N-j).
\end{aligned}
$$

Hence $\log\left(N(2 - t_N)^{3/2}\right) \leqslant \log(N-j) + \log(\frac{3\pi}{\sqrt{2}})$. As $K > 20\sqrt{2}$ and $N \geqslant 2$, $N - j \geqslant K \log N \geqslant \frac{3\pi}{\sqrt{2}}$ and $\mathrm{Var}(\mathcal{N}_{\gamma_j + u_{N,j}}) \leqslant 2c_{\tilde{\eta}, \tilde{K}} \log(N-j)$. Therefore

$$\mathrm{P}\left(\lambda_j > \gamma_j + u_{N,j}\right) \leqslant 2\exp\left(-\frac{C^2 u^2}{4c_{\tilde{\eta}, \tilde{K}} \log(N-j) + Cu}\right).$$

The proof is concluded similarly to Proposition 6. $\qquad\square$

On the basis of Proposition 10, we may then conclude the proof of Theorem 9.

*Proof of Theorem 9.* Setting $Z = N^{2/3}(N-j)^{1/3}|\lambda_j - \gamma_j|$,

$$
\begin{aligned}
\mathrm{E}[Z^2] &= \int_0^\infty \mathrm{P}(Z \geqslant v) 2v \, dv \\
&= \int_0^c \mathrm{P}(Z \geqslant v) 2v \, dv + \int_c^{\frac{C'}{C} \log(N-j)} \mathrm{P}(Z \geqslant v) 2v \, dv \\
&\quad + \int_{\frac{C'}{C} \log(N-j)}^{c'(N-j)} \mathrm{P}(Z \geqslant v) 2v \, dv + \int_{c'(N-j)}^\infty \mathrm{P}(Z \geqslant v) 2v \, dv \\
&\leqslant c^2 + J_1 + J_2 + J_3.
\end{aligned}
$$

Repeating the computations carried out with $I_2$ in the proof of Theorem 5 yields

$$
J_3 \leqslant 2AN^{11/6}(N-j)^{2/3} \exp\left( -\frac{1}{2} \frac{C^2 c'^2 (N-j)^2}{C' \log(N-j) + Cc'(N-j)} \right),
$$

where $A > 0$ is a numerical constant. For $N$ large enough (depending on $\eta$ and $K$), $C' \log(N-j) \leqslant Cc'(N-j)$ and

$$
J_3 \leqslant 2AN^{5/2} \exp\left( -\frac{Cc'}{4}(N-j) \right).
$$

Then, as $N - j \geqslant K \log N$,

$$
J_3 \leqslant 2AN^{5/2} \exp\left( -\frac{KCc'}{4} \log N \right).
$$

But $\frac{KCc'}{4} = \frac{K\alpha}{8\sqrt{2}} > \frac{5}{2}$. The right-hand side goes thus to 0 when $N$ goes to infinity. As a consequence, there exists a constant $C(\eta, K) > 0$ such that

$$
J_3 \leqslant C(\eta, K).
$$

The integral $J_1$ is handled as $I_1$, using that, in the range $v \leqslant \frac{C'}{C} \log(N-j)$,

$$
P(Z \geqslant v) \leqslant 4 \exp\left( -\frac{B}{\log(N-j)} v^2 \right),
$$

where $B = B(\eta, K) = \frac{C^2}{2C'}$ (this is due to Proposition 10). Hence, there exists a constant $C(\eta, K)$ such that

$$
J_1 \leqslant C(\eta, K) \log(N-j).
$$

Finally, $J_2$ is handled similarly. From Proposition 10, in the range $\frac{C'}{C} \log(N-j) \leqslant v \leqslant c'(N-j)$,

$$
P(Z \geqslant v) \leqslant 4 \exp\left( -\frac{C}{2} v \right).
$$

Thus

$$
J_2 \leqslant 4 \int_{\frac{C'}{C} \log(N-j)}^{c'(N-j)} \exp\left( -\frac{C}{2} v \right) 2v \, dv \leqslant 4 \int_0^\infty \exp\left( -\frac{C}{2} v \right) 2v \, dv.
$$

12

Then $J_2$ is bounded by a constant, which is independent of $\eta$ and $K$. There exists thus a constant $C > 0$ such that
$$J_2 \leqslant C.$$
Summarizing the previous steps, $\mathrm{E}\left[Z^2\right] \leqslant C(\eta, K) \log(N - j)$. Therefore
$$\mathrm{E}\left[|\lambda_j - \gamma_j|^2\right] \leqslant C(\eta, K) \frac{\log(N - j)}{N^{2/3}(N - j)^{1/3}},$$
which is the claim. $\qquad\square$

# 2 Variance bounds for Wigner Hermitian matrices

As announced, the goal of this section is to prove Theorems 1, 2 and 3 for Wigner Hermitian matrices. The eigenvalues of a Wigner Hermitian matrix do not form a determinantal process. Therefore it does not seem easy to provide deviation inequalities for the counting function and for individual eigenvalues. However the sharp non-asymptotic bounds established in the Gaussian case can still be reached by a comparison procedure.

## 2.1 Localization of the eigenvalues and the Four Moment Theorem

Two main recent theorems will be used in order to carry out this comparison procedure. First, Erdös, Yau and Yin proved in [8] a Localization Theorem which gives a high probability non-asymptotic bound on the distance between an eigenvalue $\lambda_j$ and its theoretical value $\gamma_j$. Secondly, Tao and Vu's Four Moment Theorem (see [26] and [27]) provides a very useful non-asymptotic bound on the error made by approximating a statistics of the eigenvalues of a Wigner matrix by the same statistics but with the eigenvalues of a GUE matrix.

Let $M_N$ be a Wigner Hermitian matrix. Say that $M_N$ satisfies condition $(C0)$ if the real part $\xi$ and the imaginary part $\tilde{\xi}$ of $(M_N)_{ij}$ are independent and have an exponential decay: there are two positive constants $B_1$ and $B_2$ such that
$$\mathrm{P}\left(|\xi| \geqslant t^{B_1}\right) \leqslant e^{-t} \quad \text{and} \quad \mathrm{P}\left(|\tilde{\xi}| \geqslant t^{B_1}\right) \leqslant e^{-t}$$
for all $t \geqslant B_2$.

**Theorem 11** (Localization [8]). *Let $M_N$ be a random Hermitian matrix whose entries satisfy condition $(C0)$. There are positive universal constants $c$ and $C$ such that, for any $1 \leqslant j \leqslant N$,*
$$\mathrm{P}\left(|\lambda_j - \gamma_j| \geqslant (\log N)^{C \log \log N} N^{-2/3} \min(j, N + 1 - j)^{-1/3}\right) \leqslant C e^{-(\log N)^{c \log \log N}}. \quad (21)$$

This a strong localization result and it almost yields the correct order on the bound on the variance. Indeed, by means of the Cauchy-Schwarz inequality and (13), it can be shown that, for $\eta N \leqslant j \leqslant (1 - \eta)N$,

$$\operatorname{Var}(\lambda_j) \leqslant \operatorname{E}\big[(\lambda_j - \gamma_j)^2\big] \leqslant \frac{C(\eta)}{N^2}(\log N)^{C \log \log N}.$$

In Tao and Vu's recent paper [24] on deviation inequalities, the authors proved a more precise localization result. They indeed established a bound similar to (21) but with $(\log N)^A$ instead of $(\log N)^{C \log \log N}$ (where $A > 0$ is fixed). However this more precise bound is of no help below, as the final bounds on the variances remain unchanged.

We turn now to Tao and Vu's Four Moment Theorem, in order to compare $W_N$ with a GUE matrix $W_N'$. Say that two complex random variables $\xi$ and $\xi'$ match to order $k$ if

$$\operatorname{E}\big[\Re(\xi)^m \Im(\xi)^l\big] = \operatorname{E}\big[\Re(\xi')^m \Im(\xi')^l\big]$$

for all $m, l \geqslant 0$ such that $m + l \leqslant k$.

**Theorem 12** (Four Moment Theorem [26], [27]). *There exists a small positive constant $c_0$ such that the following holds. Let $M_N = (\xi_{ij})_{1 \leqslant i, j \leqslant N}$ and $M_N' = (\xi_{ij}')_{1 \leqslant i, j \leqslant N}$ be two random Wigner Hermitian matrices satisfying condition (C0). Assume that, for $1 \leqslant i < j \leqslant N$, $\xi_{ij}$ and $\xi_{ij}'$ match to order 4 and that, for $1 \leqslant i \leqslant N$, $\xi_{ii}$ and $\xi_{ii}'$ match to order 2. Set $A_N = \sqrt{N} M_N$ and $A_N' = \sqrt{N} M_N'$. Let $G : \mathbb{R} \to \mathbb{R}$ be a smooth fonction such that:*

$$\forall\, 0 \leqslant k \leqslant 5, \quad \forall x \in \mathbb{R}, \quad \big|G^{(k)}(x)\big| \leqslant N^{c_0}. \tag{22}$$

*Then, for all $1 \leqslant i \leqslant N$ and for $N$ large enough (depending on constants $B_1$ and $B_2$ in condition (C0)),*

$$\big|\operatorname{E}\big[G(\lambda_i(A_N))\big] - \operatorname{E}\big[G(\lambda_i(A_N'))\big]\big| \leqslant N^{-c_0}. \tag{23}$$

Actually Tao and Vu proved this theorem in a more general form, involving a finite number of eigenvalues. In this work, it will only be used with one given eigenvalue. See [26] and [27] for more details.

It should be mentionned that Tao and Vu extended in [26] Gustavsson's result (see equation (1)) via this theorem. By means of a smooth bump function $G$, they compared the probability for $\lambda_j$ to be in a given interval for a non-Gaussian matrix with almost the same probability but for a GUE matrix. Applying this technique to $\operatorname{P}\big(|\lambda_j - \gamma_j| > \frac{u}{N}\big)$ in order to extend directly the deviation inequality leads to the following in the general Wigner case: for all $\eta N \leqslant j \leqslant (1 - \eta)N$ and for all $C' \leqslant u \leqslant c'N$,

$$\operatorname{P}\Big(|\lambda_j - \gamma_j| > \frac{u}{N}\Big) \leqslant C \exp\Big(-\frac{u^2}{c \log N + u}\Big) + O(N^{-c_0}),$$

where $C$, $C'$, $c$ and $c'$ are positive constants depending only on $\eta$. The bound is not exponential anymore and is not enough to conclude towards sharp bounds on the variance or higher moments.

## 2.2   Comparison with Gaussian matrices

Let $M_N$ be a Hermitian Wigner matrix and $M'_N$ be a GUE matrix such that they satisfy the hypotheses of Theorem 12. As the function $G : x \in \mathbb{R} \mapsto x^2$ does not satisfy (22), Theorem 12 will be applied to a truncation of $G$. Theorem 11 will provide a small area around the theoretical location $\gamma_j$ where the eigenvalue $\lambda_j$ is very likely to be in, so that the error due to the truncation will be well controlled. Note that this procedure is valid for eigenvalues in the bulk and at the edge of the spectrum, as well as for intermediate eigenvalues.

Let $1 \leqslant j \leqslant N$. Set $R_N^{(j)} = (\log N)^{C \log\log N} N^{1/3} \min(j, N+1-j)^{-1/3}$ and $\varepsilon_N = Ce^{-(\log N)^{c \log\log N}}$. Then Theorem 11 leads to:

$$\mathrm{P}\left(|\lambda_j - \gamma_j| \geqslant \frac{R_N^{(j)}}{N}\right) \leqslant \varepsilon_N. \tag{24}$$

Let $\psi$ be a smooth function with support $[-2, 2]$ and values in $[0, 1]$ such that $\psi(x) = \frac{1}{10}x^2$ for all $x \in [-1; 1]$. Set $G_j : x \in \mathbb{R} \mapsto \psi\left(\frac{x - N\gamma_j}{R_N^{(j)}}\right)$. We want to apply Tao and Vu's Four Moment Theorem 12 to $G_j$. As $\psi$ is smooth and has compact support, its first five derivatives are bounded by $M > 0$. Then, for all $0 \leqslant k \leqslant 5$, for all $x \in \mathbb{R}$,

$$\left|G_j^{(k)}(x)\right| \leqslant \frac{M}{(R_N^{(j)})^k} \leqslant N^{c_0},$$

where the last inequality holds for $N$ large enough (depending only on $M$ and $c_0$). Then, the Four Moment Theorem 12 yields:

$$\left|\mathrm{E}\left[G_j(\lambda_j(A_N))\right] - \mathrm{E}\left[G_j(\lambda_j(A'_N))\right]\right| \leqslant N^{-c_0} \tag{25}$$

for large enough $N$. But

$$\mathrm{E}\left[G_j(\lambda_j(A_N))\right]$$

$$= \frac{1}{10}\mathrm{E}\left[\left(\frac{\lambda_j(A_N) - N\gamma_j}{R_N^{(j)}}\right)^2 \mathbf{1}_{\frac{|\lambda_j(A_N) - N\gamma_j|}{R_N^{(j)}} \leqslant 1}\right] + \mathrm{E}\left[G_j(\lambda_j(A_N)) \mathbf{1}_{\frac{|\lambda_j(A_N) - N\gamma_j|}{R_N^{(j)}} > 1}\right]$$

$$= \frac{N^2}{10\left(R_N^{(j)}\right)^2}\mathrm{E}\left[(\lambda_j - \gamma_j)^2 \mathbf{1}_{|\lambda_j - \gamma_j| \leqslant \frac{R_N^{(j)}}{N}}\right] + \mathrm{E}\left[G_j(\lambda_j(A_N)) \mathbf{1}_{\frac{|\lambda_j(A_N) - N\gamma_j|}{R_N^{(j)}} > 1}\right].$$

On the one hand,

$$\mathrm{E}\left[G_j(\lambda_j(A_N)) \mathbf{1}_{\frac{|\lambda_j - N\gamma_j|}{R_N^{(j)}} > 1}\right] \leqslant \mathrm{P}\left(|\lambda_j(W_N) - \gamma_j| > \frac{R_N^{(j)}}{N}\right) \leqslant \varepsilon_N.$$

On the other hand,

$$\mathrm{E}\left[(\lambda_j - \gamma_j)^2 \mathbf{1}_{|\lambda_j - \gamma_j| \leqslant \frac{R_N^{(j)}}{N}}\right] = \mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] - \mathrm{E}\left[(\lambda_j - \gamma_j)^2 \mathbf{1}_{|\lambda_j - \gamma_j| > \frac{R_N^{(j)}}{N}}\right].$$

From the Cauchy-Schwarz inequality and (13),

$$\mathrm{E}\left[(\lambda_j - \gamma_j)^2 \mathbf{1}_{|\lambda_j - \gamma_j| > \frac{R_N^{(j)}}{N}}\right] \leqslant \sqrt{\mathrm{E}\left[(\lambda_j - \gamma_j)^4\right]\mathrm{P}\left(|\lambda_j - \gamma_j| > \frac{R_N^{(j)}}{N}\right)}$$

$$\leqslant A\sqrt{N\varepsilon_N}$$

15

where $A > 0$ is a numerical constant. Then

$$
\begin{aligned}
\mathrm{E}\left[G_j(\lambda_j(A_N))\right] &= \frac{N^2}{10\left(R_N^{(j)}\right)^2}\left(\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] + O\left(N^{1/2}\varepsilon_N^{1/2}\right)\right) + O(\varepsilon_N) \\
&= \frac{N^2}{10\left(R_N^{(j)}\right)^2}\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] + O\left(N^{5/2}\varepsilon_N^{1/2}\left(R_N^{(j)}\right)^{-2}\right) + O(\varepsilon_N).
\end{aligned}
$$

Repeating the same computations gives similarly

$$
\mathrm{E}\left[G_j(\lambda_j(A_N'))\right] = \frac{N^2}{10\left(R_N^{(j)}\right)^2}\mathrm{E}\left[(\lambda_j' - \gamma_j)^2\right] + O\left(N^{5/2}\varepsilon_N^{1/2}\left(R_N^{(j)}\right)^{-2}\right) + O(\varepsilon_N).
$$

Then (25) leads to

$$
\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] = \mathrm{E}\left[(\lambda_j' - \gamma_j)^2\right] + O\left(N^{1/2}\varepsilon_N^{1/2}\right) + O\left(N^{-2}\left(R_N^{(j)}\right)^2\varepsilon_N\right) + O\left(\left(R_N^{(j)}\right)^2 N^{-c_0-2}\right).
$$

As the first two error terms are smaller than the third one, the preceding equation becomes

$$
\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] = \mathrm{E}\left[(\lambda_j' - \gamma_j)^2\right] + O\left(\left(R_N^{(j)}\right)^2 N^{-c_0-2}\right). \tag{26}
$$

## 2.3  Combining the results

We distinguish between the bulk, the edge and the intermediate cases. Note that the constants $C(\eta)$ and $C$ depend on the constants $B_1$ and $B_2$ in condition $(C0)$.

### 2.3.1  Eigenvalues in the bulk of the spectrum

Let $0 < \eta \leqslant \frac{1}{2}$ and $\eta N \leqslant j \leqslant (1-\eta)N$. From Theorem 5, $\mathrm{E}\left[(\lambda_j' - \gamma_j)^2\right] \leqslant C(\eta)\frac{\log N}{N^2}$. Thus, from (26), it remains to show that the error term is smaller than $\frac{\log N}{N^2}$. But

$$
R_N^{(j)} = (\log N)^{C\log\log N} N^{1/3}\min(j, N+1-j)^{-1/3} \leqslant \eta^{-1/3}(\log N)^{C\log\log N}.
$$

Then $\left(R_N^{(j)}\right)^2 N^{-c_0-2} = o_\eta\left(\frac{\log N}{N^2}\right)$. As a consequence,

$$
\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] = \mathrm{E}\left[(\lambda_j' - \gamma_j)^2\right] + o_\eta\left(\frac{\log N}{N^2}\right)
$$

and we get the desired result

$$
\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] \leqslant C(\eta)\frac{\log N}{N^2}.
$$

### 2.3.2  Eigenvalues at the edge of the spectrum

From Corollary 8, $\mathrm{E}\left[(\lambda_N' - \gamma_N)^2\right] = \mathrm{E}\left[(\lambda_N' - 2)^2\right] \leqslant CN^{-4/3}$. By means of (26), it remains to prove that the error term is smaller than $N^{-4/3}$. We have

$$
R_N^{(N)} = (\log N)^{C\log\log N} N^{1/3}.
$$

16

Consequently $\left(R_N^{(N)}\right)^2 N^{-c_0-2} = o\left(N^{-4/3}\right)$. Then

$$\mathrm{E}\left[(\lambda_N - 2)^2\right] = \mathrm{E}\left[(\lambda_N' - 2)^2\right] + o\left(N^{-4/3}\right)$$

and

$$\mathrm{E}\left[(\lambda_N - 2)^2\right] \leqslant C N^{-4/3}.$$

As for Gaussian matrices, the same result is available for the smallest eigenvalue $\lambda_1$.

### 2.3.3 Eigenvalues between the bulk and the edge of the spectrum

Let $0 < \eta \leqslant \frac{1}{2}$, $K > 20\sqrt{2}$ and $(1-\eta)N \leqslant j \leqslant N - K \log N$. From Theorem 9, $\mathrm{E}\left[(\lambda_j' - \gamma_j)^2\right] \leqslant C(\eta, K)\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}$. Thus, from (26), it remains to show that the error term is smaller than $\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}$. But

$$R_N^{(j)} = (\log N)^{C \log \log N} N^{1/3}(N + 1 - j)^{-1/3}.$$

Then $\left(R_N^{(j)}\right)^2 N^{-c_0-2} = o\left(\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}\right)$. As a consequence,

$$\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] = \mathrm{E}\left[(\lambda_j' - \gamma_j)^2\right] + o\left(\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}\right)$$

and we get the desired result

$$\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] \leqslant C(\eta, K)\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}.$$

A similar result holds for the left-side of the spectrum.

## 3   Real matrices

The goal of this section is to prove Theorems 1, 2 and 3 for real Wigner matrices. Tao and Vu's Four Moment Theorem (Theorem 12) as well as Erdös, Yau and Yin's Localization Theorem (Theorem 11) still hold for real Wigner matrices. Section 2 is therefore valid for real matrices. The point is then to establish the results in the GOE case.

As announced in Section 1.2, the variance of eigenvalues at the edge of the spectrum is known to be bounded by $N^{-4/3}$ for GOE matrices (see [15]). The conclusion for the smallest and largest eigenvalues is then established for large families of real symmetric Wigner matrices.

$$\mathrm{Var}(\lambda_N) \leqslant \frac{\tilde{C}}{N^{4/3}} \quad \text{and} \quad \mathrm{Var}(\lambda_1) \leqslant \frac{\tilde{C}}{N^{4/3}}.$$

For eigenvalues in the bulk of the spectrum, O'Rourke proved in [18] a Central Limit Theorem which is very similar to the one established by Gustavsson in [14]. In particular, the normalisation is still of the order of $\left(\frac{\log N}{N^2}\right)^{1/2}$ and differs from the complex case only by a constant. It is therefore natural to expect the same bound on

the variance for GOE matrices. The situation is completely similar for intermediate eigenvalues. But GOE matrices do not have the same determinantal properties as GUE matrices, and it is therefore not clear that a deviation inequality (similar to (9)) holds for the eigenvalue counting function. However, as explained by O'Rourke in [18], GOE and GUE matrices are linked by interlacing formulas established by Forrester and Rains (see [9]). These formulas lead to the following relation between the eigenvalue counting functions in the complex and in the real cases: for all $t \in \mathbb{R}$,

$$\mathcal{N}_t(W_N^{\mathbb{C}}) = \frac{1}{2}\big(\mathcal{N}_t(W_N^{\mathbb{R}}) + \mathcal{N}_t(\tilde{W}_N^{\mathbb{R}})\big) + \zeta_N(t), \qquad (27)$$

where $W_N^{\mathbb{C}} = \frac{1}{\sqrt{N}}M_N^{\mathbb{C}}$ and $M_N^{\mathbb{C}}$ is from the GUE, $W_N^{\mathbb{R}} = \frac{1}{\sqrt{N}}M_N^{\mathbb{R}}$, $\tilde{W}_N^{\mathbb{R}} = \frac{1}{\sqrt{N}}\tilde{M}_N^{\mathbb{R}}$ and $M_N^{\mathbb{R}}$ and $\tilde{M}_N^{\mathbb{R}}$ are independent matrices from the GOE and $\zeta_N(t)$ takes values in $\{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$. See [18] for more details.

The aim is now to establish a deviation inequality for the eigenvalue counting function similar to (9). From (9), we know that for all $u \geqslant 0$,

$$\mathrm{P}\big(|\mathcal{N}_t(W_N^{\mathbb{C}}) - N\rho_t| \geqslant u + C_1\big) \leqslant 2\exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big).$$

Set $C_1' = C_1 + 1$ and let $u \geqslant 0$. We can then write

$$\mathrm{P}\big(\mathcal{N}_t(W_N^{\mathbb{R}}) - N\rho_t \geqslant u + C_1'\big)^2$$
$$= \mathrm{P}\Big(\mathcal{N}_t(W_N^{\mathbb{R}}) - N\rho_t \geqslant u + C_1', \ \mathcal{N}_t(\tilde{W}_N^{\mathbb{R}}) - N\rho_t \geqslant u + C_1'\Big)$$
$$\leqslant \mathrm{P}\Big(\tfrac{1}{2}\big(\mathcal{N}_t(W_N^{\mathbb{R}}) + \mathcal{N}_t(\tilde{W}_N^{\mathbb{R}})\big) - N\rho_t \geqslant u + C_1'\Big)$$
$$\leqslant \mathrm{P}\big(\mathcal{N}_t(W_N^{\mathbb{C}}) - N\rho_t \geqslant u + C_1' - 1\big)$$
$$\leqslant 2\exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big).$$

Repeating the computations for $\mathrm{P}\big(\mathcal{N}_t(W_N^{\mathbb{R}}) - N\rho_t \leqslant -u - C_1'\big)$ and combining with the preceding yield

$$\mathrm{P}\big(|\mathcal{N}_t(W_N^{\mathbb{R}}) - N\rho_t| \geqslant u + C_1'\big) \leqslant 2\sqrt{2}\exp\Big(-\frac{u^2}{4\sigma_t^2 + 2u}\Big). \qquad (28)$$

Note that $\sigma_t^2$ is still the variance of $\mathcal{N}_t(W_N^{\mathbb{C}})$ in the preceding formula.

What remains then to be proved is very similar to the complex case. From (28) and Gustavsson's bounds on the variance $\sigma_t^2$ (see (10) for the bulk case and (19) for the intermediate case), deviation inequalities for individual eigenvalues can be deduced, as was done to prove Propositions 6 and 10. It is then straightforward to derive the announced bounds on the variances for GOE matrices. The argument developed in Section 2 in order to extend the GUE results to large families of Hermitian Wigner matrices can be reproduced to reach the desired bounds on the variances of eigenvalues in the bulk and between the bulk and the edge of the spectrum for families of real Wigner matrices. Then there exists a constant $C(\eta) > 0$ such that for all $\eta N \leqslant j \leqslant (1 - \eta)N$,

$$\mathrm{Var}(\lambda_j) \leqslant \mathrm{E}\big[(\lambda_j - \gamma_j)^2\big] \leqslant C(\eta)\frac{\log N}{N^2},$$

and there exists a constant $C(\eta, K) > 0$ such that for all $(1-\eta)N \leqslant j \leqslant N - K \log N$ (and similarly for the left-side of the spectrum),

$$\text{Var}(\lambda_j) \leqslant \text{E}\left[(\lambda_j - \gamma_j)^2\right] \leqslant C(\eta, K) \frac{\log(N - j)}{N^{4/3}(N - j)^{2/3}}.$$

# 4    A corollary on the $2$-Wasserstein distance

The bounds on the variances, more exactly on $\text{E}[(\lambda_j - \gamma_j)^2]$, developed in the preceding sections lead to a bound on the rate of convergence of the empirical spectral measure $L_N$ towards the semicircle law $\rho_{sc}$ in terms of 2-Wasserstein distance. Recall that $W_2(L_N, \rho_{sc})$ is a random variable defined by

$$W_2(L_N, \rho_{sc}) = \inf \left( \int_{\mathbb{R}^2} |x - y|^2 \, d\pi(x, y) \right)^{1/2},$$

where the infimum is taken over all probability measures $\pi$ on $\mathbb{R}^2$ with respective marginals $L_N$ and $\rho_{sc}$. To achieve the expected bound, we rely on another expression of $W_2$ in terms of distribution functions, namely

$$W_2^2(L_N, \rho_{sc}) = \int_0^1 \left( F_N^{-1}(x) - G^{-1}(x) \right)^2 dx, \tag{29}$$

where $F_N^{-1}$ (respectively $G^{-1}$) is the generalized inverse of the distribution function $F_N$ (respectively $G$) of $L_N$ (respectively $\rho_{sc}$) (see for example [31]). On the basis of this representation, the following statement may be derived.

**Proposition 13.** *There exists a universal constant $C > 0$ such that for all $N \geqslant 1$,*

$$W_2^2(L_N, \rho_{sc}) \leqslant \frac{2}{N} \sum_{j=1}^{N} (\lambda_j - \gamma_j)^2 + \frac{C}{N^2}. \tag{30}$$

*Proof.* From (29),

$$W_2^2(L_N, \rho_{sc}) = \int_0^1 \left( F_N^{-1}(x) - G^{-1}(x) \right)^2 dx.$$

Then,

$$W_2^2(L_N, \rho_{sc}) = \sum_{j=1}^{N} \int_{\frac{j-1}{N}}^{\frac{j}{N}} \left( \lambda_j - G^{-1}(x) \right)^2 dx$$

$$\leqslant \frac{2}{N} \sum_{j=1}^{N} (\lambda_j - \gamma_j)^2 + 2 \sum_{j=1}^{N} \int_{\frac{j-1}{N}}^{\frac{j}{N}} \left( \gamma_j - G^{-1}(x) \right)^2 dx.$$

But $\gamma_j = G^{-1}\left(\frac{j}{N}\right)$ and $G^{-1}$ is non-decreasing. Therefore, $\left| \gamma_j - G^{-1}(x) \right| \leqslant \gamma_j - \gamma_{j-1}$ for all $x \in \left[ \frac{j-1}{N}, \frac{j}{N} \right]$. Consequently,

$$W_2^2(L_N, \rho_{sc}) \leqslant \frac{2}{N} \sum_{j=1}^{N} (\lambda_j - \gamma_j)^2 + \frac{2}{N} \sum_{j=1}^{N} (\gamma_j - \gamma_{j-1})^2. \tag{31}$$

But if $j - 1 \geqslant \frac{N}{2}$ (and therefore $\gamma_{j-1} \geqslant 0$),

$$
\begin{aligned}
\frac{1}{N} &= \int_{\gamma_{j-1}}^{\gamma_j} \frac{1}{2\pi} \sqrt{4 - x^2} \, dx \\
&\geqslant \frac{1}{\sqrt{2}\pi} \int_{\gamma_{j-1}}^{\gamma_j} \sqrt{2 - x} \, dx \\
&\geqslant \frac{\sqrt{2}}{3\pi} (2 - \gamma_{j-1})^{3/2} \left( 1 - \left( 1 - \frac{\gamma_j - \gamma_{j-1}}{2 - \gamma_{j-1}} \right)^{3/2} \right) \\
&\geqslant \frac{\sqrt{2}}{3\pi} (2 - \gamma_{j-1})^{1/2} (\gamma_j - \gamma_{j-1}) \\
&\geqslant \frac{\sqrt{2}}{3\pi} \left( \frac{3\pi}{2} \frac{N - j + 1}{N} \right)^{1/3} (\gamma_j - \gamma_{j-1}),
\end{aligned}
$$

from (12). Then

$$
\gamma_j - \gamma_{j-1} \leqslant \frac{(3\pi)^{2/3} 2^{-1/6}}{N^{2/3}(N - j + 1)^{2/3}}.
$$

It may be shown that a similar bound holds if $j - 1 \leqslant \frac{N}{2}$. As a summary, there exists a universal constant $c > 0$ such that, for all $j \geqslant 2$,

$$
\gamma_j - \gamma_{j-1} \leqslant \frac{c}{N^{2/3} \min(j, N + 1 - j)^{1/3}}. \tag{32}
$$

This yields

$$
\sum_{j=1}^{N} (\gamma_j - \gamma_{j-1})^2 \leqslant \frac{c^2}{N^{4/3}} \sum_{j=1}^{N} \frac{1}{\min(j, N + 1 - j)^{2/3}} \leqslant \frac{C}{N},
$$

where $C > 0$ is a universal constant. Then (31) becomes

$$
W_2^2(L_N, \rho_{sc}) \leqslant \frac{2}{N} \sum_{j=1}^{N} (\lambda_j - \gamma_j)^2 + \frac{C}{N^2},
$$

where $C > 0$ is a universal constant, which is the claim. $\qquad\square$

*Proof of Corollary 4.* Let $N \geqslant 2$. Due to Proposition 13,

$$
\mathrm{E}\left[ W_2^2(L_N, \rho_{sc}) \right] \leqslant \frac{2}{N} \sum_{j=1}^{N} \mathrm{E}\left[ (\lambda_j - \gamma_j)^2 \right] + \frac{C}{N^2}.
$$

We then make use of the bounds on $\mathrm{E}\left[ (\lambda_j - \gamma_j)^2 \right]$ produced in the previous sections.

Set $\eta \in (0, \frac{1}{2}]$ and $K > 20\sqrt{2}$ so that $K \log N \leqslant \eta N$. We first decompose

$$\sum_{j=1}^{N} \mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] = \sum_{j=1}^{K \log N} \mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] + \sum_{j=K \log N + 1}^{\eta N} \mathrm{E}\left[(\lambda_j - \gamma_j)^2\right]$$

$$+ \sum_{j=\eta N + 1}^{(1-\eta)N - 1} \mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] + \sum_{j=(1-\eta)N}^{N - K \log N - 1} \mathrm{E}\left[(\lambda_j - \gamma_j)^2\right]$$

$$+ \sum_{j=N - K \log N}^{N} \mathrm{E}\left[(\lambda_j - \gamma_j)^2\right]$$

$$= \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4 + \Sigma_5.$$

The sum $\Sigma_3$ will be bounded using the bulk case (Theorem 1), while Theorem 2 will be used to handle $\Sigma_2$ and $\Sigma_4$. A crude version of Theorem 11 will be enough to bound $\Sigma_1$ and $\Sigma_5$. To start with thus, from Theorem 1,

$$\Sigma_3 \leqslant \sum_{j=\eta N + 1}^{(1-\eta)N - 1} C(\eta) \frac{\log N}{N^2} \leqslant C(\eta) \frac{\log N}{N}.$$

Secondly, from Theorem 3,

$$\Sigma_2 + \Sigma_4 \leqslant \frac{C(\eta, K)}{N^{4/3}} \sum_{j=K \log N + 1}^{\eta N} \frac{\log j}{j^{2/3}} \leqslant C(\eta, K) \frac{\log N}{N}.$$

Next $\Sigma_1$ and $\Sigma_5$ have only $K \log N$ terms. If each term is bounded by $\frac{C}{N}$ where $C$ is a positive universal constant, we get that $\Sigma_1 + \Sigma_5 \leqslant \frac{2KC \log N}{N}$, which is enough to prove the desired result on $\sum_{j=1}^{N} \mathrm{E}\left[(\lambda_j - \gamma_j)^2\right]$. For $N$ large enough depending only on constant $C$ in Theorem 11, $\frac{1}{\sqrt{N}} \geqslant \frac{(\log N)^{C \log \log N}}{N^{2/3} \min(j, N+1-j)^{1/3}}$ and Theorem 11 yields

$$\mathrm{P}\left(|\lambda_j - \gamma_j| \geqslant \frac{1}{\sqrt{N}}\right) \leqslant C e^{-(\log N)^{c \log \log N}}.$$

Then, by the Cauchy-Schwarz inequality,

$$\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] \leqslant \mathrm{E}\left[(\lambda_j - \gamma_j)^2 \mathbf{1}_{|\lambda_j - \gamma_j| \leqslant \frac{1}{\sqrt{N}}}\right] + \mathrm{E}\left[(\lambda_j - \gamma_j)^2 \mathbf{1}_{|\lambda_j - \gamma_j| > \frac{1}{\sqrt{N}}}\right]$$

$$\leqslant \frac{1}{N} + \sqrt{\mathrm{E}\left[|\lambda_j - \gamma_j|^4\right]} \sqrt{\mathrm{P}\left(|\lambda_j - \gamma_j| > \frac{1}{\sqrt{N}}\right)}$$

$$\leqslant \frac{1}{N} + \sqrt{3} C N^{1/2} e^{-(\log N)^{c \log \log N}}.$$

As $\sqrt{3} C N^{1/2} e^{-(\log N)^{c \log \log N}} = o(\frac{1}{N})$, there exists a constant $C > 0$ such that $\mathrm{E}\left[(\lambda_j - \gamma_j)^2\right] \leqslant \frac{C}{N}$. Then

$$\Sigma_1 + \Sigma_5 \leqslant 2KC \frac{\log N}{N}.$$

As a consequence,

$$\sum_{j=1}^{N} \mathrm{E}\big[(\lambda_j - \gamma_j)^2\big] \leqslant C \frac{\log N}{N}.$$

Therefore

$$\mathrm{E}\big[W_2(L_N, \rho_{sc})^2\big] \leqslant C \frac{\log N}{N^2},$$

where $C > 0$ is a universal constant, which is the claim. The corollary is thus established. $\qquad\square$

The preceding Corollary 4 implies that

$$\mathrm{E}\big[W_1(L_N, \rho_{sc})\big] \leqslant C \frac{\sqrt{\log N}}{N}.$$

This is an improvement of Meckes and Meckes' rate of convergence obtained in [16]. Note however that the distance studied in [16] is the expected 1-Wasserstein distance between $L_N$ and its mean instead of $\rho_{sc}$. The rate of convergence in 1-Wasserstein distance can be furthermore compared to the rate of convergence in Kolmogorov distance. Indeed, if $\mu$ and $\nu$ are two probability measures on $\mathbb{R}$ such that $\nu$ has a bounded density function with respect to Lebesgue measure, $d_K(\mu, \nu) \leqslant c\sqrt{W_1(\mu, \nu)}$, where $c > 0$ depends on the bound on the density function. This is due to the fact that

$$W_1(\mu, \nu) = \sup_{f \, 1\text{-Lipschitz}} \left( \int_{\mathbb{R}} f d\mu - \int_{\mathbb{R}} f d\nu \right)$$

(see for example [31]) and $d_K(\mu, \nu) = \sup_{x \in \mathbb{R}} \big|\mu\big((-\infty, x]\big) - \nu\big((-\infty, x]\big)\big|$. Approximating $1_{(-\infty, x]}$ from above and from below by $\frac{1}{\varepsilon}$-Lipschitz functions and optimizing on $\varepsilon$ gives the result. Therefore, the preceding implies that

$$\mathrm{E}\big[d_K(L_N, \rho_{sc})\big] \leqslant c \frac{(\log N)^{1/4}}{N^{1/2}}$$

which is however far from Götze and Tikhomirov's recent bound [10]

$$\mathrm{E}\big[d_K(L_N, \rho_{sc})\big] \leqslant \frac{(\log N)^c}{N}.$$

# 5 Eigenvalue variance bounds for covariance matrices

This section provides the analogous non-asymptotic bounds on the variance of eigenvalues for covariance matrices. The proofs will be detailed in another redaction. Therefore, this section contains only the background and the results.

Random covariance matrices are defined by the following. Let $X$ be a $m \times n$ (real or complex) matrix, with $m \geqslant n$, such that its entries are independent, centered and have variance 1. Then $S_{m,n} = \frac{1}{n} X^* X$ is a random covariance matrix. An important example is the case when the entries of $X$ are Gaussian. Then $S_{m,n}$ belongs to the

so-called Laguerre Unitary Ensemble (LUE) if the entries of $X$ are complex and to the Laguerre Orthogonal Ensemble (LOE) if they are real. $S_{m,n}$ is Hermitian (or real symmetric) and therefore has $n$ real eigenvalues. As $m \geqslant n$, none of these eigenvalues is trivial. Furthermore, these eigenvalues are nonnegative and will be denoted by $0 \leqslant \lambda_1 \leqslant \cdots \leqslant \lambda_n$.

Similarly to Wigner's Theorem, the classical Marchenko-Pastur theorem states that, if $\frac{m}{n} \to \rho \geqslant 1$ when $n$ goes to infinity, the empirical spectral measure $L_{m,n} = \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j}$ converges almost surely to a deterministic measure, called the Marchenko-Pastur distribution of parameter $\rho$. This measure is compactly supported and is absolutely continuous with respect to Lebesgue measure, with density

$$d\mu_{MP(\rho)}(x) = \frac{1}{2\pi x}\sqrt{(b_\rho - x)(x - a_\rho)}1_{[a_\rho, b_\rho]}(x)dx,$$

where $a_\rho = (1 - \sqrt{\rho})^2$ and $b_\rho = (1 + \sqrt{\rho})^2$ (see for example [3]). Two different behaviors arise according to the value of $\rho$. If $\rho > 1$, $a_\rho > 0$ is a soft edge, which means that an eigenvalue $\lambda_j$ can be larger or smaller than $a_\rho$. On the contrary, if $\rho = 1$, $a_\rho = 0$ is called a hard edge: no eigenvalue can be less than $a_\rho$. Furthermore, the Marchenko-Pastur density function explodes at 0. It is the case in particular when $m = n$. We will denote by $\mu_{m,n}$ the approximate Marchenko-Pastur distribution whose density is defined by

$$\mu_{m,n}(x) = \frac{1}{2\pi x}\sqrt{(x - a_{m,n})(b_{m,n} - x)}1_{[a_{m,n}; b_{m,n}]}(x),$$

with $a_{m,n} = (1 - \sqrt{\frac{m}{n}})^2$ and $b_{m,n} = (1 + \sqrt{\frac{m}{n}})^2$.

The asymptotic behaviors of individual eigenvalues for LUE matrices have been known for some time and extended to more general covariance matrices in the last decade. For an eigenvalue in the bulk of the spectrum, i.e. $\lambda_j$ such that $\eta n \leqslant j \leqslant (1 - \eta)n$, a Central Limit Theorem was proved by Su (see [22]) and Tao-Vu (see [23]). From this theorem, the variance of such eigenvalues is guessed to be of the order of $\frac{\log n}{n^2}$. For right-side intermediate eigenvalues, i.e. $\lambda_j$ such that $\frac{j}{n} \to 1$ and $n - j \to \infty$, Su, and later Tao-Vu and Wang, proved a CLT (see [22], [23] and [32]). The variance appears to be of the order of $\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}$. Similar results probably hold for the left-side of the spectrum, when $\rho > 1$. Finally, for the smallest and the largest eigenvalues $\lambda_1$ and $\lambda_n$, CLTs were proved for Gaussian matrices by Borodin-Forrester (see [6]). Several authors extended these results to more general covariance matrices. The latest results are due to Tao-Vu and Wang (see [23] and [32]). Their variances are then guessed to be of the order of $n^{-4/3}$. It should be mentionned that the result for the smallest eigenvalue only holds when $\rho > 1$. When $\rho = 1$, in the case of a squared matrix, Edelman proved a CLT for the smallest eigenvalue $\lambda_1$, extended by Tao and Vu in [25], from which the variance is guessed to be of the order of $n^{-4}$.

The following statement summarizes a number of quantitative bounds on the eigenvalues of covariance matrices which are proved by methods similar to the ones developed in the preceding sections. For simplicity, we basically assume that $\rho > 1$. More precisely, we assume that $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2$ (where $A_1$ and $A_2$ are fixed constants) and that $S_{m,n}$ is a covariance matrix whose entries have an exponential

decay (condition $(C0)$) and have the same first four moments as those of a LUE matrix.

**Theorem 14.**    *1.* ***In the bulk of the spectrum.***
*Let $\eta \in (0, \frac{1}{2}]$. There exist a constant $C > 0$ (depending on $\eta$, $A_1$ and $A_2$) such that for all covariance matrix $S_{m,n}$, for all $\eta n \leqslant j \leqslant (1 - \eta)n$,*

$$\mathrm{Var}(\lambda_j) \leqslant C \frac{\log n}{n^2}.$$

*2.* ***Between the bulk and the edge of the spectrum.***
*There exists a constant $\kappa > 0$ (depending on $A_1$ and $A_2$) such that the following holds. For all $K > \kappa$, for all $\eta \in (0, \frac{1}{2}]$, there exists a constant $C > 0$ (depending on $K$, $\eta$, $A_1$ and $A_2$) such that for all covariance matrix $S_{m,n}$, for all $(1 - \eta)n \leqslant j \leqslant n - K \log n$,*

$$\mathrm{Var}(\lambda_j) \leqslant C \frac{\log(n - j)}{n^{4/3}(n - j)^{2/3}}.$$

*3.* ***At the edge of the spectrum.***
*There exists a constant $C > 0$ (depending on $A_1$ and $A_2$) such that, for all covariance matrix $S_{m,n}$,*
$$\mathrm{Var}(\lambda_n) \leqslant C n^{-4/3}.$$

As for Wigner matrices, the first two results of this theorem are first proved in the Gaussian case, using the fact that the eigenvalues of a LUE matrix form a determinantal process. It is then possible to derive a deviation inequality for the eigenvalue counting function and then for individual eigenvalues. Integrating leads to the results for LUE matrices. The result for the largest eigenvalue in the Gaussian case is already known, see [15]. These bounds are then extended to non-Gaussian matrices, relying on three recent papers. First, Pillai and Yin proved in [20] localization properties for individual eigenvalues of covariance matrices, very similar to the localization properties for Wigner matrices established by Erdös, Yau and Yin in [8]. Secondly, Tao and Vu (see [23]) and later Wang (see [32]) proved a Four Moment Theorem for these matrices. Combining these theorems as for Wigner matrices yield the theorem.

## Acknowledgements

# References

[1] G. Anderson, A. Guionnet, O. Zeitouni, An introduction to random matrices, Cambridge Studies in Advanced Mathematics 118, 2010.

[2] Z. D. Bai, Y. Q. Yin, Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix, Ann. Probab. 16 (1988), p. 1729-1741.

[3] Z. D. Bai, J. Silverstein, Spectral analysis of large dimensional random matrices. Second edition. Springer Series in Statistics. Springer, New York, 2010.

[4] J. Ben Hough, M. Krishnapur, Y. Peres, B. Virag, Determinantal processes and independence, Probability Surveys 3 (2006), p. 206-229.

[5] S. Bobkov, F. Götze, A. Tikhomirov, On concentration of empirical measures and convergence to the semi-circle law, Journal of Theoretical Probability 23 no.3 (2010), p. 792-823.

[6] A. Borodin, P. Forrester, Increasing subsequences and the hard-to-soft edge transition in matrix ensembles, J. Phys. A 36 (2003), 12, p. 2963-2981.

[7] R. M. Dudley, Real analysis and probability, Revised reprint of the 1989 original. Cambridge Studies in Advanced Mathematics, 74. Cambridge University Press, Cambridge, 2002.

[8] L. Erdös, H-T. Yau, J. Yin, Rigidity of eigenvalues of generalized Wigner matrices, arXiv:1007.4652, 2010.

[9] P. Forrester, E. Rains, Inter-relationships between orthogonal, unitary and symplectic matrix ensembles, Cambridge University Press, Cambridge, United Kingdom (2001), p. 171-208.

[10] F. Götze, A. Tikhomirov, On the rate of convergence to the semicircular law, arXiv:1109.0611, 2011.

[11] F. Götze, A. Tikhomirov, Rate of convergence to the semicircular law, Probab. Theory Relat. Fields 127 (2003), p. 228-276.

[12] F. Götze, A. Tikhomirov, The rate of convergence for spectra of GUE and LUE matrix ensembles, Cent. Eur. J. Math. 3 (2005), 4, p. 666-704.

[13] A. Guionnet, O. Zeitouni, Concentration of the spectral measure for large matrices, Electron. Comm. Probab. 5 (2000), p. 119-136.

[14] J. Gustavsson, Gaussian fluctuations of eigenvalues in the GUE, Ann. I. Poincaré 41 (2005), p. 151-178.

[15] M. Ledoux, B. Rider, Small deviations for $\beta$-ensembles, Electron. J. Probab. 15 (2010), 41, p. 1319-1343.

[16] E. Meckes, M. Meckes, Concentration and convergence rates for spectral measure of random matrices, to appear in Probab. Theory Related Fields, arXiv:1109.5997, 2011.

[17] M.L. Mehta, Random matrices, San Diego, Academic Press, second edition, 1991.

[18] S. O'Rourke, Gaussian fluctuations of eigenvalues in Wigner random matrices, J. Stat. Phys. 138 (2010), 6, p. 1045-1066.

[19] L. Pastur, M. Shcherbina, Eigenvalue distribution of large random matrices. Mathematical Surveys and Monographs, 171, American Mathematical Society, Providence, RI, 2011.

[20] N. Pillai, J. Yin, Universality of covariance matrices, 2011, arXiv:1110.2501.

[21] A. Soshnikov, Universality at the edge of the spectrum in Wigner random matrices, Comm. Math. Phys. 207 (1999), 3, p. 697-733.

[22] Z. Su, Gaussian fluctuations in complex sample covariance matrices, Electron. J. Probab. 11 (2006), 48, p. 1284-1320.

[23] T. Tao, V. Vu, Random covariance matrices: universality of local statistics of eigenvalues, Annals of Probability 40 (2012), 3, p. 1285-1315.

[24] T. Tao and V. Vu, Random matrices: sharp concentration of eigenvalues, submitted, arXiv:1201.4789, 2012.

[25] T. Tao, V. Vu, Random matrices: the distribution of the smallest singular values, Geom. Funct. Anal. 20 (2010), 1, p. 260-297.

[26] T. Tao and V. Vu, Random matrices: universality of local eigenvalues statistics, Acta Math. 206 (2011), 1, p. 127-204.

[27] T. Tao and V. Vu, Random matrices: universality of local eigenvalue statistics up to the edge, Comm. Math. Phys. 298 (2010), 2, p. 549-572.

[28] C. A. Tracy, H. Widom, Level-spacing distributions and the Airy kernel, Comm. Math. Phys. 159 (1994), 1, p. 151-174.

[29] A. W. van der Vaart, Asymptotic statistics. Cambridge Series in Statistical and Probabilistic Mathematics, 3. Cambridge University Press, Cambridge, 1998.

[30] R. Vershynin, Introduction to the non-asymptotic analysis of random matrices, Chapter 5 of Compressed sensing, theory and applications. Edited by Y. Eldar and G. Kutyniok. Cambridge University Press, 2012.

[31] C. Villani, Topics in optimal transportation, Graduate Studies in Mathematics, 58, American Mathematical Society, 2003.

[32] K. Wang, Random covariance matrices: universality of local statistics of eigenvalues up to the edge, 2011, arXiv:1104.4832.

Sandrine Dallaporta
Institut de Mathématiques de Toulouse, UMR 5219 du CNRS
Université de Toulouse, F-31062, Toulouse, France
sandrine.dallaporta@math.univ-toulouse.fr