Semicircle law for Wigner matrices and random matrices with weak but long range correlation

Joscha Henheik*

March 17, 2021

Abstract

The goal of this note is twofold: On the one hand, we give a modern proof of the local semicircle law for Wigner random matrices. It states that the resolvent, G, of an hermitian $N \times N$ random matrix H can be approximated by the solution M of a deterministic equation (Matrix Dyson Equation, MDE) in high probability sense, increasingly well, as the size N approaches infinity. The proof is based on showing that (i) G solves the MDE up to an error that is small in high-moment sense, and (ii) the MDE is stable against small additive perturbations.

On the other hand, we present a local law on perturbations of Wigner matrices by a fixed number of mutually orthogonal rank-1 random matrices, which is a basic example of an overall weak but long range correlation structure in H.

Contents

1 Introduction

In 1958, E. P. Wigner [14] noticed that the gaps of energy levels of heavy nuclei tend to follow the same universality pattern, independent of the specific material. As the appropriate theory for describing forces in the nucleus was not known at the time, he

*IST Austria, Am Campus 1, A-3400 Klosterneuburg
Mail: joscha.henheik@ist.ac.at

followed a more phenomenological approach and modelled the unknown Hamiltonian by a hermitian random matrix with independent entries (Wigner matrix), ignoring all physical details, except the symmetry type, real symmetric or complex hermitian, corresponding to physical systems with and without time reversal symmetry, respectively. Based on the observation that this modelling amazingly reproduced the experimental results, he formulated the surmise that the fluctuation of the gaps between successive eigenvalues of a random matrix follows a universal distribution, depending only on the symmetry type of matrix, analogously to the central limit theorem for scalar random variables and the universality of the Gaussian distribution. A more detailed version of this universality conjecture was formulated by Dyson and Mehta [10] in the 1960s, and is thereof called the Wigner-Dyson-Mehta (WDM) universality conjecture.

The proof of the WDM conjecture in full generality has been an open problem for more that fifty years until it got quite recently managed via the so called three step strategy developed by Erdös, Schlein, Yau, and Yin (see [7] for an overview) with parallel development by Tao and Vu [12]. At the current stage, the second and third step are very robust in the sense that they are largely independent of the details of the model under consideration and can be formulated as a "black-box" tool requiring only the input of the first step, that is a local law, which asserts that the spectral measure of the random matrix converges to a deterministic measure up to the scale of individual eigenvalues as the size of the matrix tends to infinity.

A local law has first been established for random matrices with independent and identically distributed entries (up to the symmetry ensuring hermiticity), so called Wigner matrices, where the semicircle distribution (2.4) emerges as the required deterministic probability measure. The historically first proof [14] of this celebrated Wigner semicircle law followed the moment method, although being quite elementary, not suitable for more general random matrix models, in which one relaxes (i) the identicalness of the distribution of the matrix elements, or (ii) their independentness, or both. A generalization of the first kind can be dealt with by using the Schur method, heavily using the independentness of the matrix elements. A modern approach to proving local laws, also for (slowly) correlated random matrices with not neccessarily identically distributed entries [6], builds on the analysis of a certain deterministic equation, the matrix Dyson equation (MDE), of which the resolvent of the random matrix can be shown to be an approximate solution, which yields the local law whenever the MDE is stable agains small additive perturbations. This is described in more detail below.

The first goal (Section 2) of this note is to follow the strategy described above in order to prove the local semicircle law for Wigner matrix, which has – at least to our knowledge – never been done. The second goal (Section 3) is to prove a local law for perturbations of Wigner matrices by a fixed number of mutually orthogonal rank-1 random matrices realizing a basic example for a weak but long range correlation structure in the random matrix which is generalization of both, type (i) and type (ii) above.

1.1 Proof strategy for the semicircle law

For a general hermitian $N \times N$ random matrix $H = H^{(N)}$ with ordered random eigenvalues $\lambda_1 \leq ... \leq \lambda_N$ we denote its resolvent by

$$G(z) = G^{(N)}(z) = (H - z)^{-1} \,,$$

where the spectral parameter $z = E + \mathrm{i}\eta$ is assumed to be in the upper half plane $\mathbb{H}$. Due to the quantum mechanical interpretation of H as a Hamilton operator of a system with N particles, we refer to $E = \Re z$ as energy.

The first two moments of H determine the limiting behavior of $G(z)$ in the limit of large N. More specifically, let

$$A := \mathbb{E}H \,, \qquad H =: A + \frac{1}{\sqrt{N}} W \,, \qquad \mathcal{S}[V] := \frac{1}{N}\mathbb{E}WVW \,, \tag{1.1}$$

where $\mathcal{S}$ is a linear map on the space of $N \times N$ matrices which we call the *self-energy operator* and W is a random matrix with zero expectation (see Definition 2.1). Then, the unique deterministic solution $M = M(z)$ to the matrix Dyson equation (MDE)

$$1 + (z - A + \mathcal{S}[M])M = 0 \quad \text{under the constraint} \quad \Im M = \frac{1}{2\mathrm{i}}(M - M^*) > 0 \tag{1.2}$$

approximates the random matrix $G(z)$ increasingly well as N tends to infinity, which is the general scope of a local law for random matrices as the ones formulated in Theorem 2.2 and Theorem 2.3 below.

The properties of (1.2) and its solutions have been studied in [1]. In particular, it has been shown that

$$m(z) := \frac{1}{N}\mathrm{Tr} M(z) = \int_{\mathbb{R}} \frac{1}{x - z}\mu(\mathrm{d}x) \tag{1.3}$$

is the Stieltjes transform of a measure μ on $\mathbb{R}$, which we call the *self-consistent density of states*, and whose support $\operatorname{supp}\mu$ we call the *self-consistent spectrum*. Moreover, under quite general conditions [1], one can show that μ is absolutely continuous with compactly supported Hölder continuous probability density

$$\mu(\mathrm{d}x) = \rho(x)\mathrm{d}x \qquad \text{and that} \qquad \rho(z) = \frac{1}{\pi N}\Im\, \mathrm{Tr} M(z) \tag{1.4}$$

is the harmonic extension of $\rho : \mathbb{R} \to [0, \infty)$. We note that the imaginary part of the Stieltjes transform

$$\Im m_\mu(z) = \int_{\mathbb{R}} \frac{\eta}{|x - E|^2 + \eta^2}\mu(\mathrm{d}x), \qquad z = E + \mathrm{i}\eta$$

of a measure μ can also be viewed as the convolution of μ with the Cauchy kernel P_η on scale η, i.e.

$$\Im m_\mu(E + \mathrm{i}\eta) = (P_\eta \star \mu)(E), \quad \text{with} \quad P_\eta(E) = \frac{\eta}{E^2 + \eta^2} \,.$$

Up to the normalization by $1/\pi$ (cf. (1.4)), P_η is an approximate delta function on scale η and thus $\Im m_\mu(E+\mathrm{i}\eta)$ resolves the measure μ on a scale η around the energy E.

The above notions become clear by noticing that the *empirical density of states*

$$\mu_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} \tag{1.5}$$

of H has Stieltjes transform

$$m_N(z) := m_{\mu_N}(z) = \int_{\mathbb{R}} \frac{1}{x-z}\mu_N(\mathrm{d}x) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i - z} = \frac{1}{N}\mathrm{Tr}\, G(z)\,. \tag{1.6}$$

Thus, as the convergence of the Stieltjes transforms $m_N(z) \to m(z)$ for any $z \in \mathbb{H}$ implies the convergence of the measures $\mu_N \to \mu$ in distribution [8], the study of a deterministic law for the spectrum of H for large N reduces to finding estimates on $N^{-1}\mathrm{Tr}(G-M)$. A result that controls this difference for any fixed $z \in \mathbb{H}$ is called a *global law*, whereas a *local law* allows the spectral parameter $z = E + \mathrm{i}\eta \in \mathbb{H}$ to approach the real axis while the resolution η is slightly above the typical eigenvalue spacing, which can be understood as follows (cf. [4]). In case that $\mathrm{diam}(\mathrm{spec}\, H)$ is of order one (see (2.1)), μ_N is a discrete measure on small $1/N$ scales , and thus it may behave very badly (i.e. strongly fluctuating or blow up) for η smaller than $1/N$, depending on whether there happens to be an eigenvalue in an η-vincinity of E. Thus, as the eigenvalue spacing in the bulk of $\mathrm{spec}\, H$ is typically of order $1/N$, for $\eta \lesssim 1/N$ we expect no "self-averaging behavior" of m_N. But, as long as $\eta \gg 1/N$ one might hope for a self-averaging law of large numbers phenomenon. This hope is confirmed by the local law (Theorem 2.3) down to the smallest possible (optimal) scale $\eta \gg 1/N$. Although we have a precise distinction between a global and a local law, we will often refer to the global law (as formulated in Theorem 2.2) as a local law as well.

Moreover, since (1.2) is stable against small additive perturbations, it is sufficient to study the smallness of $G - M$ by showing that the error matrix $D = D(z)$ defined by

$$D := 1 + (z - A + \mathcal{S}[G])G = (H - A + \mathcal{S}[G])G = \frac{W}{\sqrt{N}}G + \mathcal{S}[G]G \tag{1.7}$$

is small (see Section 4 for more details). Within the framework of this general strategy, choosing the correct norm to measure smallness of D is a key technical ingredient. Similarly to G, the error matrix D is very large in the usual $\ell^p \to \ell^q$ matrix norms, but its quadratic form $\langle \boldsymbol{x}, D\boldsymbol{y}\rangle$ is under control with very high probability for any fixed deterministic vectors $\boldsymbol{x}, \boldsymbol{y}$. Furthermore, to improve precision, we will distinguish two different concepts of measuring the size of D. In the first probabilistic part of the proof (Proposition 4.1) we will show that D can be bounded in *isotropic sense* as $|\langle \boldsymbol{x}, D\boldsymbol{y}\rangle| \lesssim \|\boldsymbol{x}\|\|\boldsymbol{y}\|/\sqrt{N\eta}$ for fixed deterministic vectors $\boldsymbol{x}, \boldsymbol{y}$ as well as in *averaged sense* as $N^{-1}|\mathrm{Tr} BD| \lesssim \|B\|/N\eta$ for fixed deterministic matrices B. Here, $\|\boldsymbol{x}\|$, $\|\boldsymbol{y}\|$, $\|B\|$ denote the standard Euclidean vector norm $\|\boldsymbol{x}\|^2 = \sum_a |x_a|^2$ and matrix operator norm $\|B\| := \sup_{\|\boldsymbol{x}\|,\|\boldsymbol{y}\|\leq 1} |\langle \boldsymbol{x}, B\boldsymbol{y}\rangle|$. In the second deterministic part of the proof (Proposition 4.2) we will show that since D is small, and (1.2) is stable under small additive perturbations, also $G-M$ is small in an appropriate sense. Finally, we combine these parts in a bootstrapping procedure (Proposition 4.5) in order to arrive at the local law.

1.2 Outline

In Section 2, the first part of this note, we formulate the local for the simplified class of Wigner matrices (Definition 2.1) in Theorem 2.2 and Theorem 2.3, which we prove in Section 4. Based on these theorems, we state a few classical corollaries on the completeness of the self-consistent spectrum (Corollary 2.4), the delocalization of bulk-eigenvectors (Corollary 2.5) and the rigidity of bulk-eigenvalues (Corollary 2.6).

In Section 3, the second part of this note, we will allow for perturbations of Wigner matrices by a fixed number of mutually orthogonal rank-1 random matrices, constituting a basic example of an overall weak but long range correlation structure in H, and present the corresponding local law which we formulate in Theorem 3.3 and Corollary 3.4 as well as in Theorem 3.5 and Corollary 3.6. The proofs of these results, which are given in Section 4, do not follow the general strategy sketched above, but use the local law for Wigner matrices and a simple identity for the resolvent of a hermitian rank-1 perturbation of a hermitian matrix (Proposition 3.2).

1.3 Notations

In order to formulate our results and proofs concisely, we introduce the following notation. An inequality with a subscript indicates that we allow for a constant in the bound depending only on the quantities in the subscript, e.g. $A(N,\varepsilon) \leq_\varepsilon B(N,\varepsilon)$ means that there exists a constant $C = C(\varepsilon)$, independent of N, such that $A(N,\varepsilon) \leq C(\varepsilon)B(N,\varepsilon)$ holds for all N and $\varepsilon > 0$. In many statements we will assume that N is sufficiently large, depending on any other parameters of the model. We will use boldface letters $\boldsymbol{x}, \boldsymbol{y}, ...$ from the end of the alphabet to denote vectors in $\mathbb{C}^N$ with entries $\boldsymbol{x} = (\boldsymbol{x}_i)_{i\in[N]}$ where $[N] := \{\, 1, ..., N \,\}$ and the indices are chosen from the beginning or middle of the alphabet. Elements of the set of ordered pairs $[N]^2 = [N] \times [N]$ will be called *labels* and will be denoted by by Greek letters $\alpha = (i,j) \in [N]^2$. Summations of the form $\sum_i$ and $\sum_\alpha$ are always understood to sum over all $i \in [N]$ and $\alpha \in [N]^2$.

For indices $i, j \in [N]$ and vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ we shall use the notations

$$A_{\boldsymbol{x}\boldsymbol{y}} = \langle \boldsymbol{x}, A\boldsymbol{y}\rangle\,, \qquad A_{\boldsymbol{x}i} = \langle \boldsymbol{x}, A\boldsymbol{e}_i\rangle\,, \qquad A_{i\boldsymbol{x}} = \langle \boldsymbol{e}_i, A\boldsymbol{x}\rangle\,,$$

where $\boldsymbol{e}_i$ is the i-th standard basis vector. The normalized trace is denoted by $\langle A\rangle := N^{-1}\mathrm{Tr}\, A$. Sometimes we will also use the notation $\langle z\rangle := 1+|z|$ for the complex number z, but this should no lead to confusion as it will only be used for z. Moreover, we will use the maximum norm and the normalized Hilbert-Schmidt norm

$$\|A\|_{\max} := \max_{i,j\in[N]} |A_{ij}|\,, \quad \|A\|_{\text{hs}} := \left(\frac{1}{N}\sum_{ij} |A_{ij}|^2\right)^{1/2} \quad \text{for} \quad A \in \mathbb{C}^{N\times N}\,.$$

2 Semicircle law for Wigner matrices

While the proof of a local law following the strategy sketched in the introduction has been carried out for general self-adjoint random matrices with slow correlation decay [6], we focus on giving a proof in this modern language in the prominent simplified case of H being a *Wigner matrix*, which has – to our knowledge – never been written up properly. Therefore, we assume that H is a centered random self-adjoint matrix with i.i.d. matrix elements (up to the symmetry) of variance $1/\sqrt{N}$, precisely defined as follows.

Definition 2.1. Let $H = \left(h_{i,j}\right)_{i,j=1}^N$ be a random matrix. We say that $H = A + \frac{1}{\sqrt{N}}W$ with $A := \mathbb{E}H$ is a Wigner matrix, if $A = 0$ and for $w_{ij} = \sqrt{N}h_{ij}$ the following conditions are satisfied (recall $\mathbb{E}W = 0$ from (1.1)).

(i) The matrix elements are i.i.d. random variables up to the symmetry constraint $w_{ij} = \bar{w}_{ji}$, ensuring that W is hermitian. More precicely, in the *real symmetric case*, the collection of random variables $\{\, w_{ij} : i \le j \,\}$ are independent and identically distributed, while in the *complex hermitian case* the distributions of $\{\Re w_{ij}, \Im w_{ij} : i < j\}$ and $\{\sqrt{2}h_{ii} : i = 1, 2, ..., N\}$ are independent and identical.

(ii) All moments of w_{ij} are finite, i.e. for every $p \in \mathbb{N}$ exists some $C_p > 0$ such that $\mathbb{E}|w_{ij}|^p \le C_p$ for all $i, j \in 1, ..., N$.

(iii) The common variance of the random variables w_{ij} equals 1, i.e. $\mathbb{E}|w_{ij}|^2 = 1$ for all $i, j \in 1, ..., N$.

These conditions will ensure, that the spectrum of the Wigner matrix $H = \frac{W}{\sqrt{N}}$ is real, centered around 0 and the typical size of an eigenvalue λ_i remains of order 1 even as N tends to infinity. This can be seen by computing

$$\mathbb{E}\sum_i \lambda_i^2 = \mathbb{E}\mathrm{Tr}H^2 = \frac{1}{N}\mathbb{E}\sum_{ij} |w_{ij}|^2 = N\,, \tag{2.1}$$

indicating that $\lambda_i \sim 1$ on average (see [4]).

Assuming that H is a Wigner matrix, we can determine the solution of the MDE (1.2) more explicitly as the self-energy operator reads

$$(\mathcal{S}_{\text{ch}}[V])_{ij} = \langle V\rangle\, \delta_{ij} \qquad \text{or} \qquad (\mathcal{S}_{\text{rs}}[V])_{ij} = \langle V\rangle\, \delta_{ij} + (1-\delta_{ij})N^{-1}V_{ji}\,,$$

in the complex hermitian and real symmetric case, respectively. The second summand in the formula for the real symmetric case will turn out to be irrelevant for determining the self-consistent density of states (see Section 4), such that we focus on $\mathcal{S}[V] = \langle V\rangle$ in the remainder of this paragraph. Indeed, $M(z) = m(z)I$ solves the MDE, whenever $m(z)$ is a solution of the scalar equation

$$1 + (z + m)m = 0 \quad \text{under the constraint} \quad \Im m(z) > 0 \quad \forall z \in \mathbb{H}. \tag{2.2}$$

This has a unique solution given by

$$m_{\rm sc}(z) = \frac{-z + \sqrt{z^2 - 4}}{2}\,, \tag{2.3}$$

where the square root is chosen such that the branch cut is in the segment $[-2, 2]$, ensuring that the constraint is fullfilled. An easy computation, showing that this is the Stieltjes transform of the semicircle density, i.e.

$$m_{\rm sc}(z) = \int_{\mathbb{R}} \frac{\rho_{\rm sc}(x)}{x - z}\mathrm{d}x \quad \text{where} \quad \rho_{\rm sc}(x) := \frac{1}{2\pi}\sqrt{(4 - x^2)_+}\,, \tag{2.4}$$

explains, why we call the local law for Wigner matrices the *local semicircle law.* Note that $\operatorname{supp}\rho_{\rm sc} = [-2, 2]$ and we thus call ± 2 the *edges of the self-consistent spectrum.*

2.1 Local semicircle law

In order to formulate the local semicircle law, we specify the range of spectral parameters z by defining two spectral domains characterized via any given parameters $\gamma, \delta > 0$. Outside of the self-consistent spectrum we will work on

$$\mathbb{D}^\delta_{\rm out} := \left\{z \in \mathbb{H} : |z| \le N^{C_0}, \operatorname{dist}(z, [-2, 2]) \ge N^{-\delta}\right\} \tag{2.5}$$

for some arbitrary fixed $C_0 \ge 100$, while we will use

$$\mathbb{D}^\delta_\gamma := \left\{z \in \mathbb{H} : |z| \le N^{C_0}, \Im z \ge N^{-1+\gamma}, \rho_{\rm sc}(\Re z) + \operatorname{dist}(\Re z, [-2, 2]) \ge N^{-\delta}\right\} \tag{2.6}$$

away from the edges ± 2 of the self-consistent spectrum.

Theorem 2.2. (Local law outside of the spectrum and global law)
Let H be a Wigner matrix. Then, for any $\varepsilon > 0$ there exists $\delta > 0$ such that for all $D > 0$ we have the isotropic local law away from the spectrum,

$$\mathbb{P}\left(|G_{\boldsymbol{x}\boldsymbol{y}} - m_{\rm sc}\langle \boldsymbol{x}, \boldsymbol{y}\rangle| \le \|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\,\frac{N^\varepsilon}{\langle z\rangle^2\sqrt{N}} \quad in \quad \mathbb{D}^\delta_{\rm out}\right) \ge 1 - CN^{-D} \tag{2.7}$$

for all deterministic vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ and we have the averaged local law away from the spectrum

$$\mathbb{P}\left(|\langle BG\rangle - m_{\rm sc}\langle B\rangle| \le \|B\|\,\frac{N^\varepsilon}{\langle z\rangle^2\,N} \quad in \quad \mathbb{D}^\delta_{\rm out}\right) \ge 1 - CN^{-D} \tag{2.8}$$

for all deterministic matrices $B \in \mathbb{C}^{N\times N}$. In fact, δ can be chosen such that $\delta = c\gamma$ for some absolute constant $c > 0$. Here, $C = C(D, \varepsilon, \gamma) < \infty$ is a constant, depending only on its arguments and the moment bounds C_p from Definition 2.1.

We also obtain an optimal local law away from the spectral edges, especially in the bulk.

Theorem 2.3. (Local law in the bulk of the spectrum)
Let H be a Wigner matrix. Then, for any $\gamma, \varepsilon > 0$ there exists $\delta > 0$ such that for all $D > 0$ we have the isotropic local law in the bulk,

$$\mathbb{P}\left(|G_{\boldsymbol{xy}} - m_{\mathrm{sc}}\langle \boldsymbol{x}, \boldsymbol{y}\rangle| \leq \|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\,\frac{N^{\varepsilon}}{\sqrt{N\Im z}} \quad \textit{in} \quad \mathbb{D}_{\gamma}^{\delta}\right) \geq 1 - CN^{-D} \tag{2.9}$$

for all deterministic vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ and we have the averaged law in the bulk

$$\mathbb{P}\left(|\langle BG\rangle - m_{\mathrm{sc}}\langle B\rangle| \leq \|B\|\,\frac{N^{\varepsilon}}{N\Im z} \quad \textit{in} \quad \mathbb{D}_{\gamma}^{\delta}\right) \geq 1 - CN^{-D} \tag{2.10}$$

for all deterministic matrices $B \in \mathbb{C}^{N\times N}$. In fact, δ can be chosen such that $\delta = c\min\{\varepsilon, \gamma\}$ for some absolute constant $c > 0$. Here, $C = C(D, \varepsilon, \gamma) < \infty$ is a constant, depending only on its arguments and the moment bounds C_p from Definition 2.1.

By setting $B = 1$ in (2.8) and (2.10) we get the desired comparison between the Stieltjes transform of the empirical density of states (1.5) and the Stieltjes transform of the semicircle density m_{sc} in the sense of a global and local law. Note that both theorems cover the regime of z being far away from the spectrum. In that case, the estimates from Theorem 2.2 are stronger, while Theorem 2.3 is really relevant when $\Re z$ is in the bulk of the spectrum $[-2, 2]$ and $\Im z$ is very small. Moreover, we stress the importance of the vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ and the matrix $B \in \mathbb{C}^{N\times N}$ being deterministic in the above statements.

2.2 Delocalization of eigenvectors and eigenvalue rigidity

The above theorems immediately have the following classical corollaries on completeness of the self-consistent spectrum, eigenvector delocalization and eigenvalue rigidity, as stated below. While Corollary 2.4 is obtained along the proofs of Theorem 2.2 and Theorem 2.3 in Section 4, the justification of Corollary 2.5 and Corollary 2.6 requires some additional work. Although this work focusses on giving a modern proof of the semicirle law, we provide a few arguments on the validity of these corollaries afterwards.

Corollary 2.4. (Completeness of the self-consistent spectrum)
Let H be a Wigner matrix. Then there exists $\delta > 0$ such that for any $D > 0$,

$$\mathbb{P}\Big(\operatorname{spec} H \not\subset (-N^{-\delta}, N^{-\delta}) + [-2, 2]\Big) \leq_D N^{-D}\,. \tag{2.11}$$

Corollary 2.5. (Bulk delocalization)
Let H be a Wigner matrix and $\boldsymbol{u} = (\boldsymbol{u}_i)_{i\in[N]}$ be a normalized eigenvector corresponding to a bulk eigenvalue λ of H. Then

$$\mathbb{P}\left(\max_{i\in[N]} |\boldsymbol{u}_i| \geq \frac{N^{\varepsilon}}{\sqrt{N}},\ H\boldsymbol{u} = \lambda\boldsymbol{u},\ \rho_{\mathrm{sc}}(\lambda) \geq \delta\right) \leq_{\varepsilon,\delta,D} N^{-D} \tag{2.12}$$

for any $\varepsilon, \delta, D > 0$.

Corollary 2.6. (Bulk rigidity)
Let H be a Wigner matrix with ordered eigenvalues $\lambda_1 \le \ldots \le \lambda_N$ and denote the classical position of the eigenvalue close to energy $E \in \mathbb{R}$ by

$$k(E) := \left\lceil N \int_{-\infty}^{E} \rho_{\mathrm{sc}}(x) \mathrm{d}x \right\rceil ,$$

where $\lceil \cdot \rceil$ denotes the ceiling function. It then holds that

$$\mathbb{P}\left(\sup\left\{ |\lambda_{k(E)} - E| : E \in \mathbb{R},\ \rho_{\mathrm{sc}}(E) \ge \delta \right\} \ge \frac{N^{\varepsilon}}{N} \right) \le_{\varepsilon,\delta,D} N^{-D} \tag{2.13}$$

for any $\varepsilon, \delta, D > 0$

3 Matrices with weak but long range correlation

So far we studied the case of Wigner matrices, where we are in the special situation that the matrix elements are independent and identically distributed up to the symmetry constraint (see Definition 2.1). The proof given in Section 4 follows the strategy outlined in the introduction, which was primary designed in such a way that one can also handle the case of random matrices with slow correlation decay [6]. Roughly speaking this means that the correlation between different matrix elements decays with their (natural) distance in a summable way. Here we are interested in the situation where the correlation between the matrix elements does not decay with the distance (i.e. long range correlation) but is weak in the sense that the correlation can be viewed as a perturbation of an uncorrelated matrix. More precisely, we consider the following class of examples.

Definition 3.1. (Perturbed Wigner matrices)
Let $\frac{1}{\sqrt{N}}W$ be a Wigner matrix and $(\xi_i)_{i=1,\ldots,k}$ be real random variables independent of W that satisfy

$$\mathbb{P}\left(|\xi_i| > N^{\varepsilon}\right) \le_{\varepsilon,D} N^{-D} \qquad \text{and} \qquad \mathbb{P}\left(|\xi_i| \in (0, N^{-\varepsilon})\right) \le_{\varepsilon,D} N^{-D} \tag{3.1}$$

for all $i = 1, .., k$ and $\varepsilon, D > 0$. Moreover, let $(\boldsymbol{v}_i)_{i=1,\ldots,k}$ be a deterministic orthonormal system in $\mathbb{C}^N$ and $(r_i)_{i=0,\ldots,k} \subset [0,1]$ with $r_0 > 0$ such that $\sum_{i=0}^{k} r_i = 1$ (we assume $k \in \mathbb{N}$ and $r_0 > 0$ being independent of N). Then we call

$$H = \sqrt{r_0}\, \frac{1}{\sqrt{N}}\, W + \sum_{i=1}^{k} \sqrt{N r_i}\ \xi_i\ |\boldsymbol{v}_i\rangle \langle \boldsymbol{v}_i| \tag{3.2}$$

a *(rank−k−)perturbed Wigner matrix* and $V := \mathrm{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ the *perturbed subspace* with orthogonal projection $P = \sum_{i=1}^{k} |\boldsymbol{v}_i\rangle \langle \boldsymbol{v}_i|$ and $P^{\perp} := 1 - P$. Note that $H = \frac{1}{\sqrt{N}}W$ for $r_0 = 1$.

This definition covers the special case of $k = 1$, $\boldsymbol{v}_1 = \boldsymbol{e} = N^{-1/2}(1, ..., 1)$, for which

$$H = (h_{ij})_{i,j=1}^N \quad \text{with} \quad h_{ij} = \sqrt{1-r}\, N^{-1/2}\, w_{ij} + \sqrt{r}\, N^{-1/2}\, \xi\,, \quad \text{where} \quad W = (w_{ij})_{i,j=1}^N\,,$$

realizing the aimed correlation structure described above in the most elementary way. For a perturbed Wigner matrix we expect a small modification of the semicircle law for spectral parameters z with $\Re z \sim \sqrt{N}$ due to the scale separation by $\sqrt{N}$ in the second summand in (3.2). More precisely, by Weyl's inequality [13] on eigenvalues of perturbed hermitian matrices, we expect k of the N eigenvalues to be "outliers" of the bulk spectrum of the Wigner matrix and fluctuating on the given scale. In [9] Knowles and Yin (see also, e.g., [3, 11]) showed that the (approximate) position of outliers in the spectrum of Wigner matrices deformed by *deterministic* finite rank deformations is given by $\sqrt{Nr_i}\xi_i + \frac{1}{\sqrt{Nr_i\xi_i}}$, which is consistent with our expectation. It is thus natural to distinguish the *bulk*, *ambient*, *outlier*, and *exterior domain*, defined as

$$\begin{aligned}
\mathbb{D}_{\text{bulk}}^{\gamma,\delta,\nu} &:= \left\{ z \in \mathbb{H} : \Im z \geq N^{-1+\gamma},\ \tilde{\rho}_{\text{sc}}(\Re z) + \operatorname{dist}(\Re z, [-2\sqrt{r_0}, 2\sqrt{r_0}]) \geq N^{-\delta},\ |z| \leq N^{1/2-\nu} \right\}\,, \\
\mathbb{D}_{\text{amb}}^{\delta,\nu} &:= \left\{ z \in \mathbb{H} : \operatorname{dist}(z, [-2\sqrt{r_0}, 2\sqrt{r_0}]) \geq N^{-\delta},\ |z| \leq N^{1/2-\nu} \right\}\,, \\
\mathbb{D}_{\text{out}}^{\nu} &:= \left\{ z \in \mathbb{H} : N^{1/2-\nu} \leq |z| \leq N^{1/2+\nu} \right\} \\
\mathbb{D}_{\text{ext}}^{\nu} &:= \left\{ z \in \mathbb{H} : N^{1/2+\nu} \leq |z| \leq N^{C_0} \right\}\,,
\end{aligned}$$

respectively, for (small) $\gamma, \delta, \nu > 0$. Here $C_0 \geq 100$ is some arbitrary fixed constant and $\tilde{\rho}_{\text{sc}}(x) = (2\pi r_0)^{-1}\sqrt{(4r_0 - x^2)_+}$ denotes the semicircle density with radius $2\sqrt{r_0}$ and $\tilde{m}_{\text{sc}}$ its Stieltjes transform. In order to deal with a perturbed Wigner matrix, we will employ the following key formula on inverses of rank$-1-$perturbed matrices, which can be verified by direct computation.

Proposition 3.2. (Perturbed resolvent) *Let $A \in \mathbb{C}^{N\times N}$ with $\Im A < 0$ and $\boldsymbol{v} \in \mathbb{C}^N$. Then*

$$\frac{1}{A + |\boldsymbol{v}\rangle\langle\boldsymbol{v}|} = \frac{1}{A} - \frac{1}{1 + \langle \boldsymbol{v}, \frac{1}{A}\boldsymbol{v}\rangle} \cdot \frac{1}{A}|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\frac{1}{A}\,.$$

Based on this identity and the local law from Theorem 2.2 and Theorem 2.3, we obtain two local laws in isotropic and averaged sense covering the bulk, ambient and exterior domain combined (Theorem 3.3), and the outlier domain (Theorem 3.5). In these theorems we approximate the resolvent G by a "semi-deterministic" matrix $M^{(\xi)}$, which still carries the ξ_i-randomness in the perturbed subspace but has the W-randomness removed and is defined as

$$M^{(\xi)} = \sum_{i=1}^{k} \frac{\tilde{m}_{\text{sc}}}{1 + \sqrt{Nr_i}\,\xi_i\,\tilde{m}_{\text{sc}}}\,|\boldsymbol{v}_i\rangle\langle\boldsymbol{v}_i| + \tilde{m}_{\text{sc}}P^{\perp}\,. \tag{3.3}$$

In the spectral domains $\mathbb{D}_{\text{bulk}}^{\gamma,\delta,\nu}$, $\mathbb{D}_{\text{amb}}^{\delta,\nu}$ and $\mathbb{D}_{\text{ext}}^{\nu}$, the quality of the bound on the error compared with the results in Section 2 is not affected. In $\mathbb{D}_{\text{out}}^{\nu}$ the situation is different due to the existence of outlier eigenvalues in this domain. The proofs of the statements below are all given in Section 4.

3.1 Local law away from the outliers

As in Section 2, the local laws estimate the difference between G and its approximation in high probability sense. Note that this probability $\mathbb{P}$ used below is the probability for the full randomness, i.e. both W and ξ.

Theorem 3.3. (Local law in the bulk, ambient and exterior domain)

(a) Let H be a perturbed Wigner matrix. Then, for any $\varepsilon, \gamma > 0$ there exists $\delta > 0$ such that for all $D, \nu > 0$ we have the isotropic law in the bulk domain

$$\mathbb{P}\left(|\langle \boldsymbol{x}, (G - M^{(\xi)})\boldsymbol{y}\rangle| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\| \frac{N^{\varepsilon}}{\sqrt{N\Im z}} \quad \text{in} \quad \mathbb{D}^{\gamma,\delta,\nu}_{\text{bulk}} \right) \geq 1 - CN^{-D} \tag{3.4}$$

for all deterministic vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ and we have the averaged law in the bulk domain

$$\mathbb{P}\left(|\langle B(G - M^{(\xi)})\rangle| \leq \|B\| \frac{N^{\varepsilon}}{N\Im z} \quad \text{in} \quad \mathbb{D}^{\gamma,\delta,\nu}_{\text{bulk}} \right) \geq 1 - CN^{-D} \tag{3.5}$$

for all deterministic matrices $B \in \mathbb{C}^{N\times N}$. In fact, δ can be chosen as $\delta = c\min(\varepsilon, \gamma)$ for some absolute constant $c > 0$ and $C = C(D, \varepsilon, \gamma, \nu)$ is some constant depending only on its arguments and the bounds on the moments from Definition 2.1 and Definition 3.1.

(b) Let H be a perturbed Wigner matrix. Then, for any $\varepsilon > 0$ there exists $\delta > 0$ such that for all $D, \nu > 0$ we have the isotropic law in the ambient domain

$$\mathbb{P}\left(|\langle \boldsymbol{x}, (G - M^{(\xi)})\boldsymbol{y}\rangle| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\| \frac{N^{\varepsilon}}{\sqrt{N}\langle z\rangle^2} \quad \text{in} \quad \mathbb{D}^{\delta,\nu}_{\text{amb}} \right) \geq 1 - CN^{-D} \tag{3.6}$$

for all deterministic vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ and we have the averaged law in the ambient domain

$$\mathbb{P}\left(|\langle B(G - M^{(\xi)})\rangle| \leq \|B\| \frac{N^{\varepsilon}}{N\langle z\rangle^2} \quad \text{in} \quad \mathbb{D}^{\delta,\nu}_{\text{amb}} \right) \geq 1 - CN^{-D} \tag{3.7}$$

for all deterministic matrices $B \in \mathbb{C}^{N\times N}$. In fact, δ can be chosen as $\delta = c\varepsilon$ for some absolute constant $c > 0$ and $C = C(D, \varepsilon, \nu)$ is some constant depending only on its arguments and the bounds on the moments from Definition 2.1 and Definition 3.1.

(c) Let H be a perturbed Wigner matrix. Then, for any $D, \varepsilon, \nu > 0$ we have the isotropic law in the exterior domain

$$\mathbb{P}\left(|\langle \boldsymbol{x}, (G - M^{(\xi)})\boldsymbol{y}\rangle| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\| \frac{N^{\varepsilon}}{\sqrt{N}\langle z\rangle^2} \quad \text{in} \quad \mathbb{D}^{\nu}_{\text{ext}} \right) \geq 1 - CN^{-D} \tag{3.8}$$

for all deterministic vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ and we have the averaged law in the exterior domain

$$\mathbb{P}\left(|\langle B(G - M^{(\xi)})\rangle| \le \|B\| \frac{N^\varepsilon}{N\langle z\rangle^2} \quad \text{in} \quad \mathbb{D}^\nu_{\mathrm{ext}} \right) \ge 1 - CN^{-D} \tag{3.9}$$

for all deterministic matrices $B \in \mathbb{C}^{N\times N}$. In fact, δ can be chosen as $\delta = c\varepsilon$ for some absolute constant $c > 0$ and $C = C(D, \varepsilon, \nu)$ is some constant depending only on its arguments and the bounds on the moments from Definition 2.1 and Definition 3.1.

The following corollary compares the resolvent G to the full deterministic matrix $M = \tilde{m}_{\mathrm{sc}}$ that also has the ξ-randomness removed. Its proof will be a simple application of Theorem 3.3 in combination with a suitable estimate on the difference $M - M^{(\xi)}$.

Corollary 3.4. (Full deterministic averaged law)
The averaged laws in (3.5), (3.7) and (3.9) from Theorem 3.3 remain valid in the same form if we replace $M^{(\xi)}$ by $M = \tilde{m}_{\mathrm{sc}}$ up to an additional factor of $\langle z\rangle$ in the error bound of (3.7) and (3.9) whenever $PBP \neq 0$.

3.2 Local law in the outlier domain

In the outlier domain we restrict to a rank−1−perturbed Wigner matrix. The general case could be handled similarly as in the proof of Theorem 3.3, but the error bound would be considerably complicated (see Remark 4.6).

Theorem 3.5. (Local law in the outlier domain)
Let H be a random rank−1−perturbed Wigner matrix and define the semi-deterministic control parameter

$$\Theta(z) := \left| \frac{\sqrt{N}r\xi\tilde{m}_{\mathrm{sc}}(z)}{1 + \sqrt{N}r\xi\tilde{m}_{\mathrm{sc}}(z)} \right| . \tag{3.10}$$

Moreover, let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ be deterministic vectors and $B \in \mathbb{C}^{N\times N}$ a deterministic matrix with $\|\boldsymbol{x}\|$, $\|\boldsymbol{y}\|$, $\|B\| \le 1$. Then, for any $\varepsilon > 0$ there exists $\nu > 0$ such that for all $D > 0$ we have the isotropic law in the outlier domain

$$\mathbb{P}\left(|\langle \boldsymbol{x}, (G - M^{(\xi)})\boldsymbol{y}\rangle| \le \frac{N^\varepsilon}{N^{3/2}} \left(1 + \frac{\Theta}{N} + (\|P\boldsymbol{x}\| + \|P\boldsymbol{y}\|)\,\Theta \right) \quad \text{in} \quad \mathbb{D}^\nu_{\mathrm{out}} \right) \ge 1 - CN^{-D} \tag{3.11}$$

and the averaged law in the outlier domain

$$\mathbb{P}\left(|\langle B(G - M^{(\xi)})\rangle| \le \frac{N^\varepsilon}{N^2} \left(1 + \frac{\Theta}{\sqrt{N}} + (\|PB\| + \|BP\|)\,\Theta \right) \quad \text{in} \quad \mathbb{D}^\nu_{\mathrm{out}} \right) \ge 1 - CN^{-D} \tag{3.12}$$

for all deterministic matrices. In fact, ν can be chosen as $\nu = c\varepsilon$ for some absolute constant $c > 0$ and $C = C(D, \varepsilon)$ is some constant depending only on its arguments and the bounds on the moments from Definition 2.1 and Definition 3.1.

Note that while typically $\Theta(z)$ is of order 1, in the vicinity of $z \sim \sqrt{Nr}\xi$ it is of order $\sqrt{N}/\eta$. Now, similarly to the other three domains above, the following corollary compares the resolvent G to the full deterministic matrix $M = \tilde{m}_{\mathrm{sc}}$ that also has the ξ-randomness removed. Its proof will be a simple application of Theorem 3.5 in combination with a suitable estimate on the difference $M - M^{(\xi)}$.

Corollary 3.6. (Deterministic averaged law)
If we replace $M^{(\xi)}$ with $M = \tilde{m}_{\mathrm{sc}}$ in (3.12) *the statement remains true but the bound reads*

$$\mathbb{P}\left(|\langle B(G-M)\rangle| \leq \frac{N^{\varepsilon}}{N^2}\left(1 + \frac{\Theta}{\sqrt{N}} + (\|PB\| + \|BP\|)\,\Theta + \|PBP\|\sqrt{N}\Theta\right) \quad in \ \mathbb{D}^{\nu}_{\mathrm{out}}\right) \\ \geq 1 - CN^{-D}\,. \tag{3.13}$$

4 Proofs

4.1 Local semicircle law for Wigner matrices

As described in the introduction, we will prove the local semicircle law in three steps. In the first step, we establish a high-moment bound on the error matrix D defined in (1.7). In the second step, we show that the MDE (1.2) resp. its scalar variant for Wigner matrices (2.2) is stable against small additive perturbations such as the error matrix D. In the last step, we combine the previous results in a bootstrapping procedure in order to conclude the smallness of $G - M$, i.e. the local law, on the optimal scale.

While we carry out the proof only for H being a real symmetric Wigner matrix, we will see that the only neccessary modification concerns the cumulant expansion used in Proposition 4.1, where we simply consider real and imaginary part separately to reduce it to the real case.

4.1.1 High moment bounds on the error

In this section we sketch the proof of an isotropic and averaged bound on the error matrix D defined in (1.7), in the form of high-moment estimates using a (multivariate) cumulant expansion. To formalize the bounds, we define high-moment norms for random variables X and random matrices A by

$$\|X\|_p = (\mathbb{E}|X|^p)^{1/p}\,, \qquad \|A\|_p = \sup_{\|\boldsymbol{x}\|,\|\boldsymbol{y}\|\leq 1} \|\langle \boldsymbol{x}, A\boldsymbol{y}\rangle\|_p = \left[\sup_{\|\boldsymbol{x}\|,\|\boldsymbol{y}\|\leq 1} \mathbb{E}|\langle \boldsymbol{x}, A\boldsymbol{y}\rangle|^p\right]^{1/p},$$

where the supremum is taken over deterministic vectors $\boldsymbol{x}, \boldsymbol{y}$.

Proposition 4.1. (Bound on the error matrix, Theorem 4.1 in [6])
Let H be a Wigner matrix. Then there exists a constant C_* such that for any $p \geq 1$,

$\varepsilon > 0$, $z \in \mathbb{H}$ with $\Im z \geq N^{-1}$, $B \in \mathbb{C}^{N\times N}$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ such that

$$\frac{\|\langle \boldsymbol{x}, D\boldsymbol{y}\rangle\|_p}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} \leq_{\varepsilon,p} N^\varepsilon \sqrt{\tfrac{\|\Im G\|_q}{N\Im z}}\left(1+\|G\|_q\right)^{C_*}\left(1+\tfrac{\|G\|_q}{N}\right)^{C_* p} \tag{4.1a}$$

$$\frac{\|\langle BD\rangle\|_p}{\|B\|} \leq_{\varepsilon,p} N^\varepsilon \langle z\rangle \tfrac{\|\Im G\|_q}{N\Im z}\left(1+\|G\|_q\right)^{C_*}\left(1+\tfrac{\|G\|_q}{N}\right)^{C_* p} \tag{4.1b}$$

where $q = C_* p^4/\varepsilon$, $R = 2p$.

The main tool for proving these bounds is an "integration by parts" formula (4.3) for the expectation, called cumulant expansion. Moreover, by naively counting the number of summations emerging by expanding the quantities on the left hand side of (4.1), one would not obtain the bound given above. In fact, the main gain comes from the key formula about resolvents of hermitian matrices

$$GG^* = \frac{\Im G}{\Im z} \quad \text{or alternatively} \quad \sum_j |G_{ij}|^2 = \frac{\Im G_{ii}}{\Im z}, \tag{4.2}$$

called the *Ward identity*, which directly follows from the spectral theorem and the first resolvent identity. Note, that in the second formulation a sum of order N is reduced to a $1/\Im z$ factor, so we effectively gain a factor of $1/(N\Im z)$ over naively counting the number of summations.

Sketch of the proof: In the following, we will iteratively use the expansion

$$\mathbb{E} w_{ij} f(W) = \sum_{k=1}^{R} \sum_{\boldsymbol{\alpha}\in\{ij,ji\}^k} \frac{\kappa(ij,\boldsymbol{\alpha})}{k!}\mathbb{E}\partial_{\boldsymbol{\alpha}} f(W) + \Omega_R \tag{4.3}$$

with an explicit error term Ω_R (see Proposition 3.2 from [6]). Here for a k-tuple of labels $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_k)$ we use the short-hand notation $\kappa(ij, (\alpha_1, ..., \alpha_k)) = \kappa(w_{ij}, w_{\alpha_1}, ..., w_{\alpha_k})$ for the join cumulant of $w_{ij}, w_{\alpha_1}, ..., w_{\alpha_k}$ and set $\partial_{\boldsymbol{\alpha}} = \partial_{w_{\alpha_1}} ... \partial_{w_{\alpha_k}}$, $\partial_{ij} = \partial_{w_{ij}}$. A complex variant of (4.3) required for the complex hermitian case has a similar form. For simplicity, we now assume that $p = 2$ and H has Gaussian entries, which makes all terms in (4.3) with $k \geq 2$ vanish. After performing the expansion in this setting, we will argue how to deal with the general case.

We begin with the proof of the isotropic bound (4.1a) in this simplified case. By application of (4.3) we have

$$\begin{aligned}
\mathbb{E}|D_{\boldsymbol{xy}}|^2 &= \mathbb{E}\langle \boldsymbol{x}, D\boldsymbol{y}\rangle\langle \boldsymbol{y}, D^*\boldsymbol{x}\rangle \\
&= N^{-1}\sum_{\substack{a,b,c\\ i,j,k}} \bar{\boldsymbol{x}}_a \boldsymbol{y}_c \bar{\boldsymbol{y}}_i \boldsymbol{x}_k \kappa(ab,ij)\mathbb{E}\left[G_{bc}G^*_{jk}\right] \\
&\quad + N^{-2}\sum_{\substack{a,b,c\\ i,j,k}} \bar{\boldsymbol{x}}_a \boldsymbol{y}_c \bar{\boldsymbol{y}}_i \boldsymbol{x}_k \sum_{\alpha,\beta}\kappa(ab,\alpha)\kappa(ij,\beta)\mathbb{E}\left[\left(G\Delta^\beta G\right)_{ab}\left(G^*\Delta^\alpha G^*\right)_{ij}\right].
\end{aligned}$$

Here, Δ^α is the matrix with a one at lable α and zero everywhere else. The crucial point in the above calculation lies in the effect of the term $\mathcal{S}[G]G$ in the definition of D in (1.7) which acts as the self-energy renormalization of HG in such that a way that it cancels a term of order 1 emerging from the cumulant expansion.

Now, we use the correlation structure $\kappa(ab, cd) = \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}$ of a real symmetric Wigner matrix to reduce the number of summations in both terms. With the aid of the Ward-identity (4.2) and multiple applications of the Schwartz inequality we conclude that

$$\mathbb{E}|D_{\boldsymbol{xy}}|^2 \leq C\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2\frac{\|\Im G\|_2}{N\Im z}\left(1+\|G\|_2\right)^2 ,$$

where $C > 0$ is some constant depending only on the (easy) combinatorics.

Now, we turn the proof of (4.1b) in the simplified case. Siumilarly, by application of (4.3), the Ward identity and the Schwartz inequality we obtain the bound

$$\mathbb{E}\left|\langle BD\rangle\right|^2 \leq \|B\|^2\frac{\|\Im G\|_2^2}{(N\Im z)^2}\left(1+\|G\|_2\right)^2 .$$

For the general case, we write

$$\mathbb{E}\left|D_{\boldsymbol{xy}}\right|^{2p} = \mathbb{E}\left(N^{-1/2}WG + \mathcal{S}[G]G\right)_{\boldsymbol{xy}} D_{\boldsymbol{xy}}^{p-1}\bar{D}_{\boldsymbol{xy}}^p$$

and likewise for $\mathbb{E}\left|\langle BD\rangle\right|^{2p}$, and use (4.3) in the first factor of W, considering everything else as a function f. Like in the above calculations, $\mathcal{S}[G]G$ cancels the second order cumulant and the effect of higher order cumulants will be small as the cumulant of order k is $N^{-\frac{k+1}{2}}$. However, the derivatives also act on the factors D^{p-1} and $\bar{D}^p$ and many cumulants have to be tracked. In our case of Wigner matrices and using their simple corellation structure, we could proceed, e.g., by induction. However, by doing the proof carefully [6], one arrives at the bounds given in (4.1). $\square$

4.1.2 Stability of the MDE

In this section, we prove the stability of the MDE for the simple correlation structure of Wigner matrices. Note, that the self-energy operator acts in the real symmetric case as

$$\mathcal{S}[G] = \langle G\rangle + \frac{1}{N}(G - \mathrm{diag}(G_{11}, ..., G_{NN}))^T =: \langle G\rangle + N^{-1}\tilde{G} . \tag{4.4}$$

For fixed $z \in \mathbb{H}$ define the map

$$\mathcal{J}_z[G, D] := 1 + (z + \langle G\rangle)G - D$$

on arbitrary matrices G and D. From the definition of $D(z)$ in (1.7) and the solution $M(z) = m(z)I = m_{\mathrm{sc}}(z)I$ of the MDE (2.2) it follows that $\mathcal{J}_z[M(z), 0] = 0$ and $\mathcal{J}_z[G(z), D(z) - N^{-1}\tilde{G}(z)] = 0$. Throughout the proof we will fix z and omit it from the notation, i.e. $\mathcal{J} = \mathcal{J}_z$. We will consider G as a function $G(D)$ for arbitrary error matrix D satisfying $\mathcal{J}[G(D), D] = 0$. Via the implicit function theorem, this relation defines a unique function $G(D)$ for sufficiently small D and $G(D)$ will be analytic as long as $\mathcal{J}$ is stable. The stability will be formulated in a slightly modified maximum norm accounting that the smallness of D can only be established in isotropic sense, i.e. in the sense of high-moment bounds on its quadratic form $D_{\boldsymbol{xy}}$ for deterministic vectors $\boldsymbol{x}, \boldsymbol{y}$. To define this special norm, we fix vectors $\boldsymbol{x}, \boldsymbol{y}$, define $I := \{ \boldsymbol{x}, \boldsymbol{y} \} \cup \{ \boldsymbol{e}_i \mid i \in [N] \}$ and set

$$\|G\|_* := \|G\|_*^{\boldsymbol{x},\boldsymbol{y}} := \max_{\boldsymbol{u},\boldsymbol{v} \in I} \frac{|G_{\boldsymbol{uv}}|}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} .$$

Proposition 4.2. (Stability of the MDE, Theorem 5.2 in [6])
Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ and denote the ball of radius $\delta > 0$ around M in $(\mathbb{C}^{N \times N}, \| \cdot \|_*^{\boldsymbol{x},\boldsymbol{y}})$ by $B_\delta(M)$. Then for

$$\varepsilon_1 := \frac{\Big(1 + \|M\|^2 + \|M\|^4 \, \| \left(1 - m_{\mathrm{sc}}^2 \langle \cdot \rangle\right)^{-1} \|_{\mathrm{hs} \to \mathrm{hs}}\Big)^{-2}}{10\|M\|^2}, \qquad \varepsilon_2 := \sqrt{\frac{\varepsilon_1}{10}}$$

there exists a unique function $G : B_{\varepsilon_1}(0) \to B_{\varepsilon_2}(M)$ with $G(0) = M = m_{\mathrm{sc}} I$ that satisfies $\mathcal{J}[G(D), D] = 0$. Moreover, the function G is analytic and satisfies

$$\|G(D_1) - G(D_2)\|_* \le C \left\| \left(1 - m_{\mathrm{sc}}^2 \langle \cdot \rangle\right)^{-1} \right\|_{* \to *} \|M\| \|D_1 - D_2\|_* \tag{4.5}$$

for any $D_1, D_2 \in B_{\varepsilon_1}(0)$.

Proof. As a first step, we rewrite the equation $\mathcal{J}[G, D] = 0$ in the form $\tilde{\mathcal{J}}[V, D] = 0$ where

$$\tilde{\mathcal{J}}[V, D] := \left(1 - m_{\mathrm{sc}}^2 \langle \cdot \rangle\right) V - M\langle V \rangle V + MD, \qquad V := G - M$$

and for arbitrary V and D we claim the bounds

$$\|M\langle V \rangle V\|_* \le \|M\| \, \|V\|_*^2 \tag{4.6}$$

$$\|MD\|_* \le \|M\| \, \|D\|_* \tag{4.7}$$

$$\left\| \left(1 - m_{\mathrm{sc}}^2 \langle \cdot \rangle\right)^{-1} \right\|_{* \to *} \le 1 + \|M\|^2 + \|M\|^4 \left\| \left(1 - m_{\mathrm{sc}}^2 \langle \cdot \rangle\right)^{-1} \right\|_{\mathrm{hs} \to \mathrm{hs}} . \tag{4.8}$$

For (4.6) and (4.7), note that M and $\langle V \rangle$ are diagonal matrices and that $\|\langle V \rangle\| \le \|V\|_*$. For (4.8) we use a three term geometric expansion to obtain

$$\begin{aligned} &\left\| \left(1 - m_{\mathrm{sc}}^2 \langle \cdot \rangle\right)^{-1} \right\|_{* \to *} \\ &\quad \le \; 1 + \left\| m_{\mathrm{sc}}^2 \langle \cdot \rangle \right\|_{* \to *} + \left\| m_{\mathrm{sc}}^2 \langle \cdot \rangle \right\|_{* \to \mathrm{hs}} \left\| \left(1 - m_{\mathrm{sc}}^2 \langle \cdot \rangle\right)^{-1} \right\|_{\mathrm{hs} \to \mathrm{hs}} \left\| m_{\mathrm{sc}}^2 \langle \cdot \rangle \right\|_{\mathrm{hs} \to *} \qquad (4.9) \\ &\quad \le \; 1 + \|M\|^2 \, \|\langle \cdot \rangle\|_{\max \to \|\cdot\|} + \|M\|^4 \|\langle \cdot \rangle\|_{\max \to \|\cdot\|} \left\| \left(1 - m_{\mathrm{sc}}^2 \langle \cdot \rangle\right)^{-1} \right\|_{\mathrm{hs} \to \mathrm{hs}} \|\langle \cdot \rangle\|_{\mathrm{hs} \to \|\cdot\|} \end{aligned}$$

and it remains to bound $\|\langle \cdot \rangle\|_{\mathrm{hs} \to \|\cdot\|}$ and $\|\langle \cdot \rangle\|_{\max \to \|\cdot\|}$. Therefore, we estimate

$$\|\langle B \rangle\| = |\langle B \rangle| \le \frac{1}{N} \sum_i |B_{ii}|$$

$$\le \min\left(\left(N^{-1}\sum_i |B_{ii}|^2\right)^{1/2},\ \max_{i\in[N]}|B_{ii}|\right) \le \min\left(\|B\|_{\mathrm{hs}},\ \|B\|_{\max}\right),$$

such that $\max\left(\|\langle\cdot\rangle\|_{\mathrm{hs}\to\|\cdot\|}, \|\langle\cdot\rangle\|_{\max\to\|\cdot\|}\right) \le 1$ and thus (4.8) follows in combination with (4.9). Now the statement (4.5) follows by a simple application of the implicit function theorem as formulated in Lemma A.1 applied to the equation $\tilde{\mathcal{J}}[G-M, D]$ written in the form

$$\left(1 - m_{\mathrm{sc}}^2\langle\cdot\rangle\right)V - M\langle V\rangle V = -MD$$

with $A = 1 - m_{\mathrm{sc}}^2\langle\cdot\rangle$, $B = M$ and $d = D$ in the notation from Lemma A.1. $\square$

This general stability result will be used in the following form

$$\|G-M\|_* \le_\varepsilon N^\varepsilon \langle z\rangle^{-1}\left(\|D\|_* + \frac{1}{N\eta}\right) \qquad \text{in} \quad \mathbb{D}_{\mathrm{out}}^\delta \quad \text{and} \quad \mathbb{D}_0^\delta \tag{4.10}$$

for some $\delta = \delta(\varepsilon) > 0$ as long as $\|G-M\|_* \le_\varepsilon N^{-\varepsilon}\langle z\rangle^2$ by applying it to $D_1 = 0$, $D_2 = D(z) - N^{-1}\tilde{G}$ and using that $\|G\|_* \le \|G\| = 1/\eta$, $\langle z\rangle\|M(z)\| \le C$ as well as

$$\left\|\left(1 - m_{\mathrm{sc}}^2\langle\cdot\rangle\right)^{-1}\right\|_{\mathrm{hs}\to\mathrm{hs}} = \left\|1 + \frac{m_{\mathrm{sc}}^2}{1-m_{\mathrm{sc}}^2}\langle\cdot\rangle\right\|_{\mathrm{hs}\to\mathrm{hs}} \le_\varepsilon N^\varepsilon \tag{4.11}$$

in $\mathbb{D}_{\mathrm{out}}^\delta$ and $\mathbb{D}_0^\delta$ for some $\delta = \delta(\varepsilon) > 0$, where the last bound follows since $m(z) = \mp 1$ exactly at the spectral edges $z = \pm 2$. Moreover, it will also be used with $1/\eta$ replaced by $\|G\|$.

4.1.3 Conclusion: Bootstrapping and self-improving estimates

Before going into the proof of our main theorems, we introduce a commonly used (see, e.g., [5]) notion of high-probability bound (Definition 4.3) and prepare the iteration step of our bootstrapping procedure (Proposition 4.5).

Definition 4.3. (Stochastic domination) Let

$$X = \left(X^{(N)}(u) \mid N\in\mathbb{N}, u\in U^{(N)}\right) \quad \text{and} \quad Y = \left(Y^{(N)}(u) \mid N\in\mathbb{N}, u\in U^{(N)}\right)$$

be families of non-negative random variables indexed by N, and possibly a parameter u, then we say that X is stochastically dominated by Y, if for all $\varepsilon, D > 0$ we have

$$\sup_{u\in U^{(N)}} \mathbb{P}\left(X^{(N)}(u) > N^\varepsilon Y^{(N)}(u)\right) \le N^{-D}$$

for large enough $N \ge N_0(\varepsilon, D)$. In this case we write $X \prec Y$. Moreover, if for some complex family of random variables we have $|X| \prec Y$, we also write $X = O_\prec(Y)$.

It can easily be checked (see [5]), that $\prec$ satisfies the usual arithmetic properties, i.e. if $X_1 \prec Y_1$ and $X_2 \prec Y_2$ then also $X_1+Y_1 \prec X_2+Y_2$ and $X_1Y_1 \prec X_2Y_2$. In particular, by using the Chebychev inequality, the bound on the moments from Definition 2.1 can be expressed as $w_{ij} \prec 1$. Moreover, we will say that a sequence of events $A = A^{(N)}$ holds *with very high probability* if for any $D > 0$ exists some $N_0(D) \in \mathbb{N}$ such that for all $N \geq N_0(D)$ we have $\mathbb{P}(A^{(N)}) \geq 1 - N^{-D}$.

In the following Lemma we establish that the control of the $\|\cdot\|_*^{\boldsymbol{x},\boldsymbol{y}}$-norm for all $\boldsymbol{x}, \boldsymbol{y}$ in high-probability sense is essentially equivalent to the control of the $\|\cdot\|_p$-norm for all $p \geq 1$. This translation will often be used in the proof below.

Lemma 4.4. *Let R be a random matrix and Φ a deterministic control parameter. Then the following implications hold:*

(i) If $\Phi \geq N^{-C}$, $\|R\| \leq N^C$ and $|R_{\boldsymbol{xy}}| \prec \Phi$ for all normalized $\boldsymbol{x}, \boldsymbol{y}$ and some C, then also $\|R\|_p \leq_{p,\varepsilon} N^\varepsilon \Phi$ for all $\varepsilon > 0, p \geq 1$.

(ii) If $\|R\|_p \leq_{p,\varepsilon} N^\varepsilon \Phi$ for all $\varepsilon > 0, p \geq 1$, then $\|R\|_*^{\boldsymbol{x},\boldsymbol{y}} \prec \Phi$ for any fixed $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$.

Proof. On the one hand, (i) is a direct consequence of

$$\|R\|_p \leq N^\varepsilon \Phi + \sup_{\|\boldsymbol{x}\|,\|\boldsymbol{y}\| \leq 1} \left(|R_{\boldsymbol{xy}}| \, \mathbb{P}\left(|R_{\boldsymbol{xy}}| \geq N^\varepsilon \Phi\right)^{1/p} \right)$$

in combination with $\|R\| \leq N^C$ and the definition of stochastic domination. On the other hand, we obtain by application of Chebychev's and Hölder's inequality that

$$\mathbb{P}\left(\|R\|_* > N^\sigma \Phi\right) \leq \frac{\mathbb{E}\|R\|_*^p}{N^{\sigma p}\Phi^p} \leq \frac{\left(\mathbb{E}\sum_{\boldsymbol{u},\boldsymbol{v} \in I} |R_{\boldsymbol{uv}}|^{pr}\right)^{1/r}}{N^{\sigma p}\Phi^p} \leq |I|^{2/r} \frac{\|R\|_{pr}^p}{N^{\sigma p}\Phi^p} \leq_{p,r,\varepsilon} \frac{|I|^{2/r}}{N^{(\sigma-\varepsilon)p}},$$

and since $|I| \leq N+2$ we conclude that $\|R\|_* \prec \Phi$ by choosing ε sufficiently small and p, r sufficiently large. $\square$

As mentioned above, the proof of the local law in the bulk follows a bootstrapping procedure. That is, we first prove the local law for $\eta \geq N$, and afterwards we iteratively show that if the local law holds for $\eta \geq N^{\gamma_0}$ then it also holds for $\eta \geq N^{\gamma_1}$ for some $\gamma_1 < \gamma_0$. This iteration step is formulated as follows.

Proposition 4.5. (Bootstrapping, Proposition 5.5 in [6])
Let H be a Wigner matrix, $\delta, \gamma > 0$ and $\gamma_0 > \gamma_1 \geq \gamma$ with $4(2C_*+1)(\gamma_0-\gamma_1) < \gamma < 1/2$ and suppose that

$$\|G - M\|_p \leq_{\gamma,p} \frac{N^{-\gamma/6}}{\langle z \rangle} \quad \text{in} \quad \mathbb{D}_{\gamma_0}^\delta \tag{4.12}$$

holds for all $p \geq 1$, where C_* is the constant from Proposition 4.1. Then the same inequality (4.12) *(with a possibly different implicit constant depending on γ, δ, p) holds true in $\mathbb{D}_{\gamma_1}^\delta$.*

Proof. We abbreviate the step size from γ_0 to γ_1 by $\gamma_s := \gamma_0 - \gamma_1$. By using $\|M\| \le C/\langle z\rangle$ and (4.12) we infer $\|G\|_p \le_{p,\gamma} N^{\gamma_s}\langle z\rangle^{-1}$ in $\mathbb{D}^\delta_{\gamma_0}$. Fix $E \in \mathbb{R}$ and consider the function $\eta \mapsto f(\eta) := \eta\|G(E+\mathrm{i}\eta)\|_p$. Using that

$$\eta\left|\langle \boldsymbol{x}, G^2\boldsymbol{y}\rangle\right| \le \frac{\eta}{2}\left(\langle \boldsymbol{x}, |G|^2\boldsymbol{x}\rangle + \langle \boldsymbol{y}, |G|^2\boldsymbol{y}\rangle\right) \le \frac{1}{2}\left(\langle \boldsymbol{x}, \Im G\boldsymbol{x}\rangle \langle \boldsymbol{y}, \Im G\boldsymbol{y}\rangle\right),$$

which follows by application of the Ward identity (4.2), we obtain that $\eta \mapsto f(\eta)$ is monotonically increasing since

$$\begin{aligned}\liminf_{\varepsilon\to 0}\frac{f(\eta+\varepsilon)-f(\eta)}{\varepsilon} \;&\ge\; \|G(E+\mathrm{i}\eta)\|_p - \eta\left\|\lim_{\varepsilon\to 0}\frac{G(E+\mathrm{i}(\eta+\varepsilon))-G(E+\mathrm{i}\eta)}{\varepsilon}\right\|_p \\ &=\; \|G(E+\mathrm{i}\eta)\|_p - \eta\|G^2(E+\mathrm{i}\eta)\|_p \ge 0\,. \end{aligned} \tag{4.13}$$

Therefore, $\langle z\rangle\|G\|_p \le_{\gamma,p} N^{2\gamma_s}$ in $\mathbb{D}^\delta_{\gamma_1}$ and we conclude from Proposition 4.1 and by using $\gamma_s < 1$ that

$$\|D\|_p \le_{\varepsilon,p,\gamma} N^{\varepsilon+2(C_*+1/2)\gamma_s-\gamma/2} \le N^{\varepsilon-\gamma/4} \quad \text{in} \quad \mathbb{D}^\delta_{\gamma_1}. \tag{4.14}$$

By Lemma 4.4 we can translate this bound to $\|D\|_*^{\boldsymbol{x},\boldsymbol{y}} \prec N^{-\gamma/4}$ for any fixed $\boldsymbol{x},\boldsymbol{y}$ and we drop the superscript from now on. Next, we apply (4.10) to obtain

$$\|G-M\|_*\chi\left(\|G-M\|_* \le N^{-\gamma/10}\right) \prec \frac{N^{-\gamma/5}}{\langle z\rangle} \quad \text{in} \quad \mathbb{D}^\delta_{\gamma_1}, \tag{4.15}$$

where $\gamma/10$ is some examplary small exponent. This bound shows that there is a gap in the set of possible values of $\|G-M\|_*$, i.e. whenever $\|G-M\|_* \le N^{-\gamma/10}$ we indeed have the better bound $\|G-M\|_* \prec \langle z\rangle^{-1}N^{-\gamma/5}$ in high-probability sense. The extension of (4.12) to $\mathbb{D}^\delta_{\gamma_1}$ now follows from a standard continuity argument using a fine grid of intermediate values of η. More precisely, we use a fine N^{-3}-grid and the η^{-1}-Lipschitz continuity of $\|G-M\|_*$ to connect neighboring points on the grid. See [6] for a more detailed argument. □

We are now able to complete the proof of Theorem 2.2 and Theorem 2.3.

Proof of Theorems 2.2, 2.3 and Corollary 2.4. The proof involves five steps. First, we prove an initial isotropic bound in $\mathbb{D}^\delta_\gamma$, which we improve in the second step to obtain the isotropic local law as formulated in Theorem 2.3. In the third step, we use the isotropic local law to obtain the averaged local law in the bulk, which is used in the fourth step to obtain that with very high probability there are no eigenvalues outside of the support ρ_{sc}. Finally, in the fifth step we use that there are no eigenvalues outside the support of ρ_{sc} to obtain the improved isotropic and averaged local law outside the support.

Step 1: Initial isotropic bound. We claim that

$$\|G-M\|_p \le_{p,\gamma} \frac{N^{-\gamma/6}}{\langle z\rangle} \quad \text{in} \quad \mathbb{D}^\delta_\gamma \tag{4.16}$$

for some $\delta = \delta(\gamma)$. After showing that (4.16) holds for $\eta \geq N$, i.e. in $\mathbb{D}^\delta_{\gamma=2} = \mathbb{D}^\delta_2$, we conclude (4.16) in all of $\mathbb{D}^\delta_\gamma$ by iteratively applying Proposition 4.5 starting from $\gamma_0 = 2$. Therefore, using $w_{ij} \prec 1$ we estimate

$$\|H\| = \max_{i\in[N]} |\lambda_i| \leq \sqrt{\mathrm{Tr}\,|H|^2} \prec \sqrt{N}\,,$$

which, in combination with $|z| \geq N$, implies $\|G\|_p \leq_p \langle z\rangle^{-1}$ and $\|\Im G\|_p \leq_p \eta \langle z\rangle^{-2}$ and thus from Proposition 4.1 it follows that

$$\|D\|_p \leq_{\varepsilon,p} \frac{N^\varepsilon}{\langle z\rangle \sqrt{N}} \quad \text{in} \quad \mathbb{D}^\delta_2\,.$$

We now fix normalized deterministic vectors $\boldsymbol{x}, \boldsymbol{y}$ and translate these high-moment to high-probability bounds using Lemma 4.4 to infer $\|D\|_* \prec \langle z\rangle^{-1}/\sqrt{N}$ and $\|G\|_* \prec \langle z\rangle^{-1}$. Using the stability in the form (4.10) and absorbing N^ε-factors into $\prec$ we conclude

$$\|G - M\|_* \prec \frac{1}{\langle z\rangle^2 \sqrt{N}} \quad \text{in} \quad \mathbb{D}^\delta_2\,,$$

from which follows (4.16) in $\mathbb{D}^\delta_2$ by application of Lemma 4.4 since $\boldsymbol{x}$ and $\boldsymbol{y}$ were arbitrary. At the end of this step, we note that the bound $\|H\| \prec \sqrt{N}$ obtained above is very rough, but, as we see, sufficient for the initial isotropic bound. In fact, the proof will show that $\|H\| \leq 2 + \delta$ with very high probability.

Step 2: Iterative self-improvement of the isotropic bound. We now iteratively improve the initial bound (4.16) until we reach the intermediate bound

$$\|G - M\|_p \leq_{\varepsilon,p} \frac{N^\varepsilon}{\langle z\rangle} \left(\sqrt{\frac{\|\Im M\|}{N\eta}} + \frac{1}{\langle z\rangle} \frac{1}{N\eta} \right) \quad \text{in} \quad \mathbb{D}^\delta_\gamma \tag{4.17}$$

for $\delta = \delta(\varepsilon) > 0$. From (4.16) and $\langle z\rangle \|M\| \leq C$ we conclude that $\langle z\rangle \|G\|_p$ is N^ε-bounded in $\mathbb{D}^\delta_\gamma$ for some $\delta = \delta(\varepsilon) > 0$. By application of Proposition 4.1 and (4.16) we obtain that

$$\|D\|_p \leq_{\varepsilon,p} N^\varepsilon \sqrt{\frac{\|\Im G\|_q}{N\eta}} \quad \text{and} \quad \|G - M\|_* + \|D\|_* \prec N^{-\gamma/6} \quad \text{in} \quad \mathbb{D}^\delta_\gamma. \tag{4.18}$$

Since the following estimates hold true uniformly in the spectral domain $\mathbb{D}^\delta_\gamma$, we will supress this qualifier.

In order to show (4.17) we define a sequence of successively improving control parameters $(\Psi_l)_{l=0}^L$ as $\Psi_0 = 1$ and $\Psi_{l+1} = N^{-\sigma}\Psi_l = N^{-(l+1)\sigma}$ with arbitrary $\sigma \in (0,1)$. The final parameter is implicitly determined as the largest integer L such that

$$\Psi_L \geq \frac{N^\sigma}{\langle z\rangle} \left(\sqrt{\frac{\|\Im M\|}{N\eta}} + \frac{1}{\langle z\rangle} \frac{N^\sigma}{N\eta} \right)\,. \tag{4.19}$$

For the proof of (4.17) it remains to show that

$$\|G-M\|_p \leq_{\varepsilon,p} N^{\varepsilon}\Psi_l \quad \text{for all} \quad l=0,,,.,L\,,$$

which we prove by induction. For the induction step from l to $l+1$, we write $\Im G = \Im M + \Im(G-M)$ and by using that $\sqrt{a+b} \leq \sqrt{a}+\sqrt{b}$ and $\sqrt{a/b} \leq ca + 1/(cb)$ for all $a,b,c>0$ we continue from the induction hypothesis and (4.18) such that we can estimate

$$\|D\|_p \leq_{\varepsilon,p} N^{\varepsilon}\left(\sqrt{\frac{\|\Im M\|}{N\eta}}+\sqrt{\frac{\Psi_l}{N\eta}}\right) \leq_{\varepsilon,p} N^{\varepsilon}\left(\sqrt{\frac{\|\Im M\|}{N\eta}}+\frac{1}{\langle z\rangle}\frac{N^{\sigma}}{N\eta}+\langle z\rangle\, N^{-\sigma}\Psi_l\right).$$

By application of Lemma 4.4 this translates to

$$\|D\|_* \prec \sqrt{\frac{\|\Im M\|}{N\eta}}+\frac{1}{\langle z\rangle}\frac{N^{\sigma}}{N\eta}+\langle z\rangle\, N^{-\sigma}\Psi_l$$

for all normalized $\boldsymbol{x},\boldsymbol{y}$, from which we conclude by (4.10) that

$$|(G-M)_{\boldsymbol{xy}}| \prec \frac{1}{\langle z\rangle}\left(\sqrt{\frac{\|\Im M\|}{N\eta}}+\frac{1}{\langle z\rangle}\frac{N^{\sigma}}{N\eta}\right)+N^{-\sigma}\Psi_l \leq 2N^{-\sigma}\Psi_l$$

where we used $l<L$ and (4.19) in the last step. By the definition of Ψ_{l+1} and with the aid of Lemma 4.4 this yields

$$\|G-M\|_p \leq_{\varepsilon,p} N^{\varepsilon}\Psi_{l+1}$$

completing the induction step and thereby the proof of (4.17). The isotropic local law (2.9) in Theorem 2.3 now follows since $\Im M$ is bounded and by translating (4.17) to a high-probability bound using Lemma 4.4.

Step 3: Averaged bound. Since $\tilde{\mathcal{J}}[G-M, D-N^{-1}\tilde{G}]=0$ with $\tilde{D} := D-N^{-1}\tilde{G}$ we have that $G-M$ satisfies the quadratic relation

$$G-M = \left(1-m_{\mathrm{sc}}^2\langle\cdot\rangle\right)^{-1}\left[-M\tilde{D}+M\langle G-M\rangle(G-M)\right]$$

and thus

$$\begin{aligned}&\|\langle B(G-M)\rangle\|_p\\ &\qquad\leq \left\|\langle B\left(1-m_{\mathrm{sc}}^2\langle\cdot\rangle\right)^{-1}[M\tilde{D}]\rangle\right\|_p + \left\|\langle B\left(1-m_{\mathrm{sc}}^2\langle\cdot\rangle\right)^{-1}[M\langle G-M\rangle(G-M)]\rangle\right\|_p .\end{aligned}$$

By a three-term geometric expanion, as in (4.9), we obtain

$$\left\|\left(1-m_{\mathrm{sc}}^2\langle\cdot\rangle\right)^{-1}\right\|_{\|\cdot\|\to\|\cdot\|} \leq 1+\|M\|^2+\|M\|^4\left\|\left(1-m_{\mathrm{sc}}^2\langle\cdot\rangle\right)^{-1}\right\|_{\mathrm{hs}\to\mathrm{hs}}$$

and therefore $\left\|\left((1-m_{\mathrm{sc}}^2\langle\cdot\rangle)^{-1}\right)^*[B^*]\right\| \leq_\varepsilon N^\varepsilon\|B\|$ by (4.11). Since $\|N^{-1}\widetilde{G}\| \leq 1/N\eta$, by application of (4.1b) in Proposition 4.1, where $((1-m_{\mathrm{sc}}^2\langle\cdot\rangle)^{-1})^*[B^*]$ plays the role of B, writing $\|\Im G\|_q \leq \|\Im M\| + \|G-M\|_q$ and using (4.17) we can conclude that

$$\|\langle B(G-M)\rangle\|_p \leq_{p,\varepsilon,\gamma} \frac{N^\varepsilon\|B\|}{\langle z\rangle}\left(\langle z\rangle\frac{\|\Im M\|}{N\eta} + \sqrt{\frac{\|\Im M\|}{N\eta}}\frac{1}{N\eta} + \frac{1}{(N\eta)^2}\right) \tag{4.20}$$

from Lemma A.2. The averaged local law (2.10) in Theorem 2.3 is now a direct consequence of (4.20) and the fact that $\Im M$ is bounded by translating it to a high-probability bound with the aid of Lemma 4.4.

The proof of Theorem 2.3 is now complete. For the proof of Theorem 2.2, we first work in the restricted domain $\mathbb{D}_\gamma^\delta \cap \mathbb{D}_{\mathrm{out}}^\delta$ and use that $\|\Im M\| \leq \eta \mathrm{dist}(z,[-2,2])^{-2}$ which follows from a direct computation and only adds another negligible N^ε-factors in (4.17) and (4.20) such that we have

$$\|G-M\|_p \leq_{\varepsilon,p} \frac{N^\varepsilon}{\langle z\rangle}\left(\sqrt{\frac{1}{N}} + \frac{1}{\langle z\rangle}\frac{1}{N\eta}\right) \tag{4.21}$$

and

$$\|\langle B(G-M)\rangle\|_p \leq_{p,\varepsilon,\gamma} \frac{N^\varepsilon\|B\|}{\langle z\rangle}\left(\langle z\rangle\frac{1}{N} + \sqrt{\frac{1}{N}}\frac{1}{N\eta} + \frac{1}{(N\eta)^2}\right) \tag{4.22}$$

in the restricted domain $\mathbb{D}_\gamma^\delta \cap \mathbb{D}_{\mathrm{out}}^\delta$. It requires to additional steps to prove Theorem 2.2 in all of $\mathbb{D}_{\mathrm{out}}^\delta$.

Step 4: Completeness of the self-consistent spectrum. Suppose there is an eigenvalue λ of H with $\mathrm{dist}(\lambda,[-2,2]) \geq N^{-\delta}$. By application of (4.22) with $B=1$ and a spectral decomposition of H we obtain $|\langle G(\lambda+\mathrm{i}\eta)\rangle| \geq |\langle \Im G(\lambda+\mathrm{i}\eta)\rangle| \geq 1/N\eta$ and thus

$$\mathbb{P}\left(\exists\lambda \text{ eigenv. with } \mathrm{dist}(\lambda,[-2,2]) \geq N^{-\delta}\right) \leq \mathbb{P}\left(|\langle G-M\rangle| \geq c/N\eta \text{ in } \mathbb{D}_{\mathrm{out}}^\delta \cap \mathbb{D}_{1/2}^\delta\right)$$

for some small constant $c>0$, where we used that $\langle M\rangle$ is the Stieltjes transform of ρ_{sc}. Using (4.22) again with $\varepsilon = 1/4$ and $\gamma = 1/2$, we infer by an application of Chebychevs inequality that the right hand side can be further bounded by (a constant times)

$$\inf_{\eta\geq N^{-1/2}}\left(N^{1/4}\left[\eta + \frac{1}{\sqrt{N}} + \frac{1}{N\eta}\right]\right)^p \leq_p N^{-p/4}\,, \tag{4.23}$$

from which Corollary 2.4 directly follows.

Step 5: Improved bounds outside of the support. Fix z such that $\mathrm{dist}(z,[-2,2]) \geq N^{-\delta}$ and $\eta \geq N^{-1+\gamma}$. Using that with very high probability there are no eigenvalues outside of the support, we have $\|G\| \prec \langle z\rangle^{-1}$ and $\|\Im G\| \prec \eta\langle z\rangle^{-2}$, which translate to

$\|G\|_p \le_{\varepsilon,p} N^\varepsilon \langle z\rangle^{-1}$ and $\|\Im G\|_p \le_{\varepsilon,p} N^\varepsilon \eta\langle z\rangle^{-2}$ by Lemma 4.4, so we infer from Proposition 4.1 that

$$\|D\|_p \le_{\varepsilon,p} \frac{N^\varepsilon}{\sqrt{N}\langle z\rangle} \quad \text{and therefore} \quad \|D\|_* \prec \frac{1}{\sqrt{N}\langle z\rangle}$$

by application of Lemma 4.4. Using the stability from (4.10) (the required initial smallness of $G-M$ is guaranteed by (4.21)) with $1/\eta$ replaced by $\|G\| \prec \langle z\rangle^{-1}$ we find

$$\|G-M\|_* \prec \frac{1}{\sqrt{N}\langle z\rangle^2} \quad \text{and therefore} \quad \|G-M\|_p \le_{\varepsilon,p} \frac{N^\varepsilon}{\sqrt{N}\langle z\rangle^2}$$

by Lemma 4.4. Using Lipschitz-continuity of G and M with Lipschitz constant of order 1, we extend this bound from $\eta \ge N^{-1+\gamma}$ to $\eta \ge 0$, such that the isotropic law (2.7) in Theorem 2.2 follows. Similarly to Step 3, this leads to the improved averaged local (2.8), where we use that with very high probability there are no eigenvalues outside of the support $[-2,2]$. □

4.2 Delocalization and rigidity for Wigner matrices

In this section, we sketch the proofs of delocalization of bulk eigenvectors (Corollary 2.5) and rigidity of bulk eigenvalues (Corollary 2.6) as corollaries of the local law in Theorem 2.3.

Sketch of the proof of Corollary 2.5. Let $\boldsymbol{u}^{(k)} = (\boldsymbol{u}_i^{(k)})_{i\in[N]}$ be a normalized eigenvector of H with eigenvalue λ_k and $i \in [N]$. Then we find from a spectral decomposition that

$$1 \succ \Im G_{ii} = \eta \sum_k \frac{|\boldsymbol{u}_i^{(k)}|^2}{(E-\lambda_k)^2+\eta^2} \ge \frac{|\boldsymbol{u}_i^{(k)}|^2}{\eta} \quad \text{for} \quad z = E + \mathrm{i}\eta\,,$$

where the first estimate follows from Theorem 2.3 and the boundedness of M. The last inequality is true if we assume E being η-close to λ_k. So the claim follows by going down to scale $\eta = N^{-1+\gamma}$. □

Sketch of the proof of Corollary 2.6. The main idea is to translate the bound on $\langle G - M\rangle$ from the local law in Theorem 2.3 to a bound on the difference of the probability measures

$$\mu(\mathrm{d}x) = \rho_{\mathrm{sc}}(x)\mathrm{d}x \quad \text{and} \quad \mu_N(\mathrm{d}x) = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}(\mathrm{d}x)\,,$$

with λ_i, $i = 1,...,N$ being the eigenvalues of H and for which $\langle M\rangle$ and $\langle G\rangle$ are the respective Stieltjes transforms. This translation is achieved with the aid of Lemma 5.1 in [2], from which the claim follows by application of the improved local law away from the self-consistent spectrum. See the proof of Corollary 2.9 in [1] for a more detailed argument. □

4.3 Matrices with weak but long range correlation

This section is devoted to the proof of the results from Section 3. For the proof of Theorem 3.3 and Corollary 3.4, which deal with the bulk, ambient, and exterior domain, we restrict to the case $k = 2$, assume $r_1, r_2 > 0$ to avoid redundancy, and, in order to simplify notation, we redefine $\xi_i := \sqrt{r_i}\xi_i$ for $i = 1, 2$ and set $r_0 = 1$ (see Definition 3.1). The general case then easily follows by iterating the argument. Before going into the proof, we note that, using the notion of stochastic domination from Definition 4.3, the conditions on ξ_i from (3.1) in Definition 3.1 can be rewritten as

$$|\xi_i| \prec 1 \qquad \text{and} \qquad |(\xi_i|_{\{\xi_i \neq 0\}})^{-1}| \prec 1\,, \tag{4.24}$$

leading to an arbitrarily "narrow" outlier domain, i.e. the position $\Re z$ of an outlier satisfies $N^{1/2-\nu} \lesssim |\Re z| \lesssim N^{1/2+\nu}$ for all $\nu > 0$, which corresponds to the arbitrariness of $\nu > 0$ in the ambient and exterior domain.

Proof of Theorem 3.3 and Corollary 3.4.
As mentioned in Section 3, the key formula to attack the perturbed Wigner matrix is given in Proposition 3.2, which yields an expression for the resolvent of a hermitian matrix that is perturbed by a hermitian rank−1 matrix. In order to apply Proposition 3.2 we define the following four hermitian random matrices

$$H^{(0)} := \frac{1}{\sqrt{N}}W\,, \qquad\qquad H^{(1)} := \frac{1}{\sqrt{N}}W + \xi_1 \,|\boldsymbol{v}_1\rangle\,\langle\boldsymbol{v}_1|\,,$$

$$H^{(2)} := \frac{1}{\sqrt{N}}W + \xi_2 \,|\boldsymbol{v}_2\rangle\,\langle\boldsymbol{v}_2|\,, \qquad H^{(1,2)} := \frac{1}{\sqrt{N}}W + \sum_{i=1}^{2}\xi_i \,|\boldsymbol{v}_i\rangle\,\langle\boldsymbol{v}_i|\,,$$

with resolvents $G^{(\#)} = G^{(\#)}(z) = (H^{(\#)} - z)^{-1}$ and we write $H = H^{(1,2)}$ and $G = G^{(1,2)}$. After these preparations, we split the proof in three parts, the first one concerning the isotropic laws in Theorem 3.3, the second one the averaged laws in Theorem 3.3, partly following from the isotropic laws, and the third one the averaged laws in Corollary 3.4.

Part 1: Isotropic law. The general idea in the proof of the isotropic law is to use Proposition 3.2 twice in order to incorporate both orthogonal perturbations $\xi_1 \,|\boldsymbol{v}_1\rangle\,\langle\boldsymbol{v}_1|$ and $\xi_2 \,|\boldsymbol{v}_2\rangle\,\langle\boldsymbol{v}_2|$ with the aid of the isotropic local law from Theorem 2.2 and Theorem 2.3 by using the two possible paths

$$G^{(0)} \longrightarrow G^{(1)} \longrightarrow G \quad \text{and} \quad G^{(0)} \longrightarrow G^{(2)} \longrightarrow G\,. \tag{4.25}$$

More precisely, by going the first step in both paths, we obtain

$$G^{(i)} = G^{(0)} - \frac{\sqrt{N}\xi_i}{1 + \sqrt{N}\xi_i\, G^{(0)}_{\boldsymbol{v}_i\boldsymbol{v}_i}} \cdot G^{(0)}\,|\boldsymbol{v}_i\rangle\,\langle\boldsymbol{v}_i|\,G^{(0)} \quad \text{for} \quad i = 1, 2 \tag{4.26}$$

from Proposition 3.2, for which we use the isotropic local for $G^{(0)}$ to obtain a local law for $G^{(i)}$, $i = 1, 2$. These, in turn, can then be used to obtain the desired local law for G

via two different paths. In this procedure, we will see that not every path is useful for every choice of deterministic vectors $\boldsymbol{x}, \boldsymbol{y}$ in the isotropic local law.

An important role will be played by the semi-deterministic control parameter

$$\Theta_i(z) := \left| \frac{\sqrt{N}\xi_i \tilde{m}_{\mathrm{sc}}(z)}{1+\sqrt{N}\xi_i \tilde{m}_{\mathrm{sc}}(z)} \right| , \tag{4.27}$$

similarly introduced in (3.10) for the formulation of Theorem 3.5. Since $\tilde{m}_{\mathrm{sc}}(z)$ is of order 1 for $\langle z \rangle \leq C$ and $\tilde{m}_{\mathrm{sc}}(z) \sim -1/z$ as $z \to \infty$, we infer from (4.24) that $\Theta_i(z) \prec 1$ for spectral parameters z in the domains $\mathbb{D}^{\gamma,\delta,\nu}_{\mathrm{bulk}}$, $\mathbb{D}^{\delta,\nu}_{\mathrm{amb}}$, and $\mathbb{D}^{\nu}_{\mathrm{ext}}$ for any $\nu > 0$.

In the following, we will further use the notion of stochastic domination, introduced in Definition 4.3. Therefore, we will allow the parameter $\delta > 0$ in the characterization of $\mathbb{D}^{\gamma,\delta,\nu}_{\mathrm{bulk}}$ and $\mathbb{D}^{\delta,\nu}_{\mathrm{amb}}$, to be dependent on the implicitly present power $\varepsilon > 0$ of N^{ε} in the definition of stochastic domination (and the parameter $\gamma > 0$), e.g. we write

$$G^{(0)}_{\boldsymbol{x}\boldsymbol{y}} = m_{\mathrm{sc}} \langle \boldsymbol{x}, \boldsymbol{y} \rangle + O_{\prec}\left(\frac{1}{\sqrt{N\eta}} \right) \quad \text{in} \quad \mathbb{D}^{\delta(\varepsilon,\gamma)}_{\gamma}$$

for the isotropic local law in (2.9) from Theorem 2.3 with deterministic vectors $\boldsymbol{x}, \boldsymbol{y}$ satisfying $\|\boldsymbol{x}\|, \|\boldsymbol{y}\| \leq 1$ (note that the local laws in Section 2 are homogeneous in $\|\boldsymbol{x}\|$ and $\|\boldsymbol{y}\|$ resp. $\|B\|$).

We now follow the two paths given in (4.25) and find for deterministic vectors $\boldsymbol{x}, \boldsymbol{y} \perp \boldsymbol{v}_i$, $i \in \{1, 2\}$, with the aid of (4.26), the isotropic local laws (2.7) and (2.9) in Theorem 2.2 and Theorem 2.3, respectively, and $\Theta_i \prec 1$ that

$$\begin{aligned}
G^{(i)}_{\boldsymbol{x}\boldsymbol{y}} &= \tilde{m}_{\mathrm{sc}} \langle \boldsymbol{x}, \boldsymbol{y} \rangle + O_{\prec}\left(\frac{1}{\sqrt{N\eta}} \right) & \qquad G^{(i)}_{\boldsymbol{x}\boldsymbol{y}} &= \tilde{m}_{\mathrm{sc}} \langle \boldsymbol{x}, \boldsymbol{y} \rangle + O_{\prec}\left(\frac{1}{\sqrt{N}\langle z \rangle^2} \right) \\
G^{(i)}_{\boldsymbol{v}_i\boldsymbol{v}_i} &= \frac{\tilde{m}_{\mathrm{sc}}}{1+\sqrt{N}\xi_i \tilde{m}_{\mathrm{sc}}} + O_{\prec}\left(\frac{1}{\sqrt{N\eta}} \right) & \qquad G^{(i)}_{\boldsymbol{v}_i\boldsymbol{v}_i} &= \frac{\tilde{m}_{\mathrm{sc}}}{1+\sqrt{N}\xi_i \tilde{m}_{\mathrm{sc}}} + O_{\prec}\left(\frac{1}{\sqrt{N}\langle z \rangle^2} \right) \\
G^{(i)}_{\boldsymbol{x}\boldsymbol{v}_i} &= O_{\prec}\left(\frac{1}{\sqrt{N\eta}} \right) & \qquad G^{(i)}_{\boldsymbol{x}\boldsymbol{v}_i} &= O_{\prec}\left(\frac{1}{\sqrt{N}\langle z \rangle^2} \right)
\end{aligned}$$

in $\mathbb{D}^{\gamma,\delta(\varepsilon,\gamma),\nu}_{\mathrm{bulk}}$ (left column) and $\mathbb{D}^{\delta(\varepsilon),\nu}_{\mathrm{amb}}$ (right column). The key point in this istropic law for $G^{(i)}$ is that the high-probability error estimates are of the same form as in the isotropic law for the resolvent $G^{(0)}$. This enables us to obtain the isotropic law for the full resolvent $G = G^{(1,2)}$ with high-probability error bounds of the same form from the isotropic law for $G^{(i)}$. As mentioned above, not all paths are suitable vor every choice of deterministic vectors $\boldsymbol{x}, \boldsymbol{y}$, e.g. for $G_{\boldsymbol{v}_2\boldsymbol{v}_2}$ we have to choose the first option in (4.25).

Note that these observations serve as an starting point for an iterative application of Proposition 3.2 yielding isotropic local laws with high-probability error bounds of the same form in every iteration step. In this way we can deal with the general case of a rank$-k-$perturbed Wigner matrix. As the spectral domain $\mathbb{D}^{\nu}_{\mathrm{ext}}$ can be handled completely analogously, we conclude the isotropic laws (3.4), (3.6), and (3.8) in Theorem 3.3 by the definition of stochastic domination.

Step 2: Averaged law. Fix a deterministic matrix $B \in \mathbb{C}^{N\times N}$ with $\|B\| \le 1$ and decompose

$$B = P^\perp BP^\perp + PBP^\perp + P^\perp BP + PBP\,, \tag{4.28}$$

where P is the orthogonal projection on the perturbed subspace (see Definition 3.1). While we employ the same strategy as in the proof of the isotropic law for the first term, we deal with the other three terms using the isotropic law itself. Indeed, using (4.26), the local laws (2.8) and (2.10) from Theorem 2.2 and Theorem 2.3, respectively, and $\Theta_i \prec 1$, we find

$$\begin{aligned}
\langle P^\perp BP^\perp G^{(i)}\rangle &= \tilde{m}_{\mathrm{sc}}\langle P^\perp BP^\perp\rangle + O_\prec\left(\frac{1}{N\eta}\right) &&\text{in}\qquad \mathbb{D}_{\mathrm{bulk}}^{\gamma,\delta(\varepsilon,\gamma),\nu}\,,\\
\langle P^\perp BP^\perp G^{(i)}\rangle &= \tilde{m}_{\mathrm{sc}}\langle P^\perp BP^\perp\rangle + O_\prec\left(\frac{1}{N\langle z\rangle^2}\right) &&\text{in}\qquad \mathbb{D}_{\mathrm{amb}}^{\delta(\varepsilon),\nu}\,.
\end{aligned}$$

This can again be iterated such that we conclude the same bounds with G instead of $G^{(i)}$. For the other three terms in (4.28) we use the isotropic law obtained above, the fact that the number k of rank$-1-$perturbations is fixed (and independent of N), and for the second and third term an additional Schwarz estimate to conclude the desired averaged laws (3.5) and (3.7) from Theorem 3.3 in the spectral domains $\mathbb{D}_{\mathrm{bulk}}^{\gamma,\delta,\nu}$ and $\mathbb{D}_{\mathrm{amb}}^{\delta,\nu}$ by the definition of stochastic domination. The averaged law (3.9) in $\mathbb{D}_{\mathrm{ext}}^{\nu}$ can be proven analogously.

Step 3. Corollary 3.4. Using the notation from the second step, it remains to show that

$$\begin{aligned}
|\langle B(M - M^{(\xi)})\rangle| &\prec \frac{1}{N\eta} &&\text{in}\qquad \mathbb{D}_{\mathrm{bulk}}^{\gamma,\delta(\gamma,\varepsilon),\nu}\,,\\
|\langle B(M - M^{(\xi)})\rangle| &\prec \frac{1}{N\langle z\rangle^2}\left(1 + \langle z\rangle\|PBP\|\right) &&\text{in}\qquad \mathbb{D}_{\mathrm{amb}}^{\delta(\varepsilon),\nu}\,,\\
|\langle B(M - M^{(\xi)})\rangle| &\prec \frac{1}{N\langle z\rangle^2}\left(1 + \langle z\rangle\|PBP\|\right) &&\text{in}\qquad \mathbb{D}_{\mathrm{ext}}^{\nu}\,,
\end{aligned}$$

which are all very easy to conclude from the definition of M and $M^{(\xi)}$. $\square$

Now, we turn to the proof the corresponding results for the outlier domain.

Proof of Theorem 3.5 and Corollary 3.6. In the outlier domain, we cannot use the estimate $\Theta \prec 1$ from the proof of Theorem 3.3 above. By following the same strategy as in the proof above, using that $\langle z\rangle \approx N^{1/2}$ up to a factor of $N^{\pm\nu}$, and by carefully tracking the appearances of Θ, we obtain the isotropic and averaged law in Theorem 3.5, as well as the statement of Corollary 3.6. $\square$

Remark 4.6. It is possible to deal with a general rank$-k-$perturbed Wigner matrix also in the outlier domain $\mathbb{D}_{\mathrm{out}}^{\nu}$ via the strategy used in the proof of Theorem 3.3. However, the error bounds are quite complicated and involve certain polynomials in Θ_i emerging from the iterative employment of Proposition 3.2.

Acknowledgements

I would like to thank László Erdös for many helpful discussions and for supervising this rotation project.

A Auxiliary results

Lemma A.1. (Quadratic implicit function theorem, Lemma D.1 in [6])
Let $\|\cdot\|$ be norm on $\mathbb{C}^d$, $A, B : \mathbb{C}^d \to \mathbb{C}^d$ linear maps and $Q : \mathbb{C}^d \times \mathbb{C}^d \to \mathbb{C}^d$ a bounded $\mathbb{C}^d$-valued quadratic form, i.e.,

$$\|Q\| = \sup_{x,y\in\mathbb{C}^d} \frac{\|Q(x,y)\|}{\|x\|\,\|y\|} < \infty\,.$$

Suppose that A is invertible. Then for $\varepsilon_2 := (2\|A^{-1}\|\,\|Q\|)^{-1}$ and $\varepsilon_1 := \varepsilon_2\,(2\|A^{-1}\|\,\|B\|)^{-1}$ there is a unique function $X : B_{\varepsilon_1} \to B_{\varepsilon_2}$ such that

$$AX(d) + Q(X(d), X(d)) = Bd$$

where B_ε denotes the open ε-ball around 0. Moreover, the funciton X is analytic and satisfies

$$\|X(d_1) - X(d_2)\| \le 2\|A^{-1}\|\,\|B\|\,\|d_1 - d_2\| \quad \text{for all} \quad d_1, d_2 \in B_{\varepsilon_1/2}\,.$$

Lemma A.2. *For random matrices R, T and $p \ge 1$ it holds that $\|\langle R\rangle T\|_p \le \|R\|_{2p}\|T\|_{2p}$.*

Proof. This is a direct consequence of Hölder's inequality and the simple observation that $\|\langle R\rangle\|_{2p} \le \|R\|_{2p}$. □

References

[1] O. H. Ajanki, L. Erdös, and T. Krüger. Stability of the matrix Dyson equation and random matrices with correlations. *Probab. Theory and Related Fields.* Vol. 173.1–2, 2019.

[2] O. H. Ajanki, L. Erdös, and T. Krüger. Universality for general Wigner-type matrices. *Probab. Theory and Related Fields.* Vol. 169.3, 2017.

[3] M. Capitaine, C. Donati-Martin, and D. Feral. Central limit theorems for eigenvalues of defor- mations of Wigner matrices. *Ann. Inst. H. Poinc. Probab. Stat.* Vol. 48 no. 1, 2012.

[4] L. Erdös. The matrix Dyson equation and its applications for random matrices. *arXiv:* 1903.10060, 2019.

[5] L. Erdös, A. Knowles, H.-T. Yau, and J. Yin. The local semicircle law for a general class of random matrices. *Electron. J. Probab.*, Vol. 18 no. 59, 2013.

[6] L. Erdös, T. Krüger, and D. Schröder. Random matrices with slow correlation decay. *Forum of Math., Sigma.* Vol. 7, 2019.

[7] L. Erdös and H.-T. Yau. A Dynamical Approach to Random Matrix Theory, *Courant Lecture Notes in Mathematics*, 28, 2017.

[8] J. Geronimo and T. Hill. Necessary and sufficient condition that the limit of Stieltjes transforms is a Stieltjes transform, *J. Approx. Theory* 121, 2003.

[9] A. Knowles and J. Yin. The isotropic semicircle law and deformation of Wigner matrices. *Comm. on Pure and Appl. Math.* 66.11, 2013.

[10] M. L. Mehta. *Random Matrices and the Statistical Theory of Energy Levels.* Academic Press, 1967.

[11] T. Tao. Outliers in the spectrum of iid matrices with bounded rank perturbations. *Probab. Theory and Related Fields.* Vol. 155, 2013.

[12] T. Tao and V. Vu. Random matrices: universality of local eigenvalue statistics, *Acta Math.*, 2011.

[13] H. Weyl. Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung). *Mathematische Annalen.* LXXI, 1911.

[14] E. P. Wigner. On the Distribution of the Roots of Certain Symmetric Matrices. *Ann. of Math.* Vol. 67 No. 2, 1958.