

## Random characteristics for Wigner matrices

Per von Soosten\*      Simone Warzel†

### Abstract

We extend the random characteristics approach to Wigner matrices whose entries are not required to have a normal distribution. As an application, we give a simple and fully dynamical proof of the weak local semicircle law in the bulk.

**Keywords:** local semicircle law; random characteristics; Wigner matrices.

**AMS MSC 2010:** 15B52.

Submitted to ECP on June 26, 2019, final version accepted on November 19, 2019.

Supersedes arXiv:1906.10677.

## 1 Introduction

Starting with the seminal works of Erdős, Schlein, and Yau [15, 16], a large portion of recent progress in random matrix theory rests on strong concentration of measure phenomena for the resolvent on almost microscopic scales. Such estimates are an important model-dependent step in proving Wigner-Dyson universality of the local eigenvalue statistics and uniform delocalization bounds for the eigenvectors. The mean-field setting is especially well-understood: the methods in [12, 18, 17] give strong results and the central ideas are robust enough to cover a wide range of models (see, for example, [2, 3, 5, 6, 7, 9, 14, 13] and references therein). The most basic example consists of the  $N \times N$  Wigner matrices, whose entries  $H_{ij}$  are drawn, independently up to symmetry, from some density with mean zero and variance  $N^{-1}$ . A typical result proves that the resolvent  $G(z) = (H - z)^{-1}$  has essentially deterministic entries

$$G_{ij}(z) \approx m_{sc}(z)\delta_{ij}, \quad (1.1)$$

the approximation being valid on the smallest possible scale  $\text{Im } z \gg N^{-1}$ . The function

$$m_{sc}(z) = \frac{-z + \sqrt{z^2 - 4}}{2} = \int_{-2}^2 \frac{1}{\lambda - z} \frac{\sqrt{4 - \lambda^2}}{2\pi} d\lambda$$

is the Stieltjes transform of the semicircle law.

The approximations (1.1) are usually proved by deriving an approximate self-consistent equation

$$1 + (z + \langle G(z) \rangle)G(z) \approx 0, \quad (1.2)$$

where we used the notation  $\langle A \rangle = N^{-1} \text{Tr } A$  that we retain throughout this paper. The stability of the self-consistent equation is then used to show that the approximate solution

---

\*Harvard University, United States of America. E-mail: vonsoosten@math.harvard.edu

†MCQST & Technische Universität München, Germany. E-mail: warzel@ma.tum.de

$G(z)$  is close to the solution  $m_{sc}(z)$  of the exact equation. Because it is usually not possible to prove the validity of the self-consistent equation on local scales  $\text{Im } z \gg N^{-1}$  directly, the analysis uses the stability of (1.2) again to show that rough estimates on the validity of (1.2) self-improve at finer scales. This idea enables a careful bootstrapping scheme, which allows one to successively improve the scale of the approximation (1.1).

It was noted by Pastur [22], that the self-consistent equation (1.2) can also be viewed as the terminal constraint at  $t = 1$  of the deterministic advection equation

$$\partial_t G(t, z) = \langle G(t, z) \rangle \partial_z G(t, z), \quad G(0, z) = -z^{-1}, \tag{1.3}$$

which can be easily solved by considering a suitable set of deterministic characteristic curves. In fact, if one generates a Gaussian Wigner matrix dynamically by evolving the entries with Brownian motion [11], one can derive a stochastic version of (1.3)

$$dG(t, z) = \langle G(t, z) \rangle \partial_z G(t, z) dt + dM(t, z) \tag{1.4}$$

with some explicit matrix-valued martingale  $M(t, z)$  (see, for example, [24]). This approach can be used to derive the validity of the semicircle law on global scales  $\text{Im } z = \mathcal{O}(1)$  in the infinite-volume limit [4]. Deterministic characteristic curves have also featured in the analysis on local scales in more recent random matrix literature such as [1, 8, 19, 20].

The works [25, 26] showed that the SDE (1.4) can also be analyzed on local scales by considering the evolution along the random characteristic

$$\dot{z}(t) = -\langle G(t, z(t)) \rangle, \tag{1.5}$$

yielding a simple dynamical mechanism for directly proving concentration of measure for the resolvent. The method thus allows one to completely separate any stability arguments from local concentration of measure estimates, thereby circumventing the need for a bootstrap argument.

The purpose of this paper is to show that this approach to local resolvent estimates is not limited to Wigner matrices with Gaussian entries. The basis of this is the construction of a matrix martingale  $H(t)$  whose rescaled entries follow a given density  $\varrho$  at time  $t = 1$ . This will be possible for densities  $\varrho$  satisfying the following assumption.

**Assumption 1.1.** *The density  $\varrho > 0$  is strictly positive, has zero mean and unit variance, and the function*

$$a(h) = \frac{1}{\varrho(h)} \int_h^\infty k \varrho(k) dk, \tag{1.6}$$

*is bounded and Lipschitz continuous on  $\mathbb{R}$ .*

The integral equation (1.6) is equivalent to

$$\varrho(h) = \frac{C}{a(h)} \exp \left( - \int_0^h \frac{k}{a(k)} dk \right),$$

so the condition that  $a$  be bounded essentially amounts to  $\varrho$  being sub-Gaussian, whereas the Lipschitz continuity of  $a$  is linked to the regularity of  $\varrho$ . Madan and Yor [21] used the Lipschitz continuity of  $a$  to construct a scalar martingale  $h(t)$  satisfying

$$dh(t) = a \left( \frac{h(t)}{\sqrt{t}} \right)^{1/2} db(t), \quad h(0) = 0,$$

where  $b$  is a standard Brownian motion. Inspired by an idea of Dupire [10], they then showed that Kolmogorov's forward equation implies

$$\mathbb{P}(h(t) \in dx) = t^{-1/2} \varrho(t^{-1/2}x) dx. \tag{1.7}$$

Hence, to construct the matrix process  $H(t)$ , we take an array  $(B_{ij})$  of standard Brownian motions that are independent up to the constraint  $B_{ij} = B_{ji}$  and define  $H_{ij}(t)$  as the solution of the SDE

$$dH_{ij}(t) = \frac{1}{\sqrt{N}} a \left( \sqrt{\frac{N}{t}} H_{ij}(t) \right)^{1/2} dB_{ij}(t), \quad H_{ij}(0) = 0.$$

Then  $H(t) = \sqrt{t}H$  in distribution, where  $H$  is a Wigner matrix whose rescaled entries  $\sqrt{N}H_{ij}$  have distribution  $\varrho$ . It will be important in the sequel that the quadratic variation of  $H_{ij}(t)$  satisfies

$$d[H_{ij}](t) = \sigma_{ij}(t) dt, \quad \sigma_{ij}(t) = \frac{1}{N} a \left( \sqrt{\frac{N}{t}} H_{ij}(t) \right).$$

Combining Itô's lemma with the Neumann expansion of the resolvent yields the matrix-valued SDE

$$dG(t, z) = -G(t, z) dH(t) G(t, z) + G(t, z) dH(t) G(t, z) dH(t) G(t, z)$$

for the resolvent process  $G(t, z) = (H(t) - z)^{-1}$  with  $z \in \mathbb{C}_+$ . Multiplying out the drift gives an expression of the form

$$dG(t, z) = -G(t, z) (dH(t) - \mathcal{T}[t, G(t, z)] dt) G(t, z) + G(t, z) \mathcal{S}[t, G(t, z)] G(t, z) dt. \quad (1.8)$$

The matrix-valued operator  $\mathcal{S}[t, \cdot]$ , which acts on an  $N \times N$  matrix  $A$  as

$$\mathcal{S}[t, A]_{ij} = \delta_{ij} \sum_k \sigma_{ik}(t) A_{kk},$$

is a dominant term whose presence in (1.8) provides the self-energy corrections in the resolvent. On the other hand, the operator

$$\mathcal{T}[t, A]_{ij} = (1 - \delta_{ij}) \sigma_{ij}(t) A_{ji}$$

should be thought of as a finite-volume error reflecting the lack of Hermitian symmetry in our model. Therefore, the self-energy correction coincides with the Itô correction. This is similar in spirit (and identical in the Gaussian case) to the cumulant expansion of He, Knowles, and Rosenthal [18]. However, the technical details are simpler here because the quadratic variation process naturally encodes the fluctuations around the Gaussian noise driving the dynamics.

We will illustrate our method by giving a new proof of the weak bulk local semicircle law for the normalized resolvent trace of Wigner matrices whose entries are drawn from densities satisfying Assumption 1.1. To state this result, we will make use of the stochastic domination language of [12]. Let  $\{X(u)\}_{u \in U}$  and  $\{Y(u)\}_{u \in U}$  be two non-negative  $N$ -dependent families of random variables. We will say that  $X(u)$  is stochastically dominated by  $Y(u)$  uniformly in  $u \in U$ , if, given any  $\varepsilon, p > 0$ , the inequality

$$\sup_{u \in U} \mathbb{P}(X(u) > N^\varepsilon Y(u)) \leq N^{-p}$$

is satisfied for all sufficiently large  $N \geq N_0(\varepsilon, p)$ . We express this relationship by writing  $X \prec Y$  uniformly in  $u \in U$ . Although most of the probabilities in this paper can be controlled with an explicit exponential tail, we have refrained from doing so for the sake of brevity. Our proof of the local semicircle law will be valid in a bulk spectral domain

$$D = W + i(\eta, 1)$$

where  $W = [W_1, W_2] \subset [-2 + \kappa, 2 - \kappa]$  for some  $\kappa > 0$ . For simplicity, we will assume that the minimal spectral scale is given by

$$\eta = N^{-1+\theta}$$

where  $\theta > 0$  is fixed, but arbitrarily small.

**Theorem 1.2.** *We have*

$$\sup_{z \in D} |\langle G(1, z) \rangle - m_{sc}(z)| \prec \frac{1}{\sqrt{N\eta}}.$$

We stress that this theorem first appeared in [15]. It has been extended both to Wigner matrices with minimal assumption on the distribution of the entries [2] and to random matrices with a more general spread-out variance profile [12].

Theorem 1.2 implies that  $\langle \text{Im} G(1, z) \rangle$  is bounded both above and away from zero when  $z \in D$ . The Schur complement formula and classical concentration of measure results show that

$$\sup_{z \in D} \left( \frac{1 + m_{sc}(z)}{N\eta} \right)^{-1/2} \left| \frac{1}{G_{kk}(1, z)} + z + m_{sc}(z) \right| \prec 1$$

so Taylor expansion yields

$$\sup_{z \in D} |G(1, z)_{kk} - m_{sc}(z)| \prec \frac{1}{\sqrt{N\eta}}. \tag{1.9}$$

Similar considerations apply to the off-diagonal entries

$$\sup_{z \in D} |G(1, z)_{jk}| \prec \frac{1}{\sqrt{N\eta}}, \quad j \neq k. \tag{1.10}$$

These calculations have become standard and are explained in great detail in [17]. However, given Theorem 1.2, no further bootstrapping is required to conclude (1.9) and (1.10).

The organization of this paper is as follows. In Section 2, we show in which sense the evolution (1.8) approximates (1.4). Then, in Section 3, we prove that (1.5) defines an approximate characteristic flow, which we use to derive the local semicircle law in Section 4.

## 2 The self-energy correction

In comparison the Gaussian case, the main complication in the random characteristic approach for more general Wigner matrices is that the self-energy operator  $S[t, \cdot]$  remains random and time-dependent. Nevertheless, we will prove in Theorem 2.1 that

$$S[t, G(t, z)] = \langle G(t, z) \rangle + \mathcal{O} \left( \sqrt{\frac{1 + \text{Im} \langle G(t, z) \rangle}{N \text{Im} z}} \right)$$

with very high probability so that the characteristic curve (1.5) still counteracts a large part of the term

$$G(t, z) S[t, G(t, z)] G(t, z) dt$$

in (1.8). The resulting error term is not small a-priori, but retains a specific structure that enables the Grönwall scheme of Lemma 3.3.

For technical reasons, we will prove Theorem 2.1 for spectral parameters  $z$  in a very large domain

$$D' = (W_1 - 3\eta^{-1}, W_2 + 3\eta^{-1}) + i(\eta/4, 1 + 3\eta^{-1}) \tag{2.1}$$

and for times  $t$  that are greater than a cutoff

$$t_0 = N^{-K} \tag{2.2}$$

with some fixed  $K \in \mathbb{N}$  to be specified later in the proof of Lemma 3.3. In the statement of Theorem 2.1, and throughout this paper,  $\|\cdot\|$  denotes the operator norm.

**Theorem 2.1.** *For any choice of  $K \in \mathbb{N}$  in (2.2) we have*

$$\sup_{t \in [t_0, 1], z \in D'} \left( \frac{1 + \text{Im} \langle G(t, z) \rangle}{N \text{Im} z} \right)^{-1/2} \|\mathcal{S}[t, G(t, z)] - \langle G(t, z) \rangle\| \prec 1.$$

*Proof.* We first show that

$$|\mathcal{S}[t, G(t, z)]_{kk} - \langle G(t, z) \rangle| \prec \sqrt{\frac{1 + \text{Im} \langle G(t, z) \rangle}{N \text{Im} z}} \tag{2.3}$$

uniformly in  $t \in [t_0, 1]$ ,  $z \in D'$ , and  $k \in \{1, \dots, N\}$ . Let  $H^k(t)$  denote the matrix obtained by replacing the  $k$ -th row and column of  $H(t)$  by zeros. Denoting by  $G^k(t, z)$  the resolvent of  $H^k(t)$ , the resolvent identity implies

$$G_{jj}(t, z) = G_{jj}^k(t, z) + \frac{G_{kj}(t, z)G_{jk}(t, z)}{G_{kk}(t, z)} \tag{2.4}$$

when  $j \neq k$ . Since  $\sigma_{jk}(t) \leq CN^{-1}$  by Assumption 1.1, we conclude the uniform deterministic estimate

$$\begin{aligned} \left| \sum_j \sigma_{kj}(t)(G_{jj}^k(t, z) - G_{jj}(t, z)) \right| &\leq \frac{C}{N \text{Im} z} + \frac{1}{|G_{kk}(t, z)|} \sum_j \sigma_{kj}(t) |G_{kj}(t, z)|^2 \\ &\leq \frac{C}{N \text{Im} z} \left( 1 + \frac{\text{Im} G_{kk}(t, z)}{|G_{kk}(t, z)|} \right) \leq \frac{C}{N \text{Im} z} \end{aligned}$$

using (2.4) and the trivial resolvent bound for the summand with  $j = k$ . Similarly, we have

$$|\langle G^k(t, z) \rangle - \langle G(t, z) \rangle| \leq \frac{C}{N \text{Im} z}.$$

To prove (2.3) it thus suffices to show that

$$\left| \sum_j \sigma_{kj}(t)G_{jj}^k(t, z) - \langle G^k(t, z) \rangle \right| \prec \sqrt{\frac{1 + \text{Im} \langle G^k(t, z) \rangle}{N \text{Im} z}} \tag{2.5}$$

uniformly in  $t \in [t_0, 1]$ ,  $z \in D'$ , and  $k \in \{1, \dots, N\}$ . After conditioning on  $H^k(t)$ , the random variables  $\sigma_{kj}(t)G_{jj}^k(t, z)$  are independent and bounded by  $C N^{-1}|G_{jj}^k(t, z)|$ . Using Hoeffding's inequality with respect to the conditional probability  $\mathbb{P}_k$  we get

$$\mathbb{P}_k \left( \left| \sum_j \sigma_{kj}(t)G_{jj}^k(t, z) - \mu \right| > v\alpha \right) \leq 2 \exp(-c\alpha^2)$$

with

$$\mu = \mathbb{E}_k \left[ \sum_j \sigma_{kj}(t)G_{jj}^k(t, z) \right] = \langle G^k(t, z) \rangle$$

and

$$v^2 = \frac{1}{N^2} \sum_j |G_{jj}^k(t, z)|^2 \leq \frac{1}{N^2} \sum_{i,j} |G_{ij}^k(t, z)|^2 = \frac{\langle \text{Im} G^k(t, z) \rangle}{N \text{Im} z},$$

which proves (2.5).

The extension of (2.3) to the maximum over  $k$  is by the union bound, whereas the extension to the supremum over all  $z \in D'$  and  $t \in [t_0, 1]$  beyond the cutoff  $t_0 = N^{-K}$  is by a stochastic continuity bound that we turn to in the next step. We fix  $r > 0$  and consider the neighborhood

$$B_r(t, z) = \{(s, w) \in [t_0, 1] \times D' : |t - s|^2 + |z - w|^2 \leq r^2\}$$

around some point  $(t, z) \in [t_0, 1] \times D'$ . The theorem then follows from (2.3) and

$$\sup_{(s,w) \in B_r(t,z)} \|\mathcal{S}[s, G(s, w)] - \mathcal{S}[t, G(t, z)]\| \prec N^K r^{1/2}. \tag{2.6}$$

Indeed, for every  $L \in \mathbb{N}$  and  $r = N^{-L}$ , there exists an  $L' \in \mathbb{N}$  and a finite grid  $\Lambda \subset [t_0, 1] \times D'$  of cardinality  $|\Lambda| \leq N^{L'}$  such that  $\text{dist}(\Lambda, [t_0, 1] \times D) \leq r$ . Let  $\varepsilon, p > 0$  be arbitrary. Applying the union bound to (2.3) and (2.6) shows that the events

$$\|\mathcal{S}[t, G(t, z)] - \langle G(t, z) \rangle\| \leq N^\varepsilon \left( \frac{1 + \text{Im} \langle G(t, z) \rangle}{N \text{Im} z} \right)^{1/2}$$

and

$$\sup_{(s,w) \in B_r(t,z)} \|\mathcal{S}[s, G(s, w)] - \mathcal{S}[t, G(t, z)]\| \leq N^{\varepsilon+K} r^{1/2}$$

hold simultaneously for all  $(t, z) \in \Lambda$  with probability  $1 - N^{-p}$ , when  $N$  is large enough. The claim of Theorem 2.1 follows since  $L$  can be chosen arbitrarily large.

Finally, to prove (2.6), we note first that  $\|H(t + s) - H(t)\|$  is a submartingale in  $s \geq 0$ . Doob's inequality and classical norm bounds [23] for the rescaled Wigner matrix  $H(t + r) - H(t)$  show that

$$\left( \mathbb{E} \left| \sup_{s \in [t, t+r]} \|H(s) - H(t)\| \right|^p \right)^{1/p} \leq \frac{p}{p-1} (\mathbb{E} \|H(t+r) - H(t)\|^p)^{1/p} \leq C_p r^{1/2}$$

for all  $p > 1$  with some  $p$ -dependent constant  $C_p < \infty$ . Since the resolvent is Lipschitz continuous with constant  $\eta^{-2}$  in both  $H$  and  $z$ , this implies

$$\sup_{|t-s|^2+|z-w|^2 \leq r^2} \|G(s, w) - G(t, z)\| \prec \eta^{-2} r^{1/2}.$$

Doob's inequality also implies that

$$\left( \mathbb{E} \left| \sup_{s \in [t, t+r]} |H_{kj}(s) - H_{kj}(t)| \right|^p \right)^{1/p} \leq C_p \sqrt{\frac{r}{N}}$$

for any pair of indices  $j, k$ . In combination with the Lipschitz continuity of  $a$ , this implies

$$\sup_{s \in [t, t+r]} |\sigma_{kj}(s) - \sigma_{kj}(t)| \leq \sup_{s \in [t, t+r]} \frac{C}{\sqrt{N}} \left| \frac{H_{kj}(s)}{\sqrt{s}} - \frac{H_{kj}(t)}{\sqrt{t}} \right| \prec \frac{r^{1/2}}{N t_0} = N^{K-1} r^{1/2}$$

for any  $t \geq t_0$ . Since  $\mathcal{S}[t, G(t, z)]$  is a polynomial combination of the  $\sigma_{kj}(t)$  and  $G(t, z)$ , which are bounded by  $N^{-1}$  and  $\eta^{-1} \leq N$  respectively, the bound (2.6) follows.  $\square$

### 3 Fluctuations along characteristic curves

We will show that, for fixed realizations of the randomness, the unique solutions of

$$\dot{\gamma}(t, z) = -\langle G(t, \gamma(t, z)) \rangle, \quad \gamma(0, z) = z$$

serve as approximate characteristic curves of the resolvent SDE (1.8). We start these curves at spectral parameters  $z$  in an initial spectral domain

$$D_0 = (W_1 - 2\eta^{-1}, W_2 + 2\eta^{-1}) + i(\eta/2, 1 + 2\eta^{-1}) \setminus B_\delta(0),$$

where  $\delta > 0$  is any constant satisfying

$$\frac{1}{2\delta} - 2\delta \geq 1 + \sup\{|z| : z \in D\}. \tag{3.1}$$

Thus  $G(0, z) = -z^{-1}$  is uniformly bounded and Lipschitz continuous in  $D_0$ . Given an initial point  $z \in D_0$ , we consider the process

$$R(t, z) = G(t \wedge \tau_z, \xi(t, z))$$

where

$$\xi(t, z) = \gamma(t \wedge \tau_z, z)$$

is the characteristic flow stopped at

$$\tau_z = \inf\{t > 0 : \text{Im } \xi(t, z) \leq \eta/4\}.$$

The main observation is that  $\langle R(t, z) \rangle$  is approximately constant.

**Theorem 3.1.** *The process  $R(t, z)$  satisfies*

$$\sup_{t \leq 1} |\langle R(t, z) \rangle - \langle R(0, z) \rangle| \prec \frac{1}{\sqrt{N\eta}}$$

uniformly in  $z \in D_0$ .

*Proof.* Since  $\xi$  defines a piecewise  $C^1$ -process, an application of Itô's lemma shows that  $R(t, z)$  satisfies the same SDE as  $G(t, z)$  but with an additional counter-term in the drift. More precisely, while  $t \leq \tau_z$ , the evolution consists of two terms

$$dR(t, z) = dF(t, z) + dA(t, z)$$

with

$$\begin{aligned} dF(t, z) &= -R(t, z) (dH(t) - \mathcal{T}[t, R(t, z)] dt) R(t, z), \\ dA(t, z) &= R(t, z) (\mathcal{S}[t, R(t, z)] - \langle R(t, z) \rangle) R(t, z) dt. \end{aligned}$$

When dealing with the integrated versions of these processes, we will choose the initial conditions  $F(0, z) = A(0, z) = 0$  so that  $\langle R(t, z) \rangle - \langle R(0, z) \rangle = \langle F(t, z) \rangle + \langle A(t, z) \rangle$ . The proof of Theorem 3.1 is then a direct consequence of the estimates on  $F(t, z)$  and  $A(t, z)$  provided in Lemma 3.2 and Lemma 3.3, respectively.  $\square$

The proofs of the subsequent results will repeatedly use the crucial fact that any continuously differentiable function  $f$  satisfies the identity

$$\begin{aligned} \int_a^b f'(\text{Im } \xi(s, z)) \langle \text{Im } R(s, z) \rangle ds &= - \int_a^b f'(\text{Im } \xi(s, z)) d(\text{Im } \xi(s, z)) \\ &= f(\text{Im } \xi(a, z)) - f(\text{Im } \xi(b, z)) \end{aligned}$$

provided that  $a, b \leq \tau_z$ . The term  $f(\text{Im } \xi(a, z)) - f(\text{Im } \xi(b, z))$  is then usually estimated by a trivial bound in terms of  $\eta$ . We refer to these two steps simply as the "integration trick".

**Lemma 3.2.** *The process  $F(t, z)$  satisfies*

$$\sup_{t \leq 1} |\langle F(t, z) \rangle| \prec \frac{1}{N\eta}$$

uniformly in  $z \in D_0$ .

*Proof.* The martingale part of  $F(t, z)$  is

$$dM(t) = -R(t, z) dH(t) R(t, z)$$

and the quadratic variation of its unit trace  $\langle M \rangle$  is given by

$$d[\langle M \rangle](t) = \langle R^*(t, z) dH(t) R^*(t, z) \rangle \langle R(t, z) dH(t) R(t, z) \rangle \leq \frac{C}{N^2} \|\sqrt{\sigma}(t) \odot R^2(t, z)\|_2^2 dt,$$

where  $\|\cdot\|_2 = \sqrt{\text{Tr}|\cdot|^2}$  denotes the Hilbert-Schmidt norm,  $\odot$  denotes the entrywise product, and  $\sqrt{\sigma}(t)$  is the matrix with entries  $\sqrt{\sigma_{jk}(t)}$ . If  $t \leq \tau_z$ , we conclude that

$$\begin{aligned} [\langle M \rangle](t) &\leq \frac{C}{N^2} \int_0^t \|\sqrt{\sigma}(s) \odot R^2(s, z)\|_2^2 ds \\ &\leq \frac{C}{N^3} \int_0^t \frac{\text{Tr} |R|^2(s, z)}{(\text{Im} \xi(s, z))^2} ds \\ &\leq \frac{C}{N^2} \int_0^t \frac{\langle \text{Im} R(s, z) \rangle}{(\text{Im} \xi(s, z))^3} ds = \frac{C}{N^2} \left( \frac{1}{(\text{Im} \xi(t, z))^2} - \frac{1}{(\text{Im} \xi(0, z))^2} \right) \end{aligned}$$

by the integration trick. Thus

$$\sup_{t \leq \tau_z} |\langle M(t) \rangle| \prec (N\eta)^{-1}$$

by the Burkholder-Davis-Gundy inequality and Markov's inequality. The other term in  $F(t, z)$ ,

$$\sup_{t \leq \tau_z} \left| \int_0^t \langle R(s, z) \mathcal{T}[s, R(s, z)] R(s, z) \rangle ds \right| \leq \frac{C}{N} \int_0^{\tau_z} \frac{\langle |R(s, z)|^2 \rangle}{\text{Im} \xi(s, z)} ds = \frac{C}{N} \int_0^{\tau_z} \frac{\langle \text{Im} R(s) \rangle}{(\text{Im} \xi(s, z))^2} ds,$$

is also stochastically dominated by  $(N\eta)^{-1}$  uniformly in  $z \in D_0$  because of the integration trick.  $\square$

**Lemma 3.3.** *The process  $A(t, z)$  satisfies*

$$\sup_{t \leq 1} |\langle A(t, z) \rangle| \prec \frac{1}{\sqrt{N\eta}}$$

uniformly in  $z \in D_0$ .

*Proof.* We will split the integral

$$|\langle A(t, z) \rangle| \leq \left( \int_0^{t_0} + \int_{t_0}^t \right) |\langle R(s, z) (\mathcal{S}[s, R(s, z)] - \langle R(s, z) \rangle) R(s, z) \rangle| ds \quad (3.2)$$

at the point  $t_0$  from (2.2). By the trivial bound on the resolvent, the characteristic  $\xi(t, z)$  started at  $z \in D_0$  remains in the domain  $D'$  from (2.1) for all  $t \leq 1$ . Setting

$$u(t) = 1 + \text{Im} \langle R(t, z) \rangle,$$

Theorem 2.1 shows that the second part of the integral in (3.2) is bounded by

$$N^{\theta/2} \int_{t_0}^t \frac{\langle |R(s, z)|^2 \rangle}{\sqrt{N \operatorname{Im} \xi(s, z)}} \sqrt{u(s)} ds \leq \frac{N^{\theta/2}}{\sqrt{N}} \int_0^t \frac{\operatorname{Im} \langle R(s, z) \rangle}{(\operatorname{Im} \xi(s, z))^{3/2}} u(s) ds \quad (3.3)$$

with probability  $1 - N^{-p}$  for arbitrary  $p > 0$  and large enough  $N$ . The first part of the integral in (3.2) can be bounded by  $Ct_0\eta^{-3}$  using the trivial bound on the operator norm of the resolvent. So choosing the cutoff exponent  $K$  in (2.2) large enough and using Lemma 3.2 yields

$$u(t) \leq 1 + u(0) + N^{\frac{\theta}{2}} \int_0^t \frac{\operatorname{Im} \langle R(s, z) \rangle}{(\operatorname{Im} \xi(s, z))^{3/2}} u(s) ds$$

on an event that also has probability  $1 - N^{-p}$  for arbitrary  $p > 0$  and large enough  $N$ . On this event, Grönwall's inequality implies

$$u(t) \leq (1 + u(0)) \exp \left( N^{\frac{\theta}{2}} \int_0^t \frac{\operatorname{Im} \langle R(s, z) \rangle}{N^{1/2} (\operatorname{Im} \xi(s, z))^{3/2}} ds \right)$$

and the integral inside the exponential is bounded by  $4(N\eta)^{-1/2} = 4N^{-\theta/2}$  because of the integration trick. Since  $\langle R(0, z) \rangle \leq \delta^{-1}$  for  $z \in D_0$ , we have shown that

$$\sup_{t \leq 1} u(t) \prec 1.$$

We now insert this bound on  $u$  back into the integral in (3.3). Choosing the cutoff exponent  $K$  in (2.2) large enough yields

$$\sup_{t \leq 1} |\langle A(t, z) \rangle| \prec \frac{1}{N\eta} + \frac{1}{\sqrt{N}} \int_0^1 \frac{\operatorname{Im} \langle R(s, z) \rangle}{(\operatorname{Im} \xi(s, z))^{3/2}} ds \prec \frac{1}{\sqrt{N\eta}}$$

via the integration trick. □

Having established Theorem 3.1, we now argue that the stochastic domination holds simultaneously for a continuum of points using a discretization argument.

**Theorem 3.4.** *Let  $\tau'_z = \inf\{t > 0 : \operatorname{Im} \gamma(t, z) \leq \eta/2\}$ . Then*

$$\sup_{z \in D_0} \sup_{t \leq 1 \wedge \tau'_z} |\langle R(t, z) \rangle - \langle R(0, z) \rangle| \prec \frac{1}{\sqrt{N\eta}}. \quad (3.4)$$

*Proof.* We begin by proving a Lipschitz bound for the characteristic flow  $\gamma(t, \cdot)$ . Fix  $z, w \in \mathbb{C}_+$  and let

$$u(t) = \gamma(t, z) - \gamma(t, w).$$

As long as  $\operatorname{Im} \gamma(t, z), \operatorname{Im} \gamma(t, w) > \eta/4$ , we have

$$\begin{aligned} |u(t)| &\leq |u(0)| + \int_0^t |\langle R(s, z) - R(s, w) \rangle| ds \\ &\leq |u(0)| + \frac{1}{N} \int_0^t |u(s)| \|R(s, z)\|_2 \|R(s, w)\|_2 ds \\ &\leq |u(0)| + \int_0^t \frac{|u(s)|}{2} \left( \frac{\langle \operatorname{Im} R(s, z) \rangle}{\operatorname{Im} \gamma(s, z)} + \frac{\langle \operatorname{Im} R(s, w) \rangle}{\operatorname{Im} \gamma(s, w)} \right) ds \end{aligned}$$

so Grönwall's inequality and the integration trick yield

$$\log \frac{|u(t)|}{|u(0)|} \leq \frac{1}{2} \int_0^t \left( \frac{\langle \operatorname{Im} R(s, z) \rangle}{\operatorname{Im} \gamma(s, z)} + \frac{\langle \operatorname{Im} R(s, w) \rangle}{\operatorname{Im} \gamma(s, w)} \right) ds = \frac{1}{2} \log \left( \frac{\operatorname{Im} z \operatorname{Im} w}{\operatorname{Im} \gamma(t, z) \operatorname{Im} \gamma(t, w)} \right).$$

Rearranging, we get

$$|\gamma(t, z) - \gamma(t, w)| \leq \sqrt{\frac{\operatorname{Im} z \operatorname{Im} w}{\operatorname{Im} \gamma(t, z) \operatorname{Im} \gamma(t, w)}} |z - w|. \tag{3.5}$$

To prove the theorem, we choose a finite grid  $\Lambda \subset D_0$  such that its cardinality is bounded by  $|\Lambda| \leq N^L$  for some  $L \in \mathbb{N}$  and

$$\operatorname{dist}(D_0, \Lambda) \leq \frac{\eta^4}{\sqrt{N\eta}}.$$

From Theorem 3.1 and the union bound, we have

$$\sup_{z \in \Lambda} \sup_{t \leq 1} |\langle R(t, z) \rangle - \langle R(0, z) \rangle| \prec \frac{1}{\sqrt{N\eta}}, \tag{3.6}$$

so it suffices to show that the left side of (3.6) controls the left side of (3.4). Given  $z \in D_0$ , we pick  $w \in \Lambda$  such that  $|z - w| \leq \frac{\eta^3}{\sqrt{N\eta}}$ . Then  $\tau'_w \leq \tau_w$  and  $\tau'_z \leq \tau_w$  since (3.5) guarantees that

$$|\gamma(t, z) - \gamma(t, w)| \leq C \frac{\eta^2}{\sqrt{N\eta}}$$

as long as  $\operatorname{Im} \gamma(t, z) \geq \eta/4$  and  $\operatorname{Im} \gamma(t, w) \geq \eta/4$ . The trivial  $C\eta^{-2}$ -Lipschitz continuity of the resolvent in  $D_0$  then shows that the process  $\langle R(t, z) \rangle$  stays within an error  $C/\sqrt{N\eta}$  of  $\langle R(t, w) \rangle$  for all times  $t \leq \tau'_z$ .  $\square$

#### 4 Proof of the local semicircle law

To prove Theorem 1.2 we choose an arbitrarily small  $\varepsilon > 0$  and prove that the local semicircle law is valid on the event

$$\mathcal{A}_\varepsilon = \left\{ \sup_{z \in D_0} \sup_{t \leq 1 \wedge \tau'_z} |\langle R(t, z) \rangle - \langle R(0, z) \rangle| \leq \frac{N^\varepsilon}{\sqrt{N\eta}} \right\}$$

with an error that is also of order  $N^\varepsilon/\sqrt{N\eta}$ . By Theorem 3.4 the event  $\mathcal{A}_\varepsilon$  has probability  $1 - N^{-p}$  for arbitrary  $p > 0$  when  $N$  is large enough. We begin by noting that the characteristic flow may be computed explicitly on  $\mathcal{A}_\varepsilon$ .

**Lemma 4.1.** *Let  $\varepsilon > 0$  and suppose  $\mathcal{A}_\varepsilon$  occurs. Then for every  $z \in D$ , there exists  $w \in D_0$  such that  $z = \gamma(1, w)$  and*

$$\left| w + \frac{1}{w} - z \right| \leq \frac{CN^\varepsilon}{\sqrt{N\eta}}.$$

*Proof.* Let  $\lambda$  be the time-reversal of  $\gamma$  defined by

$$\dot{\lambda}(t, \zeta) = \langle G(1 - t, \lambda(t, \zeta)) \rangle, \quad \lambda(0, \zeta) = \zeta$$

so that  $\gamma(1, \lambda(1, \zeta)) = \zeta$ . Whenever  $\zeta \in D$  and  $\lambda(1, \zeta) \in D_0$ , we necessarily have  $\tau_\zeta > 1$ , so

$$\gamma(1, \zeta) = \zeta - \langle R(0, \zeta) \rangle - \int_0^1 (\langle R(s, \zeta) \rangle - \langle R(0, \zeta) \rangle) ds = \zeta + \zeta^{-1} + \mathcal{O}\left(\frac{N^\varepsilon}{\sqrt{N\eta}}\right) \tag{4.1}$$

by the definition of  $\mathcal{A}_\varepsilon$ . The desired  $w$  in the conclusion of the lemma will be

$$w = \lambda(1, z),$$

which satisfies  $w \in D_0 \cup B_\delta(0)$  by the trivial resolvent bound. It thus remains to prove that  $w \notin B_\delta(0)$ . For this, we note that  $\lambda(1, D)$  is simply connected since  $\lambda(1, \cdot)$  is a

homeomorphism. If there were any point  $\zeta \in \lambda(1, D) \cap B_\delta(0)$ , there would also be some point  $\zeta' \in \lambda(1, D) \cap D_0$  with  $|\zeta'| = 2\delta$ . Since  $\zeta' \in D_0$ , the relation (4.1) implies

$$|\gamma(1, \zeta')| + \mathcal{O}\left(\frac{N^\varepsilon}{\sqrt{N\eta}}\right) \geq \left|\zeta' + \frac{1}{\zeta'}\right| \geq \frac{1}{2\delta} - 2\delta,$$

which leads to the contradiction  $\gamma(1, \zeta') \notin D$  by (3.1). □

Before proving Theorem 1.2, we mention that the relation

$$w' + \frac{1}{w'} = z, \quad w' \in \mathbb{C}_+ \tag{4.2}$$

is equivalent to  $-1/w' = m_{sc}(z)$ . Since the semicircle law is analytic in the bulk interval  $W$ , its Stieltjes transform  $m_{sc}$  is Lipschitz continuous in  $D$  with a constant independent of  $N$ .

*Proof of Theorem 1.2.* It suffices to prove that

$$\sup_{z \in D} |\langle G(1, z) \rangle - m_{sc}(z)| \leq \mathcal{O}\left(\frac{N^\varepsilon}{\sqrt{N\eta}}\right)$$

on the event  $\mathcal{A}_\varepsilon$  for all  $\varepsilon < \theta/2$ . Let  $z \in D$ , let  $w = w(z) \in D_0$  be the point furnished by Lemma 4.1, and let  $w' = -m_{sc}(z)^{-1}$  be the solution of (4.2). By Lemma 4.1 we have  $w + w^{-1} \in D$  for all sufficiently large  $N$ , so the Lipschitz continuous dependence of  $1/w'$  on  $z \in D$  implies

$$\left|\frac{1}{w} - \frac{1}{w'}\right| = \mathcal{O}\left(\frac{N^\varepsilon}{\sqrt{N\eta}}\right).$$

On the event  $\mathcal{A}_\varepsilon$  the bound

$$\begin{aligned} |\langle G(1, z) \rangle - m_{sc}(z)| &\leq |\langle G(1, z) \rangle - \langle G(0, w) \rangle| + \left|\frac{1}{w} - \frac{1}{w'}\right| \\ &\leq \sup_{t \leq 1 \wedge \tau'_w} |\langle R(t, w) \rangle - \langle R(0, w) \rangle| + \mathcal{O}\left(\frac{N^\varepsilon}{\sqrt{N\eta}}\right) \leq \mathcal{O}\left(\frac{N^\varepsilon}{\sqrt{N\eta}}\right) \end{aligned}$$

is valid with a uniform constant for all  $z \in D$ . □

## References

- [1] A. Adhikari and J. Huang. Dyson Brownian motion for general  $\beta$  and potential at the edge. Preprint at arXiv:1810.08308, 2018.
- [2] A. Aggarwal. Bulk universality for generalized Wigner matrices with few moments. *Probability Theory and Related Fields*, 173(1):375–432, Feb 2019. MR-3916110
- [3] J. Alt, L. Erdős, and T. Krüger. Local inhomogeneous circular law. *Ann. Appl. Probab.*, 28(1):148–203, 2018. MR-3770875
- [4] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. MR-2760897
- [5] Z. Bao, L. Erdős, and K. Schnelli. Local law of addition of random matrices on optimal scale. *Comm. Math. Phys.*, 349(3):947–990, 2017. MR-3602820
- [6] R. Bauerschmidt, J. Huang, and H.-T. Yau. Local Kesten–McKay law for random regular graphs. *Comm. Math. Phys.*, 369(2):523–636, 2019. MR-3962004
- [7] A. Bloemendal, L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Isotropic local laws for sample covariance and generalized Wigner matrices. *Electron. J. Probab.*, 19:no. 33, 53, 2014. MR-3183577

- [8] P. Bourgade. Extreme gaps between eigenvalues of Wigner matrices. Preprint at arXiv:1812.10376, 2018.
- [9] P. Bourgade, H.-T. Yau, and J. Yin. Local circular law for random matrices. *Probab. Theory Related Fields*, 159(3-4):545–595, 2014. MR-3230002
- [10] B. Dupire. Pricing with a smile. *Risk*, 7:18–20, 1994.
- [11] F. J. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.*, 3:1191–1198, 1962. MR-0148397
- [12] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. The local semicircle law for a general class of random matrices. *Electron. J. Probab.*, 18:no. 59, 58, 2013. MR-3068390
- [13] L. Erdős, T. Krüger, and Y. Nemish. Local laws for polynomials of Wigner matrices. Preprint at arXiv:1804.11340, 2018.
- [14] L. Erdős, T. Krüger, and D. Schröder. Random matrices with slow correlation decay. *Forum Math. Sigma*, 7:e8, 89, 2019. MR-3941370
- [15] L. Erdős, B. Schlein, and H.-T. Yau. Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.*, 287(2):641–655, 2009. MR-2481753
- [16] L. Erdős, B. Schlein, and H.-T. Yau. Universality of random matrices and local relaxation flow. *Invent. Math.*, 185(1):75–119, 2011. MR-2810797
- [17] L. Erdős and H.-T. Yau. *A dynamical approach to random matrix theory*, volume 28 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2017. MR-3699468
- [18] Y. He, A. Knowles, and R. Rosenthal. Isotropic self-consistent equations for mean-field random matrices. *Probab. Theory Related Fields*, 171(1-2):203–249, 2018. MR-3800833
- [19] J. Huang and B. Landon. Rigidity and a mesoscopic central limit theorem for Dyson Brownian motion for general  $\beta$  and potentials. *Probability Theory and Related Fields*, Nov 2018. MR-4009708
- [20] J. O. Lee and K. Schnelli. Edge universality for deformed Wigner matrices. *Rev. Math. Phys.*, 27(8):1550018, 94, 2015. MR-3405746
- [21] D. B. Madan and M. Yor. Making Markov martingales meet marginals: with explicit constructions. *Bernoulli*, 8(4):509–536, 2002. MR-1914701
- [22] L. A. Pastur. The spectrum of random matrices. *Teoret. Mat. Fiz.*, 10(1):102–112, 1972. MR-0475502
- [23] T. Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012. MR-2906465
- [24] P. von Soosten and S. Warzel. The phase transition in the ultrametric ensemble and local stability of Dyson Brownian motion. *Electron. J. Probab.*, 23:no. 70, 24, 2018. MR-3835476
- [25] P. von Soosten and S. Warzel. Delocalization and continuous spectrum for ultrametric random operators. *Ann. Henri Poincaré*, 20(9):2877–2898, 2019. MR-3995890
- [26] P. von Soosten and S. Warzel. Non-ergodic delocalization in the Rosenzweig-Porter model. *Lett. Math. Phys.*, 109(4):905–922, 2019. MR-3920502

**Acknowledgments.** We are grateful to two anonymous referees for their helpful suggestions. We would also like to thank Institut Mittag-Leffler for their hospitality during the early stages of this project. This work was supported by the DFG grants SO 1724/1-1 (P. S.) and EXC-2111 – 390814868 (S. W.).