

# QUENCHED UNIVERSALITY FOR DEFORMED WIGNER MATRICES

GIORGIO CIPOLLONI AND LÁSZLÓ ERDŐS

*IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria*

DOMINIK SCHRÖDER\*

*Institute for Theoretical Studies, ETH Zurich, Clausiusstr. 47, 8092 Zurich, Switzerland*

**ABSTRACT.** Following E. Wigner’s original vision, we prove that sampling the eigenvalue gaps within the bulk spectrum of a *fixed* (deformed) Wigner matrix  $H$  yields the celebrated Wigner-Dyson-Mehta universal statistics with high probability. Similarly, we prove universality for a monoparametric family of deformed Wigner matrices  $H + xA$  with a deterministic Hermitian matrix  $A$  and a fixed Wigner matrix  $H$ , just using the randomness of a single scalar real random variable  $x$ . Both results constitute *quenched* versions of bulk universality that has so far only been proven in *annealed* sense with respect to the probability space of the matrix ensemble.

## 1. INTRODUCTION

Random matrix theory in physics was originally envisioned by E. Wigner to predict statistics of gaps between the energy levels of heavy atomic nuclei. The underlying physical systems have no inherent disorder and the statistical ensemble in Wigner’s description was generated by randomly (uniformly) sampling from the experimentally measured gaps of a fixed nucleus within a large energy range. The model ensemble, the space of Hermitian random matrices with independent, identically distributed entries (Wigner ensemble), is however inherently random. Accepting the replacement of the original physical Hamiltonian with a Hermitian random matrix, one may ask whether uniform sampling within the spectrum of a *fixed, typical realisation* of a Wigner matrix also gives rise to the celebrated Wigner-Dyson-Mehta (WDM) universality. In this paper we affirmatively answer this question, in the sense that for any fixed typical Wigner matrix the empirical gap statistic is close to the Wigner surmise, see Figure 1. We thus prove a stronger version of WDM universality and confirm the applicability of Wigner’s theory even in the *quenched sense*. All previous universality proofs, see e.g. [36, 30, 31, 29, 32, 8, 16, 10, 26, 18, 37, 19, 20] (see also [3, 6, 7, 35, 14, 34] for invariant ensembles), were valid in the *annealed sense*, i.e. where the eigenvalue statistics were directly generated by the randomness of the matrix ensemble.

More generally, we consider random matrices of the form  $H^x := H + xA$ , where  $H$  is a large  $N \times N$  Wigner matrix,  $A$  is a fixed nontrivial Hermitian deterministic matrix, and  $x$  is a real random variable (in fact we can even consider more general *deformed Wigner matrices*  $H$ ). We show that for a typical but fixed (quenched)  $H$  the randomness of  $x$  alone is sufficient to generate WDM universality in the bulk of the limiting spectrum of  $H^x$ , i.e. we prove that the local statistics of  $H + xA$  are universal for *all fixed*  $H$  in a high probability set. The special case  $A = I$  and  $x$  being uniformly distributed on some small interval yields Wigner’s spectral sampling model. Another special case covered by our general result is when  $A$  itself is chosen from a Wigner ensemble. The corresponding  $H + xA$  model for the

---

*E-mail addresses:* giorgio.cipolloni@ist.ac.at, lerdos@ist.ac.at, dschroeder@ethz.ch

*Date:* June 21, 2021.

*2010 Mathematics Subject Classification.* 60B20, 15B52.

*Key words and phrases.* Global Law, Local Law, Random Matrices, Dyson Brownian motion.

\*Supported by Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zürich Foundation.

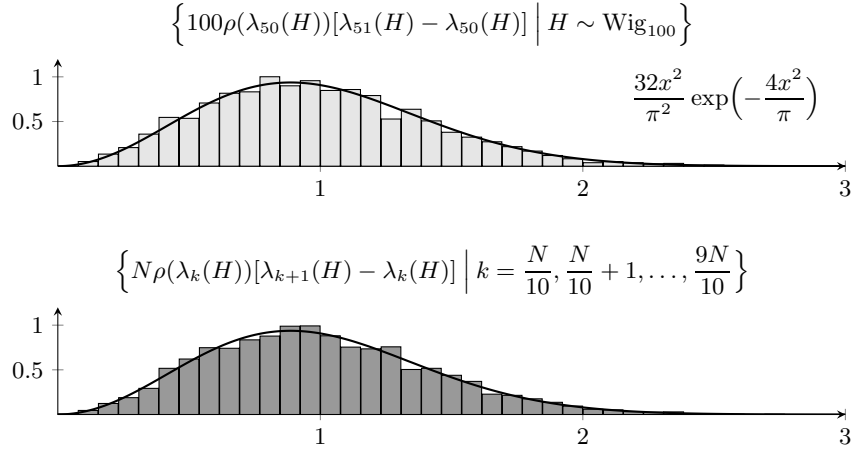


FIGURE 1. The two types of universality: The first histogram shows the normalised gaps of the two middle eigenvalues in the spectrum of 5000 complex Wigner matrices of size  $100 \times 100$ . The second histogram shows the empirical normalised bulk eigenvalue gaps of a single complex Wigner matrix of size  $5000 \times 5000$ . Both distributions asymptotically approach the Wigner surmise, drawn as solid lines.

Gaussian case was introduced by H. Gharibyan, C. Pattison, S. Shenker<sup>1</sup> and K. Wells who coined it as the *monoparametric ensemble* [23].

The basic guiding principle for establishing quenched universality of  $H^x$  is to show that near a fixed energy  $E$  the eigenvalues of  $H^x$  and  $H^{x'}$  are essentially uncorrelated whenever  $x$  and  $x'$  are not too close. This provides the sufficient (asymptotic) independence along the sampling in the space of  $x$ . Following a similar idea in [11] for a different setup, the independence of eigenvalues is proven by running the Dyson Brownian motion (DBM) for the matrix  $H$ . The corresponding stochastic differential equations for the eigenvalues of  $H^x$  and  $H^{x'}$  have almost independent stochastic differentials if the corresponding eigenvectors are asymptotically orthogonal. Therefore, independence of eigenvalues can be achieved by running the DBM already after a short time, provided we can understand eigenvector overlaps. The small Gaussian component added along the DBM flow can later be removed by fairly standard perturbation argument (*Green function comparison theorem*).

Thus the main task is to show that eigenvectors of  $H^x$  become asymptotically orthogonal for different, sufficiently distant values of  $x$ . This orthogonality can be triggered by two quite different mechanisms that we now explain.

The first mechanism is present when  $A$  is not too close to a diagonal matrix, in other words if  $\mathring{A} := A - \langle A \rangle$  is nontrivial in the sense that  $\langle \mathring{A}^2 \rangle \geq c$  with some  $N$ -independent constant  $c > 0$ . Here  $\langle A \rangle := \frac{1}{N} \text{Tr } A$  denotes the normalized trace. In this case the entire eigenbasis of  $H^x$  is *rotated*, i.e. it becomes essentially orthogonal to that of  $H^{x'}$  whenever  $x$  and  $x'$  are not too close. As a consequence, the entire spectra of  $H^x$  and  $H^{x'}$  are essentially uncorrelated. To establish this effect of *eigenbasis rotation*, we use a *multi-resolvent local law* for the resolvents of  $H^x$  and  $H^{x'}$ ; this method currently requires  $|x - x'| \geq N^{-\epsilon}$  for typical choices of  $x, x'$ . To ensure this, we assume that  $x = N^{-a}\chi$  where  $\chi$  is an  $N$ -independent real random variable with some regularity and  $a \in [0, \epsilon]$ .

The second mechanism is the most transparent when  $A = I$  and  $x = N^{-a}\chi$  where  $\chi$  is uniformly distributed on some small fixed interval; we call this mechanism the *sampling in the spectrum*. In this case the eigenbasis of  $H^x$  actually does not depend on  $x$ . However, the eigenvectors corresponding to eigenvalues close to a fixed energy are algebraically orthogonal for sufficiently distant  $x, x'$ . We also prove that distant eigenvalue gaps of  $H$ , and hence the *local spectral data* of  $H^x, H^{x'}$  are essentially uncorrelated. By the *rigidity* property of the eigenvalues, already a small change in  $x$  triggers this effect, so it works in the entire range of scales  $a \in [0, 1 - \epsilon]$ . Moreover, the proof can easily be extended

<sup>1</sup>Private communication via Stephen Shenker and Sourav Chatterjee in June 2020.

to more complicated random matrix ensembles well beyond the Wigner case. No multi-resolvent local law is needed in the proof.

A combination of these two mechanisms can be used in the situation when  $A \neq I$ , but  $A$  is still close to  $\langle A \rangle$  times the identity in the sense that  $|\langle A \rangle| \geq C \langle \mathring{A}^2 \rangle^{1/2}$  for some large  $C$ . This extension complements the main condition  $\langle \mathring{A}^2 \rangle \geq c$  needed in the first mechanism thus proving the result unconditionally for any  $A$ .

**1.1. Notations and conventions.** We introduce some notations we use throughout the paper. For integers  $k \in \mathbf{N}$  we use the notation  $[k] := \{1, \dots, k\}$ . For positive quantities  $f, g$  we write  $f \lesssim g$  and  $f \sim g$  if  $f \leq Cg$  or  $cg \leq f \leq Cg$ , respectively, for some constants  $c, C > 0$  which depend only on the *model parameters* appearing in our base Assumptions 2.3–2.2. For any two positive real numbers  $\omega_*, \omega^* \in \mathbf{R}_+$  by  $\omega_* \ll \omega^*$  we denote that  $\omega_* \leq c\omega^*$  for some small constant  $0 < c < 1/100$ . We denote vectors by bold-faced lower case Roman letters  $\mathbf{x}, \mathbf{y} \in \mathbf{C}^k$ , for some  $k \in \mathbf{N}$ . Vector and matrix norms,  $\|\mathbf{x}\|$  and  $\|A\|$ , indicate the usual Euclidean norm and the corresponding induced matrix norm. For vectors  $\mathbf{x}, \mathbf{y} \in \mathbf{C}^k$  we define

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_i \bar{x}_i y_i$$

and for any  $N \times N$  matrix  $A$  we use the notation  $\langle A \rangle := N^{-1} \text{Tr } A$  to denote the normalized trace of  $A$ . We will use the concept of “with very high probability” meaning that for any fixed  $D > 0$  the probability of an  $N$ -dependent event is bigger than  $1 - N^{-D}$  if  $N \geq N_0(D)$ . Moreover, we use the convention that  $\xi > 0$  denotes an arbitrary small constant which is independent of  $N$ .

**Acknowledgement.** The authors are indebted to Sourav Chatterjee for forwarding the very inspiring question that Stephen Shenker originally addressed to him which initiated the current paper. They are also grateful that the authors of [23] kindly shared their preliminary numerical results in June 2021.

## 2. MAIN RESULTS

In this paper we consider real and complex *Wigner matrices*, i.e. Hermitian  $N \times N$  random matrices  $H = H^*$  with independent identically distributed (i.i.d.) entries (up to Hermitian symmetry)

$$h_{ab} \stackrel{d}{=} N^{-1/2} \begin{cases} \chi_{\text{od}}, & a < b, \\ \chi_{\text{d}}, & a = b, \end{cases} \quad h_{ba} := \overline{h_{ab}} \quad (2.1)$$

having finite moments of all orders, i.e.  $\mathbf{E}|\chi_{\text{od}}|^p + \mathbf{E}|\chi_{\text{d}}|^p \leq C_p$ . The entries are normalised such that  $\mathbf{E}|\chi_{\text{od}}|^2 = 1$ , and additionally  $\mathbf{E}\chi_{\text{od}}^2 = 0$  in the complex case. The normalisation guarantees that the eigenvalues  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$  of  $H$  asymptotically follow Wigner’s semicircular distribution  $\rho_{\text{sc}}(x) := \sqrt{(4 - x^2)_+}/(2\pi)$ . In the *bulk regime*, i.e. where  $\rho_{\text{sc}} \geq c$  for some  $c > 0$ , the eigenvalue gaps are of order;  $\lambda_{i+1} - \lambda_i \sim 1/N$ .

The *Wigner-Dyson-Mehta conjecture* for the bulk of Wigner matrices  $H$  asserts that for any  $i \in [\epsilon N, (1 - \epsilon)N]$  the distribution of the rescaled eigenvalue gap converges

$$\lim_{N \rightarrow \infty} \mathbf{P} \left( N \rho_{\text{sc}}(\lambda_i) [\lambda_{i+1} - \lambda_i] \leq y \right) = \int_0^y p_\beta(t) dt \quad (2.2)$$

to a universal distribution with density  $p_1$  (for real symmetric Wigner matrices) or  $p_2$  (for complex Hermitian Wigner matrices) which can be computed explicitly from the integrable Gaussian GOE/GUE ensembles, see Section 2.3 later. This WDM conjecture was resolved in [20] while similar results with a small averaging in the index  $i$  were proven earlier [18, 37].

As a corollary to our main result Theorem 2.7 below on the *monoparametric ensemble* we prove a considerable strengthening of (2.2), namely that with high probability the *sampling* of eigenvalues within a single fixed Wigner matrix generates WDM universality.

**Corollary 2.1** (to Theorem 2.7). *Let  $H$  be a Wigner matrix and  $I \subset (-2 + \epsilon, 2 - \epsilon)$  be an interval in the bulk of  $H$  of length  $|I| \geq N^{-1+\xi}$  for some  $\epsilon, \xi > 0$ . Then there exist small  $\kappa, \alpha > 0$  and an event  $\Omega_I$  in the probability space  $\mathbf{P}_H$  of  $H$  with  $\mathbf{P}_H(\Omega_I^c) \leq N^{-\kappa}$ , such that for all  $H \in \Omega_I$  it holds that*

$$\sup_{y \geq 0} \left| \frac{1}{N \int_I \rho_{\text{sc}}} \#\{i \mid N \rho_{\text{sc}}(\lambda_i) [\lambda_{i+1} - \lambda_i] \leq y, \lambda_i \in I\} - \int_0^y p_\beta(t) dt \right| = \mathcal{O}(N^{-\alpha}), \quad (2.3)$$

where the implicit constant in (2.3) and  $\kappa, \alpha$  depend on  $\epsilon, \xi$ .

Our main results are on the *quenched (bulk) universality of monoparametric random matrices*

$$H^x := H + xA \quad (2.4)$$

for general deterministic Hermitian matrices  $A$  of the same symmetry class<sup>2</sup> as  $H$ , and independent scalar random variables  $x$ , just using the randomness of  $x$  for any fixed Wigner matrix  $H$  from a high probability set. For  $A = I$  the monoparametric universality of  $H^x$  implies the spectral sampling universality as stated in Corollary 2.1, see Section 3.3. Our results extend beyond Wigner matrices, we also allow for arbitrary additive deformations (certain results even extend to Wigner matrices with correlated entries), and cover general sufficiently regularly distributed scalar random variables  $x$ .

**Assumption 2.2** (Deformed Wigner matrix). *We consider deformed Wigner matrices of the form  $H = W + B$ , where  $W$  is a Wigner matrix as in (2.1), and  $B = B^*$  is an arbitrary deterministic matrix of bounded norm, i.e.  $\|B\| \leq C_0$  for some  $N$ -independent constant  $C_0$ .*

**Assumption 2.3.** *Assume that  $x = N^{-a}\chi$  with  $a \in [0, 1)$ , where  $\chi$  is an  $N$ -independent compactly supported real random variable such that for any small  $b_1 > 0$  there exists  $b_2 > 0$  such that for any interval  $I \subset \mathbf{R}$  with  $|I| \sim N^{-b_1}$  it holds  $\mathbf{P}(\chi \in I) \leq |I|^{b_2}$ .*

To state the result, we now introduce the self-consistent density of states of  $H^x = W + B + xA$ . It has been proven in [1, Theorem 2.7] that the resolvent  $G^x(z) = (H^x - z)^{-1}$  of  $H^x$  at a spectral parameter  $z \in \mathbf{C} \setminus \mathbf{R}$  can be well approximated by the unique deterministic matrix  $M = M^x(z)$ , solving the *Matrix Dyson Equation* (MDE) [2] (see also [24])

$$-M^{-1} = z - B - xA + \langle M \rangle, \quad \Im M(z) \Im z > 0. \quad (2.5)$$

We define the *self consistent density of states* (scDos) [2, Section 4.1] of  $H^x$  as

$$\rho^x(E) := \lim_{\eta \rightarrow 0^+} \frac{1}{\pi} \langle \Im M^x(E + i\eta) \rangle, \quad (2.6)$$

and, in particular, the scDos of  $H$  by  $\rho := \rho^0$ . It is well known that  $\rho^x$  is a probability density which is compactly supported and real analytic inside its support [2, Proposition 2.3]. For the special case  $\mathbf{E}H = B = 0$  the scDos of  $H$  is the standard Wigner semicircle law, i.e.  $\rho = \rho_{\text{sc}}$ . We say that an energy  $E \in \mathbf{R}$  lies in the *bulk of the spectrum* of  $H^x$  if  $\rho^x(E) \geq c$  for some  $N$ -independent constant  $c > 0$ . For  $E$  in the bulk, the solution  $M^x(z)$  can be continuously extended to the real line,  $M^x(E) := \lim_{\eta \rightarrow 0^+} M^x(E + i\eta)$ , and  $M^x(E + i\eta)$  for  $E$  in the bulk is *uniformly bounded*, c.f. [2, Proposition 3.5]. Finally, we define the *classical eigenvalue locations* to be the *quantiles* of  $\rho^x$ , i.e. we define  $\gamma_i^x$  by

$$\int_{-\infty}^{\gamma_i^x} \rho^x(\tau) d\tau = \frac{i}{N}, \quad i \in [N]. \quad (2.7)$$

For clarity, in this section we only present single-gap versions of both mechanisms explained in the introduction that yield quenched universality. Subsequently we will present the multi-gap analogues in Section 6.

**2.1. Monoparametric universality via eigenbasis rotation.** The main universality result for the first mechanism (*eigenbasis rotation*) is the following quenched fixed-index universality result for the *monoparametric ensemble*. We denote the probability measure and expectation of  $x$  by  $\mathbf{P}_x, \mathbf{E}_x$  in order to differentiate it from the probability measure  $\mathbf{P}_H$  of  $H$ .

**Theorem 2.4** (Quenched universality for monoparametric ensemble). *Let  $H$  be a deformed Wigner matrix satisfying Assumption 2.2, and let  $x = N^{-a}\chi$  be a scalar real random variable satisfying Assumption 2.3 with  $a \in [0, a_0]$ , where  $a_0$  is a small universal constant<sup>3</sup>. Fix any  $c_0, c_1 > 0$  small constants and assume that  $\langle \hat{A}^2 \rangle \geq c_0$ , with  $\hat{A} := A - \langle A \rangle$ . Suppose that  $i \in [N]$  is a bulk index<sup>4</sup> for  $H^x = H + xA$ , i.e. it holds that*

$$\rho^x(\gamma_i^x) \geq c_1 \quad \text{for } \mathbf{P}_x\text{-almost all } x. \quad (2.8)$$

<sup>2</sup>This restriction apparently excludes the case when  $A$  is complex Hermitian but  $H$  is real symmetric. With a slight modification of our proof (similar to the modification required in [13, Section 7] compared to [11, Section 7]), however, we can handle this case as well, but for brevity we refrain from presenting it.

<sup>3</sup>Following the explicit constants along the proof, one may choose  $a_0 = 1/100$

<sup>4</sup>To specify the  $c_1$ -dependence, we often speak of  $c_1$ -bulk index.

Then there exist small  $\alpha, \kappa > 0$  and an event  $\Omega_i = \Omega_{i,A}$  with  $\mathbf{P}_H(\Omega_i^c) \leq N^{-\kappa}$ , so that for all  $H \in \Omega_i$  the statistics of the  $i$ -th rescaled gap of the eigenvalues  $\lambda_i^x$  of  $H^x$  is universal, i.e.

$$\left| \mathbf{E}_x f\left(N\rho^x(\lambda_i^x)[\lambda_{i+1}^x - \lambda_i^x]\right) - \int_0^\infty p_\beta(t)f(t) dt \right| = \mathcal{O}(N^{-\alpha}\|f\|_{C^5}) \quad (2.9)$$

for any smooth, compactly supported function  $f$  where the implicit constant in (2.9) depends on  $a_0, c_0, c_1$  and  $\kappa, \alpha$  depend on  $a_0$ .

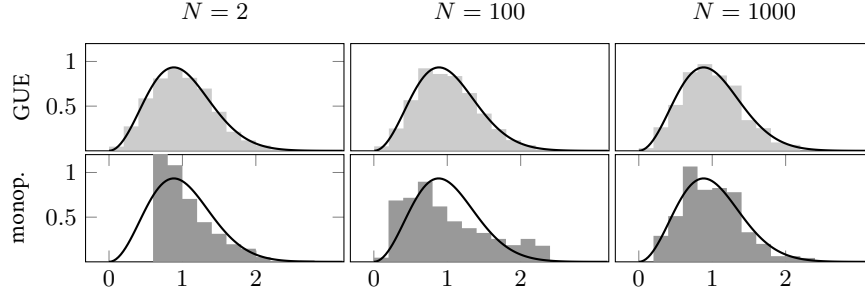


FIGURE 2. Here we show the histogram of a (rescaled) single eigenvalue gap  $\lambda_{N/2+1} - \lambda_{N/2}$  in the middle of the spectrum for  $N \in \{2, 100, 1000\}$  and for random matrices sampled from either the GUE or the monoparametric ensemble. For the GUE ensemble the histogram has been generated by sampling 2000 independent GUE matrices  $H$ . For the monoparametric ensemble only two GUE matrices  $H, A$  have been drawn at random, and the histogram has been generated by sampling 2000 standard Gaussian random variables  $x$  and considering the gaps of  $H^x = H + xA$ . The solid black line represents the theoretical limit  $p_2(s)$  which matches the empirical distribution very closely already for  $N = 2$ . In Appendix B we present numerical evidence for the speed of convergence, inspired by the observation on the slow convergence of the spectral form factor made in [23].

**Remark 2.5.** We mention a few simple observations about Theorem 2.4.

- (i) By the regularity of  $f, \rho^x$  and by rigidity of the bulk eigenvalues (see (3.14) later) we may replace the random scaling factor  $\rho^x(\lambda_i^x)$  with  $\rho^x(\gamma_i^x)$  at negligible error.
- (ii) For  $\mathbf{E}H = 0$  and  $a > 0$  the condition (2.8) can simply be replaced by  $i \in [N\epsilon', N(1 - \epsilon')]$  for some  $\epsilon' > 0$  and the argument of  $f$  in (2.9) simplifies to  $f(N\rho_{sc}(\lambda_i^x)[\lambda_{i+1}^x - \lambda_i^x])$ .
- (iii) Empirically we find that the convergence towards the universal gap statistics in (2.9) is much slower for the monoparametric ensemble compared to GUE, c.f. Figure 2. While even for  $2 \times 2$  GUE matrices the empirical gap distribution is already very close to the Gaudin-Mehta distribution (see Section 2.3), we observe the same effect only for large monoparametric matrices.

**Remark 2.6.** We mention an interesting special case of Theorem 2.4 when  $H$  is a Wigner matrix and  $A$  itself is chosen from a Wigner ensemble that is independent of  $H$  and  $x$ . In this case Theorem 2.4 implies<sup>5</sup> that for any fixed pair of Wigner matrices  $A, H$  from a high probability set, the universality of the  $i$ -th gap statistics of  $H + xA$  for  $i \in [N\epsilon', N(1 - \epsilon')]$  is solely generated by the single real random variable  $x$ , i.e.

$$\mathbf{P}_{H,A}\left(i\text{-th gap statistics of } H + xA \text{ is universal}\right) = 1 - \mathcal{O}(N^{-\epsilon}). \quad (2.10)$$

This mathematically rigorously answers to a question of Gharibyan, Pattison, Shenker and Wells [23]. While their original question referred to a standard Gaussian  $x$ , which is not compactly supported, a simple cut-off argument extends our proof to this case as well.

<sup>5</sup>The condition on  $\langle \dot{A}^2 \rangle$  is satisfied since  $\langle A^2 \rangle = 1 + o(1)$  and  $\langle A \rangle = o(1)$  with very high probability. Moreover, the scDos  $\rho^x$  is very close to a rescaled semicircle law with radius  $2\sqrt{1+x^2}$  with very high probability in the joint probability space of  $H$  and  $A$ , hence the condition (2.8) holds for all  $i \in [N\epsilon', N(1 - \epsilon')]$  for some  $\epsilon' > 0$ .

**2.2. Monoparametric universality via spectral sampling.** The main universality result for the second mechanism (*spectral sampling*) is the following quenched fixed-energy universality result for the *monoparametric ensemble*. We define  $i_0 = i_0(x, E)$  as the index such that  $\gamma_{i_0}^x$  is the quantile of  $\rho^x$  closest to  $E$ , i.e.

$$i_0(x, E) = \left\lceil N \int_{-\infty}^E \rho^x(t) dt \right\rceil \in [N] \quad (2.11)$$

with  $\lceil \cdot \rceil$  denoting rounding to the next largest integer. For the special case  $A = I$  (formulated as **Case 1**) in Theorem 2.7 below) we obtain quenched sampling universality for a much broader class of Hermitian random matrices  $H$  with slow correlation decay<sup>6</sup> defined in [17]. In the second situation, **Case 2**) in Theorem 2.7 below we consider deformed Wigner matrices  $H$  and general  $A$  with a condition complementary to the condition  $\langle \dot{A}^2 \rangle \geq c_0$  of Theorem 2.4.

**Theorem 2.7** (Quenched monoparametric universality via spectral sampling mechanism). *There is a small universal constant  $a_0$  and for any small  $c_1 > 0$  there exists a  $c_0 > 0$  such that the following hold. Let  $x = N^{-\alpha} \chi$  be a scalar random variable satisfying Assumption 2.3, and let  $H, A, a$  be such that either*

*Case 1)  $H$  is a correlated random matrix<sup>6</sup>,  $A = I$ , and  $a \in [0, 1 - a_1]$  with an arbitrary small fixed  $a_1$ ; or*

*Case 2)  $H$  is a deformed Wigner matrix (cf. Assump. 2.2),  $c_0 \langle A \rangle \geq \langle \dot{A}^2 \rangle^{1/2}$ ,  $\langle A \rangle \geq c_0$ , and  $a \in [0, a_0]$ ,*

*and fix an energy  $E$  with  $\rho^x(E) \geq c_1 > 0$  for  $\mathbf{P}_x$ -almost all  $x$ . Then there exist small  $\alpha, \kappa > 0$  and an event  $\Omega_E = \Omega_{E,A}$  with  $\mathbf{P}_H(\Omega_E^c) \leq N^{-\kappa}$  such that for all  $H \in \Omega_E$  the matrix  $H^x$  satisfies*

$$\left| \mathbf{E}_x f(N\rho^x(E)[\lambda_{i_0(x,E)+1}^x - \lambda_{i_0(x,E)}^x]) - \int p_\beta(t) f(t) dt \right| = \mathcal{O}(N^{-\alpha} \|f\|_{C^5}). \quad (2.12)$$

The exponents  $\kappa, \alpha$  depend on  $a_0, a_1$  while the implicit constant in (2.12) depends on  $a_0, a_1, c_0, c_1$ .

We remark that the condition  $c_0 \langle A \rangle \geq \langle \dot{A}^2 \rangle^{1/2}$  in **Case 2**) is not really necessary for (2.12) to hold. Indeed, if  $\langle \dot{A}^2 \rangle \geq c$  with any small positive constant, then we are back to the setup of Theorem 2.4 where the *eigenbasis rotation* mechanism is effective. One can easily see that the proof of (2.9) implies (2.12) in this case (see Remark 3.5 below). However, we kept the condition  $c_0 \langle A \rangle \geq \langle \dot{A}^2 \rangle^{1/2}$  with a sufficiently small  $c_0$  in the formulation of Theorem 2.7 since it is necessary for the *spectral sampling* mechanism to be effective which is the mechanism represented in Theorem 2.7.

Note that as long as  $\dot{A} \neq 0$ , i.e. **Case 1**) is not applicable, the eigenbasis of  $H^x$  changes with  $x$  and we have to rely on the multi-resolvent local law method. However, lacking an effective lower bound on  $\langle \dot{A}^2 \rangle$ , the effective asymptotic orthogonality still comes from the spectral sampling effect of  $\langle A \rangle$ , the nontrivial tracial part of  $A$ . So along the proof of **Case 2**), technically we follow the eigenbasis rotation mechanism, but morally the effect is similar to the spectral sampling mechanism as it still comes from a shift in the spectrum triggered by  $x \langle A \rangle$ , the leading part of  $x A$  in  $H^x = H + x A$ . Finally, a simple perturbation shows that  $x \dot{A}$  has no sizeable effect on the sampling, but its presence hinders the technically simpler orthogonality proof used in **Case 1**).

**2.3. Gaudin-Mehta distribution.** For completeness we close this section by providing explicit formulas for the universal *Gaudin-Mehta gap distributions*  $p_1, p_2$  which can either be defined as the Fredholm determinant of the sine kernel [33] or via the solution to the Painlevé V differential equation [25]. Given the solution  $\sigma$  to the nonlinear differential equation

$$(t\sigma'')^2 + 4(t\sigma' - \sigma)(t\sigma' - \sigma + (\sigma')^2) = 0, \quad \sigma(t) \sim -\frac{t}{\pi} - \frac{t^2}{\pi^2} \quad (\text{as } t \rightarrow 0), \quad (2.13)$$

we have [22]

$$p_2(s) = \frac{d^2}{ds^2} \exp\left(\int_0^{\pi s} \frac{\sigma(t)}{t} dt\right), \quad p_1(s) = \frac{d^2}{ds^2} \exp\left(\frac{1}{2} \int_0^{\pi s} \left(\frac{\sigma(t)}{t} - \sqrt{-\frac{d}{dt} \frac{\sigma(t)}{t}}\right) dt\right). \quad (2.14)$$

<sup>6</sup>These are  $N \times N$  Hermitian matrices  $H$  with covariance operator  $\mathcal{S}[R] := \frac{1}{N} \mathbf{E} W R W$ , where  $W := \sqrt{N}(H - \mathbf{E} H)$  is a correlated centred random matrix. Note that this  $W$  is  $\sqrt{N}$ -times bigger than the Wigner matrix  $W$  defined in Assumption 2.2. This notational inconsistency occurs only in this description of the correlated ensemble where we follow the convention of [17]. We assume that  $\|\mathbf{E} H\| \leq C$  and that  $W$  satisfies Assumptions (B)–(E) of [17]. We recall that Assumption (B) requires that all moments of the matrix elements of  $W$  are finite, i.e.  $\mathbf{E} |W_{ab}|^q \leq C_q$  with some constant  $C_q$  for any  $q$  integer, uniformly in the indices  $a, b \in [N]^2$ , while Assumption (E) requires that the covariance operator satisfies the so called *flatness condition*

$$c\langle R \rangle \leq \mathcal{S}[R] \leq C\langle R \rangle$$

for any positive semi-definite matrix  $R$ , where  $c, C$  are some fixed positive constants. Finally, Assumptions (C), (D) or their simplified version (CD) impose decay conditions on the cumulants of different entries of  $W$ , we refer the reader to [17, Eqs. (2.5a)–(2.5b)] for the precise condition. The self-consistent density of states  $\rho$  is defined analogously to (2.6), where  $M$  solves the MDE (2.5) with  $\langle M \rangle$  replaced by  $\mathcal{S}[M]$ .

Remarkably, the Wigner surmise

$$p_2^{\text{Wigner}}(s) := \frac{32s^2}{\pi^2} \exp\left(-\frac{4s^2}{\pi}\right), \quad p_1^{\text{Wigner}}(s) := \frac{s\pi}{2} \exp\left(-\frac{\pi s^2}{4}\right) \quad (2.15)$$

obtained by E. Wigner from explicitly computing the gap distribution for  $2 \times 2$  matrices, is very close to the large  $N$  limit  $p_2(s)$ , more precisely  $\sup_s |p_2(s) - p_2^{\text{Wigner}}(s)| \approx 0.005$  and  $\sup_s |p_1(s) - p_1^{\text{Wigner}}(s)| \approx 0.016$ .

### 3. QUENCHED UNIVERSALITY: PROOF OF THEOREM 2.4 AND THEOREM 2.7

In Section 3.1 we prove Theorem 2.4 while in Section 3.2 we present the proof of Theorem 2.7 which structurally is analogous to the argument in Section 3.1. For notational simplicity we introduce the discrete difference operator  $\delta$ , i.e. for a tuple  $\lambda$  we set

$$(\delta\lambda)_i := \delta\lambda_i := \lambda_{i+1} - \lambda_i \quad (3.1)$$

in order to express eigenvalue differences (gaps) more compactly. We also introduce the notation  $\langle f \rangle_{\text{gap}}$  for the expectation of test functions  $f$  with respect to the density  $p_\beta$  from (2.14), i.e.

$$\langle f \rangle_{\text{gap}} := \int p_\beta(t) f(t) dt. \quad (3.2)$$

**3.1. Universality via eigenbasis rotation mechanism: Proof of Theorem 2.4.** To prove Theorem 2.4 we will show that the gaps  $\lambda_{i+1}^{x_1} - \lambda_i^{x_1}, \lambda_{i+1}^{x_2} - \lambda_i^{x_2}$  for sufficiently large  $|x_1 - x_2|$  are asymptotically independent in the sense of the following proposition whose proof will be presented in Section 4. In the following we will often denote the covariance of two random variables  $X, Y$  in the  $H$ -space by

$$\text{Cov}_H(X, Y) := \mathbf{E}_H XY - (\mathbf{E}_H X)(\mathbf{E}_H Y).$$

**Proposition 3.1.** *Under the conditions of Theorem 2.4 there exists a sufficiently small  $c^* > 0$  (depending on  $c_0, c_1$ ) and for any small  $\zeta_1$  there exists  $\zeta_2 > 0$  such that the following holds. Pick real numbers  $x_1, x_2$  with  $N^{-\zeta_1} \leq |x_1 - x_2| \leq c^*$  and indices  $j_1, j_2$  with  $|j_1 - j_2| \lesssim N|x_1 - x_2|$ , such that the corresponding quantiles  $\gamma_{j_r}^{x_r}$ , are in the  $c_1$ -bulk of the spectrum of  $H^{x_r}$  for each  $r = 1, 2$ . Then the covariance  $\text{Cov}_H(X, Y)$  satisfies*

$$\text{Cov}_H\left(P_1(N\delta\lambda_{j_1}^{x_1}), P_2(N\delta\lambda_{j_2}^{x_2})\right) = \mathcal{O}\left(N^{-\zeta_2} \|P_1\|_{C^5} \|P_2\|_{C^5}\right) \quad (3.3)$$

for any  $P_1, P_2: \mathbf{R} \rightarrow \mathbf{R}$  bounded smooth compactly supported test functions, and where the implicit constant in  $\mathcal{O}(\cdot)$  may depend on  $c_0, c_1$  at most polynomially.

**Remark 3.2.** *We stated the asymptotic independence of a single gap in Proposition 3.1 and only for two  $x_1, x_2$  for notation simplicity. Exactly the same proof as in Section 4 directly gives the result in (3.3) for test functions  $P_r: \mathbf{R}^p \rightarrow \mathbf{R}$  of several gaps, for some fixed  $p \in \mathbf{N}$ . Additionally, by the same proof we can also conclude the asymptotic independence of several gaps for several  $x_1, \dots, x_q$ . For the same reason we also state Proposition 3.3 and Proposition 4.1 below only for two  $x_1, x_2$  and test functions  $P_r: \mathbf{R} \rightarrow \mathbf{R}$ .*

*Proof of Theorem 2.4.* We will first prove that without loss of generality we may assume that the linear size of the support of  $x$  is bounded by  $c^*$ , where  $c^*$  is from Proposition 3.1. This initial simplification will then allow us to use perturbation in  $x$  when proving Proposition 3.1. Suppose that Theorem 2.4 is already proved for random variables with such a small support with an error term  $N^{-\alpha}$  on sets of probability at least  $1 - N^{-2\kappa}$  and we are now given a random variable  $x$  with a larger support of size bounded by some constant  $C$ . Then we define the random variables

$$x_i := \frac{x \cdot \mathbf{1}(x \in J_i)}{\mathbf{P}_x(J_i)},$$

where  $J_i$ 's, for  $i = 1, 2, \dots, C/c^*$ , are disjoint intervals of size  $c^*$  such that  $\text{supp}(x) = \bigcup_i J_i$ . For any test function  $f$  we can then write

$$\begin{aligned} \mathbf{E}_x f(N\rho^x(\gamma_j^x)\delta\lambda_j^x) &= \sum_{i=1}^{C/c^*} \mathbf{P}_x(J_i) \mathbf{E}_{x_i} f(N\rho^{x_i}(\gamma_j^{x_i})\delta\lambda_j^{x_i}) \\ &= \langle f \rangle_{\text{gap}} \sum_{i=1}^{C/c^*} \mathbf{P}_x(J_i) + \mathcal{O}(N^{-\alpha}) = \langle f \rangle_{\text{gap}} + \mathcal{O}(N^{-\alpha}), \end{aligned} \quad (3.4)$$

on a set of probability at least  $1 - (C/c^*)N^{-2\kappa} \geq 1 - N^{-\kappa}$ , where we used Theorem 2.4 for the random variables  $x_i$  in the last step and a union bound.

From now on we assume that the linear size of the support of  $x$  is bounded by  $c^*$ . With  $\nu(dx)$  denoting the measure of  $x$  we have

$$\begin{aligned} & \mathbf{E}_H \left| \mathbf{E}_x f(N\rho^x(\gamma_j^x)\delta\lambda_j^x) - \langle f \rangle_{\text{gap}} \right|^2 \\ &= \iint_{|x_1 - x_2| \geq N^{-\epsilon_2}} \nu(dx_1)\nu(dx_2) \mathbf{E}_H \left[ \prod_{r=1}^2 f(N\rho^{x_r}(\gamma_j^{x_r})\delta\lambda_j^{x_r}) \right] - \langle f \rangle_{\text{gap}}^2 + \mathcal{O}\left(N^{-c(\epsilon_2)} \|f\|_{C^5}^2\right), \end{aligned} \quad (3.5)$$

for some sufficiently small  $\epsilon_2$  so that we can apply Proposition 3.1 with  $\zeta_1 = \epsilon_2$ . In (3.5) we used that the regime  $|x_1 - x_2| \leq N^{-\epsilon_2}$  can be removed at the price of a negligible error by the regularity assumption on the distribution of  $x = N^{-a}\chi$ , with  $\chi$  satisfying Assumption 2.3. For the cross-term in (3.5) we used that by gap universality for the deformed Wigner matrix  $H^x$  with a fixed  $x$  (see e.g. [1, Corollary 2.11]) it follows that

$$\mathbf{E}_H f(N\rho^x(\gamma_j^x)\delta\lambda_j^x) = \langle f \rangle_{\text{gap}} + \mathcal{O}\left(N^{-\zeta_3} \|f\|_{C^5}\right) \quad (3.6)$$

for some small fixed  $\zeta_3 > 0$  depending only on the model parameters and on the constants  $a_0, c_1$ .

Applying Proposition 3.1 to the first term in (3.5) with  $P_r(t) := f(\rho^{x_r}(\gamma_j^{x_r})t)$  noting that  $\rho^{x_r}$  is uniformly bounded, so that for  $N^{-\epsilon_2} \leq |x_1 - x_2| \leq c^*$ , we get

$$\mathbf{E}_H \left[ \prod_{r=1}^2 f(N\rho^{x_r}(\gamma_j^{x_r})\delta\lambda_j^{x_r}) \right] = \prod_{r=1}^2 \mathbf{E}_H f(N\rho^{x_r}(\gamma_j^{x_r})\delta\lambda_j^{x_r}) + \mathcal{O}\left(N^{-\zeta_2} \|f\|_{C^5}^2\right). \quad (3.7)$$

By using (3.7) and (3.6) in (3.5) it follows that

$$\mathbf{E}_H \left| \mathbf{E}_x f(N\rho^x(\gamma_j^x)\delta\lambda_j^x) - \langle f \rangle_{\text{gap}} \right|^2 \leq \left(N^{-c(\epsilon_2)} + N^{-\zeta_2} + N^{-\zeta_3}\right) \|f\|_{C^5}^2. \quad (3.8)$$

From (3.8) and the Chebyshev inequality we obtain events  $\Omega_{j,f}$  on which (2.9) holds with probability  $\mathbf{P}_H(\Omega_{j,f}^c) \leq N^{-\kappa}$  for some suitably chosen  $\kappa, \alpha > 0$ .  $\square$

**3.2. Universality via spectral sampling mechanism: Proof of Theorem 2.7.** The mechanism behind the proof of Theorem 2.7 is quite different compared to Theorem 2.4. In particular, in order to prove Theorem 2.7 we will first show that under the assumptions of Theorem 2.4 the gaps  $\delta\lambda_i^x, \delta\lambda_j^x$  are asymptotically independent for any fixed  $x$  in the probability space of  $H$  as long as  $|i - j|$  is sufficiently large. This independence property for the  $A = I$  case has already been used as a heuristics without proof, e.g. in [4, 5] (a related result for not too distant gaps for local log-gases can be deduced from the De Giorgi-Nash-Moser Hölder regularity estimate, see [20, [Section 8.1)]). More precisely, we have the following proposition:

**Proposition 3.3.** *Under the conditions of Theorem 2.7 there exists a sufficiently small  $c_* > 0$  (depending on  $c_0, c_1$ ) and for any sufficiently small  $\zeta_1 > 0$  there exists  $\zeta_2 > 0$  such that the following hold. Pick indices  $j_1, j_2$  and real numbers  $x_1, x_2$  such that the corresponding quantiles  $\gamma_{j_r}^{x_r}$  are in the  $c_1$ -bulk of the spectrum of  $H^{x_r}$ , i.e.  $\rho^{x_r}(\gamma_{j_r}^{x_r}) \geq c_1$ , for  $r = 1, 2$ . In the two different cases listed in Theorem 2.7 we additionally assume the following:*

*Case 1)  $|j_1 - j_2| \geq N^{\zeta_1}$ ;*

*Case 2)  $N^{1-\zeta_1} \leq |j_1 - j_2| \leq c_*N$  and  $N|x_1 - x_2| \lesssim |j_1 - j_2|$ .*

*Then in both cases it holds that*

$$\text{Cov}_H \left( P_1(N\delta\lambda_{j_1}^{x_1}), P_2(N\delta\lambda_{j_2}^{x_2}) \right) = \mathcal{O} \left( N^{-\zeta_2} \prod_{r=1}^2 \|P_r\|_{C^5} \right) \quad (3.9)$$

for  $P_r: \mathbf{R} \rightarrow \mathbf{R}$  bounded, smooth, compactly supported test functions. The implicit constant in  $\mathcal{O}(\cdot)$  may depend on  $c_0, c_1$  at most polynomially.

*Proof of Theorem 2.7.* We present the proof only for the more involved Case 2), the Case 1) is much easier and omitted. Similarly to (3.8) in the proof of Theorem 2.4 it is enough to consider the case when the linear size of the support of  $x$  is bounded by some  $\tilde{c} > 0$  (determined later) and prove that

$$\mathbf{E}_H \left| \mathbf{E}_x f(N\rho^x(E)\delta\lambda_{i_0(x,E)}) - \langle f \rangle_{\text{gap}} \right|^2 \leq N^{-2\alpha - \kappa} \|f\|_{C^5}^2, \quad (3.10)$$



for some small  $\alpha, \kappa > 0$ .

Note that under the assumptions of **Case 2**) in Theorem 6.2 for any  $x_1, x_2$  in the support of the random variable  $x$  it holds

$$\theta|i_0(x_1, E) - i_0(x_2, E)| \leq N|x_1 - x_2| \leq \Theta|i_0(x_1, E) - i_0(x_2, E)| \quad (3.11)$$

with some  $\theta, \Theta$  (depending on  $c_0, c_1$ ) as long as  $|x_1 - x_2| \gg N^{-1}$ . The bound in (3.11) is a direct consequence of the following Lemma 3.4 (assuming that  $\tilde{c} \leq c^*$ ) whose proof is postponed to Appendix A.

**Lemma 3.4.** *For any  $c_1 > 0$ , there exists a  $c^* = c^*(c_1) > 0$  such that for  $x_1, x_2$  with  $|x_1 - x_2| \leq c^*$  it holds that*

$$\gamma_i^{x_1} = \gamma_i^{x_2} + (x_1 - x_2)\langle A \rangle + \mathcal{O}\left(|x_1 - x_2|\langle \hat{A}^2 \rangle^{1/2} + |x_1 - x_2|^2\right), \quad (3.12)$$

where  $\gamma_i^{x_r}$  are the quantiles of  $\rho^{x_r}$  and  $i$  is in the  $c_1$ -bulk, i.e.  $\rho^{x_r}(\gamma_i^{x_r}) \geq c_1$ , for  $r \in [2]$ .

Indeed, (3.11) follows from  $|\langle A \rangle| \geq c_0$  and from the inequality

$$\begin{aligned} \left| \frac{i_0(x_1, E)}{N} - \frac{i_0(x_2, E)}{N} \right| &= \left| \int_{\gamma_{i_0(x_2, E)}^{x_1}}^{\gamma_{i_0(x_1, E)}^{x_1}} \rho^{x_1}(E) \, dE \right| \geq c_1 \left| \gamma_{i_0(x_2, E)}^{x_1} - \gamma_{i_0(x_1, E)}^{x_1} \right| \\ &\geq \frac{c_1}{2} \left| \gamma_{i_0(x_2, E)}^{x_1} - \gamma_{i_0(x_2, E)}^{x_2} \right| + \mathcal{O}(N^{-1}) \geq \frac{c_1}{4} |x_1 - x_2| |\langle A \rangle|, \end{aligned}$$

where to go from the first to the second line we used that  $\gamma_{i_0(x_2, E)}^{x_2} = \gamma_{i_0(x_1, E)}^{x_1} + \mathcal{O}(N^{-1})$  by the definition of  $i_0(x, E)$  and that we are in the bulk. In the last inequality we used (3.12) and that its error terms are negligible by  $c_0|\langle A \rangle| \geq \langle \hat{A}^2 \rangle^{1/2}$  and  $|x_1 - x_2| \leq \tilde{c}$  assuming that  $\tilde{c} \leq c_0/10$ . Then by a similar chain of inequalities, and using (3.12) once more, we get the matching upper bound in (3.11).

To prove (3.10), we use the counterpart of (3.5) and that we can neglect the regime  $|x_1 - x_2| \leq N^{-\epsilon_2}$  for some sufficiently small  $\epsilon_2 > 0$  so that we can apply Proposition 3.3 with  $\zeta_1 = \epsilon_2$ . We remove this regime to ensure that on its complement  $|i_0(x_1, E) - i_0(x_2, E)|$  is sufficiently large by (3.11). More precisely, for any  $x_1, x_2$  with  $N^{-\epsilon_2} \leq |x_1 - x_2| \leq \tilde{c}$ , we have  $\Theta^{-1}N^{1-\epsilon_2} \leq |i_0(x_1, E) - i_0(x_2, E)| \leq \theta^{-1}\tilde{c}N$  from (3.11). Assuming  $\tilde{c} \leq c_*\theta$ , we can apply **Case 2**) of Proposition 3.3 by choosing  $j_1 = i_0(x_1, E)$ ,  $j_2 = i_0(x_2, E)$  and with exponent  $\zeta_2$  to factorise the expectation in the equivalent of (3.5). Using again the gap universality (3.6), similarly to (3.8) we conclude (2.12) choosing  $\alpha, \kappa > 0$  appropriately.  $\square$

**Remark 3.5.** *The proof of (2.12) in the case  $\langle \hat{A}^2 \rangle \geq c$  is analogous to the proof of Theorem 2.4 above. We note that Proposition 3.1 allows to also conclude the asymptotic independence of  $\delta\lambda_{i_0(x_1, E)}$  and  $\delta\lambda_{i_0(x_2, E)}$  since  $|i_0(x_1, E) - i_0(x_2, E)| \lesssim N|x_1 - x_2|$  due to Lemma 3.4.*

**3.3. Proof of Corollary 2.1.** Picking  $E = 0$  and the test function<sup>7</sup>  $f(u) = \mathbf{1}(0 \leq u \leq y)$  in **Case 1**) of Theorem 2.7, and choosing the random variable  $x$  such that  $-x$  has density proportional to  $\rho|_I$ , with  $\rho = \rho_{sc}$ , it follows that with very high probability in the space of  $H$  it holds that

$$\begin{aligned} &\mathbf{E}_x F(N\rho^x(0)\delta\lambda_{i_0(x,0)}^x) \\ &= \mathbf{E}_x F(N\rho(\lambda_{i_0(0,-x)})\delta\lambda_{i_0(0,-x)}) + \mathcal{O}(N^{-1+\xi}) \\ &= \left( \int_I \rho \right)^{-1} \sum_i F(N\rho(\lambda_i)\delta\lambda_i) \int_{I \cap (\gamma_{i-1}, \gamma_i]} \rho(t) \, dt + \mathcal{O}(N^{-1+\xi}) \\ &= \left( N \int_I \rho \right)^{-1} \# \left\{ i \left| \lambda_{i+1} - \lambda_i \leq \frac{y}{N\rho(\lambda_i)}, \lambda_i \in I \right. \right\} + \mathcal{O}(N^{-1+\xi}|I|^{-1}), \end{aligned} \quad (3.13)$$

where in the first and third step we used rigidity (see e.g. [15, Lemma 7.1, Theorem 7.6] or [21, Section 5]), i.e. that for any small  $\xi > 0$  we have

$$|\lambda_i - \gamma_i| \leq \frac{N^\xi}{N}, \quad (3.14)$$

for all  $\gamma_i$  in the bulk, with very high probability.

<sup>7</sup>While  $f$  does not literally satisfy the regularity condition, one can easily extend the validity of (2.12) to interval characteristic functions  $f$  by a standard approximation argument.

## 4. DBM ANALYSIS: PROOF OF PROPOSITION 3.1 AND PROPOSITION 3.3

In this section we first present the proof of Proposition 3.1 in details and later in Section 4.3 we explain the very minor changes that are required to prove Proposition 3.3.

4.1. **Proof of Proposition 3.1.** By standard Green function comparison (GFT) argument (see e.g. [19, Section 15]) it is enough to prove Proposition 3.1 only for matrices with a small Gaussian component. More precisely, consider the DBM flow

$$d\widehat{H}_t = \frac{d\widehat{B}_t}{\sqrt{N}}, \quad \widehat{H}_0 = H^\#, \quad (4.1)$$

with  $\widehat{B}_t$  being a real symmetric or complex Hermitian standard Brownian motion (see e.g. [19, Section 12.1] for the precise definition) independent of the initial condition  $\widehat{H}_0 = H^\#$ , where  $H^\#$  is a deformed Wigner matrix specified later. Throughout this section we fix  $T > 0$  and analyse the DBM for times  $0 \leq t \leq T$ .

We denote the ordered collection of eigenvalues of  $\widehat{H}_t + xA$  by  $\boldsymbol{\lambda}^x(t) = \{\lambda_i^x(t)\}_{i \in [N]}$ . The main result of this section is the asymptotic independence of  $\boldsymbol{\lambda}^{x_1}(t_1)$ ,  $\boldsymbol{\lambda}^{x_2}(t_1)$  for  $|x_1 - x_2| \geq N^{-\zeta_1}$  and  $t_1 \geq N^{-1+\omega_1}$ , for some  $\omega_1 > 0$ . We note that in this entire section we do not use the randomness of  $x$ , in the statement of Propositions 3.1 and 3.3  $x_1, x_2$  are fixed parameters. Hence all probabilistic statements, such as covariances etc., are understood in the probability space of the random matrices and the driving Brownian motions in (4.1).

**Proposition 4.1.** *Let  $H^\#$  be a deformed Wigner matrix satisfying Assumption 2.2, let  $\widehat{H}_t$  be the solution of (4.1), and let  $A$  be a deterministic matrix such that  $\langle \widehat{A}^2 \rangle \geq c_0$  and  $\|A\| \lesssim 1$ . Then there exists a small  $c^* > 0$  (depending on  $c_0, c_1$ ) and for any small  $\zeta_1, \omega_1 > 0$  there exists some  $\zeta_2 > 0$  such that the following hold. Fix  $x_1, x_2$  with  $N^{-\zeta_1} \leq |x_1 - x_2| \leq c^*$  and indices  $j_1, j_2$  such that  $|j_1 - j_2| \lesssim N|x_1 - x_2|$ , and the corresponding quantiles  $\gamma_{j_r}^{x_r}$  are in the bulk of the spectrum of  $H^\# + x_r A$  for  $r = 1, 2$ . Then for the eigenvalues of  $\widehat{H}_t + x_r A$  it holds that*

$$\text{Cov} \left( P \left( N\delta\lambda_{j_1}^{x_1}(t_1) \right), Q \left( N\delta\lambda_{j_2}^{x_2}(t_1) \right) \right) = \mathcal{O} \left( N^{-\zeta_2} \|P\|_{C^1} \|Q\|_{C^1} \right), \quad (4.2)$$

with  $t_1 = N^{-1+\omega_1}$  for any  $P, Q: \mathbf{R} \rightarrow \mathbf{R}$  bounded smooth compactly supported test functions.

Using Proposition 4.1 as an input we readily conclude Proposition 3.1.

*Proof of Proposition 3.1.* Let  $H$  be the deformed Wigner matrix from Proposition 3.1, and consider the Ornstein-Uhlenbeck flow

$$dH_t = -\frac{1}{2}(H_t - \mathbf{E} H_0) dt + \frac{dB_t}{\sqrt{N}}, \quad H_0 = H, \quad (4.3)$$

with  $B_t$  being a real symmetric or complex Hermitian standard Brownian motion independent of  $H_0$ .

Let  $H_{t_1}^\#$ , with  $t_1$  from Proposition 4.1, be such that

$$H_{t_1} \stackrel{d}{=} H_{t_1}^\# + \sqrt{c(t_1)t_1}U, \quad (4.4)$$

with  $U$  a GOE/GUE matrix independent of  $H_{t_1}^\#$  and  $c = c(t_1) = 1 + \mathcal{O}(t_1)$  is an appropriate constant very close to one. Then by (4.4) it follows that

$$\widehat{H}_{ct_1} \stackrel{d}{=} H_{t_1}, \quad (4.5)$$

with  $\widehat{H}_{ct_1}$  being the solution of (4.1) with initial condition  $\widehat{H}_0 = H_{t_1}^\#$ .

Then, by a standard GFT argument [19, Section 15], we have that

$$\text{Cov} \left( P \left( N\delta\lambda_{j_1}^{x_1} \right), Q \left( N\delta\lambda_{j_2}^{x_2} \right) \right) = \text{Cov} \left( P \left( N\delta\lambda_{j_1}^{x_1}(ct_1) \right), Q \left( N\delta\lambda_{j_2}^{x_2}(ct_1) \right) \right) + \mathcal{O} \left( N^{-\zeta_2} \|P\|_{C^5} \|Q\|_{C^5} \right). \quad (4.6)$$

Finally, by (4.6) together with (4.5) and Proposition 4.1 applied to  $H^\# := H_{t_1}^\#$  we conclude the proof of Proposition 3.1.  $\square$

**4.2. Proof of Proposition 4.1.** This proof is an adaptation of the proof of [11, Proposition 7.2] (which itself is based upon [28]) with two minor differences. First, the DBM in this paper is for eigenvalues (see (4.7) below) while in [11, Eq. (7.15)] it was for singular values. Second, in [11, Section 7] it was sufficient to consider singular values close to zero hence the base points  $j_1$  and  $j_2$  were fixed to be 0; here they are arbitrary. Both changes are simple to incorporate, so we present only the backbone of the proof that shows the differences, skipping certain steps that remain unaffected.

The flow (4.1) induces the following flow on the eigenvalues of  $\widehat{H}_t + x_t A$ :

$$d\lambda_i^{x_r}(t) = \sqrt{\frac{2}{\beta N}} db_i^{x_r}(t) + \frac{1}{N} \sum_{j \neq i} \frac{1}{\lambda_i^{x_r}(t) - \lambda_j^{x_r}(t)} dt, \quad (4.7)$$

with  $r \in [2]$  and  $\beta = 1, 2$  in the real and complex case, respectively. Here (omitting the time dependence) we used the notation

$$db_i^{x_r} = \sqrt{\frac{2}{\beta}} \sum_{a,b=1}^N \mathbf{u}_i^{x_r}(a) d\widehat{B}_{ab}(t) \mathbf{u}_i^{x_r}(b), \quad (4.8)$$

with  $\mathbf{u}_i^{x_r}(t)$  being the orthonormal eigenvectors of  $\widehat{H}_t + x_r A$ . The collection  $\mathbf{b}^{x_r} := \{b_i^{x_r}\}_{i \in [N]}$ , for fixed  $r$ , consists of i.i.d standard real Brownian motions. However, the families  $\mathbf{b}^{x_1}, \mathbf{b}^{x_2}$  are not independent for different  $r$ 's, in fact their joint distribution is not necessarily Gaussian. The quadratic covariation of these two processes is given by

$$d[b_i^{x_1}(t), b_j^{x_2}(t)] = \langle \mathbf{u}_i^{x_1}(t), \mathbf{u}_j^{x_2}(t) \rangle^2 dt. \quad (4.9)$$

We remark that in (4.9) we used a different notation for the quadratic covariation compared to [11, Section 7.2.1].

**4.2.1. Definition of the comparison processes for  $\lambda^{x_r}$ .** To make the notation cleaner we only consider the real case ( $\beta = 1$ ). To prove the asymptotic independence of the processes  $\lambda^{x_1}, \lambda^{x_2}$ , realized on the probability space  $\Omega_b$ , we will compare them with two completely independent processes  $\boldsymbol{\mu}^{(r)}(t) = \{\mu_i^{(r)}(t)\}_{i=1}^N$  realized on a different probability space  $\Omega_\beta$ . The processes  $\boldsymbol{\mu}^{(r)}(t)$  are the unique strong solution of

$$d\mu_i^{(r)}(t) = \sqrt{\frac{2}{N}} d\beta_i^{(r)} + \frac{1}{N} \sum_{j \neq i} \frac{1}{\mu_i^{(r)}(t) - \mu_j^{(r)}(t)} dt, \quad \mu_i^{(r)}(0) = \mu_i^{(r)}, \quad (4.10)$$

with  $\mu_i^{(r)}$  being the eigenvalues of two independent GOE matrices  $H^{(r)}$ , and  $\beta^{(r)} = \{\beta_i^{(r)}(t)\}_{i=1}^N$  being independent vectors of standard i.i.d. Brownian motions.

We now define two intermediate processes  $\widetilde{\lambda}^{(r)}(t), \widetilde{\mu}^{(r)}(t)$  so that for  $t \gg N^{-1}$  the particles  $\widetilde{\lambda}_i^{(r)}(t), \widetilde{\mu}_i^{(r)}(t)$  will be close to  $\lambda_i^{x_r}(t)$  and  $\mu_i^{x_r}(t)$ , respectively, for indices  $i$  close to  $j_r$ , with very high probability (see Lemmas 4.2–4.3 below). Additionally, the processes  $\widetilde{\lambda}^{(r)}(t), \widetilde{\mu}^{(r)}(t)$ , which will be realized on two different probability spaces, will have the same joint distribution:

$$\left( \widetilde{\lambda}^{(1)}(t), \widetilde{\lambda}^{(2)}(t) \right)_{0 \leq t \leq T} \stackrel{d}{=} \left( \widetilde{\mu}^{(1)}(t), \widetilde{\mu}^{(2)}(t) \right)_{0 \leq t \leq T}. \quad (4.11)$$

Fix any small  $\omega_A > 0$  (later  $\omega_A$  will be chosen smaller than  $\omega_B$  from (4.27)) and define the process  $\widetilde{\lambda}^{(r)}(t)$  to be the unique strong solution of

$$d\widetilde{\lambda}_i^{(r)}(t) = \frac{1}{N} \sum_{j \neq i} \frac{1}{\widetilde{\lambda}_i^{(r)}(t) - \widetilde{\lambda}_j^{(r)}(t)} dt + \begin{cases} \sqrt{\frac{2}{N}} db_i^{x_r} & \text{if } |i - j_r| \leq N^{\omega_A}, \\ \sqrt{\frac{2}{N}} d\widetilde{b}_i^{(r)} & \text{if } |i - j_r| > N^{\omega_A}, \end{cases} \quad (4.12)$$

with initial data  $\widetilde{\lambda}^{(r)}(0)$  being the eigenvalues of independent GOE matrices, which are also independent of  $H^\#$  in (4.1). Here the Brownian motions

$$\underline{\mathbf{b}}^{\text{in}} = (b_{j_1 - N^{\omega_A}}^{x_1}, \dots, b_{j_1 + N^{\omega_A}}^{x_1}, b_{j_2 - N^{\omega_A}}^{x_2}, \dots, b_{j_2 + N^{\omega_A}}^{x_2}). \quad (4.13)$$

for indices close to  $j_r$  are exactly the ones in (4.7). For indices away from  $j_r$  we define the driving Brownian motions to be an independent family

$$\underline{\mathbf{b}}^{\text{out}} = \left\{ \widetilde{b}_i^{(r)} \mid |i - j_r| > N^{\omega_A}, r \in [2] \right\}. \quad (4.14)$$

of standard real i.i.d. Brownian motions which are also independent of  $\underline{\mathbf{b}}^{\text{in}}$ . The Brownian motions  $\underline{\mathbf{b}}^{\text{out}}$  are defined on the same probability space of  $\underline{\mathbf{b}}^{\text{in}}$ , which we will still denote by  $\Omega_b$ , with a slight abuse of notation.

For any  $i, j \in [4N^{\omega_A} + 2]$  we use the notation

$$i = (r-1)N^{\omega_A} + i, \quad j = (m-1)N^{\omega_A} + j \quad (4.15)$$

with  $r, m \in [2]$  and  $i, j \in [2N^{\omega_A} + 1]$ . The covariance matrix  $C(t)$  of the increments of  $\underline{\mathbf{b}}^{\text{in}}$ , consisting of four blocks of size  $2N^{\omega_A} + 1$ , is given by

$$C_{ij}(t) dt := d[\underline{\mathbf{b}}_i^{\text{in}}, \underline{\mathbf{b}}_j^{\text{in}}] = \Theta_{ij}^{x_r, x_m}(t) dt, \quad (4.16)$$

where

$$\Theta_{ij}^{x_r, x_m}(t) := |\langle \mathbf{u}_{i+j_r-N^{\omega_A}-1}^{x_r}(t), \mathbf{u}_{j+j_m-N^{\omega_A}-1}^{x_m}(t) \rangle|^2 \quad (4.17)$$

and  $\{\mathbf{u}_i^{x_r}(t)\}_{i=1}^N$  are the orthonormal eigenvectors of  $\widehat{H}_t + x_r A$ . Note that  $\{\mathbf{u}_i^{x_r}(t)\}_{i=1}^N$  are not well-defined if  $\widehat{H}_t + x_r A$  has multiple eigenvectors, however, without loss of generality, we can assume that almost surely  $\widehat{H}_t + x_r A$  does not have multiple eigenvectors for any  $r \in [2]$  for almost all  $t \in [0, T]$  by [9, Proposition 2.3] together with Fubini's theorem. By Doob's martingale representation theorem [27, Theorem 18.12] there exists a real standard Brownian motion  $\boldsymbol{\theta}(t) \in \mathbf{R}^{4N^{\omega_A}+2}$  such that

$$d\underline{\mathbf{b}}^{\text{in}} = \sqrt{C} d\boldsymbol{\theta}. \quad (4.18)$$

Similarly, on the probability space  $\Omega_\beta$  we define the comparison process  $\widetilde{\boldsymbol{\mu}}^{(r)}(t)$  to be the solution of

$$d\widetilde{\boldsymbol{\mu}}_i^{(r)}(t) = \frac{1}{N} \sum_{j \neq i} \frac{1}{\widetilde{\boldsymbol{\mu}}_i^{(r)}(t) - \widetilde{\boldsymbol{\mu}}_j^{(r)}(t)} dt + \begin{cases} \sqrt{\frac{2}{N}} d\zeta_i^{(r)} & \text{if } |i-j_r| \leq N^{\omega_A}, \\ \sqrt{\frac{2}{N}} d\widetilde{\zeta}_i^{(r)} & \text{if } |i-j_r| > N^{\omega_A}, \end{cases} \quad (4.19)$$

with initial data  $\widetilde{\boldsymbol{\mu}}^{(r)}(0)$  being the eigenvalues of independent GOE matrices defined on the probability space  $\Omega_\beta$ , which are also independent of  $H^{(r)}$ . We now construct the driving Brownian motions in (4.19) so that (4.11) is satisfied. For indices away from  $j_r$ , the standard real Brownian motions

$$\underline{\zeta}^{\text{out}} = \left\{ \widetilde{\zeta}_i^{(r)} \mid |i-j_r| > N^{\omega_A}, r \in [2] \right\} \quad (4.20)$$

are i.i.d. and they are independent of  $\boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(2)}$  in (4.10). For indices  $|i-j_r| \leq N^{\omega_A}$  the collections

$$\underline{\zeta}^{\text{in}} = (\zeta_{j_1-N^{\omega_A}}^{(1)}, \dots, \zeta_{j_1+N^{\omega_A}}^{(1)}, \zeta_{j_2-N^{\omega_A}}^{(2)}, \dots, \zeta_{j_2+N^{\omega_A}}^{(2)}) \quad (4.21)$$

will be constructed from the independent families<sup>8</sup>

$$\underline{\boldsymbol{\beta}}^{\text{in}} := (\beta_{j_1-N^{\omega_A}}^{(1)}, \dots, \beta_{j_1+N^{\omega_A}}^{(1)}, \beta_{j_2-N^{\omega_A}}^{(2)}, \dots, \beta_{j_2+N^{\omega_A}}^{(2)}), \quad (4.22)$$

as follows. Since the original process  $\boldsymbol{\lambda}^{x_r}(t)$  and the comparison processes  $\boldsymbol{\mu}^{(r)}(t)$  are realized on two different probability spaces, we construct a matrix valued process  $C^\#(t)$  and a vector-valued Brownian motion  $\underline{\boldsymbol{\beta}}^{\text{in}}$  on the probability space  $\Omega_\beta$  such that  $(C^\#(t), \underline{\boldsymbol{\beta}}^{\text{in}}(t))$  have the same joint distribution as  $(C(t), \boldsymbol{\theta}(t))$  with  $C, \boldsymbol{\theta}$  from (4.18). This  $\underline{\boldsymbol{\beta}}^{\text{in}}$  is the driving Brownian motion of the  $\boldsymbol{\mu}^{(r)}(t)$  process in (4.10). Define the process

$$\underline{\zeta}^{\text{in}}(t) := \int_0^t \sqrt{C^\#(s)} d\underline{\boldsymbol{\beta}}^{\text{in}}(s) \quad (4.23)$$

on the probability space  $\Omega_\beta$ . By construction we see that the processes  $\underline{\mathbf{b}}^{\text{in}}$  and  $\underline{\zeta}^{\text{in}}$  have the same distribution, and that the two collections  $\underline{\mathbf{b}}^{\text{out}}$  and  $\underline{\zeta}^{\text{out}}$  are independent of  $\underline{\mathbf{b}}^{\text{in}}, \underline{\boldsymbol{\beta}}^{\text{in}}$  and among each other. Hence we conclude that

$$(\underline{\mathbf{b}}^{\text{in}}(t), \underline{\mathbf{b}}^{\text{out}}(t))_{0 \leq t \leq T} \stackrel{d}{=} (\underline{\zeta}^{\text{in}}(t), \underline{\zeta}^{\text{out}}(t))_{0 \leq t \leq T}. \quad (4.24)$$

Finally, by the definitions in (4.12), (4.19) and by (4.24), we conclude that the processes  $\widetilde{\boldsymbol{\lambda}}^{(r)}(t), \widetilde{\boldsymbol{\mu}}^{(r)}(t)$  have the same joint distribution (see (4.11)), since their initial conditions and their driving processes (4.24) agree in distribution.

<sup>8</sup>The families  $\underline{\boldsymbol{\beta}}^{\text{in}}, \underline{\mathbf{b}}^{\text{in}}$  were denoted by  $\underline{\boldsymbol{\beta}}$  and  $\underline{\mathbf{b}}$  respectively, in [11, Eqs. (7.22)-(7.23)].

4.2.2. *Proof of the asymptotic independence of the eigenvalues.* In this section we use that the processes  $\lambda^{x_r}(t)$ ,  $\tilde{\lambda}^{(r)}(t)$  and  $\tilde{\mu}^{(r)}(t)$ ,  $\mu^{(r)}(t)$  are close pathwise at time  $t_1 = N^{-1+\omega_1}$  as stated below in Lemma 4.2 and Lemma 4.3, respectively, to conclude the proof of Proposition 4.1. The proof of these lemmas is completely analogous to the proof in [11, Lemmas 7.6–7.7], [28, Eq. (3.7), Theorem 3.1], hence we will only explain the very minor differences required in this paper. First, we compare the processes  $\lambda^{x_r}(t)$ ,  $\tilde{\lambda}^{(r)}(t)$ , in particular this lemma shows that for  $i$  far away from  $j_1, j_2$  the Brownian motions  $b_i^{x_1}, b_i^{x_2}$  can be replaced by the independent Brownian motions from  $\underline{b}^{\text{out}}$  at a negligible error.

**Lemma 4.2.** *Let  $\lambda^{x_r}(t)$ ,  $\tilde{\lambda}^{(r)}(t)$ , with  $r \in [2]$ , be the processes defined in (4.7) and (4.12), respectively. For any small  $\omega_1 > 0$  there exists  $\omega > 0$ , with  $\omega \ll \omega_1$ , such that it holds*

$$\left| \rho^{x_r}(\gamma_{j_r}^{x_r}) \delta \lambda_{j_r}^{x_r}(t_1) - \rho_{\text{sc}}(\gamma_{j_r}) \delta \tilde{\lambda}_{j_r}^{(r)}(t_1) \right| \leq N^{-1-\omega}, \quad (4.25)$$

for any  $j_r$  in the  $c_1$ -bulk, with very high probability on the probability space  $\Omega_b$ , where  $t_1 := N^{-1+\omega_1}$ . Here by  $\gamma_{j_r}$  we denoted the  $j_r$ -quantile of the semicircular law.

Second, we compare the processes  $\tilde{\mu}^{(r)}(t)$ ,  $\mu^{(r)}(t)$ , i.e. we control the error made by replacing the weakly correlated Brownian motions  $\underline{z}^{\text{in}}$  by the independent Brownian motions  $\underline{\beta}^{\text{in}}$ .

**Lemma 4.3.** *Let  $\mu^{(r)}(t)$ ,  $\tilde{\mu}^{(r)}(t)$ , with  $r \in [2]$ , be the processes defined in (4.10) and (4.19), respectively. For any small  $\omega_1, \zeta_1 > 0$  there exists  $\omega > 0$ , with  $\omega \ll \omega_1$ , such that for any  $N^{-\zeta_1} \leq |x_1 - x_2| \leq c^*$  it holds*

$$\left| \delta \mu_{j_r}^{(r)}(t_1) - \delta \tilde{\mu}_{j_r}^{(r)}(t_1) \right| \leq N^{-1-\omega}, \quad (4.26)$$

with very high probability on the probability space  $\Omega_\beta$ , where  $t_1 := N^{-1+\omega_1}$ .

The key ingredient for the proof of Lemma 4.3 is the following fundamental bound on the eigenvector overlaps in (4.27) proven in Section 5, which ensures that the correlation  $\Theta_{ij}^{x_r, x_m}$  in (4.17) is small.

**Proposition 4.4.** *Given  $c_0, c_1$  as in Proposition 3.1, assume  $\langle \hat{A}^2 \rangle \geq c_0$ ,  $\|A\| \lesssim 1$ . There exists  $c^*$  depending on  $c_0, c_1$  such that the following holds for any small  $\zeta_1 > 0$ . Pick  $x_1, x_2$  such that  $N^{-\zeta_1} \leq |x_1 - x_2| \leq c^*$ , and let  $\{\mathbf{u}_i^{x_r}\}_{i \in [N]}$ , for  $r \in [2]$ , be the orthonormal eigenbasis of the matrices  $H + x_r A$ . Then there exists  $\omega_E > 0$  such that*

$$|\langle \mathbf{u}_{j_1}^{x_1}, \mathbf{u}_{j_2}^{x_2} \rangle| \leq N^{-\omega_E} \quad (4.27)$$

with very high probability for any  $j_1, j_2$  in the  $c_1$ -bulk with  $|j_1 - j_2| \lesssim N|x_1 - x_2|$ .

Using Lemmas 4.2–4.3 as an input we conclude Proposition 4.1.

*Proof of Proposition 4.1.* By Lemma 4.2 we readily conclude that

$$\mathbf{E} \left[ P \left( N \delta \lambda_{j_1}^{x_1}(t_1) \right) Q \left( N \delta \lambda_{j_2}^{x_2}(t_1) \right) \right] = \mathbf{E} \left[ P \left( N \rho_1 \delta \tilde{\lambda}_{j_1}^{(1)}(t_1) \right) Q \left( N \rho_2 \delta \tilde{\lambda}_{j_2}^{(2)}(t_1) \right) \right] + \mathcal{O} \left( N^{-\omega} \|P\|_{C^1} \|Q\|_{C^1} \right), \quad (4.28)$$

where we denoted  $\rho_r := \rho_{\text{sc}}(\gamma_{j_r}) / \rho^{x_r}(\gamma_{j_r}^{x_r})$  and used the uniform boundedness of  $\rho_{\text{sc}}, \rho^x$ . Then, by (4.11), it follows that

$$\mathbf{E} \left[ P \left( N \rho_1 \delta \tilde{\lambda}_{j_1}^{(1)}(t_1) \right) Q \left( N \rho_2 \delta \tilde{\lambda}_{j_2}^{(2)}(t_1) \right) \right] = \mathbf{E} \left[ P \left( N \rho_1 \delta \tilde{\mu}_{j_1}^{(1)}(t_1) \right) Q \left( N \rho_2 \delta \tilde{\mu}_{j_2}^{(2)}(t_1) \right) \right]. \quad (4.29)$$

Moreover, by Lemma 4.3, we have that

$$\mathbf{E} \left[ P \left( N \rho_1 \delta \tilde{\mu}_{j_1}^{(1)}(t_1) \right) Q \left( N \rho_2 \delta \tilde{\mu}_{j_2}^{(2)}(t_1) \right) \right] = \mathbf{E} \left[ P \left( N \rho_1 \delta \mu_{j_1}^{(1)}(t_1) \right) Q \left( N \rho_2 \delta \mu_{j_2}^{(2)}(t_1) \right) \right] + \mathcal{O} \left( N^{-\omega} \|P\|_{C^1} \|Q\|_{C^1} \right). \quad (4.30)$$

Additionally, by the definition of the processes  $\mu^{(r)}(t)$  in (4.10) it follows that  $\mu^{(1)}(t)$ ,  $\mu^{(2)}(t)$  are independent, and so that

$$\mathbf{E} \left[ P \left( N \rho_1 \delta \mu_{j_1}^{(1)}(t_1) \right) Q \left( N \rho_2 \delta \mu_{j_2}^{(2)}(t_1) \right) \right] = \mathbf{E} \left[ P \left( N \rho_1 \delta \mu_{j_1}^{(1)}(t_1) \right) \right] \mathbf{E} \left[ Q \left( N \rho_2 \delta \mu_{j_2}^{(2)}(t_1) \right) \right] \quad (4.31)$$

Combining (4.28)–(4.31) we get

$$\mathbf{E} \left[ P \left( N \rho_1 \delta \lambda_{j_1}^{x_1}(t_1) \right) Q \left( N \rho_2 \delta \lambda_{j_2}^{x_2}(t_1) \right) \right] = \mathbf{E} \left[ P \left( N \rho_1 \delta \mu_{j_1}^{(1)}(t_1) \right) \right] \mathbf{E} \left[ Q \left( N \rho_2 \delta \mu_{j_2}^{(2)}(t_1) \right) \right] + \mathcal{O} \left( N^{-\omega} \|P\|_{C^1} \|Q\|_{C^1} \right). \quad (4.32)$$

Proceeding similarly to (4.28)–(4.30), but for  $\mathbf{E}P$  and  $\mathbf{E}Q$  separately, we also conclude that

$$\mathbf{E} \left[ P \left( N \delta \lambda_{j_1}^{x_1}(t_1) \right) \right] \mathbf{E} \left[ Q \left( N \delta \lambda_{j_2}^{x_2}(t_1) \right) \right] = \mathbf{E} \left[ P \left( N \rho_1 \delta \mu_{j_1}^{(1)}(t_1) \right) \right] \mathbf{E} \left[ Q \left( N \rho_2 \delta \mu_{j_2}^{(2)}(t_1) \right) \right] + \mathcal{O} \left( N^{-\omega} \|P\|_{C^1} \|Q\|_{C^1} \right). \quad (4.33)$$

Finally, combining (4.32)–(4.33), we conclude the proof of (4.2).  $\square$

Before concluding this section with the proof of Lemmas 4.2–4.3, in Proposition 4.7 below we state the main technical result used in their proofs. The proofs of these lemmas rely on extending the homogenisation analysis of [28, Theorem 3.1] to two DBM processes with weakly coupled driving Brownian motions. We used a very similar idea in [11, Section 7.4] for DBM processes for singular values. We now first present the general version of this idea before applying it to prove Lemmas 4.2–4.3.

In Proposition 4.7 below we compare the evolution of two DBMs whose driving Brownian motions are nearly the same for indices close to a fixed index  $i_0$  and are independent for indices away from  $i_0$ . Proposition 4.7 is the counterpart of [11, Proposition 7.14], where a similar analysis is performed for DBMs describing the evolution of particles satisfying slightly different DBMs.

Define the processes  $s_i(t), r_i(t)$  to be the solution of

$$ds_i(t) = \sqrt{\frac{2}{N}} db_i^s(t) + \frac{1}{2N} \sum_{j \neq i} \frac{1}{s_i(t) - s_j(t)} dt, \quad i \in [N], \quad (4.34)$$

and

$$dr_i(t) = \sqrt{\frac{2}{N}} db_i^r(t) + \frac{1}{2N} \sum_{j \neq i} \frac{1}{r_i(t) - r_j(t)} dt, \quad i \in [N], \quad (4.35)$$

with initial conditions  $s_i(0) = s_i$  being the eigenvalues of a deformed Wigner matrix  $H$  satisfying Assumption 2.2, and  $r_i(0) = r_i$  being the eigenvalues of a GOE matrix. Here we used the same notations of [11, Eqs. (7.44)–(7.45)] to make the comparison with [11] easier. For simplicity in (4.34)–(4.35) we consider the DBMs only in the real case (the complex case is completely analogous).

**Remark 4.5.** In [11, Eqs. (7.44)] we assumed that the initial condition  $s_i(0) = s_i$  were general points satisfying [11, Definition 7.12], and not necessary the singular values of a matrix. Here we choose  $s_i(0) = s_i$  to be the eigenvalues of a deformed Wigner matrix to make the presentation shorter and simpler, however Proposition 4.7 clearly holds also for collections of particles satisfying similar assumptions to [11, Definition 7.12].

We now formulate the assumptions on the driving Brownian motions in (4.34)–(4.35). Set an  $N$ -independent parameter  $K = K_N := N^{\omega_K}$ , for some small fixed  $\omega_K > 0$ .

**Assumption 4.6.** Suppose that the families  $\{\mathbf{b}_i^s\}_{i \in [N]}$ ,  $\{\mathbf{b}_i^r\}_{i \in [N]}$  in (4.34) and (4.35) are realised on a common probability space. Let

$$L_{ij}(t) dt := d[\mathbf{b}_i^s(t) - \mathbf{b}_i^r(t), \mathbf{b}_j^s(t) - \mathbf{b}_j^r(t)] \quad (4.36)$$

denote their quadratic covariation (in [11, Eqs. (7.46)] we used a different notation to denote the covariation). Fix an index  $i_0$  in the bulk of  $H$ , and let the processes satisfy the following assumptions:

- (a)  $\{\mathbf{b}_i^s\}_{i \in [N]}$ ,  $\{\mathbf{b}_i^r\}_{i \in [N]}$  are two families of i.i.d. standard real Brownian motions.
- (b)  $\{\mathbf{b}_i^r\}_{|i-i_0| > K}$  is independent of  $\{\mathbf{b}_i^s\}_{i=1}^N$ , and  $\{\mathbf{b}_i^s\}_{|i-i_0| > K}$  is independent of  $\{\mathbf{b}_i^r\}_{i=1}^B$ .
- (c) Fix  $\omega_Q > 0$  so that  $\omega_K \ll \omega_Q$ . We assume that the subfamilies  $\{\mathbf{b}_i^s\}_{|i-i_0| \leq K}$ ,  $\{\mathbf{b}_i^r\}_{|i-i_0| \leq K}$  are very strongly dependent in the sense that for any  $|i-i_0|, |j-i_0| \leq K$  it holds

$$|L_{ij}(t)| \leq n^{-\omega_Q} \quad (4.37)$$

with very high probability for any fixed  $t \geq 0$ .

Let  $\rho$  denote the self-consistent density of  $H$ , and recall that  $\rho_{sc}$  denotes the semicircular density. By  $\rho_t, \rho_{sc,t}$  we denote the evolution of  $\rho$  and  $\rho_{sc}$ , respectively, along the semicircular flow (see e.g. [10, Eq. (4.1)] and let  $\hat{\gamma}_i(t), \gamma_i(t)$  denote the quantiles of  $\rho_t$  and  $\rho_{sc,t}$ .

**Proposition 4.7.** *Let the processes  $\mathbf{s}(t) = \{s_i(t)\}_{i \in [N]}$ ,  $\mathbf{r}(t) = \{r_i(t)\}_{i \in [N]}$  be the solutions of (4.34) and (4.35), and assume that the driving Brownian motions in (4.34)–(4.35) satisfy Assumption 4.6. Let  $i_0$  be the index fixed in Assumption 4.6. Then for any small  $\omega_1, \omega_\ell > 0$  such that  $\omega_1 \ll \omega_\ell \ll \omega_K \ll \omega_Q$  there exist  $\omega, \widehat{\omega} > 0$  with  $\widehat{\omega} \ll \omega \ll \omega_1$ , and such that it holds*

$$\begin{aligned} & \rho(\widehat{\gamma}_{i_0})[s_{i_0+i}(t_1) - \widehat{\gamma}_{i_0}(t_1)] - \rho_{\text{sc}}(\gamma_{i_0})[r_{i_0+i}(t_1) - \gamma_{i_0}(t_1)] \\ &= \sum_{|j| \leq N^{2\omega_1}} \frac{1}{N} p_{t_1} \left( 0, \frac{-j}{N \rho_{\text{sc}}(\gamma_{i_0})} \right) \left[ \rho(\widehat{\gamma}_{i_0})(s_{i_0+j}^{x_r}(0) - \widehat{\gamma}_{i_0}(0)) - \rho_{\text{sc}}(\gamma_{i_0})(r_{i_0+j}^{(r)}(0) - \gamma_{i_0}(0)) \right] + \mathcal{O}(N^{-1-\omega}), \end{aligned} \quad (4.38)$$

for any  $|i| \leq N^{\widehat{\omega}}$ , with very high probability, where  $t_1 := n^{-1+\omega_1}$  and  $p_t(x, y)$  is the fundamental solution (heat kernel) of the parabolic equation

$$\partial_t f(x) = \int_{|x-y| \leq \eta_\ell} \frac{f(y) - f(x)}{(x-y)^2} \rho_{\text{sc}}(\gamma_{i_0}) \, dy, \quad (4.39)$$

with  $\eta_\ell := N^{-1+\omega_\ell} \rho_{\text{sc}}(\gamma_{i_0})^{-1}$  (see [28, Eqs. (3.88)–(3.89)]) for more details.

*Proof.* The proof of (4.38) is nearly identical to that of [28, Theorem 3.1] up to a straightforward modification owing to the fact that the driving Brownian motions in (4.34)–(4.35) are not exactly identical but they are very strongly correlated, see (4.37). A similar modification to handle this strong correlation was explained in details in a closely related context in [11, Proof of Proposition 7.14 in Section 7.6], with the difference that in [11] singular values were considered instead of eigenvalues hence the corresponding DBMs are slightly different. Furthermore, Proposition 4.7 is stated in a simpler form than [11, Proposition 7.14] since the initial conditions are already eigenvalues and not arbitrary points hence they automatically satisfy certain regularity assumptions. The precise changes due to this simplification are described in the technical Remark 4.8 below.  $\square$

**Remark 4.8.** *There are a few differences in the setup of Proposition 4.7 and [11, Proposition 7.14]. These are caused by the fact that we now consider  $s_i(0) = s_i$  to be the eigenvalues of a deformed Wigner matrix  $H$ , instead of a collection of particles satisfying [11, Definition 7.12]. In particular,  $\nu$  in [11, Definition 7.12] can be chosen equal to zero, then, since the eigenvalues of  $H$  are regular ([11, Eq. (7.48)]) on an order one scale, we can choose  $g = N^{-1+\xi}$ , for an arbitrary small  $\xi > 0$ , and  $G = 1$  in [11, Definition 7.12]. Additionally,  $t_f = N^{-1+\omega_f}$  is replaced by  $t_1 = N^{-1+\omega_1}$ , and  $\rho_{\text{fc}, t_f}$  in is replaced by  $\rho$ . Finally, we remark that in [11, Proposition 7.14] for  $\omega_f$  we required that  $\omega_K \ll \omega_f \ll \omega_Q$ , instead in Proposition 4.7 we required that  $\omega_1 \ll \omega_K \ll \omega_Q$ . This discrepancy is caused by the fact that in the proof of [11, Proposition 7.14] we first needed to run the DBM for  $s_i(t)$  for an initial time  $t_0 = N^{-1+\omega_0}$  to regularise the particles  $s_i(0) = s_i$ , with  $\omega_K \ll \omega_0 \ll \omega_Q$ , and then run both DBMs for an additional time  $N^{-1+\omega_1}$ , with  $\omega_1 \ll \omega_K \ll \omega_0 \ll \omega_Q$  (see below [11, Eq. (7.56)]). Finally, in [11, Proposition 7.14] we have  $t_f := t_0 + t_1 \sim t_0$ , hence the reader can think  $\omega_f = \omega_0$ . In the current case we do not need to run (4.34) for an initial time  $t_0$  since  $s_i(0) = s_i$  are already regular being the eigenvalues of a deformed Wigner matrix.*

We are now ready to prove Lemmas 4.2–4.3.

*Proof of Lemmas 4.2–4.3.* By construction the processes  $\boldsymbol{\lambda}^{x_r}(t)$ ,  $\widetilde{\boldsymbol{\lambda}}^{(r)}(t)$  satisfy the assumptions of Proposition 4.7 with  $i_0 = j_r$ ,  $i = 0$ ,  $\rho = \rho^{x_r}$  and  $\omega_K = \omega_A$ . Hence, by Proposition 4.7, we get

$$\begin{aligned} & \rho^{x_r}(\gamma_{j_r}^{x_r})[\lambda_{j_r}^{x_r}(t_1) - \gamma_{j_r}^{x_r}(t_1)] - \rho_{\text{sc}}(\gamma_{j_r})[\widetilde{\lambda}_{j_r}^{(r)}(t_1) - \gamma_{j_r}(t_1)] \\ &= \sum_{|j| \leq N^{2\omega_1}} \frac{1}{N} p_{t_1} \left( 0, \frac{-j}{N \rho_{\text{sc}}(\gamma_{j_r})} \right) \left[ \rho^{x_r}(\gamma_{j_r}^{x_r})(\lambda_{j_r+j}^{x_r}(0) - \gamma_{j_r}^{x_r}(0)) - \rho_{\text{sc}}(\gamma_{j_r})(\widetilde{\lambda}_{j_r+j}^{(r)}(0) - \gamma_{j_r}(0)) \right] + \mathcal{O}(N^{-1-\kappa}), \end{aligned} \quad (4.40)$$

with very high probability, for some small fixed  $\kappa > 0$ . Here  $\gamma_{j_r}^{x_r}(t)$ ,  $\gamma_{j_r}(t)$  denote the quantiles of  $\rho_t^{x_r}$  and  $\rho_{\text{sc}, t}$ , respectively, with  $\rho_t^{x_r}$ ,  $\rho_{\text{sc}, t}$  the evolution of  $\rho^{x_r}$ ,  $\rho_{\text{sc}}$  along the semicircular flow (see e.g. [10, Eq. (4.1)]) and  $p_t(x, y)$  is defined in (4.39).

Analogously, we observe that the processes  $\widetilde{\boldsymbol{\mu}}^{(r)}(t)$ ,  $\boldsymbol{\mu}^{(r)}(t)$  satisfy the assumptions of Proposition 4.7 with  $i_0 = j_r$ ,  $i = 0$ ,  $\rho = \rho_{\text{sc}}$ ,  $\omega_K = \omega_A$ ,  $\omega_Q = \omega_E$  due to (4.27) (in particular (4.27) is needed to check Assumption 4.6–(c)), and thus

we obtain

$$\mu_{j_r}^{(r)}(t_1) - \tilde{\mu}_{j_r}^{(r)}(t_1) = \sum_{|j| \leq N^{2\omega_1}} \frac{1}{N} p_{t_1} \left( 0, \frac{-j}{N\rho_{\text{sc}}(\gamma_{j_r})} \right) [\mu_{j_r+j}^{(r)}(0) - \tilde{\mu}_{j_r+j}^{(r)}(0)] + \mathcal{O}(N^{-1-\kappa}). \quad (4.41)$$

From now on we focus only on the precesses  $\lambda^{x_r}(t_1)$ ,  $\tilde{\lambda}^{(r)}(t_1)$  and so on the proof of Lemma 4.2. The proof to conclude Lemma 4.3 is completely analogous and so omitted. Combining (4.40) with another application of Proposition 4.7, this time for  $i_0 = j_r$  and  $i = 1$ , we readily conclude that

$$\begin{aligned} & \rho^{x_r}(\gamma_{j_r}^{x_r}) [\lambda_{j_r+1}^{x_r}(t_1) - \lambda_{j_r}^{x_r}(t_1)] - \rho_{\text{sc}}(\gamma_{j_r}) [\tilde{\lambda}_{j_r+1}^{(r)}(t_1) - \tilde{\lambda}_{j_r}^{(r)}(t_1)] \\ &= \sum_{|j| \leq N^{2\omega_1}} \frac{1}{N} \left[ p_{t_1} \left( 0, \frac{1-j}{N\rho_{\text{sc}}(\gamma_{j_r})} \right) - p_{t_1} \left( 0, \frac{-j}{N\rho_{\text{sc}}(\gamma_{j_r})} \right) \right] \\ & \quad \times [\rho^{x_r}(\gamma_{j_r}^{x_r})(\lambda_{i_0+j}^{x_r}(0) - \gamma_{j_r}^{x_r}) - \rho_{\text{sc}}(\gamma_{j_r})(\tilde{\lambda}_{i_0+j}^{(r)}(0) - \gamma_{j_r})] + \mathcal{O}(N^{-1-\kappa}) \\ &= \mathcal{O}(N^{-1-\kappa+\xi} + N^{-1-\omega_1+\xi}), \end{aligned} \quad (4.42)$$

with very high probability, where we used rigidity

$$|\lambda_i^{x_r} - \gamma_i^{x_r}| \leq \frac{N^\xi}{N}, \quad (4.43)$$

a similar rigidity bound for  $\tilde{\lambda}_i^{(r)}$ . Additionally, to go to the last line of (4.42) we used the following properties of the heat kernel  $p_{t_1}(x, y)$ :

$$\begin{aligned} \left| p_{t_1} \left( 0, \frac{1-j}{N\rho_{\text{sc}}(\gamma_{j_r})} \right) - p_{t_1} \left( 0, \frac{-j}{N\rho_{\text{sc}}(\gamma_{j_r})} \right) \right| &\leq \frac{1}{N\rho_{\text{sc}}(\gamma_{j_r})} \int_{i_r}^{i_r+1} \left| \partial_y p_{t_1} \left( 0, \frac{\tau-j}{N\rho_{\text{sc}}(0)} \right) \right| d\tau \\ &\lesssim \frac{1}{Nt_1} \int_{i_r}^{i_r+1} p_{t_1} \left( 0, \frac{\tau-j}{N\rho_{\text{sc}}(\gamma_{j_r})} \right) d\tau \\ \frac{1}{N} \sum_{|j| \leq N^{2\omega_1}} p_{t_1} \left( 0, \frac{\tau-j}{N\rho_{\text{sc}}(0)} \right) &= 1 + \mathcal{O}(N^{-\omega_1}). \end{aligned} \quad (4.44)$$

The bound in the second line of (4.44) follows by [28, Eq. (3.96)]. The second relation of (4.44) follows by [28, Eqs. (3.90), (3.103)]. The bound in (4.42) concludes the proof of Lemma 4.2.  $\square$

**4.3. Proof of Proposition 3.3.** We now turn to the proof of Proposition 3.3. We first present **Case 2**) which is structurally very similar to the proof of Proposition 3.3. Afterwards we turn to **Case 1**) which is easier but additionally requires to modify the flow (4.3) to account for the correlations among entries of  $H$ .

**4.3.1. Case 2).** Proceeding as in (4.1)–(4.5) and using the notations and assumptions from **Case 2**) of Proposition 3.3, it is enough to prove

$$\text{Cov} \left( P(N\delta\lambda_{j_1}^{x_1}(t_1)), Q(N\delta\lambda_{j_2}^{x_2}(t_1)) \right) = \mathcal{O} \left( N^{-\zeta_2} \|P\|_{C^1} \|Q\|_{C^1} \right), \quad (4.45)$$

with  $t_1 = N^{-1+\omega_1}$ , for some small  $\omega_1 > 0$ . Here  $\lambda(t)$  are the eigenvalues of  $\widehat{H}_t$ , which is the solution of (4.1) with initial condition  $\widehat{H}_0 = H_{t_1}^\#$ , where  $H_{t_1}^\#$  is from in (4.4).

The proof of (4.45) follows by a DBM analysis very similar to the one in Section 4.2. More precisely, all the processes  $\lambda^{x_r}(t)$ ,  $\tilde{\lambda}^{(r)}(t)$ ,  $\tilde{\mu}^{(r)}(t)$ , and  $\mu^{(r)}(t)$  are defined exactly in the same way; the only difference is that Proposition 4.4 has to be replaced by the following bound on the eigenvector overlap (its proof will be given at the end of Section 5).

**Proposition 4.9.** *We are in the setup of **Case 2**) of Proposition 3.3. For any small  $c_1 > 0$  there exists a  $c_0 > 0$  and a  $c_*$  depending on  $c_0, c_1$  such that the following hold for any  $\zeta_1 > 0$  sufficiently small. Assume  $c_0|\langle A \rangle| \geq \langle \dot{A}^2 \rangle^{1/2}$ ,  $|\langle A \rangle| \geq c_0$ ,  $\|A\| \lesssim 1$ . Pick indices  $j_1, j_2$  with  $N^{1-\zeta_1} \leq |j_1 - j_2| \leq c_*N$  and choose  $x_1, x_2$  with  $|x_1 - x_2| \lesssim |j_1 - j_2|/N$  such that  $\rho^{x_r}(\gamma_{j_r}^{x_r}) \geq c_1$ . Let  $\{\mathbf{u}_i^{x_r}\}_{i \in [N]}$ , for  $r \in [2]$ , be the orthonormal eigenbasis of the matrices  $H + x_r A$ . Then there exists  $\omega_E > 0$  such that*

$$|\langle \mathbf{u}_{j_1}^{x_1}, \mathbf{u}_{j_2}^{x_2} \rangle| \leq N^{-\omega_E} \quad (4.46)$$

with very high probability.



Then, using (4.46), instead of (4.27), as an input we readily conclude the analogous versions of Lemmas 4.2–4.3. Finally, by Lemmas 4.2–4.3 we conclude the proof of (4.45) proceeding exactly as in (4.28)–(4.33).

4.3.2. *Case 1).* In this case we consider the following Ornstein-Uhlenbeck (OU) flow instead of (4.3):

$$dH_t = -\frac{1}{2}(H_t - \mathbf{E} H_0) dt + \frac{\Sigma^{1/2}[dB_t]}{\sqrt{N}}, \quad H_0 = H, \quad \Sigma[\cdot] := \frac{\beta}{2} \mathbf{E} W \text{Tr}[W \cdot] \quad (4.47)$$

Here  $W = \sqrt{N}(H - \mathbf{E} H)$  and note that the OU flow is chosen to keep the expectation and the covariance structure of  $H_t$  invariant under the time evolution. As usual, the parameter  $\beta = 1$  in the real case and  $\beta = 2$  in the complex case. Here  $\Sigma^{1/2}$  denotes the square root of the positive operator  $\Sigma$  acting on  $N \times N$  matrices equipped with the usual Hilbert-Schmidt scalar product structure.

Then proceeding as in (4.1)–(4.5), after replacing (4.3) with (4.47), we find that to conclude the proof of this proposition it is enough to prove

$$\text{Cov} \left( P(N\delta\lambda_{j_1}^{x_1}(t_1)), Q(N\delta\lambda_{j_2}^{x_2}(t_1)) \right) = \mathcal{O} \left( N^{-\zeta_2} \|P\|_{C^1} \|Q\|_{C^1} \right) \quad (4.48)$$

with  $t_1 = N^{-1+\omega_1}$ , for some small  $\omega_1 > 0$ . Here  $\lambda(t)$  are the eigenvalues of  $\widehat{H}_t$ , which is the solution of (4.1) with initial condition  $\widehat{H}_0 = H_{t_1}^\#$ , where  $H_{t_1}^\#$  is from in (4.4) with  $H_{t_1}$  coming from (4.47).

Note that for  $A = I$  the gaps  $\lambda_{i+1}^x - \lambda_i^x = \lambda_{i+1} - \lambda_i$  do not depend on  $x$ . In particular (4.48) simplifies since

$$\delta\lambda_{j_r}^{x_r} = \lambda_{j_r+1} - \lambda_{j_r}, \quad (4.49)$$

where we recall that  $\{\lambda_i\}_{i \in [N]}$  are the eigenvalues of  $H$ ,  $\rho$  is its limiting density of states, and  $\{\gamma_i\}_{i \in [N]}$  are the corresponding quantiles.

By (4.49), the proof of (4.48) is a much simpler version of the proof of Propositions 4.1 presented in Section 4.2 for general  $A$ 's. More precisely, since for  $A = I$  the gaps are independent of  $x$ , it is enough to consider the DBM for the evolution of the eigenvalues of  $H$  instead of  $H^x$ :

$$d\lambda_i(t) = \sqrt{\frac{2}{N}} db_i(t) + \frac{1}{N} \sum_{j \neq i} \frac{1}{\lambda_i(t) - \lambda_j(t)} dt, \quad (4.50)$$

with  $\{b_i\}_{i \in [N]}$  a family of standard i.i.d. real Brownian motion (we wrote up the real symmetric case for simplicity). The fact that

$$d[b_i(t), b_j(t)] = \delta_{ij} dt, \quad (4.51)$$

follows by the orthogonality of the eigenvectors of  $H$ . Note that (4.50) does not depend on  $x$ , unlike (4.7) in Section 4. In particular by (4.51) it follows that  $C_{ij}(t) \equiv I$  in (4.16); indeed Proposition 4.4 is trivially satisfied by orthogonality since  $j_1$  and  $j_2$  are sufficiently away from each other by assumption. Additionally, it is not necessary to define the comparison processes  $\tilde{\lambda}, \tilde{\mu}$  since the driving Brownian motions in (4.50) are completely independent among each other, hence the processes  $\lambda(t)$  with indices close to  $j_r$  and  $\mu^{(r)}(t)$  can be compared directly (see e.g. [28, Section 3]).

## 5. BOUND ON THE EIGENVECTOR OVERLAP

The overlap in (4.27) and in (4.46) will be estimated by a local law involving the trace of the product of the resolvents of  $H^{x_1}$  and  $H^{x_2}$  for any fixed  $x_1, x_2$ . Individual resolvents can be approximated by the solution  $M$  of the MDE (2.5) but the deterministic approximation of products of resolvents are not simply products of  $M$ 's. Local laws are typically proven by deriving an approximate self-consistent equation and then effectively controlling its stability. In Proposition 5.1 we formulate a more accurate form of the overlap bounds (4.27)–(4.46) in terms of the stability factor of the self-consistent equation for the product of two resolvents. In the subsequent Lemma 5.2 we give an effective control on this stability factor. Proposition 5.1 will be proven in this section while the proof of Lemma 5.2 is postponed to Appendix A.

For notational convenience we introduce the commonly used notion of *stochastic domination*. For some family of non-negative random variables  $X = X(N) \geq 0$  and a deterministic control parameter  $\psi > 0$  we write  $X \prec \psi$  if for each  $\epsilon > 0, D > 0$  there exists some constant  $C$  such that

$$\mathbf{P}(X > N^\epsilon \psi) \leq CN^{-D}.$$

**Proposition 5.1.** Let  $\{\mathbf{u}_i^{x_r}\}_{i \in [N]}$ , for  $r \in [2]$ , be the orthonormal eigenbasis of the matrices  $H + x_r A$  and fix indices  $i_1, i_2$  in the bulk i.e. with  $\langle \Im M^{x_r}(\gamma_{i_r}^{x_r}) \rangle \gtrsim 1$ . Then it holds that<sup>9</sup>

$$|\langle \mathbf{u}_{i_1}^{x_1}, \mathbf{u}_{i_2}^{x_2} \rangle|^2 \prec N^{-1/15} |1 - \langle M^{x_1}(\gamma_{i_1}^{x_1}) M^{x_2}(\gamma_{i_2}^{x_2})^{(*)} \rangle|^{-16/15}. \quad (5.1)$$

**Lemma 5.2.** For any  $c_1 > 0$  there is a  $c^*$  such that for any  $x_1, x_2, E_1, E_2$  such that  $|x_1 - x_2| + |E_1 - E_2| \leq c^*$  and  $\rho^{x_r}(E_r) \geq c_1$ ,  $r = 1, 2$ , it holds

$$\begin{aligned} |1 - \langle M^{x_1}(E_1) M^{x_2}(E_2)^{(*)} \rangle| &\gtrsim |E_1 - x_1 \langle A \rangle - E_2 + x_2 \langle A \rangle|^2 + |x_1 - x_2|^2 \langle \hat{A}^2 \rangle \\ &+ \mathcal{O}(|x_1 - x_2|^3 + |E_1 - E_2|^3). \end{aligned} \quad (5.2)$$

For  $z_1, z_2 \in \mathbf{C} \setminus \mathbf{R}$  we abbreviate

$$G_i := (H^{x_i} - z_i)^{-1}, \quad M_i := M^{x_i}(z_i), \quad M_{ij} := \frac{M_i M_j}{1 - \langle M_j M_j \rangle} \quad (5.3)$$

and will prove the following  $G_1 G_2$  local law.

**Proposition 5.3.** For  $\delta := |1 - \langle M_1 M_2 \rangle|$  and any  $z_1, z_2 \in \mathbf{C}$  with  $|\Im z_1| = |\Im z_2| = \eta \gg N^{-1}$  in the bulk, i.e.  $\langle \Im M^{x_i}(z_i) \rangle \gtrsim 1$ , we have

$$|\langle G_1 G_2 A - M_{12} A \rangle| \prec \frac{\|A\|}{\delta N \eta^2} \left( \eta^{1/12} + \frac{1}{\sqrt{N} \eta} + \left( \frac{\eta}{\delta} \right)^{1/4} + \frac{1}{(\delta N \eta)^{1/3}} \right) \quad (5.4)$$

uniformly in deterministic matrices  $A$ .

*Proof of Proposition 5.1.* By spectral decomposition we obtain for  $z_r = E_r + i\eta$  with  $E_r := \gamma_{i_r}^{x_r}$  and setting  $\eta := (N\delta)^{-4/5}$

$$\begin{aligned} &\sum_{|\lambda_{j_1}^{x_1} - E_1| \lesssim \eta} \sum_{|\lambda_{j_2}^{x_2} - E_2| \lesssim \eta} |\langle \mathbf{u}_{j_1}^{x_1}, \mathbf{u}_{j_2}^{x_2} \rangle|^2 \\ &\lesssim \sum_{j_1, j_2} |\langle \mathbf{u}_{j_1}^{x_1}, \mathbf{u}_{j_2}^{x_2} \rangle|^2 \frac{\eta^4}{[(\lambda_{j_1}^{x_1} - E_1)^2 + \eta^2][(\lambda_{j_2}^{x_2} - E_2)^2 + \eta^2]} \\ &= \eta^2 \operatorname{Tr} \Im G_1 \Im G_2 \prec N^{-1/16} \delta^{-16/15} \end{aligned} \quad (5.5)$$

from Proposition 5.3. By rigidity (4.43) the sums in the l.h.s. of (5.5) contain the term  $|\langle \mathbf{u}_{i_1}^{x_1}, \mathbf{u}_{i_2}^{x_2} \rangle|^2$  as long as  $\eta \geq N^{-1+\xi}$ . This relation clearly holds with our choice since  $\delta \lesssim 1$ , concluding the proof.  $\square$

*Proof of Proposition 5.3.* This proof is an adaptation of a similar argument from [u], so here we only give a short explanation. From (2.5) obtain

$$\begin{aligned} (1 - M_1 M_2 \langle \cdot \rangle) [G_1 G_2 - M_{12}] &= \Delta := -M_1 \underline{W} G_1 G_2 + M_1 (G_2 - M_2) \\ &+ M_1 \langle G_1 G_2 \rangle (G_2 - M_2) + M_1 \langle G_1 - M_1 \rangle G_1 G_2, \end{aligned} \quad (5.6)$$

where

$$\underline{W} G_1 G_2 := W G_1 G_2 + \langle G_1 \rangle G_1 G_2 + \langle G_1 G_2 \rangle G_2.$$

Thus we have

$$\langle G_1 G_2 A - M_{12} A \rangle = \langle \Delta A \rangle + \frac{\langle M_1 M_2 A \rangle \langle \Delta \rangle}{1 - \langle M_1 M_2 \rangle}. \quad (5.7)$$

Recall that it was proven in [u, Proposition 5.3] that if  $|\langle G_1 G_2 A \rangle| \prec \|A\| \theta$  for some constant  $\theta \leq \eta^{-1}$  uniformly in  $A$ , then also

$$|\langle \underline{W} G_1 G_2 A \rangle| \prec \frac{1}{N \eta^2} \left( (\theta \eta)^{1/4} + \frac{1}{(N \eta)^{1/2}} + \eta^{1/12} \right), \quad (5.8)$$

again uniformly in  $A$ . Strictly speaking [u, Proposition 5.3] was stated in the context of Hermitized i.i.d. random matrices. However, a simpler version of the same proof clearly applies to deformed Wigner matrices. The main simplification compared to [u] is that due to the constant variance profile of Wigner matrices summations as the one in [u, Eq. (5.28a)]

<sup>9</sup>Star in bracket  $M^{(*)}$  indicates that the statement holds both  $M$  and its adjoint  $M^*$

can be directly performed, without introducing the block matrices  $E_1, E_2$ . The remainder of the proof apart from the simplified *resummation step* verbatim applies to the present case. Using (5.8) in (5.7) and  $\theta \leq \eta^{-1}, \eta \lesssim 1$  it follows that

$$\begin{aligned} |\langle G_1 G_2 A - M_{12} A \rangle| &< \frac{1}{\delta} \left( \frac{1}{N\eta} + \frac{\theta}{N\eta} + \frac{1}{N\eta^2} \left( \eta^{1/12} + \frac{1}{(N\eta)^{1/2}} + (\theta\eta)^{1/4} \right) \right) \\ &\lesssim \frac{1}{\delta N\eta^2} \left( \eta^{1/12} + \frac{1}{(N\eta)^{1/2}} + (\theta\eta)^{1/4} \right) \end{aligned} \quad (5.9)$$

and therefore

$$|\langle G_1 G_2 A \rangle| < \theta' := \frac{1}{\delta} + \frac{1}{\delta N\eta^2} \left( \eta^{1/12} + \frac{1}{(N\eta)^{1/2}} + (\theta\eta)^{1/4} \right). \quad (5.10)$$

By iterating (5.10) it follows that after finitely many steps starting from  $\theta_0 = 1/\eta$  we obtain

$$|\langle G_1 G_2 A \rangle| < \theta_*, \quad (5.11)$$

where  $\theta_*$  is the unique positive solution to the equation

$$\theta_* = \frac{1}{\delta} \left( 1 + \frac{1}{N\eta^2} \left( \eta^{1/12} + \frac{1}{(N\eta)^{1/2}} \right) \right) + \frac{\theta_*^{1/4}}{\delta N\eta^{7/4}}. \quad (5.12)$$

Asymptotically we have

$$\theta_* \sim \frac{1}{\delta} \left( 1 + \frac{1}{N\eta^2} \left( \eta^{1/12} + \frac{1}{(N\eta)^{1/2}} + \frac{1}{(\delta N\eta)^{1/3}} \right) \right) \quad (5.13)$$

and using (5.9) once more with  $\theta_*$  concludes the proof.  $\square$

**5.1. Proof of Propositions 4.4 and 4.9.** Both proofs rely on Proposition 5.1 and proving that the lower bound on the stability factor given in Lemma 5.2 with  $E_r = \gamma_{i_r}^{x_r}, r = 1, 2$ , is bounded from below by  $N^{-\epsilon}$  with some small  $\epsilon$ . This will be done separately for the two propositions.

For Proposition 4.4 we use that  $|E_1 - E_2| \lesssim |x_1 - x_2| \leq c^*$  with a small  $c^*$  and that  $\langle \mathring{A}^2 \rangle \gtrsim 1$ , hence

$$|1 - \langle M^{x_1}(\gamma_{i_1}^{x_1}) M^{x_2}(\gamma_{i_2}^{x_2})^* \rangle| \gtrsim |x_1 - x_2|^2 \langle \mathring{A}^2 \rangle \gtrsim N^{-2\zeta_1}.$$

The relation  $|E_1 - E_2| \lesssim |x_1 - x_2|$  follows from

$$|E_1 - E_2| = |\gamma_{i_1}^{x_1} - \gamma_{i_2}^{x_2}| \leq |\gamma_{i_1}^{x_1} - \gamma_{i_1}^{x_2}| + |\gamma_{i_1}^{x_2} - \gamma_{i_2}^{x_2}| \lesssim |x_1 - x_2| + |i_1 - i_2|/N$$

and the fact that  $|i_1 - i_2|/N \lesssim |x_1 - x_2|$  from the conditions of Propositions 4.4. The estimate on  $|\gamma_{i_1}^{x_1} - \gamma_{i_1}^{x_2}|$  comes from Lemma 3.4.

For Proposition 4.9 we have

$$\begin{aligned} |E_1 - x_1 \langle A \rangle - E_2 + x_2 \langle A \rangle| &\geq |\gamma_{i_1}^{x_1} - \gamma_{i_2}^{x_1}| - |\gamma_{i_2}^{x_1} - \gamma_{i_2}^{x_2} - (x_1 - x_2) \langle A \rangle| \\ &\geq c_1 |i_1 - i_2|/N - C |x_1 - x_2| \left( \langle \mathring{A}^2 \rangle^{1/2} + |x_1 - x_2| \right). \end{aligned} \quad (5.14)$$

In estimating the first term we used that  $\gamma_{i_1}^{x_1}, \gamma_{i_2}^{x_1}$  are in the bulk, while we used (3.12) for the second term. Notice that

$$C |x_1 - x_2| \left( \langle \mathring{A}^2 \rangle^{1/2} + |x_1 - x_2| \right) \leq c_0 \|A\| N^{-\zeta_1}$$

by the bound  $\langle \mathring{A}^2 \rangle^{1/2} \leq c_0 \langle A \rangle \leq c_0 \|A\|$ . Choosing  $c_0$  sufficiently small, depending on  $c_1$ , and recalling that  $|i_1 - i_2| \geq N^{1-\zeta_1}$ , we can achieve that

$$C |x_1 - x_2| \left( \langle \mathring{A}^2 \rangle^{1/2} + |x_1 - x_2| \right) \leq \frac{1}{2} c_1 |i_1 - i_2|/N$$

in particular

$$|E_1 - x_1 \langle A \rangle - E_2 + x_2 \langle A \rangle| \geq \frac{1}{2} c_1 |i_1 - i_2|/N \gtrsim N^{-\zeta_1} \quad (5.15)$$

from (5.14). This shows the required lower bound for the leading (first) term in (5.2). The second term is non-negative. The first error term is negligible,  $|x_1 - x_2|^3 \leq N^{-3\zeta_1}$ . For the second error term we have

$$|E_1 - E_2| \leq |\gamma_{i_1}^{x_2} - \gamma_{i_2}^{x_2}| + |\gamma_{i_1}^{x_1} - \gamma_{i_1}^{x_2}| \lesssim |i_1 - i_2|/N + |x_1 - x_2| \lesssim |i_1 - i_2|/N$$

using the upper bound on the density  $\rho^{x_2}$  in the first term and (3.12) in the second term. In the last step we used  $|x_1 - x_2| \lesssim |i_1 - i_2|/N$  from the conditions of Proposition 4.9. This shows that the error term  $|E_1 - E_2|^3 \lesssim$

$(|i_1 - i_2|/N)^3$  is negligible compared with the main term A.4 of order at least  $(|i_1 - i_2|/N)^2$  since we also assumed  $|i_1 - i_2|/N \leq c_*$  which is small.

This proves that

$$|1 - \langle M^{x_1}(\gamma_{i_1}^{x_1}) M^{x_2}(\gamma_{i_2}^{x_2})^{(*)} \rangle| \gtrsim |E_1 - x_1 \langle A \rangle - E_2 + x_2 \langle A \rangle|^2 \gtrsim N^{-2\zeta_1}.$$

in the setup of Proposition 4.9 as well.

## 6. MULTI-GAP QUENCHED UNIVERSALITY

The following results are the multi-gap versions of Theorems 2.4 and 2.7. The gaps will be tested by functions of  $k$  variables, so we define the set

$$\mathcal{F}_k = \mathcal{F}_{k,L,B} := \left\{ F: \mathbf{R}^k \rightarrow \mathbf{R} \mid \text{supp}(F) \subset [0, L]^k, \|F\|_{C^5} \leq B \right\} \quad (6.1)$$

of  $k$ -times differentiable and compactly supported test functions  $F$  with some large constants  $L, B > 0$ . In the following we will often use the notation  $\mathbf{i} := (i_1, \dots, i_k)$  for a  $k$ -tuple of integer indices  $i_1, \dots, i_k$ . The gap distribution for  $H^x$  will be compared with that of the Gaussian Wigner matrices, we therefore let  $\{\mu_i\}_{i \in [N]}$  denote the eigenvalues of a GOE/GUE matrix corresponding to the symmetry class of  $H$ .

**Theorem 6.1** (Quenched universality via eigenbasis rotation mechanism). *Under the conditions of Theorem 2.4 for any  $c_1$ -bulk-index  $i_0$  we have the following multi-gap version of Wigner-Dyson universality. There exist  $\epsilon = \epsilon(a_0, c_0, c_1) > 0$  and an event  $\Omega_{i_0, A}$  with  $\mathbf{P}_H(\Omega_{i_0, A}^c) \leq N^{-\epsilon}$  such that for all  $H \in \Omega_{i_0, A}$  the matrix  $H^x = H + xA$  satisfies*

$$\max_{\|\mathbf{i}\|_\infty \leq K} \sup_{F \in \mathcal{F}_k} \left| \mathbf{E}_x F \left( \left( N \rho^x(\gamma_{i_0}^x) \delta \lambda_{i_0 + i_j}^x \right)_{j \in [k]} \right) - \mathbf{E}_\mu F \left( \left( N \rho_{\text{sc}}(0) \delta \mu_{N/2 + i_j} \right)_{j \in [k]} \right) \right| \leq CN^{-c}, \quad (6.2)$$

for  $K := N^\zeta$  and some  $\zeta = \zeta(a_0, c_0, c_1) > 0$ , and  $c = c(k) > 0$ . The constant  $C$  in (6.2) may depend on  $k, L, B, a_0, c_0, c_1$  and all constants in Assumptions 2.2 and 2.3 at most polynomially, but it is independent of  $N$ .

**Theorem 6.2** (Quenched universality via spectral sampling mechanism). *Under the conditions of Theorem 2.7 for any  $c_1$ -bulk-energy  $E$  we have the following multi-gap version of Wigner-Dyson universality. There exists  $\epsilon = \epsilon(a_0, c_0, c_1) > 0$  and an event  $\Omega_{E, A}$  with  $\mathbf{P}(\Omega_{E, A}^c) \leq N^{-\epsilon}$  such that for all  $H \in \Omega_{E, A}$  the matrix  $H^x$  satisfies*

$$\max_{\|\mathbf{i}\|_\infty \leq K} \sup_{F \in \mathcal{F}_k} \left| \mathbf{E}_x F \left( \left( N \rho^x(E) \delta \lambda_{i_0(x, E) + i_j}^x \right)_{j \in [k]} \right) - \mathbf{E}_\mu F \left( \left( N \rho_{\text{sc}}(0) \delta \mu_{N/2 + i_j} \right)_{j \in [k]} \right) \right| \leq CN^{-c}, \quad (6.3)$$

where  $K := N^\zeta$ , and some  $\zeta = \zeta(a_0, c_0, c_1) > 0$ ,  $c = c(k) > 0$ . The constant  $C$  in (6.3) may depend on  $k, L, B, a_0, c_0, c_1$  and all constants in Assumptions 2.2 and 2.3 at most polynomially, but it is independent of  $N$ .

First, to handle the supremum over the uncountable family  $\mathcal{F}_{k,L,B}$  of test functions  $F$  we reduce the problem to a finite set of test functions so that the union bound can be taken. Notice that for sufficiently smooth test functions  $F$ , which are compactly supported on some box  $[0, L]^k$  of size  $L$ , we can expand  $F$  in partial Fourier series as (see e.g. [38, Remark 3] and [12, Eq. (30)])

$$F(x_1, \dots, x_k) = \sum_{|n_1|, \dots, |n_k| \in [N^{\zeta_*}]} C_F(n_1, \dots, n_k) \prod_{j=1}^k e^{in_j x_j / L} \varphi(x_j) + \mathcal{O}(N^{-c(\zeta_*)}), \quad (6.4)$$

$$\sum_{n_1, \dots, n_k} |C_F(n_1, \dots, n_k)| \lesssim 1,$$

with integer  $n_1, \dots, n_k$ , where  $\varphi: \mathbf{R} \rightarrow \mathbf{R}$  is a smooth cut-off function such that it is equal to one on  $[0, L]$  and it is equal to zero on  $[-L/2, 3L/2]^c$ . Here  $\zeta_* > 0$  is a small fixed constant that will be chosen later. Introduce the notation

$$f_n(x) := e^{inx/L} \varphi(x), \quad n \in \mathbf{Z}. \quad (6.5)$$

*Proof of Theorem 6.1.* By (6.4) we get that

$$\begin{aligned} & \sup_{\|i\|_\infty \leq N^\zeta} \sup_F \left| \mathbf{E}_x F \left( (N\rho^x(\gamma_{i_0}^x) \delta \lambda_{i_0+i_j}^x)_{j \in [k]} \right) - \mathbf{E}_\mu F \left( (N\rho_{\text{sc}}(0) \delta \mu_{N/2+i_j})_{j \in [k]} \right) \right| \\ & \lesssim \sup_{\substack{\|i\|_\infty \leq N^\zeta, \\ \|\mathbf{n}\|_\infty \leq N^{\zeta_*}}} \left| \mathbf{E}_x \prod_{j \in [k]} f_{n_j} \left( N\rho^x(\gamma_{i_0}^x) \delta \lambda_{i_0+i_j}^x \right) - \mathbf{E}_\mu \prod_{j \in [k]} f_{n_j} \left( N\rho_{\text{sc}}(0) \delta \mu_{N/2+i_j} \right) \right| + \mathcal{O} \left( N^{-c(\zeta_*)} \right), \end{aligned} \quad (6.6)$$

with  $f_{n_j}$  defined in (6.5) and  $\mathbf{n} := (n_1, \dots, n_k)$ .

Proceeding exactly as in the proof of Theorem 2.4 in Section 3.1, and using the fact that (3.3) holds for test functions  $P_1, P_2$  of  $k$  variables (see Remark 3.2), we conclude that for any fixed  $i_1, \dots, i_k$  and  $n_1, \dots, n_k$  there exists a probability event  $\Omega_{i_0, \mathbf{i}, \mathbf{n}}$ , with  $\mathbf{P}(\Omega_{i_0, \mathbf{i}, \mathbf{n}}^c) \leq N^{-\kappa}$ , on which

$$\left| \mathbf{E}_x \prod_{j \in [k]} f_{n_j} \left( N\rho^x(\gamma_{i_0}^x) \delta \lambda_{i_0+i_j}^x \right) - \mathbf{E}_\mu \prod_{j \in [k]} f_{n_j} \left( N\rho_{\text{sc}}(0) \delta \mu_{N/2+i_j} \right) \right| \lesssim N^{-\alpha} \prod_{j \in [k]} \|f_{n_j}\|_{C^5}. \quad (6.7)$$

Then choosing  $\zeta, \zeta_* \leq \kappa(10k)^{-1}$  we define the event

$$\Omega_{i_0} := \bigcap_{\|i\|_\infty \leq N^\zeta} \bigcap_{\|\mathbf{n}\|_\infty \leq N^{\zeta_*}} \Omega_{i_0, \mathbf{i}, \mathbf{n}}, \quad \mathbf{P}_H(\Omega_{i_0}^c) \lesssim N^{-\kappa} N^{k(\zeta+\zeta_*)} \leq N^{-\kappa/2}. \quad (6.8)$$

Finally, by (6.7)–(6.8), for all  $H \in \Omega_{i_0}$ , choosing  $\zeta_* \leq \alpha(10k)^{-1}$ , the claim (6.2) follows with exponent  $c = \min\{\alpha - 5k\zeta_*, c(\zeta_*)\}$  using that  $\|f_{n_j}\|_{C^5} \leq N^{5\zeta_*}$ , for any  $j \in [k]$ , and where  $c(\zeta_*)$  is from (6.4).  $\square$

*Proof of Theorem 6.2.* Given (6.6), the proof of Theorem 6.2, following Section 3.2 instead of Section 3.1 and using that Proposition 3.3 holds for  $P_1, P_2$  of  $k$  variables (see Remark 3.2), is completely analogous and so omitted.  $\square$

#### APPENDIX A. BOUND FOR THE STABILITY OPERATOR

*Proof of Lemma 5.2.* Note that

$$|1 - \langle M_1 M_2^* \rangle| \geq \Re[1 - \langle M_1 M_2^* \rangle] = \frac{1}{2} \langle (M_1 - M_2)(M_1 - M_2)^* \rangle + \mathcal{O}(\eta),$$

where we used that  $\langle M_i M_i^* \rangle = 1 + \mathcal{O}(\eta)$ , which follows taking the imaginary in the MDE (2.5). The using Taylor expansion in the  $x_2$  and the  $E_2$  variables we get that

$$\begin{aligned} & \langle (M_1 - M_2)(M_1 - M_2)^* \rangle \\ & = \langle (\partial_{x_1} M_1(x_2 - x_1) + \partial_{E_1} M_1(E_2 - E_1)) (\partial_{x_1} M_1(x_2 - x_1) + \partial_{E_1} M_1(E_2 - E_1))^* \rangle \\ & \quad + \mathcal{O}(|x_1 - x_2|^3 + |E_1 - E_2|^3). \end{aligned} \quad (\text{A.1})$$

To estimate the error term in (A.1) we used the following bounds for  $E = \Re z$  in the bulk of the spectrum for any  $x \in [x_1, x_2]$  and  $E \in [E_1, E_2]$ , a condition that is guaranteed by  $|x_1 - x_2| + |E_1 - E_2| \leq c^*$  is small. By [2, Corollary 5.3, Lemma 5.7] we have

$$\left\| \partial_x^\alpha \partial_E^\beta M^x(E + i\eta) \right\| \leq C_{\alpha, \beta}, \quad \left\| \frac{1}{1 - M^x(z) \langle \cdot \rangle M^x(z)} \right\|_{\|\cdot\| \rightarrow \|\cdot\|} \leq \frac{C}{\rho^x(z) [\rho^x(z) + |\sigma^x(z)|]}, \quad (\text{A.2})$$

for any  $\alpha, \beta \in \mathbf{N}$ , for any fixed  $x$ , where  $|\sigma^x(z)| \geq c$  unless  $\rho^x(z)$  has a near-cusp singularity and  $E = \Re z$  is close to this cusp point. Recall that the norm  $\|\cdot\|$  denotes the standard euclidean matrix norm on  $N \times N$  matrices. Here  $1 - M^x(z) \langle \cdot \rangle M^x(z)$  is a linear operator acting on such matrices  $R$  as  $(1 - M^x(z) \langle \cdot \rangle M^x(z))[R] = R - M^x(z) \langle R \rangle M^x(z)$ . Finally, the second formula in (A.2) involves the norm induced by the euclidean matrix norm.

Then differentiating the MDE in  $x$  and  $E$  we find that

$$\partial_{x_1} M_1 = -\frac{1}{1 - M_1 \langle \cdot \rangle M_1} [M_1 A M_1], \quad \partial_{E_1} M_1 = \frac{1}{1 - M_1 \langle \cdot \rangle M_1} [M_1 M_1].$$

Hence, by

$$\left( \frac{1}{1 - M_1 \langle \cdot \rangle M_1} \right) \left( \frac{1}{1 - M_1 \langle \cdot \rangle M_1} \right)^* \geq c, \quad M_1 M_1^* \geq c,$$

we conclude

$$\begin{aligned} & \langle (\partial_{x_1} M_1(x_2 - x_1) + \partial_{E_1} M_1(E_2 - E_1)) (\partial_{x_1} M_1(x_2 - x_1) + \partial_{E_1} M_1(E_2 - E_1))^* \rangle \\ &= \left\langle \left( \frac{1}{1 - M_1 \langle \cdot \rangle M_1} M_1 [ - (x_1 - x_2) A + (E_1 - E_2) ] M_1 \right)^* \frac{1}{1 - M_1 \langle \cdot \rangle M_1} M_1 [ - (x_1 - x_2) A + (E_1 - E_2) ] M_1 \right\rangle \\ &\gtrsim \langle [(E_1 - E_2) - (x_1 - x_2) A]^2 \rangle = |(E_1 - x_1 \langle A \rangle - E_2 + x_2 \langle A \rangle)|^2 + |x_1 - x_2|^2 \langle \dot{A}^2 \rangle, \end{aligned}$$

where in the last equality we wrote  $A = \langle A \rangle + \dot{A}$ . This concludes the proof of (5.2) in case when the adjoint is present. The estimate of  $|1 - \langle M_1 M_2 \rangle|$  is much easier, it follows directly from (A.2).  $\square$

*Proof of Lemma 3.4.* To make the presentation clearer we just consider the case  $x_1 = x$  and  $x_2 = 0$ , the general case is analogous and so omitted.

For any fixed real parameters  $x, y$  consider the MDE

$$M^{-1} = z + B + x \langle A \rangle + y \dot{A} + \langle M \rangle, \quad \Im M \Im z > 0. \quad (\text{A.3})$$

Note that for  $y = 0$  (A.3) is the MDE for  $H$  and for  $y = x$  (A.3) is the one for  $H^x = H + xA$ . We denote the unique solution of (A.3) by  $M^{x,y} = M^{x,y}(z)$ , the associated scDos by  $\rho^{x,y}$  and the corresponding quantiles by  $\gamma_i^{x,y}$ . We will use that

$$\gamma_i^{x_1} - \gamma_i^{x_2} = \gamma_i^x - \gamma_i^0 = \int_0^x \partial_s \gamma_i^{s,y} ds = \int_0^x [(\partial_x \gamma_i^{x,s})|_{x=s} + (\partial_y \gamma_i^{s,y})|_{y=s}] ds. \quad (\text{A.4})$$

For the first term we use that  $\partial_x \gamma_i^{x,s} = \langle A \rangle$ , giving the leading term  $x \langle A \rangle$  in Lemma 3.4. To estimate  $\partial_y \gamma_i^{s,y}$ , we differentiate the defining equation of the quantiles

$$\int_{-\infty}^{\gamma_i^{s,y}} \langle \Im M^{s,y}(E) \rangle dE = \frac{i}{N}$$

with respect to  $y$ . We obtain

$$\partial_y \gamma_i^{s,y} \langle \Im M^{s,y}(\gamma_i^{s,y}) \rangle + \int_{-\infty}^{\gamma_i^{s,y}} \partial_y \langle \Im M^{s,y}(E) \rangle dE = 0$$

for any  $s, y \in [0, x]$ . Then, using that in the bulk  $|\langle \Im M^{s,y}(\gamma_i^{s,y}) \rangle| \geq c$ , we conclude

$$|\partial_y \gamma_i^{s,y}| \lesssim \int_{-C}^{\gamma_i^{s,y}} \left\langle \frac{1}{1 - M^{s,y}(E) \langle \cdot \rangle M^{s,y}(E)} M^{s,y}(E) \dot{A} M^{s,y}(E) \right\rangle dE \lesssim \langle \dot{A}^2 \rangle^{1/2}, \quad (\text{A.5})$$

where we used Schwarz inequality and the bounds in (A.2). The important fact about the second bound in (A.2) is that it is integrable in  $E$  since it has a  $|E - E_0|^{-1/2}$  singularity near an edge point  $E_0$  and a  $|E - E_0|^{-2/3}$  singularity near a cusp point  $E_0$ . Here we also used that  $|x| = |x_1 - x_2| \leq c^*$  is sufficiently small so that  $\gamma_i^{s,y}$  is in the bulk not only for  $s = y = x$ , but for all  $s, y \in [0, x]$ . From (A.4) and (A.5) we readily conclude (3.12).  $\square$

## APPENDIX B. NUMERICS

Here we present numerical evidence quantifying the speed of convergence of the single gap distribution to its theoretical limit for the monoparametric ensemble, c.f. Figure 3. This numerics was inspired by the observation made in [23]<sup>10</sup> on the slow convergence of the spectral form factor.

## REFERENCES

- <sup>1</sup>O. H. Ajanki, L. Erdős, and T. Krüger, *Stability of the matrix Dyson equation and random matrices with correlations*, *Probab. Theory Related Fields* **173**, 293–373 (2019), [MR3916109](#).
- <sup>2</sup>J. Alt, L. Erdős, and T. Krüger, *The Dyson equation with linear self-energy: spectral bands, edges and cusps*, *Doc. Math.* **25**, 1421–1539 (2020), [MR4164728](#).
- <sup>3</sup>F. Bekerman, A. Figalli, and A. Guionnet, *Transport maps for  $\beta$ -matrix models and universality*, *Comm. Math. Phys.* **338**, 589–619 (2015), [MR3351052](#).

<sup>10</sup>We thank Stephen Shenker for communicating preliminary numerical results supporting this observation in June 2021.

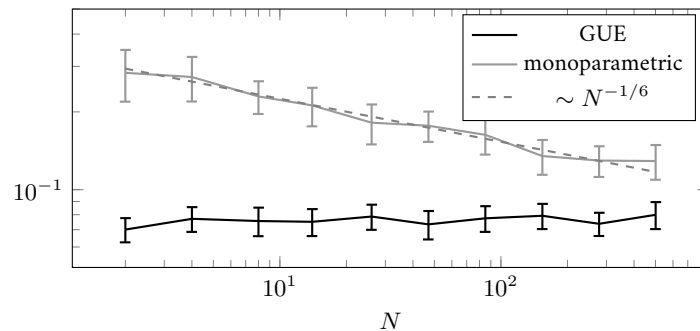


FIGURE 3. The figure shows the Kolmogorov-Smirnov distance  $D(F, F') := \sup_s |F(s) - F'(s)|$  of the empirical cumulative distribution function (CDF) of the (rescaled) eigenvalue gap  $\lambda_{N/2+1} - \lambda_{N/2}$  to the CDF  $F_2$  corresponding to  $p_2$  for various values of  $N$  for both GUE and the monoparametric ensemble. The empirical CDF for the GUE has been generated by sampling 100 GUE matrices  $H$ . For the monoparametric ensemble  $H^x = H + xA$  typical GUE random matrices  $H$ ,  $A$  have been fixed and 100 Gaussian random variables  $x$  have been sampled. The error bars represent the standard deviation of the obtained Kolmogorov-Smirnov distance for 50 independent repetitions. In accordance with Figure 2 we find that the gap distribution for GUE matches its theoretical limit very well for any value of  $N$ , while for the monoparametric ensemble the KS-distance seems to decay only slowly with  $N$ .

- <sup>4</sup>G. Ben Arous and P. Bourgade, *Extreme gaps between eigenvalues of random matrices*, *Ann. Probab.* **41**, 2648–2681 (2013), [MR3112927](#).
- <sup>5</sup>P. Bourgade, *Extreme gaps between eigenvalues of Wigner matrices*, preprint (2018), [arXiv:1812.10376](#).
- <sup>6</sup>P. Bourgade, L. Erdős, and H.-T. Yau, *Edge universality of beta ensembles*, *Comm. Math. Phys.* **332**, 261–353 (2014), [MR3253704](#).
- <sup>7</sup>P. Bourgade, L. Erdős, and H.-T. Yau, *Universality of general  $\beta$ -ensembles*, *Duke Math. J.* **163**, 1127–1190 (2014), [MR3192527](#).
- <sup>8</sup>P. Bourgade, L. Erdős, H.-T. Yau, and J. Yin, *Fixed energy universality for generalized Wigner matrices*, *Comm. Pure Appl. Math.* **69**, 1815–1881 (2016), [MR3541852](#).
- <sup>9</sup>Z. Che and B. Landon, *Local spectral statistics of the addition of random matrices*, *Probab. Theory Related Fields* **175**, 579–654 (2019), [MR4009717](#).
- <sup>10</sup>G. Cipolloni, L. Erdős, T. Krüger, and D. Schröder, *Cusp universality for random matrices, II: The real symmetric case*, *Pure Appl. Anal.* **1**, 615–707 (2019), [MR4026551](#).
- <sup>11</sup>G. Cipolloni, L. Erdős, and D. Schröder, *Central limit theorem for linear eigenvalue statistics of non-Hermitian random matrices*, preprint (2019), [arXiv:1912.04100](#).
- <sup>12</sup>G. Cipolloni, L. Erdős, and D. Schröder, *Edge universality for non-Hermitian random matrices*, *Probab. Theory Related Fields* **179**, 1–28 (2021), [MR4221653](#).
- <sup>13</sup>G. Cipolloni, L. Erdős, and D. Schröder, *Fluctuation around the circular law for random matrices with real entries*, *Electron. J. Probab.* **26**, Paper No. 24, 61 (2021), [MR4235475](#).
- <sup>14</sup>P. Deift and D. Gioev, *Universality at the edge of the spectrum for unitary, orthogonal, and symplectic ensembles of random matrices*, *Comm. Pure Appl. Math.* **60**, 867–910 (2007), [MR2306224](#).
- <sup>15</sup>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), [MR3068390](#).
- <sup>16</sup>L. Erdős, T. Krüger, and D. Schröder, *Cusp universality for random matrices I: local law and the complex Hermitian case*, *Comm. Math. Phys.* **378**, 1203–1278 (2020), [MR4134946](#).
- <sup>17</sup>L. Erdős, T. Krüger, and D. Schröder, *Random matrices with slow correlation decay*, *Forum Math. Sigma* **7**, e8, 89 (2019), [MR3941370](#).
- <sup>18</sup>L. Erdős, B. Schlein, and H.-T. Yau, *Universality of random matrices and local relaxation flow*, *Invent. Math.* **185**, 75–119 (2011), [MR2810797](#).

- <sup>19</sup>L. Erdős and H.-T. Yau, *A dynamical approach to random matrix theory*, Vol. 28, Courant Lecture Notes in Mathematics (Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2017), pp. ix+226, [MR3699468](#).
- <sup>20</sup>L. Erdős and H.-T. Yau, *Gap universality of generalized Wigner and  $\beta$ -ensembles*, *J. Eur. Math. Soc. (JEMS)* **17**, 1927–2036 (2015), [MR3372074](#).
- <sup>21</sup>L. Erdős, H.-T. Yau, and J. Yin, *Rigidity of eigenvalues of generalized Wigner matrices*, *Adv. Math.* **229**, 1435–1515 (2012), [MR2871147](#).
- <sup>22</sup>P. J. Forrester and N. S. Witte, *Exact Wigner surmise type evaluation of the spacing distribution in the bulk of the scaled random matrix ensembles*, *Lett. Math. Phys.* **53**, 195–200 (2000), [MR1808248](#).
- <sup>23</sup>H. Gharibyan, C. Pattison, S. Shenker, and K. Wells, *Work in preperation*, (2021).
- <sup>24</sup>J. W. Helton, R. Rashidi Far, and R. Speicher, *Operator-valued semicircular elements: solving a quadratic matrix equation with positivity constraints*, *Int. Math. Res. Not. IMRN, Art. ID rnm086*, 15 (2007), [MR2376207](#).
- <sup>25</sup>M. Jimbo, T. Miwa, Y. Mōri, and M. Sato, *Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent*, *Phys. D* **1**, 80–158 (1980), [MR573370](#).
- <sup>26</sup>K. Johansson, *Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices*, *Comm. Math. Phys.* **215**, 683–705 (2001), [MR1810949](#).
- <sup>27</sup>O. Kallenberg, *Foundations of modern probability*, Second Edition, Probability and its Applications (New York) (Springer-Verlag, New York, 2002), pp. xx+638, [MR1876169](#).
- <sup>28</sup>B. Landon, P. Sosoë, and H.-T. Yau, *Fixed energy universality of Dyson Brownian motion*, *Adv. Math.* **346**, 1137–1332 (2019), [MR3914908](#).
- <sup>29</sup>B. Landon and H.-T. Yau, *Convergence of local statistics of Dyson Brownian motion*, *Comm. Math. Phys.* **355**, 949–1000 (2017), [MR3687212](#).
- <sup>30</sup>B. Landon and H.-T. Yau, *Edge statistics of Dyson Brownian motion*, preprint (2017), [arXiv:1712.03881](#).
- <sup>31</sup>J. O. Lee and K. Schnelli, *Edge universality for deformed Wigner matrices*, *Rev. Math. Phys.* **27**, 1550018, 94 (2015), [MR3405746](#).
- <sup>32</sup>J. O. Lee, K. Schnelli, B. Stetler, and H.-T. Yau, *Bulk universality for deformed Wigner matrices*, *Ann. Probab.* **44**, 2349–2425 (2016), [MR3502606](#).
- <sup>33</sup>M. L. Mehta, *Random matrices and the statistical theory of energy levels* (Academic Press, New York-London, 1967), pp. x+259, [MR0220494](#).
- <sup>34</sup>L. Pastur and M. Shcherbina, *On the edge universality of the local eigenvalue statistics of matrix models*, *Mat. Fiz. Anal. Geom.* **10**, 335–365 (2003), [MR2012268](#).
- <sup>35</sup>M. Shcherbina, *Change of variables as a method to study general  $\beta$ -models: bulk universality*, *J. Math. Phys.* **55**, 043504, 23 (2014), [MR3390602](#).
- <sup>36</sup>A. Soshnikov, *Universality at the edge of the spectrum in Wigner random matrices*, *Comm. Math. Phys.* **207**, 697–733 (1999), [MR1727234](#).
- <sup>37</sup>T. Tao and V. Vu, *Random matrices: universality of local eigenvalue statistics*, *Acta Math.* **206**, 127–204 (2011), [MR2784665](#).
- <sup>38</sup>T. Tao and V. Vu, *Random matrices: universality of local spectral statistics of non-Hermitian matrices*, *Ann. Probab.* **43**, 782–874 (2015), [MR3306005](#).