The Annals of Probability
2017, Vol. 45, No. 6A, 3626–3663
DOI: 10.1214/16-AOP1145
© Institute of Mathematical Statistics, 2017

BULK EIGENVALUE STATISTICS FOR RANDOM REGULAR GRAPHS

BY ROLAND BAUERSCHMIDT*, JIAOYANG HUANG†, ANTTI KNOWLES‡,1 AND HORNG-TZER YAU†,2

University of Cambridge, Harvard University† and University of Geneva‡*

We consider the uniform random d-regular graph on N vertices, with $d \in [N^\alpha, N^{2/3-\alpha}]$ for arbitrary $\alpha > 0$. We prove that in the bulk of the spectrum the local eigenvalue correlation functions and the distribution of the gaps between consecutive eigenvalues coincide with those of the Gaussian orthogonal ensemble.

1. Introduction and results.

1.1. *Introduction*. The universality of local eigenvalue statistics is one of the central questions in random matrix theory. Random matrix statistics are believed to apply to very general complex systems, including the zeros of the Riemann ζ-function on the critical line. However, proofs of random matrix statistics have so far been limited mostly to matrix ensembles, with the notable exception [32]. There are two classes of matrix ensembles for which random matrix statistics have been established under very general conditions: invariant ensembles and ensembles with independent entries. For ensembles of random matrices that are invariant under the unitary or orthogonal group (invariant ensembles), much has been understood via the method of orthogonal polynomials (see, e.g., [6, 14, 38]) and, more recently, general results have been obtained by direct comparision of ensembles with different potentials [4, 9, 40]. For Wigner matrices and generalized Wigner matrices, whose entries are independent and typically nonzero, the universality problem has also essentially been solved completely [8, 18–20, 23–25, 29, 44]. For random sparse matrices with independent entries, significant progress has been made as well. In particular, for Erdős–Rényi graphs, in which each edge is chosen independently with probability p, random matrix statistics for both the bulk eigenvalues and the second largest eigenvalue of the adjacency matrix were established in [18] under the condition $pN \geq N^{2/3+\alpha}$ with any $\alpha > 0$. For the bulk eigenvalues, the lower bound on pN was recently extended to $pN \geq N^\alpha$ for any $\alpha > 0$ in [27], and

Received June 2015; revised August 2016.
[1]Supported in part by Swiss National Science Foundation Grant 144662.
[2]Supported in part by NSF Grant DMS-13-07444 and a Simons Investigator fellowship.
MSC2010 subject classifications. 05C80, 05C50, 60B20, 15B52.
Key words and phrases. Random regular graphs, spectral statistics, universality, GOE, switchings, Dyson Brownian motion.

GOE statistics for the eigenvalue gaps was also established. Finally, aside from the approaches discussed above, supersymmetry has been used to obtain results on the local eigenvalue statistics for some special classes of distributions with independent entries (see, e.g., [41]). In addition, local random matrix statistics have been established by an analysis of transfer matrices (see, e.g., [46]).

In this paper, we study random regular graphs, which are not invariant and do not have independent entries. We show that the eigenvalues of their adjacency matrices obey random matrix statistics in the bulk of the spectrum. The universality of local eigenvalue statistics for noninvariant matrix ensembles with correlated entries has recently been studied in a few other cases. In particular, after the appearance of this paper, GOE eigenvalue statistics were proved for the Laplacian matrix of sparse Erdős–Rényi graphs in [26], and for random matrices with certain short-range correlations in [1, 11]; these results do not cover the hard constraints of random regular graphs.

1.2. *Main results*. Let A be the adjacency matrix of the uniform random d-regular graph (RRG) on N vertices, that is, a uniformly chosen symmetric matrix with entries in $\{0, 1\}$ such that all rows and columns have sum equal to d and all diagonal entries vanish. For $d \to \infty$ as $N \to \infty$, it is known [3, 15, 45] that the eigenvalue density of $(d-1)^{-1/2}A$ converges to the Wigner *semicircle law* whose density is $\varrho(x) := \frac{1}{2\pi}\sqrt{[4-x^2]_+}$. For d at least $(\log N)^4$, three of the authors recently proved a *local semicircle law* for random regular graphs [3], giving precise estimates on the Green's function and the eigenvalue density, down to spectral scales comparable with the typical eigenvalue spacing (up to a logarithmic correction). In this paper, we consider the local eigenvalue statistics of random regular graphs in the bulk of the spectrum.

As the adjacency matrix of a d-regular graph, the matrix A has the trivial uniform eigenvector $\mathbf{e} := N^{-1/2}(1, \dots, 1)^*$ with eigenvalue d. We denote by $\lambda_1 \geq \cdots \geq \lambda_{N-1}$ the ordered nontrivial eigenvalues of $(d-1)^{-1/2}A$, and by $\mathbb{E}_{\mathrm{RRG}}$ the expectation with respect to the induced law on $\lambda_1 \geq \cdots \geq \lambda_{N-1}$. By comparison, we denote by $\mathbb{E}_{\mathrm{GOE}}$ the expectation with respect to the law of the ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_{N-1}$ of the Gaussian Orthogonal Ensemble (GOE) on $\mathbb{R}^{(N-1)\times(N-1)}$, normalized so that the off-diagonal entries have variance N^{-1}.

The typical locations γ_i of the eigenvalues under the semicircle law are defined by

$$\frac{i}{N} = \int_{\gamma_i}^{2} \varrho(x)\,\mathrm{d}x. \tag{1.1}$$

THEOREM 1.1. *Fix $\alpha > 0$, and suppose that $d \in [N^\alpha, N^{2/3-\alpha}]$. Then, in the limit $N \to \infty$, the bulk gap statistics of the random d-regular graph coincide with those of the GOE. More precisely, for any fixed $\kappa > 0$, $n \in \mathbb{N}$, and $\phi \in C_c^\infty(\mathbb{R}^n)$,*

we have

$$(\mathbb{E}_{\mathrm{RRG}} - \mathbb{E}_{\mathrm{GOE}})\phi\big(N\varrho(\gamma_i)(\lambda_i - \lambda_{i+1}), \dots, N\varrho(\gamma_i)(\lambda_i - \lambda_{i+n})\big) = o(1) \tag{1.2}$$

as $N \to \infty$, *uniformly in* $i \in [\![\kappa N, (1-\kappa)N]\!]$.

Next, let $p_\# \equiv p_{\#,N}$ denote the symmetrized joint law of the eigenvalues of the ensemble $\# = \mathrm{RRG}, \mathrm{GOE}$. The correlation functions are defined for $n \in [\![1, N-1]\!]$ by

$$p_\#^{(n)}(\mathrm{d}\lambda_1, \dots, \mathrm{d}\lambda_n) := p_\#\big(\mathrm{d}\lambda_1, \dots, \mathrm{d}\lambda_n, \mathbb{R}^{N-1-n}\big). \tag{1.3}$$

THEOREM 1.2. *Fix* $\alpha > 0$, *and suppose that* $d \in [N^\alpha, N^{2/3-\alpha}]$. *Then, in the limit* $N \to \infty$, *the locally averaged local correlation functions of the random* d*-regular graph coincide with those of the GOE. More precisely, fix a small enough constant* $c > 0$, *and define* $b \equiv b_N := N^{-1+c}$. *Then for any fixed* $n \in \mathbb{N}$, $\phi \in C_c^\infty(\mathbb{R}^n)$, *and* $E \in (-2, 2)$ *we have*

$$\begin{aligned} &\frac{1}{2b}\int_{E-b}^{E+b} \mathrm{d}E' \int_{\mathbb{R}^n} \phi(x_1, \dots, x_n) \\ &\qquad \times N^n\big(p_{\mathrm{RRG}}^{(n)} - p_{\mathrm{GOE}}^{(n)}\big)\bigg(E' + \frac{\mathrm{d}x_1}{N\varrho(E)}, \dots, E' + \frac{\mathrm{d}x_n}{N\varrho(E)}\bigg) = o(1). \end{aligned} \tag{1.4}$$

For the GOE, the eigenvalue correlation functions are known explicitly; see, for example, [35]. Hence, the quantities for the GOE appearing on the left-hand sides of (1.2) and (1.4) can be computed explicitly. In fact, the eigenvalue gap distribution has only been computed in the sense of averages over the gap index; for the GUE, the computation for a fixed gap was performed in [43].

The proofs of Theorems 1.1–1.2 follow the general three-step strategy developed in [19, 20, 23]; see, for example, [21] for a survey. In our setup, the strategy is formulated precisely in Section 2. The general idea is to study the convergence of eigenvalue statistics under Dyson Brownian motion (DBM) [16]. The three steps consist of (i) a *local law* providing precise estimates on the eigenvalue density down to the scale of individual eigenvalues, as well as the complete delocalization of the eigenvectors; (ii) the universality of the local eigenvalue statistics after the short time $t = N^{-1+\delta}$; and (iii) effective approximation of the local eigenvalue statistics of the original matrix ensemble at $t = 0$ by the one evolved up to time $t = N^{-1+\delta}$.

In all previous instances of the three-step strategy outlined above, the independence of the matrix entries was crucial for steps (i) and (iii). For the random regular graph, a new approach is required for both of these steps, the last one of which is the main content of this paper. The local law for random regular graphs was recently established in [3], thus performing step (i). As for step (ii), the convergence

of the local eigenvalue statistics under DBM with deterministic initial data was recently established in [33], under the sole assumption that the eigenvalue density be bounded at the scale $N^{-1+\delta}$. Therefore, the local semicircle law provides sufficient control on the eigenvalues so that using [3, 33] we can perform step (ii).

Thus, the main difficulty is step (iii). There are several known methods for performing this step, including Lindeberg's proof of the central limit theorem combined with higher moment matching conditions [44], or the Green's function comparison theorem [25]. For short times, a more direct method is to prove the stability of the eigenvalues under the DBM by analysing the dynamics of the individual matrix entries [10]. In all of these approaches, the independence of the matrix entries is used in an essential way. In contrast, the entries of random regular graphs are subject to hard constraints, and are therefore not independent. Tracking carefully the dependence of the matrix entries (using the methods from [3]), we find that the eigenvalue evolution is stable under a *constrained DBM*, for times $t \le N^{-1+\delta}$. Here, by *stability*, we mean that the changes in the local eigenvalue statistics are negligible.

This stability can also be interpreted as follows: there is a class of reasonably well-behaved observables, which completely characterize the local bulk eigenvalue statistics, and whose time evolution under the constrained DBM can be well approximated by a *switching dynamics* of random regular graphs. We note that it has been proposed that, for random regular graphs, the dynamics provided by DBM should be replaced with a switching dynamics; see, in particular, [31]. However, obtaining rigorous results on the local eigenvalue statistics using only a switching dynamics is difficult, because the induced eigenvalue process is neither continuous nor autonomous [30]. Our strategy crucially relies on the fact that the eigenvalue process under DBM is continuous and satisfies an autonomous system of SDEs.

Theorem 1.1 and Theorem 1.2 hold also for sparse random matrices with independent entries; see [27]. We will use parts of that analysis which are applicable here. The main effort and novelty of this paper is in the control of eigenvalues under constrained DBM up to time $t = N^{-1+\delta}$ using switchings.

1.3. *Further related results*. Large regular graphs have been proposed as a testing ground for quantum chaos [42]. It is conjectured [7] that chaotic quantum systems (i.e., quantum systems obtained by quantization of ergodic classical systems) exhibit random matrix statistics. Regular graphs are random matrices with a local structure, and as such a step in the direction of understanding highly structured systems. It is believed that the eigenvalues of random d-regular graphs obey random matrix statistics for any $d \ge 3$. Indeed, there is numerical evidence that the local spectral statistics in the bulk of the spectrum are governed by those of the GOE [28, 37], and further that the distribution of the appropriately rescaled second largest eigenvalue converges to the Tracy–Widom distribution of the GOE [36]. Our assumption $d \ge N^{\alpha}$ for arbitrary $\alpha > 0$ is purely technical, since some of the results used in our proof have only been established up to multiplicative errors

of order N^c (with arbitrary $c > 0$). We believe that our results can be extended to $d \geq (\log N)^{O(1)}$ with the same method. Furthermore, our results also extend to the other models of random regular graphs considered in [3], such as the permutation model. Other results about the eigenvalue and eigenvector distribution of d-regular graphs on mesoscopic and macroscopic scales, with $d \to \infty$ and with d fixed, are discussed in [3].

Our proof relies on *switchings* that leave the random regular graph invariant. Switchings of random regular graphs were introduced to obtain enumeration results in [34]; see [47] for a survey of subsequent developments. They are also used for simulation of random regular graphs; see, for example, [13] and references therein. Recently, switchings were used to bound the singularity probability of directed random regular graphs [12]. They also played an important role in our recent proof of the local semicircle law for random regular graphs [3].

Notation. We use $a = O(b)$ to mean that there exists an absolute constant $C > 0$ such that $|a| \leq Cb$, and $a \gg b$ to mean that $a \geq Cb$ for some sufficiently large absolute constant $C > 0$. We use c for an arbitrarily small positive constant that may change from line to line. Moreover, we abbreviate $[\![a, b]\!] := [a, b] \cap \mathbb{Z}$. We use the standard notation $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. Every quantity that is not explicitly a constant may depend on N, which we almost always omit from our notation. Throughout the paper, we tacitly assume $N \gg 1$. Unless otherwise stated, all sums of indices are over the set $[\![1, N]\!]$.

2. Strategy of proof. Our goal is to prove that, in the bulk on the spectrum, the local eigenvalue statistics of $A/\sqrt{d-1}$ are the same as those of the GOE. As mentioned in Section 1, in order to show this, we interpolate between the RRG and the GOE using Dyson Brownian motion, or more precisely its Ornstein–Uhlenbeck version.

2.1. *Constrained Dyson Brownian motion.* The adjacency matrix A of a regular graph is subject to the hard constraints that its rows and columns have fixed sum [i.e., it has the eigenvector $\mathbf{e} = N^{-1/2}(1, \dots, 1)^*$]. Therefore, instead of the usual Dyson Brownian motion, we use Dyson Brownian motion constrained to the subspace of symmetric matrices whose row and column sums vanish.

We begin with the notion of an Ornstein–Uhlenbeck process on a general finite-dimensional space.

DEFINITION 2.1. Let $\mathcal{H}$ be a real finite-dimensional Hilbert space. Let $(\mathbf{f}_\alpha)_\alpha$ be an orthonormal basis of $\mathcal{H}$.

(i) Let $(w_\alpha)_\alpha$ be i.i.d. standard normal random variables. Then we define the *standard Gaussian measure on* $\mathcal{H}$ as $W := \sum_\alpha w_\alpha \mathbf{f}_\alpha$.

(ii) Let $(h_\alpha)_\alpha$ be i.i.d. Ornstein–Uhlenbeck processes satisfying

$$\mathrm{d}h_\alpha = \mathrm{d}B_\alpha - \frac{1}{2}h_\alpha \,\mathrm{d}t,$$

where $(B_\alpha)_\alpha$ is a family of i.i.d. standard Brownian motions. Then we define the *standard Ornstein–Uhlenbeck process on* $\mathcal{H}$ as $H(t) := \sum_\alpha h_\alpha(t)\mathbf{f}_\alpha$.

It is easy to verify that the laws of W and the process H do not depend on the choice of the orthonormal basis $(\mathbf{f}_\alpha)$, and that the standard Gaussian measure is invariant under the standard Ornstein–Uhlenbeck process. We use these properties tacitly from now on.

For example, let $\mathcal{H} := \{H \in \mathbb{R}^{N\times N} : H = H^*\}$ be the Hilbert space of real symmetric $N \times N$ matrices with inner product

$$\langle X, Y\rangle := \frac{N}{2}\operatorname{Tr}(XY). \tag{2.1}$$

Then the usual N-dimensional Dyson Brownian motion is the standard Ornstein–Uhlenbeck process $H(t)$ on $\mathcal{H}$. More explicitly, $H(t)$ is the Markov process satisfying the SDE

$$\mathrm{d}H = \frac{1}{\sqrt{N}}\,\mathrm{d}B - \frac{1}{2}H\,\mathrm{d}t, \tag{2.2}$$

where $B(t)$ is Brownian motion on the space of $N \times N$ real symmetric matrices with quadratic covariation $\langle B_{ij}, B_{kl}\rangle(t) = (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})t$.

More intrinsically, given a finite-dimensional Hilbert space V, we denote the Hilbert space of symmetric linear maps on V with inner product (2.1) by $\mathcal{H}(V)$. Then we define *Dyson Brownian motion* (*DBM*) *on* V to be the standard Ornstein–Uhlenbeck process on $\mathcal{H}(V)$. With this point of view, the usual N-dimensional DBM is the DBM on $V = \mathbb{R}^N$, and the constrained DBM is the DBM on $V = \mathbf{e}^\perp$. Note that the normalization N in (2.1) does not need to agree with the dimension of V, which is $N - 1$ for $V = \mathbf{e}^\perp$. We make the convention to always normalize the inner product (2.1) by N, no matter the dimension of V, and always denote the dimension of V by M. Finally, we denote the inner product on V by $\mathbf{v}\cdot\mathbf{w}$ for $\mathbf{v}, \mathbf{w} \in V$.

DEFINITION 2.2 (Constrained DBM and GOE). The *constrained DBM* is the DBM on $\mathbf{e}^\perp$, that is, the standard Ornstein–Uhlenbeck process on $\mathcal{H}(\mathbf{e}^\perp)$ with inner product (2.1). The *constrained GOE* is the standard Gaussian measure on $\mathcal{H}(\mathbf{e}^\perp)$ with inner product (2.1).

Thus, up to a change of basis, the constrained DBM is equivalent to the usual $(N-1)$-dimensional DBM, with the minor difference of normalization by N rather than $N - 1$. However, since the definition of the d-regular graph is tied to the

standard basis of $\mathbb{R}^N$, it is frequently convenient to work with the constrained DBM in the standard basis of $\mathbb{R}^N$.

Next, in accordance with the decomposition $\mathbb{R}^N = \mathbf{e}^\perp \oplus \operatorname{span}(\mathbf{e})$, we have a canonical isomorphism $H \mapsto \tilde{H} := H \oplus 0$ from $\mathcal{H}(\mathbf{e}^\perp)$ to the set of matrices

$$\mathcal{M} := \{H \in \mathbb{R}^{N\times N} : H = H^*, H\mathbf{e} = 0\}. \tag{2.3}$$

Throughout this paper, we tacitly identify H and $\tilde{H}$.

We denote by $\mathcal{C}^n(\mathcal{M})$ the space of functions $F : \mathcal{M} \to \mathbb{C}$ with continuous bounded derivatives up to order n. Sometimes it will be convenient to compute derivatives of functions $F \in \mathcal{C}^n(\mathcal{M})$ in directions of $\mathbb{R}^{N\times N}$ that do not lie in $\mathcal{M}$, which is made possible by the following convention.

DEFINITION 2.3. Let $P = I - \mathbf{e}\mathbf{e}^*$ be the orthogonal projection from $\mathbb{R}^N$ onto $\mathbf{e}^\perp$. We extend any function $F \in \mathcal{C}^n(\mathcal{M})$ to a $\mathcal{C}^n$-function on $\mathbb{R}^{N\times N}$ through

$$H \longmapsto F\left(\frac{1}{2}P(H + H^*)P\right),$$

and denote this extended function also by F. Finally, for any $F \in \mathcal{C}^1(\mathcal{M})$ and $i, j \in [\![1, N]\!]$, we use the abbreviation $\partial_{ij}F(H) \equiv \frac{\partial F}{\partial H_{ij}}(H)$.

From now on, we take W to be the constrained GOE and $H \equiv H(t)$ to be the constrained DBM, with initial condition

$$H(0) := \frac{1}{\sqrt{d-1}}(A - d\mathbf{e}\mathbf{e}^*) \in \mathcal{M}. \tag{2.4}$$

Here, A is the adjacency matrix of the random d-regular graph. In particular, the eigenvalues of $H(0)$ as an element of $\mathcal{H}(\mathbf{e}^\perp)$ are the rescaled nontrivial eigenvalues of A.

2.2. *Switchings*. Simple switchings are an especially convenient generating set of $\mathcal{M}$; they play a central role throughout this paper. For any $i, j, k, l \in [\![1, N]\!]$, we define the *simple switching* $\xi_{ij}^{kl} \in \mathcal{M}$ by

$$\xi_{ij}^{kl} := \Delta_{ij} + \Delta_{kl} - \Delta_{ik} - \Delta_{jl} \qquad \text{where } (\Delta_{ij})_{pq} := \delta_{ip}\delta_{jq} + \delta_{iq}\delta_{jp}. \tag{2.5}$$

The action of a simple switching ξ_{ij}^{kl} on an adjacency matrix, given by $A \mapsto A + \xi_{ij}^{kl}$, amounts to adding the edges $\{i, j\}, \{k, l\}$ and removing the edges $\{i, k\}, \{j, l\}$; this is illustrated in Figure 1 and made precise in (3.14) below. In this section, the four vertices need not be distinct.

Next, we define the abbreviations

$$H_{ij}^{kl} := \operatorname{Tr}(\xi_{ij}^{kl} H), \qquad \partial_{ij}^{kl} := \partial_{\xi_{ij}^{kl}} = \operatorname{Tr}(\xi_{ij}^{kl}\,\partial), \tag{2.6}$$

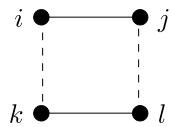

FIG. 1. *A simple switching is given by replacing the solid edges by the dashed edges.*

for all $i, j, k, l \in [\![1, N]\!]$. Here ∂_X denotes the directional derivative in the direction X. Explicitly, expressed in the standard basis on $\mathbb{R}^N$, we have

$$H_{ij}^{kl} = 2(H_{ij} + H_{kl} - H_{ik} - H_{jl}), \tag{2.7}$$

$$\partial_{ij}^{kl} F(H) = 2(\partial_{ij} + \partial_{kl} - \partial_{ik} - \partial_{jl}) F(H), \tag{2.8}$$

where $F \in \mathcal{C}^1(\mathcal{M})$. With these abbreviations, the generator of the constrained DBM can be expressed in terms of switchings as stated in the following proposition.

PROPOSITION 2.4. *The generator of the constrained DBM from Definition 2.2 is*

$$L := \frac{1}{16N^3} \sum_{i,j,k,l} (\partial_{ij}^{kl})^2 - \frac{1}{32N^2} \sum_{i,j,k,l} H_{ij}^{kl}\, \partial_{ij}^{kl}. \tag{2.9}$$

This means that for any $F \in \mathcal{C}^2(\mathcal{M})$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}\big[F\big(H(t)\big)\big] = \mathbb{E}\big[LF\big(H(t)\big)\big]. \tag{2.10}$$

PROOF. Let $\hat{H}(t)$ be the standard Ornstein–Uhlenbeck process from Definition 2.1 on the space $\mathcal{H}(\mathbb{R}^{N-1})$ with inner product (2.1). As in the example (2.2), we obtain the quadratic covariation

$$\langle \hat{H}_{ij}, \hat{H}_{kl} \rangle(t) = \frac{1}{N} (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) t. \tag{2.11}$$

Next, let $R \in O(N)$ satisfy $R\mathbf{e}_N = \mathbf{e}$. Then, since the inner product (2.1) is invariant under orthogonal conjugations, we can express the constrained DBM as $H(t) = R(\hat{H}(t) \oplus 0)R^*$. We abbreviate $H \equiv H(t)$ and write for $F \in \mathcal{C}^2(\mathcal{M})$, using Itô calculus,

$$\begin{aligned} \mathrm{d}\,\mathbb{E}F(H) = &-\frac{1}{2} \sum_{i,j} \mathbb{E}\big[H_{ij} (\partial_{ij} F)(H)\big]\,\mathrm{d}t \\ &+ \frac{1}{2} \sum_{i,j,k,l} \mathbb{E}\big[(\partial_{ij}\, \partial_{kl} F)(H)\, \mathrm{d}\langle H_{ij}, H_{kl} \rangle\big]. \end{aligned}$$

By definition of R, we have $R_{iN} = \frac{1}{\sqrt{N}}$ for all i, so that (2.11) yields

$$\begin{aligned} \mathrm{d}\langle H_{ij}, H_{kl}\rangle &= \frac{1}{N}\sum_{m,n=1}^{N-1} (R_{im}R_{jn}R_{km}R_{ln} + R_{im}R_{jn}R_{kn}R_{lm})\,\mathrm{d}t \\ &= \frac{1}{N}\left(\delta_{ik} - \frac{1}{N}\right)\left(\delta_{jl} - \frac{1}{N}\right)\mathrm{d}t + \frac{1}{N}\left(\delta_{il} - \frac{1}{N}\right)\left(\delta_{jk} - \frac{1}{N}\right)\mathrm{d}t. \end{aligned}$$

Thus, for any $F \in \mathcal{C}^2(\mathcal{M})$ we have (2.10) with

$$L = \frac{1}{N^3}\sum_{i,j,k,l} \partial_{ij}(\partial_{ij} + \partial_{kl} - \partial_{il} - \partial_{jk}) - \frac{1}{2}\sum_{i,j} H_{ij}\,\partial_{ij}. \tag{2.12}$$

Finally, using $\sum_j H_{ij} = \sum_j H_{ji} = 0$ for $H \in \mathcal{M}$, we observe that L from (2.12) can be rewritten as (2.9). $\square$

2.3. *Outline of proof of Theorems* 1.1–1.2. Theorems 1.1–1.2 are an immediate consequence of the following two propositions. As in [3], we set

$$D := d \wedge \frac{N^2}{d^3}. \tag{2.13}$$

We always assume $d \in [N^\alpha, N^{2/3-\alpha}]$, which implies $D \geq N^\alpha$. To state the two propositions concisely, we introduce the following definition. It will also be convenient in the proofs.

DEFINITION 2.5. Given $H \in \mathcal{M}$, we denote by $\lambda_1 \geq \cdots \geq \lambda_{N-1}$ the eigenvalues of $H|_{\mathbf{e}^\perp}$. Consider two random matrix ensembles H_1 and H_2 in $\mathcal{M}$. Then we say that:

(i) the *bulk eigenvalue gap statistics of H_1 and H_2 coincide* if for any $n \in \mathbb{N}$, $\phi \in C_c^\infty(\mathbb{R}^n)$, and $\kappa > 0$, we have

$$(\mathbb{E}_{H_1} - \mathbb{E}_{H_2})\phi\big(N\varrho(\gamma_i)(\lambda_i - \lambda_{i+1}), \dots, N\varrho(\gamma_i)(\lambda_i - \lambda_{i+n})\big) = o(1) \tag{2.14}$$

as $N \to \infty$, uniformly in $i \in [\![\kappa N, (1-\kappa)N]\!]$;

(ii) the *averaged bulk eigenvalue correlation functions of H_1 and H_2 coincide* if for any $n \in \mathbb{N}$, $\phi \in C_c^\infty(\mathbb{R}^n)$, $c > 0$ small enough, and $E \in (-2, 2)$, we have for $b := N^{-1+c}$

$$\begin{aligned} &\frac{1}{2b}\int_{E-b}^{E+b} \mathrm{d}E' \int_{\mathbb{R}^n} \phi(x_1, \dots, x_n) N^n \big(p_{H_1}^{(n)} - p_{H_2}^{(n)}\big) \\ &\qquad \times \left(E' + \frac{\mathrm{d}x_1}{N\varrho(E)}, \dots, E' + \frac{\mathrm{d}x_n}{N\varrho(E)}\right) = o(1), \end{aligned} \tag{2.15}$$

where the correlation functions $p_{H_i}^{(n)}$ are defined as in (1.3).

Moreover, we say that the *bulk eigenvalue statistics of H_1 and H_2 coincide* if (i) and (ii) hold.

PROPOSITION 2.6. *For any fixed $\delta > 0$ and $t \le N^{-1-\delta} D^{1/2}$, the bulk eigenvalue statistics of $H(0)$ and $H(t)$ coincide.*

PROPOSITION 2.7. *For any fixed $\delta > 0$ and $t \ge N^{-1+\delta}$, the bulk eigenvalue statistics of $H(t)$ and $H(\infty) \overset{\mathrm{d}}{=} W$ coincide.*

Propositions 2.6–2.7 are proved in Section 5. As mentioned in Section 1, our main effort and novelty is in proving Proposition 2.6. Proposition 2.7 is essentially a consequence of general results on universality of local eigenvalue statistics with small Gaussian component [33]. The local semicircle law of [3] is an important input in the proofs of both propositions.

PROOF OF THEOREM 1.2. The proof is immediate from Propositions 2.6–2.7, with $\delta \le \alpha/4$. $\square$

3. Switchings and short-time comparision. The main result of this section is Proposition 3.1 below. To state it, we introduce the following Sobolev-type seminorms, whereby the derivatives are taken in the directions of all switchings

$$\mathcal{X} := \{\xi_{ij}^{kl} \in \mathbb{R}^{N\times N} : i, j, k, l \in [\![1, N]\!]\}. \tag{3.1}$$

First, for $r \ge 1$, we define an L^r-seminorm on $\mathcal{C}^0(\mathcal{M})$ through

$$\|F\|_{r,t} := \big(\mathbb{E}\big|F\big(H(t)\big)\big|^r\big)^{1/r}. \tag{3.2}$$

Then we extend this seminorm to include derivatives: for $F \in \mathcal{C}^n(\mathcal{M})$ we define

$$\big\|\partial^n F\big\|_{r,t} := \Big\| \sup_{\theta\in[0,1]^n} \sup_{X\in\mathcal{X}^n} \big|\partial_{X_1}\cdots\partial_{X_n} F\big(\cdot + (d-1)^{-1/2}\theta\cdot X\big)\big| \Big\|_{r,t}, \tag{3.3}$$

where ∂_Y denotes the directional derivative in the direction Y, and for $\theta \in [0, 1]^n$ and $X \in \mathcal{X}^n$ we abbreviate

$$\theta \cdot X := \theta_1 X_1 + \cdots + \theta_n X_n.$$

PROPOSITION 3.1. *Let $H(t)$ be the constrained Dyson Brownian motion from Definition 2.2 with initial condition (2.4). Fix $\varepsilon > 0$ and let $r \equiv r(\varepsilon)$ be large enough depending on ε. Then for any $F \in \mathcal{C}^4(\mathcal{M})$ we have*

$$\mathbb{E}F\big(H(t)\big) - \mathbb{E}F\big(H(0)\big) = O\Big(D^{-1/2} N^{1+\varepsilon} \max_{1\le i\le 4} \int_0^t \|\partial^i F\|_{r,s}\,\mathrm{d}s\Big). \tag{3.4}$$

In the applications in Section 5, we will use functions F satisfying $\|\partial^i F\|_{r,s} \le N^c$ for $i \le 4$ and a constant $c > 0$ that can be chosen arbitrarily small. Thus, for $t \le N^{-1-\delta} D^{1/2}$ the right-hand side of (3.4) will be $O(N^{-\delta+\varepsilon+c})$ which is $o(1)$ provided that $c + \varepsilon < \delta$.

The starting point for the proof of Proposition 3.1 is the idea of [10], Lemma A.1, namely to estimate the left-hand side of (3.4) by estimating $\mathbb{E}(LF(H(t)))$. However, since the entries of $H(t)$ are not independent, a different approach from [10] is needed to control $\mathbb{E}(LF(H(t)))$. We do this by approximating the constrained DBM by a Markovian jump process induced by switchings. This process is defined as follows.

3.1. *Switching dynamics.* We introduce a Markovian jump process on simple regular graphs by defining its generator

$$Qf(A) := \frac{1}{8Nd} \sum_{i,j,m,n} I_{ij}^{mn}(A)\big(f\big(A - \xi_{ij}^{mn}\big) - f(A)\big), \tag{3.5}$$

where we recall the definition of a switching from (2.5) and introduce the indicator function

$$I_{ij}^{mn}(A) := A_{ij}A_{mn}(1 - A_{im})(1 - A_{in})(1 - A_{jm})(1 - A_{jn}). \tag{3.6}$$

The indicator function $I_{ij}^{mn}(A)$ ensures that the graph encoded by A contains the edges $\{i, j\}$ and $\{m, n\}$ but no other edges between the four vertices $\{i, j, m, n\}$ (i.e., its restriction to $\{i, j, m, n\}$ is 1-regular).

Thus, the process generated by Q is a Markovian jump process whose jump times are the events of a Poisson clock with a constant rate; at each event of the clock, four vertices are selected uniformly at random, and a switching as in Figure 1 is performed on the graph if the four vertices are connected by exactly two edges. It is not hard to show that the uniform measure on d-regular graphs is invariant under this jump process.

PROPOSITION 3.2. *The uniform measure on simple d-regular graphs is invariant under Q. This means that for any function f on the set of simple d-regular graphs we have $\mathbb{E}(Qf(A)) = 0$.*

The proof of the proposition is given in Section 3.2, in a slightly more general context. The following proposition shows that the switching jump process generated by Q is well approximated by the constrained DBM generated by L.

The generator L acts naturally on functions of H (denoted henceforth by an uppercase F), and the generator Q on functions of A (denoted henceforth by a lowercase f). It is therefore convenient to introduce, for any $F \in \mathcal{C}^n(\mathcal{M})$, the abbreviations

$$H = H_A := \frac{1}{\sqrt{d-1}}(A - d\mathbf{e}\mathbf{e}^*), \qquad f(A) = f_F(A) := F(H_A). \tag{3.7}$$

PROPOSITION 3.3. *For any $F \in \mathcal{C}^4(\mathcal{M})$ and using the notation (3.7), we have*

$$Qf(A) = LF(H) + R, \tag{3.8}$$

where

$$\mathbb{E}R = O\big(D^{-1/2}N^{1+\varepsilon}\big) \max_{1\le i\le 4} \big\|\partial^i F\big\|_{r,0}. \tag{3.9}$$

Here, $\mathbb{E}$ *denotes expectation with respect to the uniform measure on random* d-*regular graphs* A.

The proof of this proposition is also deferred to Section 3.2 below. Roughly, the idea of the proof is as follows. By Taylor's expansion, we obtain

$$Qf(A) \approx \frac{1}{8Nd} \sum_{i,j,m,n} A_{ij}A_{mn}\Big(-\partial_{ij}^{mn} f(A) + \frac{1}{2}\big(\partial_{ij}^{mn}\big)^2 f(A)\Big) \tag{3.10}$$

with high probability. Now $\mathbb{E}A_{ij} = \frac{d}{N}$ if $i \neq j$. By expanding $A_{ij}A_{mn} = (\frac{d}{N} + (A_{ij} - \frac{d}{N}))(\frac{d}{N} + (A_{mn} - \frac{d}{N}))$, and keeping only the leading terms, we find that the right-hand side of (3.10) becomes by $LF(H)$. Here, for the second-order term on the right-hand side of (3.10), the leading term from $A_{ij}A_{mn}$ is $\frac{d^2}{N^2}$; for the first-order term on the right-hand side of (3.10), the leading term from $A_{ij}A_{mn}$ is $\frac{d}{N}(A_{ij} - \frac{d}{N}) + \frac{d}{N}(A_{mn} - \frac{d}{N})$. Further error terms result from the dependence of the entries of the adjacency matrix.

Before giving the proofs of Propositions 3.2–3.3, we deduce Proposition 3.1 from them.

PROOF OF PROPOSITION 3.1. By (2.10), it suffices to estimate $\mathbb{E}[LF(H(t))]$. By explicit solution of the constrained DBM, $H(t)$, we find for any fixed $t \ge 0$ that

$$H(t) \overset{\mathrm{d}}{=} \mathrm{e}^{-t/2}H(0) + \big(1 - \mathrm{e}^{-t}\big)^{1/2} W, \tag{3.11}$$

where W is a copy of the constrained GOE independent of $H(0)$. For the remainder of the proof, we identify the right-hand side with $H(t)$, abbreviate $H \equiv H(0)$, and introduce the two functions

$$F_W(H) = F_H(W) := F\big(\mathrm{e}^{-t/2}H + \big(1 - \mathrm{e}^{-t}\big)^{1/2} W\big),$$

where the choice of the argument determines the variables on which the generator L acts. We recall the generator L from (2.9),

$$L = \frac{1}{16N^3} \sum_{i,j,k,l} \big(\partial_{ij}^{kl}\big)^2 - \frac{1}{32N^2} \sum_{i,j,k,l} H_{ij}^{kl}\, \partial_{ij}^{kl}.$$

From $\partial^2 = (\mathrm{e}^{-t} + (1 - \mathrm{e}^{-t}))\,\partial^2$, $\mathrm{e}^{-t/2}\,\partial F = \partial F_W$, and $(1 - \mathrm{e}^{-t})^{1/2}\,\partial F = \partial F_H$, we then deduce that $LF(\mathrm{e}^{-t/2}H + (1 - \mathrm{e}^{-t})^{1/2}W) = LF_W(H) + LF_H(W)$. We therefore get

$$\mathbb{E}\big[LF\big(H(t)\big)\big] = \mathbb{E}\big[LF_W(H)\big] + \mathbb{E}\big[LF_H(W)\big] = \mathbb{E}\big[LF_W(H)\big],$$

FIG. 2. *Given four vertices i, j, k, l with two edges between them, there are two possible switchings. By equipping the edges with directions, one of these two switchings can be selected canonically.*

where in the second step we used that the constrained GOE, W, is invariant with respect to the generator L.

Next, we define $f_W(A) := F_W(H)$ where $H \equiv H_A$ is defined as (3.7). By Proposition 3.2, the random d-regular graph A is invariant with respect to the generator Q, and Proposition 3.3 therefore yields

$$\begin{aligned}\mathbb{E}[LF_W(H)] &= \mathbb{E}[Qf_W(A)] + O(D^{-1/2}N^{1+\varepsilon}) \max_{1\le i\le 4} \|\partial^i F_W\|_{r,0} \\ &= O(D^{-1/2}N^{1+\varepsilon}) \max_{1\le i\le 4} \|\partial^i F\|_{r,t}.\end{aligned}$$

Thus, with (2.10), we have shown that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[F(H(t))] = O(D^{-1/2}N^{1+\varepsilon}) \max_{1\le i\le 4} \|\partial^i F\|_{r,t},$$

and the claim follows by integrating over t. $\square$

3.2. *Proofs of Propositions* 3.2–3.3. Propositions 3.2–3.3 concern switchings of regular graphs. Switchings played an important role in the proof of the local semicircle law for random regular graphs [3]. Here, we use simple switchings instead of the double switchings needed in [3].

Given two disjoint edges of a regular graph such that the graph has no other edges between the vertices incident to these two edges, there are two possible switchings; see Figure 2. To specify one of these two switchings, it is convenient to assign to each of the edges to be switched a direction; there is then a canonical choice between the two possible switchings. We write ij for the edge $\{i, j\}$ directed from i to j.

We consider sets S of two directed edges of the complete graph, which we write in the form $S = \{ij, kl\}$. We denote by $[S] = \{i, j, k, l\}$ the set of vertices incident to the edges of S. For two such sets S and S', we define the indicator functions

$$I(S) \equiv I(S; A) := \mathbf{1}\big(|[S]| = 4 \text{ and } E|_{[S]} \text{ is 1-regular}\big), \tag{3.12}$$

$$J(S, S') \equiv J(S, S'; A) := \mathbf{1}\big([S] \cap [S'] = \varnothing\big), \tag{3.13}$$

where $E \equiv E(A) := \{\{i, j\} : A_{ij} = 1\}$ is the set of (undirected) edges of the graph encoded by A, and $E|_B := \{e \in E : e \subset B\}$ is the restriction of the graph E to the subset of vertices B. The indicator functions are illustrated in Figure 3. Note that $I_{ij}^{mn} = A_{ij}A_{mn}I(\{ij, mn\})$ [recall (3.6)].

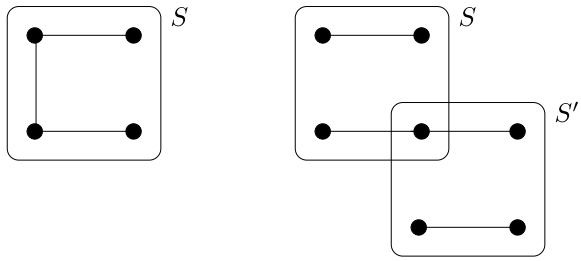

FIG. 3. *In the left diagram, $I(S) = 0$ since the restricted graph is not 1-regular. In the right diagram, $J(S, S') = 0$ since the two sets of vertices intersect.*

REMARK 3.4. The definitions (3.12)–(3.13) are similar to those given in [3], Section 6, with the following differences. First, the current set S consists of two directed edges instead of the three undirected edges in [3]. Because of the directions contained in the current set S, it effectively incorporates the extra parameter s of [3], Section 6. Second, the edges in S are edges of the complete graph, and we do not assume that they are contained in some regular graph A; we will ultimately define the switching associated with the set S to act trivially unless S is contained in the edges E of the given graph.

For a set $S = \{ij, kl\}$ of two directed edges, we define the *switching*

$$T_S(A) := \begin{cases} A - \xi_{ij}^{kl}, & \text{if } I(S) = 1, A_{ij} = 1, A_{kl} = 1, \\ A + \xi_{ij}^{kl}, & \text{if } I(S) = 1, A_{ik} = 1, A_{jl} = 1, \\ A, & \text{otherwise,} \end{cases} \tag{3.14}$$

where we recall the definition of ξ_{ij}^{kl} from (2.5). In words, $T_S(A)$ switches the edges S if they are contained in A and are switchable in the sense that the switching results again in a d-regular graph. Moreover, for S, S' as above, we define

$$T_{S,S'}(A) := \begin{cases} T_{S'}(T_S(A)), & \text{if } J(S, S') = 1, \\ A, & \text{otherwise.} \end{cases} \tag{3.15}$$

In words, $T_{S,S'}(A)$ switches the edges in S and S' if they are contained in A and the two switchings do not interfere with each other.

LEMMA 3.5. *For any fixed S, S' we have $A \overset{\mathrm{d}}{=} T_S(A)$ and $A \overset{\mathrm{d}}{=} T_{S,S'}(A)$.*

PROOF. It is easy to check that $T_S(A)$ is a d-regular graph if and only if A is. Moreover, $T_S(T_S(A)) = A$, so T_S is a bijection on the set of d-regular graphs. Since the distribution of A is uniform, we obtain $A \overset{\mathrm{d}}{=} T_S(A)$. The second claim follows similarly from $T_{S,S'}(T_{S,S'}(A)) = A$. $\square$

Now Proposition 3.2 follows easily.

PROOF OF PROPOSITION 3.2. For any f, we get

$$\begin{aligned}\sum_{i,j,m,n} \mathbb{E}\big(I_{ij}^{mn}(A) f(A)\big) &= \sum_{i,j,m,n} \mathbb{E}\big(A_{ij} A_{mn} I\big(\{ij, mn\}; A\big) f(A)\big) \\ &= \sum_{i,j,m,n} \mathbb{E}\big(A_{im} A_{jn} I\big(\{ij, mn\}; A + \xi_{ij}^{mn}\big) f\big(A + \xi_{ij}^{mn}\big)\big) \\ &= \sum_{i,j,m,n} \mathbb{E}\big(A_{ij} A_{mn} I\big(\{ij, mn\}; A\big) f\big(A - \xi_{ij}^{mn}\big)\big) \\ &= \sum_{i,j,m,n} \mathbb{E}\big(I_{ij}^{mn}(A) f\big(A - \xi_{ij}^{mn}\big)\big),\end{aligned}$$

where the first and last steps follows from the definition of I_{ij}^{mn}, the second step from Lemma 3.5, and the third step from the exchangeability of i, j, m, n and using $I(S; A) = I(S; T_S(A))$. This completes the proof. □

For the proof of Proposition 3.3, we shall need estimates on the moments of entries of the adjacency matrix, as well as estimates on such moments restricted to low-probability events where the indicator functions (3.12)–(3.13) are zero. These estimates are collected in the following sequence of lemmas.

The following two lemmas show that moments of the entries of the adjacency matrix behave roughly like those of an Erdős–Rényi graph.

LEMMA 3.6. *Let $b \ll N$ and $i_1, j_1, \dots, i_b, j_b \in [\![1, N]\!]$. Then for any $p \in [\![1, N]\!]$ and $q \in [\![1, N]\!] \setminus \{i_1, j_1, \dots, i_b, j_b\}$, we have*

$$\mathbb{E}(A_{i_1 j_1} \cdots A_{i_b j_b} A_{pq}) = O\left(\frac{d}{N}\right) \mathbb{E}(A_{i_1 j_1} \cdots A_{i_b j_b}), \tag{3.16}$$

where we use the convention $\mathbb{E}(A_{i_1 j_1} \cdots A_{i_b j_b}) = 1$ if $b = 0$.

PROOF. Set $X := A_{i_1 j_1} \cdots A_{i_b j_b}$ and $I := \{i_1, j_1, \dots, i_b, j_b, p\}$. Then, since $\sum_n A_{pn} = d$ for any p, we find for any $q \notin I$ that

$$\begin{aligned}\mathbb{E}(X) &= \frac{1}{d} \sum_n \mathbb{E}(X A_{pn}) = \frac{1}{d} \sum_{n \notin I} \mathbb{E}(X A_{pn}) + \frac{1}{d} \sum_{n \in I} \mathbb{E}(X A_{pn}) \\ &\geq \frac{1}{d} \sum_{n \notin I} \mathbb{E}(X A_{pn}) = \frac{N - |I|}{d} \mathbb{E}(X A_{pq}),\end{aligned}$$

where in the third step we used that $X A_{pn} \geq 0$ and in the last step that the law of A is invariant under permutation of vertices. Using that $|I| \leq N/2$ by assumption on b, the claim now follows. □

As a consequence of Lemma 3.6, we obtain the following explicit bounds.

LEMMA 3.7. *Suppose that* $|\{i, j, m, n\}| = 4 - a$ *and* $|\{i, j, k, l, m, n, p, q\}| = 8 - b$. *Then*

$$\mathbb{E}(A_{ij}A_{mn}) = O\left(\frac{d}{N}\right)^{2-\lfloor a/2\rfloor}, \tag{3.17}$$

$$\mathbb{E}(A_{ij}A_{mn}A_{kl}A_{pq}) = O\left(\frac{d}{N}\right)^{4-\lfloor b/2\rfloor}. \tag{3.18}$$

PROOF. Since $A_{ss} = 0$ for all s, we can assume that $i \neq j$, $m \neq n$, $k \neq l$, and $p \neq q$, and thus $a \leq 2$ and $b \leq 4$. Then (3.17)–(3.18) follow easily from Lemma 3.6. □

In the next two lemmas we estimate moments restricted to low-probability events where the indicator functions (3.12)–(3.13) vanish, that is, we estimate the contribution of graphs A that are not switchable. Throughout the rest of this section, for given indices i, j, k, l, m, n, p, q we use the abbreviations

$$I_1 := I(\{ij, mn\}; A), \qquad I_2 := I(\{kl, pq\}; A), \tag{3.19}$$

$$J_{12} := J(\{ij, mn\}, \{kl, pq\}; A), \qquad I_{12} := I_1 I_2 J_{12}, \tag{3.20}$$

with I and J defined in (3.12)–(3.13).

LEMMA 3.8. *Let* $|\{i, j, m, n\}| = 4 - a$ *and* $|\{i, j, k, l, m, n, p, q\}| = 8 - b$. *Then*

$$\mathbb{E}\big((A_{ij}A_{mn} + A_{im}A_{jn})(1 - I_1)\big) = O\left(\frac{d}{N}\right)^{3-a}, \tag{3.21}$$

$$\mathbb{E}\big((A_{ij}A_{mn} + A_{im}A_{jn})(A_{kl}A_{pq} + A_{kp}A_{lq})(1 - I_{12})\big) = O\left(\frac{d}{N}\right)^{5-b}. \tag{3.22}$$

PROOF. First, assume that i, j, k, l, m, n, p, q are all distinct, that is, we consider the case $a = b = 0$. Then, since $|\{i, j, m, n\}| = 4$ and $I_1 = 0$ implies that the graph A restricted to $\{i, j, m, n\}$ is not 1-regular, we find

$$\mathbb{E}\big(A_{ij}A_{mn}(1 - I_1)\big) \leq \mathbb{E}\big(A_{ij}A_{mn}(A_{im} + A_{in} + A_{jm} + A_{jn})\big),$$
$$\mathbb{E}\big(A_{im}A_{jn}(1 - I_1)\big) \leq \mathbb{E}\big(A_{im}A_{jn}(A_{ij} + A_{mn} + A_{in} + A_{jm})\big),$$

and Lemma 3.6 implies that the right-hand sides are bounded by $O(d/N)^3$. The proof of (3.22) for $b = 0$ is analogous. We only consider the term $A_{ij}A_{kl}A_{mn}A_{pq}$; the others dealt with similarly. First, note that $J_{12} = 1$ if $|\{i, j, k, l, m, n, p, q\}| = 8$. Since $|\{i, j, k, l, m, n, p, q\}| = 8$ and $I_1 I_2 = 0$ imply that $E|_{\{i,j,m,n\}}$ or $E|_{\{k,l,p,q\}}$

has at least three edges, we find

$$\begin{aligned}
&\mathbb{E}\big(A_{ij}A_{mn}A_{kl}A_{pq}(1-I_1I_2J_{12})\big)\\
&\quad=\mathbb{E}\big(A_{ij}A_{mn}A_{kl}A_{pq}(1-I_1I_2)\big)\\
&\quad\le\mathbb{E}\big(A_{ij}A_{mn}A_{kl}A_{pq}(A_{im}+A_{in}+A_{jm}+A_{jn}+A_{kp}+A_{kq}+A_{lp}+A_{lq})\big)\\
&\quad=O\left(\frac{d}{N}\right)^5,
\end{aligned}$$

where the last step follows from Lemma 3.6.

Finally, if $a>0$ we have $I_1=0$, and if $b>0$ we have $I_{12}=0$. In these cases, we can directly apply (3.17) and (3.18), respectively, and the claim follows since $2-\lfloor a/2\rfloor\ge 3-a$ if $a>0$ and $4-\lfloor b/2\rfloor\ge 5-b$ if $b>0$. $\square$

As a consequence of Lemma 3.8, we obtain the following averaged estimates.

LEMMA 3.9. *If* $|\{i,j\}|=2-a$ *and* $|\{i,j,k,l\}|=4-b$, *then*

$$\frac{1}{N^2}\sum_{m,n}\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})(1-I_1)\big)=O\left(\frac{d}{N}\right)^{3-a}, \tag{3.23}$$

$$\begin{aligned}
&\frac{1}{N^4}\sum_{m,n}\sum_{p,q}\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})(A_{kl}A_{pq}+A_{kp}A_{lq})(1-I_{12})\big)\\
&\qquad=O\left(\frac{d}{N}\right)^{5-b}.
\end{aligned} \tag{3.24}$$

Moreover,

$$\frac{1}{N^4}\sum_{i,j,m,n}\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})(1-I_1)\big)=O\left(\frac{d}{N}\right)^{3}, \tag{3.25}$$

$$\begin{aligned}
&\frac{1}{N^8}\sum_{i,j,m,n}\sum_{k,l,p,q}\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})(A_{kl}A_{pq}+A_{kp}A_{lq})(1-I_{12})\big)\\
&\qquad=O\left(\frac{d}{N}\right)^{5}.
\end{aligned} \tag{3.26}$$

PROOF. To prove (3.23), we split the summation over m,n by fixing $|\{i,j,m,n\}|=4-a-s$ where $s\in[\![0,2]\!]$; there are $O(N^{2-s})$ terms corresponding to each $s\in[\![0,2]\!]$. By (3.21), the left-hand side of (3.23) is bounded by

$$O\left(\frac{d}{N}\right)^{3-a}+\sum_{s=1}^{2}O(N^{-s})O\left(\frac{d}{N}\right)^{3-a-s}=O\left(\frac{d}{N}\right)^{3-a}.$$

The proofs of (3.24)–(3.26) are analogous. $\square$

Finally, as a consequence of Lemmas 3.6–3.9 and the Hölder inequality, we obtain the following estimates incorporating an arbitrary function $f(A)$. These and the remainder of the proof of Proposition 3.3 are simplest to state in terms of versions of the seminorms (3.2)–(3.3) for $t = 0$ without rescaling by $(d-1)^{-1/2}$. Thus, instead of (3.2) and (3.3), we use the seminorms

$$\|f\|_r := \left(\mathbb{E}\left|f(A)\right|^r\right)^{1/r}$$

and

$$\|\partial^n f\|_r := \Big\| \sup_{\theta\in[0,1]^n} \sup_{X\in\mathcal{X}^n} \left|\partial_{X_1}\cdots\partial_{X_n} f(\cdot+\theta\cdot X)\right| \Big\|_r.$$

LEMMA 3.10. *Fix $\varepsilon > 0$ and let $r \equiv r(\varepsilon)$ be large enough depending on ε. Let $f \in \mathcal{C}^0(\mathcal{M})$ satisfy $\|f\|_r \le 1$. Then if $|\{i,j\}| = 2-a$ and $|\{i,j,k,l\}| = 4-b$, we have*

$$\frac{1}{N^2}\sum_{m,n}\mathbb{E}\big(A_{ij}A_{mn}f(A)\big) = O\left(\frac{d}{N}\right)^{2-\lfloor a/2\rfloor-\varepsilon}, \tag{3.27}$$

$$\frac{1}{N^4}\sum_{m,n,p,q}\mathbb{E}\big(A_{ij}A_{mn}A_{kl}A_{pq}f(A)\big) = O\left(\frac{d}{N}\right)^{4-\lfloor b/2\rfloor-\varepsilon}, \tag{3.28}$$

$$\frac{1}{N^2}\sum_{m,n}\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})\bar{I}_1 f(A)\big) = O\left(\frac{d}{N}\right)^{3-a-\varepsilon}, \tag{3.29}$$

$$\begin{aligned}\frac{1}{N^4}\sum_{m,n}\sum_{p,q}&\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})(A_{kl}A_{pq}+A_{kp}A_{lq})\bar{I}_{12}f(A)\big)\\ &= O\left(\frac{d}{N}\right)^{5-b-\varepsilon},\end{aligned} \tag{3.30}$$

$$\frac{1}{N^4}\sum_{i,j,m,n}\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})\bar{I}_1 f(A)\big) = O\left(\frac{d}{N}\right)^{3-\varepsilon}, \tag{3.31}$$

$$\begin{aligned}\frac{1}{N^8}\sum_{i,j,m,n}\sum_{k,l,p,q}&\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})(A_{kl}A_{pq}+A_{kp}A_{lq})\bar{I}_{12}f(A)\big)\\ &= O\left(\frac{d}{N}\right)^{5-\varepsilon},\end{aligned} \tag{3.32}$$

where $\bar{I}_1 := 1 - I_1$, $\bar{I}_{12} := 1 - I_{12}$, and the indicator functions I_1 and I_{12} were defined in (3.19)–(3.20).

PROOF. We only prove (3.31); the other estimates are proved similarly and we comment on the differences at the end of the proof. By Hölder's inequality, applied

twice, first to $\mathbb{E}(\cdot)$ and then to the sum over m, n, we obtain from (3.25) that

$$\begin{aligned}
&\frac{1}{N^4}\sum_{i,j,m,n}\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})(1-I_1)f(A)\big)\\
&\quad\le \frac{1}{N^4}\sum_{i,j,m,n}\big[\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})(1-I_1)\big)\big]^{1-1/r}\|f\|_r\\
&\quad\le \bigg(\frac{1}{N^4}\sum_{i,j,m,n}\mathbb{E}\big((A_{ij}A_{mn}+A_{im}A_{jn})(1-I_1)\big)\bigg)^{1-1/r}\|f\|_r\\
&\quad\le O\bigg(\frac{d}{N}\bigg)^{3-3/r}\|f\|_r = O\bigg(\frac{d}{N}\bigg)^{3-\varepsilon}\|f\|_r,
\end{aligned}$$

where we chose r large enough that $3/r \le \varepsilon$.

To prove (3.32), we use (3.26) instead of (3.25), and to prove (3.27)–(3.28) we apply (3.16) instead of (3.21). To prove (3.29)–(3.30), we use (3.23)–(3.24). This completes the proof. $\square$

The next lemma estimates the effect of replacing A_{ij} by its mean d/N, or, equivalently, of conditioning on $\{A_{ij}=1\}$. Since the entries of A are not independent, we use switchings to analyse such a conditioning.

LEMMA 3.11. *Fix $\varepsilon > 0$ and let $r \equiv r(\varepsilon)$ be large enough depending on ε. For any $f \in \mathcal{C}^2(\mathcal{M})$ and any i, j, k, l with $|\{i, j\}| = 2-a$ and $|\{i, j, k, l\}| = 4-b$, we have*

$$\mathbb{E}\bigg(f(A)\bigg(A_{ij}-\frac{d}{N}\bigg)\bigg) = O\bigg(\frac{d}{N}\bigg)^{1-\varepsilon}\|\partial f\|_r + O\bigg(\frac{d}{N}\bigg)^{2-a-\varepsilon}\|f\|_r, \tag{3.33}$$

$$\begin{aligned}
&\mathbb{E}\bigg(f(A)\bigg(A_{ij}-\frac{d}{N}\bigg)\bigg(A_{kl}-\frac{d}{N}\bigg)\bigg)\\
&\qquad = O\bigg(\frac{d}{N}\bigg)^{2-\varepsilon}\|\partial^2 f\|_r + O\bigg(\frac{d}{N}\bigg)^{3-b-\varepsilon}\|f\|_r.
\end{aligned} \tag{3.34}$$

PROOF. We begin with (3.33). Since $A \in \mathcal{M} + d\mathbf{ee}^*$, we have $\sum_{m,n} A_{mn} = Nd$ and $\sum_m A_{im} = \sum_n A_{jn} = d$ for all i, and the left-hand side of (3.33) is therefore equal to

$$\mathbb{E}\bigg(f(A)\bigg(A_{ij}-\frac{d}{N}\bigg)\bigg) = \frac{1}{Nd}\sum_{m,n}\mathbb{E}\big(f(A)(A_{ij}A_{mn}-A_{im}A_{jn})\big). \tag{3.35}$$

Using (3.29), using the notation from (3.19), we therefore find

$$\begin{aligned}
&\mathbb{E}\bigg(f(A)\bigg(A_{ij}-\frac{d}{N}\bigg)\bigg)\\
&\qquad = \frac{1}{Nd}\sum_{m,n}\mathbb{E}\big(f(A)(A_{ij}A_{mn}-A_{im}A_{jn})I_1\big) + O\bigg(\frac{d}{N}\bigg)^{2-a-\varepsilon}\|f\|_r.
\end{aligned}$$

Because of the indicator function I_1, the first term on the right-hand side vanishes unless $a = 0$. Therefore, we may assume that $a = 0$ when estimating it. By Lemma 3.5, and since $I_1(A) = I_1(T_S(A))$ with $S = \{ij, mn\}$, the first term on the right-hand side equals

$$\frac{1}{Nd} \sum_{m,n} \mathbb{E}\big((f(A) - f(A - \xi_{ij}^{mn}))A_{ij}A_{mn}I_1\big). \tag{3.36}$$

The difference of the f's is bounded in absolute value by

$$\sup_{\theta \in [0,1]} \sup_{X \in \mathcal{X}} \big|\partial_X f(A + \theta X)\big|.$$

Hence, (3.27) implies that (3.36) is bounded by

$$O\left(\frac{d}{N}\right)^{1-\varepsilon} \|\partial f\|_r.$$

This completes the proof of (3.33).

The proof of (3.34) is similar. As in (3.35), we write

$$\begin{aligned}\left(A_{ij} - \frac{d}{N}\right)&\left(A_{kl} - \frac{d}{N}\right)\\ &= \frac{1}{(Nd)^2} \sum_{m,n,p,q} (A_{ij}A_{mn} - A_{im}A_{jn})(A_{kl}A_{pq} - A_{kp}A_{lq}).\end{aligned}$$

As above, we write $1 = I_{12} + (1 - I_{12})$ inside the expectation on the left-hand side of (3.34). The second term yields a contribution of order $O(\frac{d}{N})^{3-b-\varepsilon}\|f\|_r$, by (3.30). The first term is zero unless $b = 0$ because of the factor J_{12} in I_{12}. We may therefore assume that $b = 0$ for the estimate of the first term. Using Lemma 3.5, as in (3.36), we find that the first term is equal to

$$\begin{aligned}\frac{1}{(Nd)^2} \sum_{m,n,p,q} \mathbb{E}\big((&f(A) - f(A - \xi_{ij}^{mn}) - f(A - \xi_{kl}^{pq})\\ &+ f(A - \xi_{ij}^{mn} - \xi_{kl}^{pq}))A_{ij}A_{mn}A_{kl}A_{pq}I_{12}\big).\end{aligned} \tag{3.37}$$

The difference of the four f's is bounded in absolute value by

$$\sup_{\theta_1,\theta_2 \in [0,1]} \sup_{X_1,X_2 \in \mathcal{X}} \big|\partial_{X_1}\partial_{X_2} f(A + \theta_1 X_1 + \theta_2 X_2)\big|.$$

By (3.28), we therefore find that (3.37) is bounded in absolute value by

$$O\left(\frac{d}{N}\right)^{2-\varepsilon} \|\partial^2 f\|_r.$$

This completes the proof. □

Finally, with the preparations provided by the previous lemmas, we now complete the proof of Proposition 3.3.

PROOF OF PROPOSITION 3.3. First, note that $I_{ij}^{mn} = A_{ij}A_{mn}I_1$. By Taylor's expansion, and writing $I_1 = 1 + (I_1 - 1)$, we therefore have

$$\begin{aligned} Qf(A) = {} & \frac{1}{8Nd}\sum_{i,j,m,n} A_{ij}A_{mn}\Big(-\partial_{ij}^{mn} f(A) + \frac{1}{2}(\partial_{ij}^{mn})^2 f(A)\Big) \\ & + N^2(R_1 + R_2), \end{aligned} \tag{3.38}$$

where

$$R_1 = O\Big(\frac{N}{d}\Big)\frac{1}{N^4}\sum_{i,j,m,n} A_{ij}A_{mn}(1 - I_1)\sup_{\theta\in[0,1]}\sup_{X\in\mathcal{X}}\big|\partial_X f(A + \theta X)\big|,$$

$$R_2 = O\Big(\frac{N}{d}\Big)\frac{1}{N^4}\sum_{i,j,m,n} A_{ij}A_{mn}\sup_{\theta\in[0,1]^3}\sup_{X\in\mathcal{X}^3}\big|\partial_{X_1}\partial_{X_2}\partial_{X_3} f(A + \theta\cdot X)\big|.$$

By (3.31) and (3.27), respectively, the two error terms are estimated by

$$\mathbb{E}R_1 = O\Big(\frac{d}{N}\Big)^{2-\varepsilon}\|\partial f\|_r, \qquad \mathbb{E}R_2 = O\Big(\frac{d}{N}\Big)^{1-\varepsilon}\|\partial^3 f\|_r.$$

Next, we estimate the main terms in (3.38), which we write as

$$\frac{1}{8Nd}\sum_{i,j,k,l} A_{ij}A_{kl}\Big(-\partial_{ij}^{kl} f(A) + \frac{1}{2}(\partial_{ij}^{kl})^2 f(A)\Big). \tag{3.39}$$

The idea is to write $A_{ij} = \frac{d}{N} + (A_{ij} - \frac{d}{N})$ and likewise for A_{kl}. For the second-order term in (3.39), the term obtained by selecting both factors $\frac{d}{N}$ yields the main contribution. More precisely, we write

$$\frac{1}{16Nd}\sum_{i,j,k,l} A_{ij}A_{kl}(\partial_{ij}^{kl})^2 f(A) = \frac{d}{16N^3}\sum_{i,j,k,l}(\partial_{ij}^{kl})^2 f(A) + N^2(R_3 + R_4),$$

where

$$R_3 = \frac{N}{8d}\frac{1}{N^4}\sum_{i,j,k,l}\Big(((\partial_{ij}^{kl})^2 f(A))\Big(A_{ij} - \frac{d}{N}\Big)\frac{d}{N}\Big),$$

$$R_4 = \frac{N}{16d}\frac{1}{N^4}\sum_{i,j,k,l}\Big(((\partial_{ij}^{kl})^2 f(A))\Big(A_{ij} - \frac{d}{N}\Big)\Big(A_{kl} - \frac{d}{N}\Big)\Big).$$

By (3.33) and (3.34), respectively, with f replaced by $(\partial_{ij}^{kl})^2 f$, we obtain

$$\mathbb{E}(R_3 + R_4) = O\Big(\frac{d}{N}\Big)^{1-\varepsilon}(\|\partial^3 f\|_r + \|\partial^4 f\|_r) + O\Big(\frac{d}{N}\Big)^{2-\varepsilon}\|\partial^2 f\|_r.$$

Next, we estimate the first-order term in (3.39) using a similar argument. Here, the term obtained by selecting both factors $\frac{d}{N}$ from A_{ij} and A_{kl} vanishes because $\sum_{i,j,k,l} \partial_{ij}^{kl} = 0$; the main contribution is given by the mixed term. More precisely, we write

$$\begin{aligned}
&\frac{1}{8Nd} \sum_{i,j,k,l} A_{ij} A_{kl}\, \partial_{ij}^{kl} f(A) \\
&\quad = \frac{d}{8N^3} \sum_{i,j,k,l} \partial_{ij}^{kl} f(A) + \frac{1}{4N^2} \sum_{i,j,k,l} \left(A_{ij} - \frac{d}{N} \right) \partial_{ij}^{kl} f(A) + N^2 R_5 \\
&\quad = \frac{\sqrt{d-1}}{4N^2} \sum_{i,j,k,l} H_{ij}\, \partial_{ij}^{kl} f(A) + N^2 R_5 \\
&\quad = \frac{\sqrt{d-1}}{32N^2} \sum_{i,j,k,l} H_{ij}^{kl}\, \partial_{ij}^{kl} f(A) + N^2 R_5,
\end{aligned}$$

where

$$R_5 = \frac{N}{8d} \frac{1}{N^4} \sum_{i,j,k,l} \left((\partial_{ij}^{kl} f(A)) \left(A_{ij} - \frac{d}{N} \right) \left(A_{kl} - \frac{d}{N} \right) \right).$$

By (3.34), with f replaced by $\partial_{ij}^{kl} f$, we obtain

$$\mathbb{E} R_5 = O\left(\frac{d}{N} \right)^{1-\varepsilon} \|\partial^3 f\|_r + O\left(\frac{d}{N} \right)^{2-\varepsilon} \|\partial f\|_r.$$

We conclude that

$$\begin{aligned}
Qf(A) &= \frac{d}{16N^3} \sum_{i,j,k,l} (\partial_{ij}^{kl})^2 f(A) - \frac{\sqrt{d-1}}{32N^2} \sum_{i,j,k,l} H_{ij}^{kl}\, \partial_{ij}^{kl} f(A) + N^2 \sum_{i=1}^{5} R_i \\
&= \frac{d-1}{16N^3} \sum_{i,j,k,l} (\partial_{ij}^{kl})^2 f(A) - \frac{\sqrt{d-1}}{32N^2} \sum_{i,j,k,l} H_{ij}^{kl}\, \partial_{ij}^{kl} f(A) + N^2 \sum_{i=1}^{6} R_i,
\end{aligned}$$

where we defined

$$R_6 := \frac{1}{16N} \frac{1}{N^4} \sum_{i,j,k,l} (\partial_{ij}^{kl})^2 f(A).$$

Clearly, $\mathbb{E} R_6 = O(\frac{1}{N}) \|\partial^2 f\|_r$.

Using the notation introduced in (3.7), we have $\sqrt{d-1}\, \partial f(A) = \partial F(H)$. Hence, we obtain (3.8) with $R := N^2 \sum_{i=1}^{6} R_i$. The error term R is estimated, using the above estimates on $\mathbb{E} R_i$, as

$$\begin{aligned}
\mathbb{E} R = O(N^{2+\varepsilon}) \Bigg[&\left(\frac{d}{N} \right)^2 (\|\partial f\|_r + \|\partial^2 f\|_r) + \frac{1}{N} \|\partial^2 f\|_r \\
&+ \frac{d}{N} (\|\partial^3 f\|_r + \|\partial^4 f\|_r) \Bigg]
\end{aligned}$$

$$= O(D^{-1/2}N^{1+\varepsilon})\big[\|\partial F\|_{r,0} + D^{-1/2}\|\partial^2 F\|_{r,0} + \|\partial^3 F\|_{r,0} + D^{-1/2}\|\partial^4 F\|_{r,0}\big],$$

as claimed. □

4. Stability of eigenvectors and eigenvalues. In this section we derive basic stability properties for the eigenvalues and eigenvectors of the Dyson Brownian motion $H(t)$. These allow us to deduce estimates on the eigenvalues and eigenvectors of $H(t)$, assuming similar estimates have been proved for $H(0)$.

As discussed in Section 2.1, we consider a general Dyson Brownian motion $H(t)$ on an M-dimensional Hilbert space V, with normalization constant N as in (2.1). For the usual DBM we have $N = M$, while for the constrained DBM we have $M = N - 1$; we always assume that N and M are comparable. We denote by $\lambda_1(t) \geq \cdots \geq \lambda_M(t)$ the eigenvalues of $H(t)$, and by $\mathbf{v}_1(t), \dots, \mathbf{v}_M(t) \in V$ the associated normalized eigenvectors of $H(t)$. Moreover, we define the Stieltjes transform of the empirical spectral measure of $H(t)$ by $s(t; z) := \frac{1}{M}\sum_{i=1}^{M} \frac{1}{\lambda_i(t)-z}$.

Throughout the rest of the paper, we use the following notion of high probability events and high probability bounds introduced in [17].

DEFINITION 4.1. (i) We say that an event Ξ has *high probability* if for every $\zeta > 0$ there exists an $N_0(\zeta) > 0$ such that $\mathbb{P}(\Xi^c) \leq N^{-\zeta}$ for $N \geq N_0(\zeta)$.

(ii) For nonnegative random variables A, B, we write $A \prec B$ or $A = O_{\prec}(B)$ if for any $\zeta > 0$ there exists an $N_0(\zeta)$ such that $\mathbb{P}(A > N^{1/\zeta}B) \leq N^{-\zeta}$ for $N \geq N_0(\zeta)$.

If the event Ξ from (i) and the random variables A and B from (ii) depend on some additional parameter $u \in U$ in some possibly N-dependent set U, we we say that (i) and (ii) hold uniformly in u if $N_0(\zeta)$ does not depend on u.

Throughout the following, the definitions (i) and (ii) will always be uniform in all parameters, such as z, any matrix indices, and deterministic vectors. Note that $\prec$ is compatible with the usual algebraic operations, so that for instance we have $\sum_i A_i \prec \sum_i B_i$ provided that $A_i \prec B_i$ for all i and the size of the index set for i is $N^{O(1)}$.

4.1. *Delocalization of eigenvectors.* The following result shows that if all eigenvectors of $H(0)$ are uniformly delocalized in some direction $\mathbf{q} \in V$, then with high probability they remain delocalized in this direction under the DBM on V, for any time $t > 0$.

LEMMA 4.2. *Suppose that $H(t)$ is the DBM on an M-dimensional space V. Let $\mathbf{q} \in V$ and suppose that* $\max_i |\mathbf{q} \cdot \mathbf{v}_i(0)| \leq B$. *Then, for any $t > 0$, any $i \in [\![1, M]\!]$, and $\xi \gg 1$,*

$$\mathbb{P}\big(|\mathbf{q} \cdot \mathbf{v}_i(t)| \geq \xi B\big) \leq \mathrm{e}^{-\frac{1}{2}\xi^2}. \tag{4.1}$$

In particular,

$$\left|\mathbf{q}\cdot\mathbf{v}_i(t)\right| \prec B. \tag{4.2}$$

Lemma 4.2 is a simple consequence of the *eigenvector moment flow* (EMF) introduced in [10]. Suppose for simplicity that the eigenvalues of $H(0)$ are distinct. Then the eigenvalue process $(\lambda_i(t))$ is almost surely continuous and simple for all $t > 0$; see [10] for more details. We study the dynamics of the eigenvectors $\mathbf{v}_i(t)$ by conditioning on the eigenvalue process; see again [10] for a precise construction. Hence, for the following argument, we condition on $(\lambda_i(t))$ and regard the eigenvalue process as deterministic.

We give the definition of the EMF restricted to moments of a fixed order $p \in \mathbb{N}$. The configuration space is $\Omega_p := \{\eta = (\eta_i)_{i=1}^M \in \mathbb{N}^M : \sum_{i=1}^M \eta_i = p\}$. The configurations $\eta \in \Omega_p$ are interpreted as configurations of p particles on the lattice $[\![1, M]\!]$, whereby a single site of $[\![1, M]\!]$ may be occupied by multiple particles. We denote by $\eta^{i,j} := \eta + \mathbf{1}(\eta_i > 0)(\mathbf{e}_j - \mathbf{e}_i)$ the configuration obtained from η by moving one particle from i to j. The time-dependent generator $R(t)$ of the EMF is defined by

$$\big(R(t)f\big)(\eta) := \sum_{i\neq j} W_{ij}(t) 2\eta_i(1+2\eta_i)\big(f\big(\eta^{i,j}\big) - f(\eta)\big),$$

where

$$W_{ij}(t) := \frac{1}{N(\lambda_i(t)-\lambda_j(t))^2}.$$

For our purposes, the precise form of the coefficients $W_{ij}(t)$ is not important; we only use that they are nonnegative and continuous in t. The p-particle EMF is given by the equation

$$\partial_t f_t(\eta) = \big(R(t)f_t\big)(\eta), \qquad f_0 : \Omega_p \to \mathbb{R} \text{ given.} \tag{4.3}$$

This is a linear (time-dependent) ODE on a finite dimensional vector space, and thus well-posed. It is also easy to see that it is contractive on $L^\infty(\Omega)$ in the sense that $\|f_t\|_{L^\infty(\Omega_p)} \le \|f_0\|_{L^\infty(\Omega_p)}$.

Next, for deterministic $\eta \in \Omega_p$ and $\mathbf{q} \in V$, we define

$$f_t(\eta) := \mathbb{E}\left[\prod_{i=1}^M \frac{1}{(2\eta_i-1)!!}\big(\mathbf{q}\cdot\mathbf{v}_i(t)\big)^{2\eta_i}\Big|\big(\lambda_i(t) : i \in [\![1, M]\!], t \ge 0\big)\right], \tag{4.4}$$

where $n!! := n\cdot(n-2)\cdots 3\cdot 1$ for odd n, and by convention $(-1)!! = 1$. In [10], Theorem 3.1, it is shown that f_t solves (4.3).

REMARK 4.3. In [10], Dyson Brownian motion is defined without the Ornstein–Uhlenbeck drift term in the SDE (2.2), and the SDEs for the eigenvalues

and eigenvectors are stated in [10], Definition 2.2. In the present case, with drift term, the SDEs for eigenvalue and eigenvector flows are given by

$$\mathrm{d}\lambda_i = \frac{\mathrm{d}B_{ii}}{\sqrt{N}} + \frac{1}{N}\sum_{j:j\neq i}\frac{1}{\lambda_i-\lambda_j}\,\mathrm{d}t - \frac{\lambda_i}{2}\,\mathrm{d}t,$$

$$\mathrm{d}\mathbf{v}_i = \frac{1}{\sqrt{N}}\sum_{j:j\neq i}\frac{\mathrm{d}B_{ij}}{\lambda_i-\lambda_j}\mathbf{v}_j - \frac{1}{2N}\sum_{j:j\neq i}\frac{\mathrm{d}t}{(\lambda_i-\lambda_j)^2}\mathbf{v}_i,$$

for $i = 1, 2, \dots, M$, and with $B(t)$ a Brownian motion on the space of $M \times M$ real symmetric matrices with quadratic covariation $\langle B_{ij}, B_{kl}\rangle(t) = (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})t$. Thus, the SDEs for the eigenvectors are the same with or without the drift term. Therefore the arguments of [10], Section 3, apply verbatim in our setting as well.

PROOF OF LEMMA 4.2. Suppose first that $H(0)$ has simple spectrum. Let f_t be the given by (4.4), which solves (4.3) as remarked above. Then, since the EMF (4.3) is a contraction on $L^\infty(\Omega_p)$, we obtain from the assumption of Lemma 4.2 that

$$\max_{\eta\in\Omega_p}|f_t(\eta)| \le \max_{\eta\in\Omega_p}|f_0(\eta)| \le B^{2p}.$$

Therefore, choosing $\eta = p\mathbf{e}_i$, we get

$$\mathbb{E}\big[(\mathbf{q}\cdot\mathbf{v}_i(t))^{2p}\big] = (2p-1)!!\mathbb{E}\big[f_t(\eta)\big] \le (2p-1)!!B^{2p},$$

from which the claim follows. Finally, if $H(0)$ does not have simple spectrum, the same estimate holds by a simple approximation argument using the continuity of the eigenvectors as functions of the matrix. $\square$

4.2. *Stability of eigenvalues*. The following result shows that if the empirical spectral measure at $t = 0$ is close to the semicircle law, this remains true for $t > 0$. For its statement, recall that $s(t, z)$ denotes the Stieltjes transform of the empirical spectral measure of $H(t)$. We denote the Stieltjes transform of the semicircle law by m. It can be characterized as the unique holomorphic function $m : \mathbb{C}_+ \to \mathbb{C}_+$ such that $m^2 + mz + 1 = 0$ and $m(z) \sim 1/z$ as $|z| \to \infty$; see, for example, [2].

LEMMA 4.4. *Suppose that $C^{-1}M \le N \le CM$. Fix $\varepsilon > 0$. If for some $B \le N^{-\varepsilon}$ we have*

$$|s(0;z) - m(z)| \prec B + \frac{1}{(N\eta)^{1/4}} \tag{4.5}$$

uniformly for $z = E + \mathrm{i}\eta$ with $\eta \ge N^{-1+\varepsilon}$, then for any $t \le B$ we have

$$|s(t;z) - m(z)| \prec B + \frac{1}{(N\eta)^{1/4}}, \tag{4.6}$$

uniformly for $z = E + \mathrm{i}\eta$ with $\eta \in [N^{-1+\varepsilon}, 1]$.

PROOF. Define $s_{\mathrm{fc},t}(z)$ as the unique solution $\mathbb{C}_+ \to \mathbb{C}_+$ of the self-consistent equation

$$s_{\mathrm{fc},t}(z) = \frac{1}{M}\sum_{i=1}^{M} \frac{1}{\mathrm{e}^{-t/2}\lambda_i(0) - z - (1-\mathrm{e}^{-t})s_{\mathrm{fc},t}(z)}. \tag{4.7}$$

Thus, $s_{\mathrm{fc},t}(z)$ is the Stieltjes transform of the free convolution of the empirical eigenvalue distribution of $\mathrm{e}^{-t/2}H(0)$ and the semicircle law rescaled by $(1-\mathrm{e}^{-t})^{1/2}$. We refer to [5] for the existence and uniqueness of $s_{\mathrm{fc},t}(z)$ and relative properties on the free convolution with semicircle law.

As in (3.11), we find that $H(t) \overset{\mathrm{d}}{=} \mathrm{e}^{-t/2}H(0) + (1-\mathrm{e}^{-t})^{1/2}W$, where W is the standard Gaussian measure on $\mathcal{H}(V)$ with inner product (2.1). Under the assumptions of the lemma, [33], Corollary 7.11, implies that for $t \le N^{-\varepsilon}$ we have

$$\bigl|s(t;z) - s_{\mathrm{fc},t}(z)\bigr| \prec \frac{1}{(N\eta)^{1/3}} \tag{4.8}$$

uniformly for $z = E + \mathrm{i}\eta$ with $\eta \ge N^{-1+\varepsilon}$. (Note that in [33], the Stieltjes transform is denoted by m_V instead of s, and that $s_{\mathrm{fc},t}$ is denoted $m_{\mathrm{fc},t}$. Moreover, [33], Corollary 7.11, is stated for a diagonal matrix $H(0)$; however, since W is invariant under orthogonal transformations which diagonalize $H(0)$, the results of [33] trivially apply to any symmetric matrix $H(0)$.)

Set $\vartheta_t := 1 - e^{-t} \le t$. Note that the Stieltjes transform of the empirical eigenvalue distribution of $\mathrm{e}^{-t/2}H(0)$ is given by $\mathrm{e}^{t/2}s(0, \mathrm{e}^{t/2}z)$, and that (4.7) can be rephrased as

$$s_{\mathrm{fc},t}(z) = \mathrm{e}^{t/2}s\bigl(0, \mathrm{e}^{t/2}\bigl(z + \vartheta_t s_{\mathrm{fc},t}(z)\bigr)\bigr).$$

For any $z = E + \mathrm{i}\eta$ such that $\eta \ge N^{-1+\varepsilon}$, we have $\operatorname{Im}\mathrm{e}^{t/2}(z + \vartheta_t s_{\mathrm{fc},t}(z)) \ge \operatorname{Im}\mathrm{e}^{t/2}z \ge N^{-1+\varepsilon}$, where we used that $\operatorname{Im} s_{\mathrm{fc},t}(z) > 0$. From the assumption (4.5), we therefore get

$$s_{\mathrm{fc},t}(z) = \mathrm{e}^{t/2}m\bigl(\mathrm{e}^{t/2}\bigl(z + \vartheta_t s_{\mathrm{fc},t}(z)\bigr)\bigr) + O_{\prec}\biggl(B + \frac{1}{(N\eta)^{1/4}}\biggr). \tag{4.9}$$

Next, note that

$$m(z) = \mathrm{e}^{t/2}m\bigl(\mathrm{e}^{t/2}\bigl(z + \vartheta_t m(z)\bigr)\bigr). \tag{4.10}$$

(This may be interpreted as the fact that the semicircle law with variance t is a semigroup with respect to free convolution.) Moreover, from the definition of $m(z)$ it is easy to deduce the continuity estimate

$$\bigl|m(z) - m(w)\bigr| \le 2|z-w|^{1/2}, \tag{4.11}$$

for any $z, w \in \mathbb{C}_+$.

By (4.11), and using that $t = O(1)$, the difference between (4.9) and (4.10) is

$$\begin{aligned}
&\left|s_{\mathrm{fc},t}(z) - m(z)\right| \\
&\quad = \mathrm{e}^{t/2}\left|m\left(\mathrm{e}^{t/2}\left(z + \vartheta_t s_{\mathrm{fc},t}(z)\right)\right) - m\left(\mathrm{e}^{t/2}\left(z + \vartheta_t m(z)\right)\right)\right| \\
&\qquad + O_{\prec}\left(B + \frac{1}{(N\eta)^{1/4}}\right) \\
&\quad \le O\left(t^{1/2}\right)\left|s_{\mathrm{fc},t}(z) - m(z)\right|^{1/2} + O_{\prec}\left(B + \frac{1}{(N\eta)^{1/4}}\right) \\
&\quad \le \max\left\{O\left(t^{1/2}\right)\left|s_{\mathrm{fc},t}(z) - m(z)\right|^{1/2}, O_{\prec}\left(B + \frac{1}{(N\eta)^{1/4}}\right)\right\}.
\end{aligned}$$

Therefore, either $|s_{\mathrm{fc},t}(z) - m(z)| = O(t)$ or $|s_{\mathrm{fc},t}(z) - m(z)| \prec B + (N\eta)^{-1/4}$, and we get

$$\left|s_{\mathrm{fc},t}(z) - m(z)\right| \prec B + \frac{1}{(N\eta)^{1/4}} + t \prec B + \frac{1}{(N\eta)^{1/4}}, \tag{4.12}$$

where we used $t \le B$. Combining (4.8) and (4.12) and using $\eta \le 1$, the claim (4.6) follows. □

5. Proof of Propositions 2.6–2.7. With the preparations provided by Sections 3–4, and using results of [3, 27, 33], we now complete the proofs of Propositions 2.6–2.7. First, recall that $\alpha > 0$ is fixed, and that we always assume $D \ge N^{\alpha}$. We also use the notation $z = E + \mathrm{i}\eta$ for the real and imaginary parts of the spectral parameter $z \in \mathbb{C}_+$.

Throughout this section, $H(t)$ denotes the constrained DBM from Definition 2.2 with $H(0)$ given by (2.4). We use the notation of Section 4 applied to the constrained DBM. In particular,

$$M := N - 1$$

is the dimension of the space $V := \mathbf{e}^{\perp}$.

5.1. *A priori estimates on eigenvalues and eigenvectors.* We begin by collecting some results on the eigenvalues and eigenvectors of $H(t)|_{\mathbf{e}^{\perp}}$ required for the proofs of Propositions 2.6–2.7.

For any $H \in \mathcal{M}$, we denote the eigenvalues of $H|_{\mathbf{e}^{\perp}}$ by $\lambda_1(H) \ge \cdots \ge \lambda_M(H)$, and the corresponding normalized eigenvectors by $\mathbf{v}_1(H), \dots, \mathbf{v}_M(H)$. The components of the eigenvectors in the standard basis on $\mathbb{R}^N$ are denoted $v_k(H; i) := \mathbf{e}_i \cdot \mathbf{v}_k(H)$, $i \in [\![1, N]\!]$, $k \in [\![1, M]\!]$. Moreover, for $H \in \mathcal{M}$, we denote by $G_{ij}(H; z)$ the entries of the Green's function of H restricted to $\mathbf{e}^{\perp}$ in the standard basis of $\mathbb{R}^N$, and by $s(H; z)$ the Stieltjes transform of the empirical spectral measure.

Explicitly,

$$G_{ij}(H;z) := \sum_{k=1}^{M} \frac{v_k(H;i)v_k(H;j)}{\lambda_k(H)-z}, \tag{5.1}$$

$$s(H;z) := \frac{1}{M}\operatorname{Tr} G(H;z) = \frac{1}{M}\sum_{k=1}^{M}\frac{1}{\lambda_k(H)-z}. \tag{5.2}$$

Finally, we set

$$\Gamma(H) \equiv \Gamma(H;z) := \max_{i,j}\bigl|G_{ij}(H;z)\bigr| \vee 1. \tag{5.3}$$

We also recall the definition of the typical location γ_i of the ith eigenvalue from (1.1).

The following proposition summarizes the input we need from the local semicircle law of [3]. The local semicircle law, as proved in [3], only applies for $t=0$, and the extension to $t>0$ is provided by the results of Section 4.

PROPOSITION 5.1. *For any $z \in \mathbb{C}_+$, $i \in [\![1,N]\!]$, $k \in [\![1,M]\!]$, and $0 \le t \le D^{-1/4}$, we have*

$$\bigl|v_k\bigl(H(t);i\bigr)\bigr| \prec N^{-1/2}, \qquad \Gamma\bigl(H(t);z\bigr) \prec 1 + \frac{1}{N\eta}. \tag{5.4}$$

Moreover, for any fixed $\kappa > 0$ and any $i \in [\![\kappa N, (1-\kappa)N]\!]$, we also have

$$\bigl|\lambda_i\bigl(H(t)\bigr) - \gamma_i\bigr| \prec D^{-1/4}. \tag{5.5}$$

PROOF. First, as special cases of [3], Theorem 1.1 and Corollary 1.2, for any $z = E + \mathrm{i}\eta$ with $E \in \mathbb{R}$ and $\eta \ge N^{-1+\varepsilon}$, for arbitrary $\varepsilon > 0$, we have

$$\bigl|s\bigl(H(0);z\bigr) - m(z)\bigr| \prec \frac{1}{D^{1/4}} + \frac{1}{(N\eta)^{1/4}}, \qquad \bigl|v_k\bigl(H(0);i\bigr)\bigr| \prec N^{-1/2}. \tag{5.6}$$

(Note that the local semicircle law from [3] also includes the trivial eigenvalue at 0; it is easy to see that its contribution to s is negligible compared to the error bounds in (5.6).)

Next, we extend these bounds from $t=0$ to $t>0$. For $i \in [\![1,N]\!]$ define $\hat{\mathbf{e}}_i = \mathbf{e}_i - (\mathbf{e}_i \cdot \mathbf{e})\mathbf{e} \in \mathbf{e}^{\perp}$. Since $v_k(H(t);i) = \hat{\mathbf{e}}_i \cdot \mathbf{v}_k(H(t))$, from (5.6) and Lemma 4.2, applied to the constrained DBM with $\mathbf{q} = \hat{\mathbf{e}}_i$, we find $|v_k(H(t);i)| \prec N^{-1/2}$, for any $t > 0$. Similarly, for $t \le D^{-1/4}$, the extension of the bound on the Stieltjes transform follows immediately from Lemma 4.4 with $B = D^{-1/4}$. Summarizing, for any $\eta \ge N^{-1+\varepsilon}$ and $0 \le t \le D^{-1/4}$, we have

$$\bigl|s\bigl(H(t);z\bigr) - m(z)\bigr| \prec \frac{1}{D^{1/4}} + \frac{1}{(N\eta)^{1/4}}, \qquad \bigl|v_k\bigl(H(t);i\bigr)\bigr| \prec N^{-1/2}. \tag{5.7}$$

This proves the first estimate of (5.4).

In order to prove the second estimate of (5.4), we use a dyadic decomposition (see, e.g., [25], (8.2)) to obtain, for any matrix $H \in \mathcal{M}$,

$$|G_{ij}(z)| \le \sum_{k=1}^{M} \frac{|v_k(i)v_k(j)|}{|\lambda_k - E + \mathrm{i}\eta|} \le 4N \max_{k,l} |v_k(l)|^2 \left(1 + \sum_{n=0}^{\lceil \log_2 \eta^{-1} \rceil} \operatorname{Im} s(E + \mathrm{i}2^n \eta)\right).$$

We apply this estimate to the matrix $H(t)$. By (5.7), we have $\max_{k,l} |v_k(l)|^2 \prec 1/N$. Moreover, since $\eta \operatorname{Im} s(E + \mathrm{i}\eta)$ is increasing in η [as may be easily seen from the right-hand side of (5.2)], and since $|m| \le 1$, the first bound in (5.7) implies $\operatorname{Im} s(z) \prec 1 + 1/(N\eta)$ for any $\eta > 0$, and thus $\operatorname{Im} s(E + \mathrm{i}2^n \eta) \prec 1 + 2^{-n}/N\eta$. For $\eta \ge 1/N^{O(1)}$, we then have $\log \eta^{-1} \prec 1$ and obtain $\Gamma(z) \prec 1$ as desired. For arbitrary $\eta > 0$, the claim then follows by [3], Lemma 2.1. [In fact, we shall only need (5.4) with $\eta \ge 1/N^{O(1)}$.]

Finally, we deduce (5.5) from the bound on the Stieltjes transform in (5.7). We abbreviate $\lambda_k \equiv \lambda_k(H(t))$, and denote by

$$\varrho(I) := \int_I \varrho(x)\,\mathrm{d}x, \qquad \nu(I) := \frac{1}{M} \sum_{k=1}^{M} \mathbf{1}(\lambda_k \in I)$$

the semicircle and empirical spectral measures, respectively, applied to an interval I. Then, following a standard application of the Helffer–Sjöstrand functional calculus along the lines of [22], Section 8.1, we find from (5.7) and $D \le N$ that for any interval $I \subseteq [-3, 3]$ we have

$$|\nu(I) - \varrho(I)| \prec \frac{1}{D^{1/4}} + \frac{1}{N^{1/4}} \prec \frac{1}{D^{1/4}}. \tag{5.8}$$

(We note that previously (5.8) for $t = 0$ was given in [3], Corollary 1.3.) Using (5.8), we may estimate $\lambda_i - \gamma_i$ as follows. By (5.8) applied to $I = [-3, 3]$, we find that there are at most $O_\prec(ND^{-1/4})$ eigenvalues outside $[-3, 3]$. Defining $f(E) := \varrho([E, \infty))$, we therefore find from (1.1) and (5.8) that

$$\begin{aligned} f(\gamma_i) = \frac{i}{N} &= \nu([\lambda_i, \infty)) + O\left(\frac{1}{N}\right) = \nu([\lambda_i, 3)) + O_\prec\left(\frac{1}{D^{1/4}}\right) \\ &= \varrho([\lambda_i, 3)) + O_\prec\left(\frac{1}{D^{1/4}}\right) = f(\lambda_i) + O_\prec\left(\frac{1}{D^{1/4}}\right). \end{aligned}$$

Since $i \in [\![\kappa N, (1 - \kappa N)]\!]$, we have $|f'| \ge c > 0$ in a neighbourhood of γ_i, and we therefore get (5.5). This completes the proof. $\square$

The next result shows that the suprema in (3.3) do not essentially change the size of Γ.

COROLLARY 5.2. *Fix $n \in \mathbb{N}$. For any $z \in \mathbb{C}_+$ and $0 \le t \le D^{-1/4}$, we have*

$$\sup_{\theta \in [0,1]^n} \sup_{X \in \mathcal{X}^n} \Gamma\big(H(t) + (d-1)^{-1/2}\theta \cdot X; z\big) \prec 1 + \frac{1}{N\eta}. \tag{5.9}$$

Moreover, for any $i \in [\![1, N]\!]$ *and* $k \in [\![1, M]\!]$, *we have*

$$\sup_{\theta \in [0,1]^n} \sup_{X \in \mathcal{X}^n} \left| v_k\big(H(t) + (d-1)^{-1/2}\theta \cdot X; i\big) \right| \prec N^{-1/2}. \tag{5.10}$$

PROOF. We abbreviate $H \equiv H(t)$. Without loss of generality, by an argument analogous to [3], Lemma 2.1, we may assume that $\eta \geq 1/N$. Hence, by (5.4), we have $\Gamma(H; z) \prec 1$. It therefore suffices to show that if $\Gamma(H; z) \leq (d-1)^{1/2}/(16n)$ then for any $\theta \in [0, 1]^n$ and $X \in \mathcal{X}^n$ we have

$$\Gamma\big(H + (d-1)^{-1/2}\theta \cdot X; z\big) \leq 2\Gamma(H; z). \tag{5.11}$$

To show (5.11), we use the resolvent identity to obtain (omitting the argument z for brevity)

$$\begin{aligned}
&\left|G_{ij}\big(H + (d-1)^{-1/2}\theta \cdot X\big)\right| \\
&\quad = \left|G_{ij}(H) - (d-1)^{-1/2}\big(G(H)(\theta \cdot X)G\big(H + (d-1)^{-1/2}\theta \cdot X\big)\big)_{ij}\right| \\
&\quad \leq \Gamma(H) + 8n(d-1)^{-1/2}\Gamma(H)\Gamma\big(H + (d-1)^{-1/2}\theta \cdot X\big) \\
&\quad \leq \Gamma(H) + \Gamma\big(H + (d-1)^{-1/2}\theta \cdot X\big)/2.
\end{aligned}$$

Taking the maximum over i and j yields (5.11). Finally, (5.10) follows from (5.9), as in the proof of [3], Corollary 1.2. □

Note that since $G_{ij}(H; \bar{z}) = \overline{G_{ij}(H; z)}$, the estimates (5.4) and (5.9) for Γ also hold with $\eta < 0$ if η is replaced by $|\eta|$ on the right-hand sides. We shall use this tacitly in the following.

5.2. *Proof of Proposition 2.6*: *Eigenvalue correlation functions*. We now prove that the locally averaged local correlation functions of the matrix $H(0)|_{\mathbf{e}^\perp}$ converge to those of $H(t)|_{\mathbf{e}^\perp}$ for times $t \leq N^{-1-\delta}D^{1/2}$. The main ingredient of the proof is the following lemma comparing functions of Green's functions with spectral parameter η slightly smaller than $1/N$. Its proof follows easily from Proposition 3.1 and Lemma 5.1. For random matrices with independent entries, analogous results were previously proved by the Green's function comparison theorem [25], and by direct analysis of the evolution of the matrix entries under Dyson Brownian motion [10]. We also remark that, in [44], eigenvalues are compared directly without involving the Green's function.

LEMMA 5.3. *Fix* $n \in \mathbb{N}$, *and let* $\kappa, \gamma, \delta > 0$ *be sufficiently small. Then the following holds for any* $\eta \in [N^{-1-\gamma}, N^{-1}]$, *any sequence of positive integers* $k_1, k_2, \dots, k_n$, *any set of complex parameters* $z_j^m = E_j^m \pm \mathrm{i}\eta$, *where* $j \in [\![1, k_m]\!]$, $m \in [\![1, n]\!]$, $|E_j^m| \leq 2 - \kappa$, *and the* $\pm$ *signs are arbitrary. Let* $\phi \in C^\infty(\mathbb{R}^n)$ *be a test*

function such that for any multi-index $m = (m_1, \dots, m_n)$ with $1 \le |m| \le 4$ and for any $\omega > 0$ sufficiently small,

$$\max\{|\partial^m \phi(x)| : |x| \le N^{\omega}\} \le N^{O(\omega)}, \tag{5.12}$$

$$\max\{|\partial^m \phi(x)| : |x| \le N^{2}\} \le N^{O(1)}. \tag{5.13}$$

Then, with the notation $G_1(z) := G(H(0); z)$ and $G_2(z) := G(H(t); z)$, for any $t \le D^{1/2} N^{-1-\delta}$, we have

$$\begin{aligned} &\left| \mathbb{E}\phi\left(N^{-k_1} \operatorname{Tr}\left(\prod_{j=1}^{k_1} G_1(z_j^1) \right), \dots, N^{-k_n} \operatorname{Tr}\left(\prod_{j=1}^{k_n} G_1(z_j^n) \right) \right) - \mathbb{E}\phi(G_1 \to G_2) \right| \\ &\qquad = O\big(N^{-\delta/2 + O(\gamma)}\big). \end{aligned} \tag{5.14}$$

Here, $\phi(G_1 \to G_2)$ is the expression obtained from the one to its left by replacing G_1 with G_2. The implicit constants depend on $n, k_1, \dots, k_n, m_1, \dots, m_n$, and the constants in (5.12)–(5.13).

PROOF. For simplicity of notation, we show (5.14) only for $n = 1$ and $k_1 = 1$; the general case is analogous. We then write z instead of z_1^1. To show the claim, it then suffices to show that

$$\begin{aligned} &\big|\mathbb{E}\phi\big(N^{-1} \operatorname{Tr} G\big(H(t); z\big)\big) - \mathbb{E}\phi\big(N^{-1} \operatorname{Tr} G\big(H(0); z\big)\big)\big| \\ &\qquad = O\big(t D^{-1/2} N^{1+\delta/2} N^{O(\gamma)}\big). \end{aligned} \tag{5.15}$$

Set $F(H) := \phi(N^{-1} \operatorname{Tr} G(H; z))$. We claim that if r and n are fixed (arbitrarily, independently of N), and if $t \le D^{-1/4}$, for any sufficiently large N (depending on r, n, δ), we have

$$\sup_{0 \le s \le t} \|\partial^n F\|_{r,s} \le N^{\delta/4 + O(\gamma)}. \tag{5.16}$$

Given (5.16), Proposition 3.1 with $\varepsilon = \delta/4$ yields (5.15).

Thus, it only remains to show (5.16). Recall that the derivative of the Green's function in the direction of a matrix $X \in \mathcal{X}$ is given by $\partial_X G = -GXG$ (using that elements in $\mathcal{X}$ act on $\mathbf{e}^{\perp}$). Therefore, by the Leibniz rule, for any $X_1, \dots, X_n \in \mathcal{X}$ and any $H \in \mathcal{M}$, we have

$$\partial_{X_1} \cdots \partial_{X_n} G = (-1)^n \sum_{\sigma \in S_n} G X_{\sigma(1)} G \cdots G X_{\sigma(n)} G,$$

where S_n is the set of permutations of n elements, and we omit the dependence on H on both sides in our notation. Since (with respect to the standard basis of $\mathbb{R}^N$) each $X \in \mathcal{X}$ has at most 8 nonvanishing entries, and since these are in $\{\pm 1\}$, by definition of Γ it follows that

$$\big|N^{-1} \operatorname{Tr} \partial_{X_1} \cdots \partial_{X_n} G\big| \le N^{-1} \sum_{i=1}^{N} n! \max_{\sigma \in S_n} \big|(G X_{\sigma(1)} G \cdots G X_{\sigma(n)} G)_{ii}\big| \le 8^n n! \Gamma^{n+1}.$$

From this and the chain rule, we obtain that there exist constants C_n such that

$$\left|\partial_{X_1}\cdots\partial_{X_n}\phi\left(N^{-1}\operatorname{Tr}G\right)\right| \le C_n\Gamma^{2n}\max_{0\le m\le n}\left|\phi^{(m)}\right|. \tag{5.17}$$

By Corollary 5.2 and since $|\eta| \ge N^{-1-\gamma}$, we have $\sup_{\theta\in[0,1]^n}\sup_{X\in\mathcal{X}^n}\Gamma(H(s)+(d-1)^{-1/2}\theta\cdot X) \prec N^{\gamma}$, for any $0\le s\le t$. For $n\le 4$, by assumption (5.12) and (5.17), therefore,

$$\begin{aligned}\sup_{\theta\in[0,1]^n}\sup_{X\in\mathcal{X}^n}&\left|\partial_{X_1}\cdots\partial_{X_n}\phi\left(N^{-1}\operatorname{Tr}G\left(H(s)\right.\right.\right.\\ &\left.\left.\left.+(d-1)^{-1/2}\theta\cdot X\right)\right)\right| \prec N^{O(\gamma)}.\end{aligned} \tag{5.18}$$

On the complement of the high-probability event of $\prec$ in (5.18), we use the trivial bound $\Gamma \le \eta^{-1} \le N^{1+\gamma}$ and (5.13). We obtain

$$\begin{aligned}\sup_{\theta\in[0,1]^n}\sup_{X\in\mathcal{X}^n}&\left|\partial_{X_1}\cdots\partial_{X_n}\phi\left(N^{-1}\operatorname{Tr}G\left(H(s)+(d-1)^{-1/2}\theta\cdot X\right)\right)\right|\\ &\le C_n\eta^{-2n}N^{O(1)} \le N^{O(1)},\end{aligned} \tag{5.19}$$

for any $0\le s\le t$. By combining the estimates (5.18)–(5.19), for any constant $r = O(1)$, we have

$$\left\|\partial^n F\right\|_{r,s} \le N^{1/\zeta+O(\gamma)} + N^{-\zeta/r+O(1)} \le N^{\delta/4+O(\gamma)}, \tag{5.20}$$

where ζ is as in Definition 4.1 and chosen sufficiently large, depending on r. This completes the proof. □

The following lemma is essentially [25], Theorem 6.4. It transforms the statement about the Green's function of Lemma 5.3 to a statement about the local correlation functions.

LEMMA 5.4. *Consider two random matrix ensembles H_1 and H_2 with Green's functions $G_1(z)$ and $G_2(z)$. Suppose that, for all ϕ and parameters as in the statement of Lemma 5.3, the estimate (5.14) holds. Then the local bulk eigenvalue correlation functions of H_1 and H_2 coincide.*

PROOF OF PROPOSITION 2.6: CORRELATION FUNCTIONS. The proof follows directly by combining Lemmas 5.3–5.4, with δ given as in the assumption of Proposition 2.6. □

5.3. *Proof of Proposition 2.6: Eigenvalue gap statistics.* To prove that the eigenvalue gap statistics are stable for short times, we require a weak level repulsion estimate. Such an estimate was derived in [27], Theorem 4.1, for sparse matrices with independent entries, using a level repulsion estimate for $t \ge N^{-1+c}$ established in [33]. Here, we adapt the proof of [27], Theorem 4.1, to random regular graphs. The nontrivial dependence is dealt with by Proposition 3.1.

If $\lambda_i(H)$ is a simple eigenvalue of $H|_{\mathbf{e}^\perp}$, we define

$$Q_i(H) := \frac{1}{N^2} \sum_{j:j\neq i, j\leq M} \frac{1}{(\lambda_j(H) - \lambda_i(H))^2}, \tag{5.21}$$

and extend this definition by $Q_i(H) := \infty$ if $\lambda_i(H)$ is not a simple eigenvalue. This quantity plays an important role in [44], where it is observed that it captures the singularities of the derivatives of $\lambda_i(H)$. In [27], it is found that Q_i is stable under DBM and can thus be used to show weak level repulsion from such an estimate for larger times (when a Gaussian component is present).

PROPOSITION 5.5 (Level repulsion). *Fix $\kappa > 0$. Then for any sufficiently small $\tau > 0$, any $i \in [\![\kappa N, (1-\kappa)N]\!]$, and any $s \geq 0$, we have*

$$\mathbb{P}\big(Q_i\big(H(s)\big) \geq N^{2\tau}\big) = O\big(N^{-\tau/2}\big). \tag{5.22}$$

In particular,

$$\mathbb{P}\big(\lambda_i\big(H(s)\big) - \lambda_{i+1}\big(H(s)\big) \leq N^{-1-\tau}\big) = O\big(N^{-\tau/2}\big). \tag{5.23}$$

PROOF. The proof is analogous to that of [27], Theorem 4.1, with $H|_{\mathbf{e}^\perp}$ instead of H. We here focus on the differences. These result from the replacement of [27], Lemma 4.3, by Proposition 3.1, which takes into account the nontrivial correlation structure of the random regular graph. As in [27], if $\lambda_i(H)$ is a simple eigenvalue of $H|_{\mathbf{e}^\perp}$, we define the matrix

$$\begin{aligned} R_i(H) &:= \sum_{j:j\neq i, j\leq M} \frac{1}{\lambda_i(H) - \lambda_j(H)} \mathbf{v}_j(H)\mathbf{v}_j(H)^* \\ &= \frac{1}{2\pi i} \oint_{|z-\lambda_i(H)|=\omega} \frac{G(H;z)}{\lambda_i(H) - z}\, \mathrm{d}z, \end{aligned}$$

where ω is chosen such that the contour $|z - \lambda_i(H)| = \omega$ encloses only $\lambda_i(H)$. Then we have

$$Q_i(H) = \frac{1}{N^2} \operatorname{Tr}\big(R_i(H)^2\big).$$

Given $\tau > 0$, define a cutoff function χ satisfying the following two properties: (1) χ is smooth, and the first four derivatives are bounded, that is, $|\chi^{(k)}(x)| = O(1)$, for $k = 1, 2, 3, 4$; (2) On the interval $[0, N^{2\tau}]$, $|\chi(x) - x| \leq 1$, and for $x \geq N^{2\tau}$, $\chi(x) = N^{2\tau}$. Then $\chi \circ Q_i$ extends to a smooth function on the space of symmetric matrices.

The proof of (5.22) consists of three steps. The first step is the estimate

$$\mathbb{E}\big[\chi\big(Q_i\big(H(s)\big)\big)\big] = O\big(N^{3\tau/2}\big), \tag{5.24}$$

for $s \geq t := N^{-1+c}$. This estimate follows from [33], Theorem 3.6, whose assumptions are satisfied with high probability for the random d-regular graph by Proposition 5.1. In particular, independence of the entries of H is not used.

In the second step, we derive the comparison estimate

$$\left|\mathbb{E}\left[\chi\left(Q_i\left(H(t)\right)\right)\right] - \mathbb{E}\left[\chi\left(Q_i\left(H(s)\right)\right)\right]\right| \leq 1, \tag{5.25}$$

for $s \in [0, t]$. Instead of using [27], Lemma 4.3, which requires that the entries of the random matrix $H(s)$ are independent, we use Proposition 3.1, which takes into account the nontrivial correlation structure of the random regular graph. By Proposition 3.1 with $F(H) := \chi(Q_i(H))$, it suffices to bound

$$\begin{aligned} \left\|\partial^n F\right\|_{r,s} = \mathbb{E}\Big[\sup_{\theta\in[0,1]^n} \sup_{X\in\mathcal{X}^n} \big|\partial_{X_1}\partial_{X_2}\cdots\partial_{X_n} F\big(H(s) \\ + (d-1)^{-1/2}\theta\cdot X\big)\big|^r\Big]^{1/r}, \end{aligned} \tag{5.26}$$

for any (large) fixed integer r and $n = 1, 2, 3, 4$. To this end, the computation of the proof of [27], Proposition 4.6, applies, by simply replacing the derivatives $\partial_{ab}^{(n)}$ by $\partial_{X_1}\cdots\partial_{X_n}$ with $X_l \in \mathcal{X}$. Here, the formulas (4.16)–(4.18) in [27] remain valid after replacing V by the X_l appropriately, and similarly the formula below [27], (4.18), remains valid after replacing V_{ij} by $\mathbf{v}_i^*(H)X_l\mathbf{v}_j(H)$. Moreover, an analogous formula holds for $n = 4$; see, for example, [39], page 8. The same formulas are valid with H replaced by $H + (d-1)^{-1/2}\theta\cdot X$. Since the X_l have only 8 nonvanishing entries (in the standard basis on $\mathbb{R}^N$), and these are equal to ± 1, Corollary 5.2 then implies

$$\begin{aligned} \sup_{\theta\in[0,1]^n} \sup_{X\in\mathcal{X}^n} \big|\mathbf{v}_i^*\big(H(s) + (d-1)^{-1/2}\theta\cdot X\big)X_l\mathbf{v}_j\big(H(s) \\ + (d-1)^{-1/2}\theta\cdot X\big)\big| \prec N^{-1} \end{aligned}$$

for any $s \in [0, t]$. As in the proof of [27], Proposition 4.6, we therefore get

$$\sup_{\theta\in[0,1]^n} \sup_{X\in\mathcal{X}^n} \big|\partial_{X_1}\partial_{X_2}\cdots\partial_{X_n} F\big(H(s) + (d-1)^{-1/2}\theta\cdot X\big)\big| \prec N^{(n+2)\tau}.$$

From this, bounding (5.26) as in (5.20), we obtain

$$\left\|\partial^n F\right\|_{r,s} \leq N^{c+(n+2)\tau}, \tag{5.27}$$

for arbitrarily small $c > 0$ and N large enough. Then (5.25) follows from Proposition 3.4 since $O(tD^{-1/2}N)N^{c+6\tau} \leq O(N^{-\alpha/2+2c+6\tau}) \leq 1$ for $t \leq N^{-1+c}$ and $D \geq N^{\alpha}$, by choosing c and τ sufficiently small.

In the last step, we combine (5.24) and (5.25), and thus obtain

$$\mathbb{E}\left[\chi\left(Q_i\left(H(s)\right)\right)\right] = O\left(N^{3\tau/2}\right),$$

for any $s \geq 0$. Then (5.22) follows easily by Markov's inequality and the definition of χ. $\square$

PROOF OF PROPOSITION 2.6: GAP STATISTICS. Throughout the proof, we use the abbreviation $\lambda_i(t) \equiv \lambda_i(H(t))$. Fix $\kappa > 0$, $\delta > 0$, and $t \le N^{-1-\delta} D^{1/2}$. Since $\varrho(\gamma_i)$ is bounded above and below for $i \in [\![\kappa N, (1-\kappa)N]\!]$, it suffices to prove (2.14) with $\varrho(\gamma_i)$ replaced by 1. Moreover, for any $n \in \mathbb{N}$ and $\phi \in C^\infty(\mathbb{R}^n)$ with bounded first four derivatives, it suffices to show the stronger claim

$$\mathbb{E}\phi\big(N\lambda_i(0), \dots, N\lambda_{i+n}(0)\big) = \mathbb{E}\phi\big(N\lambda_i(t), \dots, N\lambda_{i+n}(t)\big) + o(1) \tag{5.28}$$

as $N \to \infty$, uniformly in $i \in [\![\kappa N, (1-\kappa)N]\!]$. For simplicity of notation, we only prove (5.28) for $n = 1$; the general case is analogous and we comment on the differences at the end of the proof. Thus, for any $i \in [\![\kappa N, (1-\kappa)N]\!]$ and $\phi \in C^\infty(\mathbb{R})$ with bounded first four derivatives, we show

$$\mathbb{E}\big[\phi\big(N\lambda_i(0)\big)\big] - \mathbb{E}\big[\phi\big(N\lambda_i(t)\big)\big] = o(1). \tag{5.29}$$

Given a small constant $\tau > 0$, we choose a cutoff function ρ such that $\rho(x) = 1$ for $x \le N^{2\tau}$ and $\rho(x) = 0$ for $x \ge 2N^{2\tau}$. Using (5.22), we can first remove a bad event on which Q_i is large:

$$\begin{aligned}
&\big|\mathbb{E}\big[\phi\big(N\lambda_i(0)\big)\big] - \mathbb{E}\big[\phi\big(N\lambda_i(t)\big)\big]\big| \\
&\quad \le \big|\mathbb{E}\big[\phi\big(N\lambda_i(0)\big)\rho\big(Q_i\big(H(0)\big)\big)\big] - \mathbb{E}\big[\phi\big(N\lambda_i(t)\big)\rho\big(Q_i\big(H(t)\big)\big)\big]\big| \\
&\qquad + \|\phi\|_\infty\big(\mathbb{P}\big(Q_i\big(H(0)\big) \ge N^{2\tau}\big) + \mathbb{P}\big(Q_i\big(H(t)\big) \ge N^{2\tau}\big)\big) \\
&\quad \le \big|\mathbb{E}\big[\phi\big(N\lambda_i(0)\big)\rho\big(Q_i\big(H(0)\big)\big)\big] - \mathbb{E}\big[\phi\big(N\lambda_i(t)\big)\rho\big(Q_i\big(H(t)\big)\big)\big]\big| \\
&\qquad + O\bigg(\frac{\|\phi\|_\infty}{N^{\tau/2}}\bigg).
\end{aligned}$$

To estimate the right-hand side, we apply Proposition 3.1 with $F(H) := \phi(N\lambda_i(H))\rho(Q_i(H))$. By an argument analogous to that used to obtain (5.27), for any r and $n = 1, 2, 3, 4$, we find the bound

$$\big\|\partial^n F\big\|_{r,s} \le N^{c+O(\tau)} \tag{5.30}$$

for arbitrarily small $c > 0$ (and N sufficiently large). More precisely, by the product rule, the derivatives act either on $\phi(N\lambda_i)$ or $\rho \circ Q_i$. In the bound of any of these derivatives, by definition of ρ, we can assume that $Q_i \le 2N^{2\tau}$. Then the derivatives of $\rho \circ Q_i$ are bounded exactly as in the proof of Proposition 5.5. For the derivatives of $\phi(N\lambda_i)$, by the chain rule and since ϕ is smooth, it suffices to bound the derivatives of the eigenvalues λ_i. This is again done similarly to the bounds on the derivatives of Q_i. Indeed, the derivatives of the eigenvalues can be expressed in terms of the eigenvalues and eigenvectors as done in [27], (4.16)–(4.18) (and with [39], page 8, for $n = 4$). The latter expressions are bounded using the delocalization of eigenvectors (5.4), and using that

$$\sum_{j:j\ne i} \frac{1}{|\lambda_i(s) - \lambda_j(s)|} \prec N Q_i^{1/2}\big(H(s)\big),$$

$$\sum_{j:j\neq i} \frac{1}{|\lambda_i(s)-\lambda_j(s)|^k} \le N^k Q_i^{k/2}\big(H(s)\big),$$

as in [27], (4.11)–(4.12).

As a consequence of Proposition 3.1 and (5.30), with $t \le N^{-1-\delta} D^{1/2}$, we finally obtain

$$\big|\mathbb{E}\big[\phi\big(N\lambda_i(0)\big)\rho\big(Q_i\big(H(0)\big)\big)\big] - \mathbb{E}\big[\phi\big(N\lambda_i(t)\big)\rho\big(Q_i\big(H(t)\big)\big)\big]\big| = O\big(N^{c+O(\tau)-\delta}\big),$$

and (5.29) then follows by taking c and τ small enough that $c + O(\tau) < \delta$.

In the general case of a test function $\phi(N\lambda_i, \dots, N\lambda_{i+n})$, we use the product of cutoff functions $(\rho \circ Q_i) \cdots (\rho \circ Q_{i+n})$ instead of $\rho \circ Q_i$, and proceed otherwise analogously. $\square$

5.4. *Proof of Proposition* 2.7. Given the estimates (5.4)–(5.5), the same argument as in [27], Section 3, applies.

Acknowledgments. We thank Ben Landon for helpful discussions, Peter Sarnak for informing us about [28, 36] and Michael Aizenman for bringing the works [30, 31] to our attention.

REFERENCES

[1] AJANKI, O., ERDŐS, L. and KRÜGER, T. (2016). Local eigenvalue statistics for random matrices with general short range correlations. Available at arXiv:1604.08188.

[2] ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge.

[3] BAUERSCHMIDT, R., KNOWLES, A. and YAU, H.-T. (2015). Local semicircle law for random regular graphs. Preprint. Available at arXiv:1503.08702.

[4] BEKERMAN, F., FIGALLI, A. and GUIONNET, A. (2015). Transport maps for β-matrix models and universality. *Comm. Math. Phys.* **338** 589–619. MR3351052

[5] BIANE, P. (1997). On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.* **46** 705–718. MR1488333

[6] BLEHER, P. and ITS, A. (1999). Semiclassical asymptotics of orthogonal polynomials, Riemann–Hilbert problem, and universality in the matrix model. *Ann. of Math.* (2) **150** 185–266. MR1715324

[7] BOHIGAS, O., GIANNONI, M. J. and SCHMIT, C. (1984). Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Phys. Rev. Lett.* **52** 1–4.

[8] BOURGADE, P., ERDŐS, L., YAU, H.-T. and YIN, J. (2016). Fixed energy universality for generalized Wigner matrices. *Comm. Pure Appl. Math.* **69** 1815–1881. MR3541852

[9] BOURGADE, P., ERDŐS, L. and YAU, H.-T. (2014). Universality of general β-ensembles. *Duke Math. J.* **163** 1127–1190. MR3192527

[10] BOURGADE, P. and YAU, H.-T. (2017). The eigenvector moment flow and local quantum unique ergodicity. *Comm. Math. Phys.* 1–48. MR3606475

[11] CHE, Z. (2017). Universality of random matrices with correlated entries. *Electron. J. Probab.* **22** Paper No. 30.

[12] COOK, N. A. (2014). On the singularity of adjacency matrices for random regular digraphs. Preprint. Available at arXiv:1411.0243.

[13] COOPER, C., DYER, M. and GREENHILL, C. (2007). Sampling regular graphs and a peer-to-peer network. *Combin. Probab. Comput.* **16** 557–593. MR2334585

[14] DEIFT, P., KRIECHERBAUER, T., MCLAUGHLIN, K. T.-R., VENAKIDES, S. and ZHOU, X. (1999). Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.* **52** 1335–1425. MR1702716

[15] DUMITRIU, I. and PAL, S. (2012). Sparse regular random graphs: Spectral density and eigenvectors. *Ann. Probab.* **40** 2197–2235. MR3025715

[16] DYSON, F. J. (1962). A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.* **3** 1191–1198. MR0148397

[17] ERDŐS, L., KNOWLES, A. and YAU, H.-T. (2013). Averaging fluctuations in resolvents of random band matrices. *Ann. Henri Poincaré* **14** 1837–1926. MR3119922

[18] ERDŐS, L., KNOWLES, A., YAU, H.-T. and YIN, J. (2012). Spectral statistics of Erdős–Rényi graphs II: Eigenvalue spacing and the extreme eigenvalues. *Comm. Math. Phys.* **314** 587–640. MR2964770

[19] ERDŐS, L., RAMÍREZ, J. A., SCHLEIN, B. and YAU, H.-T. (2010). Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation. *Electron. J. Probab.* **15** 526–603. MR2639734

[20] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2011). Universality of random matrices and local relaxation flow. *Invent. Math.* **185** 75–119. MR2810797

[21] ERDŐS, L. and YAU, H.-T. (2012). Universality of local spectral statistics of random matrices. *Bull. Amer. Math. Soc.* (*N.S.*) **49** 377–414. MR2917064

[22] ERDŐS, L., KNOWLES, A., YAU, H.-T. and YIN, J. (2013). Spectral statistics of Erdős–Rényi graphs I: Local semicircle law. *Ann. Probab.* **41** 2279–2375. MR3098073

[23] ERDŐS, L., PÉCHÉ, S., RAMÍREZ, J. A., SCHLEIN, B. and YAU, H.-T. (2010). Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.* **63** 895–925. MR2662426

[24] ERDŐS, L. and YAU, H.-T. (2015). Gap universality of generalized Wigner and β-ensembles. *J. Eur. Math. Soc.* (*JEMS*) **17** 1927–2036. MR3372074

[25] ERDŐS, L., YAU, H.-T. and YIN, J. (2012). Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields* **154** 341–407. MR2981427

[26] HUANG, J. and LANDON, B. (2015). Spectral statistics of sparse Erdős–Rényi graph Laplacians. Preprint. Available at arXiv:1510.06390.

[27] HUANG, J., LANDON, B. and YAU, H.-T. (2015). Bulk universality of sparse random matrices. *J. Math. Phys.* **56** 123301. MR3429490

[28] JAKOBSON, D., MILLER, S. D., RIVIN, I. and RUDNICK, Z. (1999). Eigenvalue spacings for regular graphs. In *Emerging Applications of Number Theory* (*Minneapolis*, *MN*, 1996). *IMA Vol. Math. Appl.* **109** 317–327. Springer, New York. MR1691538

[29] JOHANSSON, K. (2001). Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.* **215** 683–705. MR1810949

[30] JOYNER, C. H. and SMILANSKY, U. (2015). Dyson's Brownian-motion model for random matrix theory—revisited. With an Appendix by Don Zagier. Preprint. Available at arXiv:1503.06417.

[31] JOYNER, C. H. and SMILANSKY, U. (2015). Spectral statistics of Bernoulli matrix ensembles—A random walk approach (I). *J. Phys. A* **48** 255101, 30. MR3355258

[32] KATZ, N. M. and SARNAK, P. (1999). *Random Matrices*, *Frobenius Eigenvalues*, *and Monodromy*. *American Mathematical Society Colloquium Publications* **45**. Amer. Math. Soc., Providence, RI. MR1659828

[33] LANDON, B. and YAU, H.-T. (2015). Convergence of local statistics of Dyson Brownian motion. Preprint. Available at arXiv:1504.03605.

[34] MCKAY, B. D. (1985). Asymptotics for symmetric 0–1 matrices with prescribed row sums. *Ars Combin.* **19** 15–25. MR0790916

[35] MEHTA, M. L. (2004). *Random Matrices*, 3rd ed. *Pure and Applied Mathematics* (*Amsterdam*) **142**. Elsevier, Amsterdam. MR2129906

[36] MILLER, S. J. and NOVIKOFF, T. (2008). The distribution of the largest nontrivial eigenvalues in families of random regular graphs. *Exp. Math.* **17** 231–244. MR2433888

[37] OREN, I. and SMILANSKY, U. (2010). Trace formulas and spectral statistics for discrete Laplacians on regular graphs (II). *J. Phys. A* **43** 225205, 13.

[38] PASTUR, L. and SHCHERBINA, M. (1997). Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles. *J. Stat. Phys.* **86** 109–147. MR1435193

[39] REED, M. and SIMON, B. (1978). *Methods of Modern Mathematical Physics. IV. Analysis of Operators*. Academic Press, New York. MR0493421

[40] SHCHERBINA, M. (2014). Change of variables as a method to study general β-models: Bulk universality. *J. Math. Phys.* **55** 043504, 23.

[41] SHCHERBINA, T. (2014). Universality of the local regime for the block band matrices with a finite number of blocks. *J. Stat. Phys.* **155** 466–499.

[42] SMILANSKY, U. (2013). Discrete graphs—A paradigm model for quantum chaos. In *Chaos. Prog. Math. Phys.* **66** 97–124. Birkhäuser, Basel. MR3204183

[43] TAO, T. (2013). The asymptotic distribution of a single eigenvalue gap of a Wigner matrix. *Probab. Theory Related Fields* **157** 81–106. MR3101841

[44] TAO, T. and VU, V. (2011). Random matrices: Universality of local eigenvalue statistics. *Acta Math.* **206** 127–204.

[45] TRAN, L. V., VU, V. H. and WANG, K. (2013). Sparse random graphs: Eigenvalues and eigenvectors. *Random Structures Algorithms* **42** 110–134. MR2999215

[46] VALKÓ, B. and VIRÁG, B. (2014). Random Schrödinger operators on long boxes, noise explosion and the GOE. *Trans. Amer. Math. Soc.* **366** 3709–3728. MR3192614

[47] WORMALD, N. C. (1999). Models of random regular graphs. In *Surveys in Combinatorics*, 1999 (*Canterbury*). *London Mathematical Society Lecture Note Series* **267** 239–298. Cambridge Univ. Press, Cambridge. MR1725006

R. BAUERSCHMIDT
UNIVERSITY OF CAMBRIDGE
DPMMS, CENTRE FOR MATHEMATICAL SCIENCES
WILBERFORCE ROAD
CAMBRIDGE CB3 0WB
UNITD KINGDOM
E-MAIL: brt@math.harvard.edu

J. HUANG
H.-T. YAU
DEPARTMENT OF MATHEMATICS, FAS
HARVARD UNIVERSITY
ONE OXFORD STREET
CAMBRIDGE, MASSACHUSETTS 02138
USA
E-MAIL: jiaoyang@math.harvard.edu
htyau@math.harvard.edu

A. KNOWLES
UNIVERSITY OF GENEVA
SECTION DE MATHÉMATIQUES
2-4 RUE DU LIÈVRE
CASE POSTALE 64
1211 GENÈVE 4
SWITZERLAND
E-MAIL: knowles@math.ethz.ch