Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 150, 1–19.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1203

# Fluctuations of eigenvector overlaps and the Berry conjecture for Wigner matrices

Lucas Benigni* Giorgio Cipolloni†

**Abstract**

We prove that any finite collection of quadratic forms (overlaps) of general deterministic matrices and eigenvectors of an $N \times N$ Wigner matrix has joint Gaussian fluctuations. This can be viewed as the random matrix analogue of the Berry random wave conjecture.

**Keywords:** random matrix theory; quantum unique ergodicity; Berry conjecture.
**MSC2020 subject classifications:** 60B20; 15B52.
Submitted to EJP on November 8, 2023, final version accepted on September 10, 2024.

## 1 Introduction

The groundbreaking discovery of Wigner [Wig57] was that gap statistics of energy levels of heavy nuclei are universal and can be modeled by eigenvalue statistics of large random Hermitian matrices called *Wigner matrices*. We recall that such matrices are $N \times N$ Hermitian matrices $W^* = W$ with centered i.i.d. entries (modulo symmetry) of finite variance. While the analysis of eigenvectors were not considered in Wigner's original study, they were examined by Porter and Thomas [PT56] to understand nuclear reaction widths. In general, $W$ is a tractable mathematical model for a Hamiltonian of a disordered quantum system and in particular, its eigenvectors, as stationary states of the system, are important quantities to understand.

Let $\{\mathbf{u}_i\}_{i=1}^N$ denote the $\ell^2$-normalized eigenvectors of $W$. Several properties of $\mathbf{u}_i$ can be examined to better understand disordered quantum systems. *Delocalization* consists in bounding their individual entries and it can be shown [ESY09b, ESY09a, VW15, BL22b] that $\sqrt{N}\|\mathbf{u}_i\|_\infty \leqslant C\sqrt{\log N}$ with very high probability[1]. Proving a *Porter–Thomas law* consists in showing the convergence of $\sqrt{2N}\mathbf{u}_i(\alpha)$ to a standard normal variable; this has been proved in the series of work [KY13, TV12, BY17, MY22]. For a deterministic *traceless* matrix $A$, i.e. $\mathrm{Tr} A = 0$, the *Eigenstate Thermalization Hypothesis* (ETH),

*Université de Montréal, Canada. E-mail: lucas.benigni@umontreal.ca
†Princeton University, United States of America. E-mail: gc4233@princeton.edu
[1] We say that an event $\Omega$ holds with very high probability if $\mathbb{P}(\Omega^c) \leqslant N^{-D}$ for any $D > 0$.

introduced in [Deu18], consists in examining the quadratic form $\langle \mathbf{u}_i, A\mathbf{u}_j\rangle$ and proving that with very high probability and any $\xi > 0$,

$$|\langle \mathbf{u}_i, A\mathbf{u}_j\rangle| \leqslant N^{\xi}\sqrt{\frac{\langle AA^*\rangle}{N}}. \tag{1.1}$$

Here $\langle\cdot\rangle := N^{-1}\mathrm{Tr}[\cdot]$ denotes the normalized trace. For Wigner matrices the ETH conjecture was proven in [CES22b]. Partial previous results for specific observables $A$ were given in [BY17, BYY20, CES21b, BL22a].

The universal properties of Wigner matrices for truly interacting quantum systems are yet to be proved, but it is conjectured to be true for many different models. For instance, in quantum chaos theory, the celebrated Bohigas–Giannoni–Schmit conjecture [BGS84] states that for a classically chaotic quantum billiard, eigenvalue statistics of the Laplacian should be given by random matrix statistics. The comparison does not stop at eigenvalues since the ETH can also be compared with the Quantum Unique Ergodicity (QUE) conjecture introduced by Rudnick and Sarnak [RS94] which states that if $\{\psi_k\}_{k\geqslant 1}$ is an orthonormal basis of eigenfunctions (with increasing eigenvalues) of the Laplace–Beltrami operator on a negatively curved manifold $(\mathcal{M},\mu)$ then for any open set $A\subset\mathcal{M}$

$$\int_A |\psi_k(x)|^2\mu(\mathrm{d}x) \xrightarrow[k\to\infty]{} \mu(A).$$

Quantum ergodicity, which proves this convergence along a subsequence, has been proved in greater generality in [Shn74, Zel87, CDV85, AS19], while QUE has only been proved for some arithmetic surfaces in [Lin06, HS10, Hol10] using advanced tools from number theory and ergodic theory. For negatively curved manifolds, Berry's *random wave conjecture* from [Ber77] asserts that the wave functions $\psi_i$, as $i\to\infty$, tend to exhibit Gaussian random behavior. In particular, they should locally "*converge*" to the *isotropic Gaussian field*, a centered stationary Gaussian field $\Psi$ on $\mathbb{R}^d$ with covariance

$$\mathbb{E}\left[\Psi(x)\Psi(y)\right] = \int_{\mathbb{S}^{d-1}} \mathrm{e}^{\mathrm{i}\langle x-y,\theta\rangle}\mathrm{d}\theta.$$

Recently, [Ing21, ABLM18] gave a rigorous statement of the Berry conjecture in this setting.

Figure 1: High energy eigenfunction of the Laplacian on the "Barnett stadium" which is conjectured to be quantum unique ergodic and the random plane wave on a square.

In this article, we study joint fluctuations of the overlaps $\langle \mathbf{u}_i, A\mathbf{u}_j\rangle$. For simplicity, consider $A$ and $B$ to be two deterministic *traceless* Hermitian matrices. The first results in this direction are in [Ben21] where it was shown that for $A=\sum_{\alpha\in I}|\mathbf{q}_\alpha\rangle\langle\mathbf{q}_\alpha| - |I|/N$, with $(\mathbf{q}_\alpha)$ being an orthonormal family of vectors in $\ell^2$ and $1\ll |I|\ll N$,

$$\mathbb{E}\left[\frac{N^2}{|I|}\langle \mathbf{u}_i, A\mathbf{u}_i\rangle\langle \mathbf{u}_j, A\mathbf{u}_j\rangle\right] \ll 1 \quad \text{and} \quad \mathbb{E}\left[\frac{N^2}{|I|}\langle \mathbf{u}_i, A\mathbf{u}_j\rangle^2\right] \xrightarrow[N\to\infty]{} 1.$$

The main difficulty for this result at the time was to understand correlations between different entries of different eigenvectors and was handled by introducing new moment observables along the Dyson Brownian motion. For the same specific family of observables, and introducing again other symmetrized observables, convergence of the diagonal overlap to the Gaussian distribution was proved in [BL22a] for all eigenvectors (including the edge). Simultaneously, for general full-rank observables $A$ and indices in the bulk, the same convergence was proved in [CES22a] using multi-resolvent local laws proved in [CES21b],

$$\sqrt{\frac{N}{\langle |A|^2\rangle}}\langle \mathbf{u}_i, A\mathbf{u}_i\rangle \xrightarrow[N\to\infty]{} \mathcal{N}(0,1).$$

This convergence was then extended to all observables $A$ such that $A$ has effective rank $\gg 1$ in [CES22b] by refining these multi-resolvent local laws. Note that the condition is necessary to obtain some averaging in light of the central limit theorem making the result optimal. Our goal in this article is to obtain similar convergence, not just for the diagonal overlaps, but jointly for any finite collections of eigenvectors and observables $A$ using the machinery set up developed in [MY22] and [CES22a].

We prove that $(\langle \mathbf{u}_i, A\mathbf{u}_j\rangle)_{i,j,A}$, for bulk indices $i, j$ and deterministic matrices $A$, forms a Gaussian field with covariance structure given by (see Theorem 2.2 below)

$$\mathbb{E}\left[N\langle \mathbf{u}_i, A\mathbf{u}_j\rangle\langle \mathbf{u}_k, B\mathbf{u}_l\rangle\right] \approx \left[\delta_{kj}\delta_{li} + \left(\frac{2}{\beta}-1\right)\delta_{ki}\delta_{lj}\right]\langle AB\rangle, \tag{1.2}$$

where $\beta$ is the Dyson index such that $\beta = 1$ in the symmetric case and $\beta = 2$ in the Hermitian case. In particular, note that $\langle \mathbf{u}_i, A\mathbf{u}_i\rangle, \langle \mathbf{u}_j, B\mathbf{u}_j\rangle$ are asymptotically independent for any $j \neq i$ and, even more interestingly, that $\langle \mathbf{u}_i, A\mathbf{u}_i\rangle, \langle \mathbf{u}_i, B\mathbf{u}_i\rangle$ are also asymptotically independent as long as $\langle AB\rangle \ll \sqrt{\langle A^2\rangle\langle B^2\rangle}$. The limiting statement in (1.2) can be easily checked for $W$ drawn from the GOE/GUE ensembles by Weingarten calculus as the eigenbasis of these matrices are distributed according to the Haar measure on the orthogonal or unitary group, respectively. Thus (1.2) can be seen as a universality statement for finite collection of eigenvectors of Wigner matrices: any finite number of eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ behaves as if they were Haar distributed on the corresponding compact group.

Wigner matrices are often seen as a toy probabilistic model for quantum chaos, as can be seen by the celebrated Bohigas–Gianonni–Schmidt conjecture [BGS84] linking eigenvalue statistics of quantized Laplacian of chaotic systems to universal random matrix statistics, or for eigenvectors through the Quantum Unique Ergodicity conjecture [RS94, BY17, CES21b, BL22a]. A next natural step is thus to consider fluctuations around the QUE property, in the context of quantum chaos, this is given by the Berry conjecture [Ber77] and in the context of Wigner matrices, by the result above. We thus prove a probabilistic version of the Berry conjecture. While the covariance structure between the two models strongly differ due to eigenvectors of Wigner matrices being close to Haar distributed on unitary or orthogonal group (and thus simpler), the presence of a Gaussian field shows another similarity between quantum chaos and random matrices.

Our result, and many mentioned above, rely on a careful analysis of the Dyson vector flow, the flow of eigenvectors induced by the Dyson Brownian motion (DBM). This flow was first computed in [Bru89] and first used in this setting in the seminal paper [BY17] where it was used to compute asymptotic moments of individual eigenvector entries. The dynamical method to prove universality results as introduced in the groundbreaking paper [ESY11] consists in three steps. The first step establishes an (almost) optimal a priori bound for the quantity of interest and the second step use the DBM, combined with our a priori bound, to prove our joint fluctuations of eigenvectors along the dynamics.

Finally, the third step is based on a comparison step between the dynamical model and Wigner matrices by an application of the by now standard *Green's function comparison theorem* (see Appendix A).

Given the a priori bound (1.1) as an input, to prove Theorem 2.2, we compute arbitrary moments of several eigenvector overlaps $\langle \mathbf{u}_i, A\mathbf{u}_j\rangle$. Following Marcinek–Yau [MY22] one can see that arbitrary moments of eigenvectors overlaps, which are encoded by

$$\mathbb{E}\left[\prod_{i=1}^{n}\langle \mathbf{u}_{x_{2i-1}}, A_i \mathbf{u}_{x_{2i}}\rangle \Big| \boldsymbol{\lambda}\right], \tag{1.3}$$

evolve according to the *colored* eigenvector moment flow (see (3.3)–(3.4) below). Note that the conditioning is with respect to the whole trajectories of the eigenvalues for times in $[0,1]$.

Then to study relaxation properties of this flow, we rely on energy estimates for multi–indexed DBM developed in [MY22] for finite rank $A$'s, and extended to the analysis of general observables $A$ in [CES22a]. The observable considered in [CES22a] is different compared to (1.3) since the authors studied the evolution of *perfect matching observables* introduced in [BYY20], which enabled to obtain a central limit theorem only for diagonal overlaps. The main novelty in the present article is that the deterministic approximation of $f(\mathbf{x})$ depends on the configuration $\mathbf{x}$ and $f$ follows the *colored* eigenvector moment flow while the perfect matching observable evolve according to the original one from [BY17]. The *colored* version from [MY22] loses its parabolicity due to the exchange term $\mathscr{E}$ in (3.4). In particular, we need to introduce the correct *ansatz observable* $F(\mathbf{x})$, the deterministic limit of $f(\mathbf{x})$, and prove that it is stationary along the dynamics which is done in Subsection 3.1.

## 2 Main result

We consider $N \times N$ Wigner matrices, i.e. real symmetric or complex Hermitian matrices $W = W^*$ with entries distributed as $w_{ab} \stackrel{\mathrm{d}}{=} n^{-1/2}\chi_{\mathrm{od}}$, for $a > b$, and $w_{aa} \stackrel{\mathrm{d}}{=} n^{-1/2}\chi_{\mathrm{d}}$. On the random variables $\chi_{\mathrm{d}}, \chi_{\mathrm{od}}$ we assume:

**Assumption 2.1.** *$\chi_{\mathrm{od}}$ is a real or complex random variable with $\mathbb{E}\chi_{\mathrm{od}} = 0$, $\mathbb{E}|\chi_{\mathrm{od}}|^2 = 1$; additionally, in the complex case we also assume that $\mathbb{E}\chi_{\mathrm{od}}^2 = 0$. $\chi_{\mathrm{d}}$ is a real random variable with $\mathbb{E}\chi_{\mathrm{d}} = 0$.*

*Furthermore, we assume that all the moments of $\chi_{\mathrm{d}}, \chi_{\mathrm{od}}$ exist, i.e. that for any $p \in \mathbb{N}$ there exist constants $C_p > 0$ such that*

$$\mathbb{E}|\chi_{\mathrm{od}}|^p + \mathbb{E}|\chi_{\mathrm{d}}|^p \leqslant C_p. \tag{2.1}$$

By $\lambda_1 \leqslant \ldots \leqslant \lambda_N$ we denote the eigenvalues of $W$, and by $\{\mathbf{u}_i\}_{i\in[N]}$ the corresponding orthonormal eigenvectors. Since the eigenvectors are defined only modulo a phase, we assume that the eigenvectors are always multiplied by a uniformly random phase $e^{\mathrm{i}\theta_i}\mathbf{u}_i$ where $\theta_i \sim \mathrm{Unif}([0,2\pi])$ in the Hermitian case and $\theta_i \sim \mathrm{Unif}(\{0,\pi\})$ in the symmetric case; this choice will simplify the statement of our main result in Theorem 2.2 below[2]. To make the notation more concise, given a deterministic matrix $A \in \mathbb{C}^{N\times N}$, we denote its *traceless* part by $\mathring{A} := A - \langle A\rangle$ where $\langle A\rangle = \frac{1}{N}\operatorname{Tr} A$. From now on we only consider the case when $A$ is a real symmetric or a complex Hermitian matrices, the general case is discussed below Theorem 2.2.

Before stating our main result we introduce the following rescaled quantity:

$$\Phi_N(A,i,j) := \sqrt{\frac{N}{\langle \mathring{A}^2\rangle}}\big[\langle e^{\mathrm{i}\theta_i}\mathbf{u}_i, Ae^{\mathrm{i}\theta_j}\mathbf{u}_j\rangle - \langle A\rangle\delta_{ij}\big]. \tag{2.2}$$

[2] Note that we could remove this uniform phase at the price of considering $|\langle \mathbf{u}_i, A\mathbf{u}_j\rangle - \langle A\rangle\delta_{ij}|$ in (2.2).

**Theorem 2.2.** *Let $W$ be a Wigner matrix satisfying Assumption 2.1. Fix any $p \in \mathbb{N}$, any small $\delta, \delta' > 0$, and any real symmetric or complex Hermitian deterministic matrices $A_1, \dots, A_p$ with $\|A_i\| \leqslant 1$, $\langle \mathring{A}_i^2 \rangle \geqslant N^{-1+\delta'}$. Then, for any $i_1, \dots, i_{2p} \in [\![\delta N, (1-\delta)N]\!]$, the collection $(\Phi_N(A_1, i_1, i_2), \dots, \Phi_N(A_p, i_{2p-1}, i_{2p}))$ is approximated in the sense of moments[3] by a Gaussian field*

$$\Big(\Phi(A_1, i_1, i_2), \dots, \Phi(A_p, i_{2p-1}, i_{2p})\Big), \tag{2.3}$$

*with covariance structure*

$$\mathbb{E}\Phi(A, i, j)\Phi(B, k, l) = \left[\delta_{kj}\delta_{li} + \left(\frac{2}{\beta} - 1\right)\delta_{ki}\delta_{lj}\right] \frac{\langle \mathring{A}\mathring{B} \rangle}{\sqrt{\langle \mathring{A}^2 \rangle \langle \mathring{B}^2 \rangle}}. \tag{2.4}$$

The assumption $\langle \mathring{A}_i^2 \rangle \geqslant N^{-1+\delta'}$ in Theorem 2.2 is not a restriction since it ensures that the the observables $A_i$ are not finite rank, and for finite rank matrices (2.3) does not hold. This case, corresponding to finitely many entries, has been studied in [BY17, MY22]. Furthermore, in Theorem 2.2 we stated the convergence (2.3) only for real symmetric or complex Hermitian matrices $A$. The general case immediately follows by polarization and by the fact that any matrix can be written as the sum of symmetric or Hermitian matrices.

**Remark 2.3** (Special cases of Theorem 2.2)**.** We now specialize Theorem 2.2 to some simple relevant cases:

(i) By (2.3) for $p = 1$ and $i \neq j$ it follows that (here $\mathbb{F} = \mathbb{R}$ for $\beta = 1$ and $\mathbb{F} = \mathbb{C}$ for $\beta = 2$)

$$\sqrt{\frac{N}{\langle \mathring{A}^2 \rangle}} \big|\langle \mathbf{u}_i, A\mathbf{u}_j \rangle\big| \Longrightarrow \big|\mathcal{N}_{\mathbb{F}}(0,1)\big|.$$

(ii) By (2.3) for $p = 2$ it follows that, after rescaling, $\langle \mathbf{u}_i, A\mathbf{u}_i \rangle, \langle \mathbf{u}_j, A\mathbf{u}_j \rangle$ are asymptotically independent for any $i \neq j$. Interestingly, we find out that $\langle \mathbf{u}_i, A\mathbf{u}_i \rangle, \langle \mathbf{u}_i, B\mathbf{u}_i \rangle$ are also asymptotically independent as long as

$$\langle \mathring{A}\mathring{B} \rangle \ll \sqrt{\langle \mathring{A}^2 \rangle \langle \mathring{B}^2 \rangle}.$$

## 3 Short-time relaxation

In this section we analyze the mixed moments of (2.2) for different indices and observable matrices through the matrix Dyson Brownian motion. This dynamics is defined as

$$\mathrm{d}W_t = \frac{\mathrm{d}B_t}{\sqrt{N}}, \qquad W_0 = W, \tag{3.1}$$

and the eigenvalues $\boldsymbol{\lambda}(t)$ and eigenvectors $\mathbf{u}(t)$ of $W_t$ are the solutions to a system of couple stochastic differential equations [BY17, Definition 2.2]. It was observed in the seminal paper [BY17] that moments of eigenvectors follow a dynamics which can be represented as a random walk in a random environment given by the eigenvalue process of $W_t$. We now introduce the set of observables we will analyze which follow the *colored* eigenvector moment flow, a variant of the process from [BY17], introduced in [MY22].

Consider $\Lambda^n \subset [N]^{2n}$ be defined as

$$\Lambda^n := \big\{\mathbf{x} \in [N]^{2n} \,:\, n_i(\mathbf{x}) \text{ even } \forall\, i \in [N]\big\}$$

where $n_i(\mathbf{x}) = |\{a \in [2n] \,:\, x_a = i\}|$ denotes the number of particles at site $i$. Note that for $\mathbf{x} \in [N]^{2n}$ we denote its entries by $x_a$, with $a \in [2n]$.

[3] For two possilby $N$-dependent random variables $X_N$, $Y_N$ we say that $X_N$ is approximated in the sense of moments by $Y_N$ if $\mathbb{E}|X_N|^k = \mathbb{E}|Y_N|^k + \mathcal{O}(N^{-c(k)})$, with $c(k) > 0$ a small constant depending on $k$.

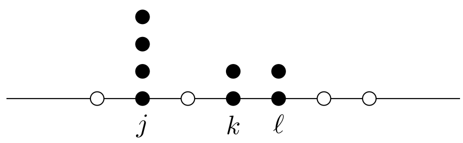


Figure 2: A visualization of the vertices of the configuration $\mathbf{x} = (j, j, j, j, k, \ell, k, \ell)$ corresponding to the conditional moment of $\langle \mathbf{u}_j, A_1\mathbf{u}_j\rangle\langle \mathbf{u}_j, A_2\mathbf{u}_j\rangle\langle \mathbf{u}_k, A_3\mathbf{u}_\ell\rangle\langle \mathbf{u}_k, A_4\mathbf{u}_\ell\rangle$.

**Definition 3.1.** *For a given family of traceless symmetric matrices $\mathbf{A} = (A_i)_{1\leqslant i\leqslant n}$, we define*

$$f_{\boldsymbol{\lambda},t}(\mathbf{x},\mathbf{A}) := f_t(\mathbf{x}) = \frac{1}{\prod_{i=1}^N (n_i(\mathbf{x})-1)!!}\mathbb{E}\left[\prod_{i=1}^n \langle \mathbf{u}_{x_{2i-1}}(t), A_i\mathbf{u}_{x_{2i}}(t)\rangle \middle| \boldsymbol{\lambda}\right] \tag{3.2}$$

*where the conditioning is on the whole path of eigenvalues from $0$ to $1$.*

This observable follow the colored eigenvector moment flow from [MY22].

**Theorem 3.2** ([MY22, Theorem 4.8])**.** *It follows that*

$$\partial_t f_t(\mathbf{x}) = \sum_{1\leqslant i<j\leqslant N} c_{ij}(t)\mathscr{L}_{ij} f_t(\mathbf{x}), \qquad c_{ij}(t) := \frac{1}{N(\lambda_i-\lambda_j)^2}, \tag{3.3}$$

*where $\mathscr{L}_{ij} = \mathscr{M}_{ij} - \mathscr{E}_{ij}$, with*

$$\begin{aligned} \mathscr{M}_{ij}h(\mathbf{x}) &= \frac{n_j(\mathbf{x})+1}{n_i(\mathbf{x})-1}\sum_{a\neq b}\Big(h(m_{ab}^{ij}\mathbf{x}) - h(\mathbf{x})\Big) + \frac{n_i(\mathbf{x})+1}{n_j(\mathbf{x})-1}\sum_{a\neq b}\Big(h(m_{ab}^{ji}\mathbf{x}) - h(\mathbf{x})\Big) \\ \mathscr{E}_{ij}h(\mathbf{x}) &= \sum_{a\neq b}\Big(h(s_{ab}^{ij}\mathbf{x}) - h(\mathbf{x})\Big) \end{aligned} \tag{3.4}$$

*for $h$ any function in $L^2(\Lambda^n)$. Here we used the definition*

$$m_{ab}^{ij}\mathbf{x} = \mathbf{x} + \mathbf{1}_{x_a=x_b=i}(j-i)(\mathbf{e}_a+\mathbf{e}_b), \qquad s_{ab}^{ij}\mathbf{x} = \mathbf{x} + \mathbf{1}_{x_a=i}\mathbf{1}_{x_b=j}(j-i)(\mathbf{e}_a-\mathbf{e}_b). \tag{3.5}$$

*Besides, the reversible measure for the generator $\mathscr{L}$ is given by*

$$\pi(\mathbf{x}) = \prod_{i=1}^N \left[(n_i(\mathbf{x})-1)!!\right]^2.$$

### 3.1 Ansatz observable

We now define the observable $F(\mathbf{x},\mathbf{A})$ which corresponds to the asymptotic deterministic behavior of $f_{\boldsymbol{\lambda},t}(\mathbf{x},\mathbf{A})$. It is the analogue of the *anzats observable* from [MY22]. In particular since we want to show that the family of observables of the form $\langle \mathbf{u}_i, A_j\mathbf{u}_k\rangle$ form a Gaussian process with a specific covariance structure, we are going to use Wick's theorem to describe $F(\mathbf{x},\mathbf{A})$.

From (3.2), we see that the configuration of particles is not enough to describe $f_{\boldsymbol{\lambda},t}(\mathbf{x},\mathbf{A})$ since we need to add an ordering of the particles for it to be well defined. This can correspond to performing a perfect matching on the configuration $\mathbf{x}$ (note that we consider configurations such that $n_i(\mathbf{x})$ is even for all $i$). Thus, we have the bijective property that $f_{\boldsymbol{\lambda},t}(\mathbf{x},\mathbf{A}) = f_{\boldsymbol{\lambda},t}(\mathcal{G},\mathbf{A})$ where $\mathcal{G} = (\mathcal{V},\mathcal{E})$. with $\mathcal{V} = \{x_a\}_{a\in[2n]}$ and $\mathcal{E} = \{\{x_{2a-1},x_{2a}\}_{a\in[2n]}\}$. Thus we rewrite,

$$f_{\boldsymbol{\lambda},t}(\mathcal{G},\mathbf{A}) = \frac{1}{\prod_{i=1}^N (n_i(\mathcal{G})-1)!!}\mathbb{E}\left[\prod_{e=\{k,\ell\}\in\mathcal{E}} \langle \mathbf{u}_k(t), A_e\mathbf{u}_\ell(t)\rangle \middle| \boldsymbol{\lambda}\right] \tag{3.6}$$

where $n_i(\mathcal{G}) = n_i(\mathcal{V})$ is the number of particles at site $i$ and the reindexing of the matrices from $\mathbf{A}$ is straightforward. For each edge $e \in \mathcal{E}$, we denote $n(e)$ the number of times it occurs in the graphs,

$$n(\{k,\ell\}) = |\{e \in \mathcal{E}, e = \{k,\ell\}\}|,$$

and we denote by $e^{(1)}, \dots, e^{(n(e))}$ the occurrences of the edges in the graph, note that it is important to differentiate them since while they connect the same eigenvector indices, they may be labeled by different matrices from $\mathbf{A}$. We define an equivalence relation between edges $e \sim e'$ if and only if $e = e'$ so that we can define the set $\widetilde{\mathcal{E}} = \mathcal{E}/\sim$ where each element is a representative of the corresponding edge. The graph representation is illustrated in Figure 3

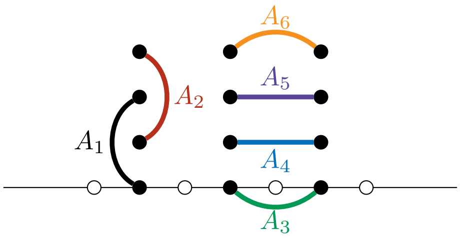


Figure 3: An example of the graph representation of an ordered configuration of 12 particles. We have $n(\{j,j\}) = 2$ and $n(\{k,\ell\}) = 4$ and each edge is colored by a matrix from the set $\mathbf{A}$. For this graph, the ansatz observable is given by $F(\mathbf{x},\mathbf{A}) = \frac{2}{27N^3}\langle A_1A_2\rangle\left(\langle A_3A_4\rangle\langle A_5A_6\rangle + \langle A_3A_5\rangle\langle A_4A_6\rangle + \langle A_3A_6\rangle\langle A_4A_5\rangle\right)$.

**Definition 3.3.** *We define the (time-independent)* ansatz observable *to be,*

$$F(\mathcal{G},\mathbf{A}) = \frac{\displaystyle\prod_{e=\{k,\ell\}\in\widetilde{\mathcal{E}}} 2^{\frac{n(e)}{2}\mathbb{1}_{k=\ell}} \sum_{\Pi\in \mathrm{PM}[n(e)]} \prod_{\{i,j\}\in\Pi} \langle A_{e^{(i)}} A_{e^{(j)}}\rangle}{N^{n/2}\prod_{i=1}^N (n_i(\mathcal{G})-1)!!}$$

*where $n_i(\mathcal{G}) = n_i(\mathcal{V})$ is the number of particles at site $i$, $\mathrm{PM}[i]$ is the set of perfect matchings on $\{1,\dots i\}$ with the convention that the empty sum is zero. In the rest of the paper, we may use the equivalence $F(\mathcal{G},\mathbf{A}) = F(\mathbf{x},\mathbf{A}) = F(\mathbf{x})$.*

Before starting the analysis of the eigenvector observable, it is important to check that our ansatz observable $F$ is an actual one, in other words that it is in the kernel of $\mathscr{L}$. We first recall a corollary from [MY22] which describes the kernel of $\mathscr{L}$.

**Lemma 3.4** ([MY22, Corollary 6.28])**.** *Consider a configuration $\mathbf{x}$ over $2n$ points, for a given perfect matching $\Pi$ over $2n$ points, we say that $\Pi \in \mathrm{Stab}(\mathbf{x})$ if*

$$\Pi \cdot \mathbf{x} := (x_{\Pi(1)}, \dots x_{\Pi(2n)}) = (x_1, \dots, x_{2n}) = \mathbf{x}.$$

*We then have*

$$\bigcap_{1\leqslant i<j\leqslant N} \mathrm{Ker}\,\mathscr{L}_{ij} = \mathrm{Span}\{\chi_\Pi | \Pi \in \mathrm{PM}[2n]\}, \quad \chi_\Pi = \frac{\mathbb{1}_{\Pi\in\mathrm{Stab}(\mathbf{x})}}{\sqrt{\pi(\mathbf{x})}}.$$

Using this lemma we prove that $F$ is in the kernel of the operator $\mathscr{L}$:

**Lemma 3.5.** *It holds that $F \in \bigcap_{1\leqslant i<j\leqslant N} \mathrm{Ker}\,\mathscr{L}_{ij}$.*

*Proof.* The proof is based on rewriting the ansatz observable $F$ as a linear combination of $\chi_\Pi$. For now, it is defined as a product of sums of perfect matchings on the edges of $\mathcal{G}$.

We first need to rewrite these perfect matchings as acting on the vertices of the graph rather than the edges. To do this, we use the first convention of ordered particles.

For each edge $e$ of the graph, we consider $B_e$ to be the subgraph containing all and only occurrences of the edge $e$ from the initial graph. In particular, we see that $\mathcal{G} = \cup_{e\in\mathcal{E}} B_e$ where the union is not disjoint (for two occurrences of the same edge $e_1$ and $e_2$ we have $B_{e_1} = B_{e_2}$). The blocks of a configuration are illustrated in Figure 4.

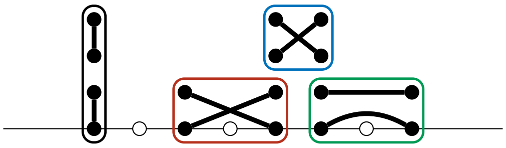

Figure 4: Illustration of blocks of a given configuration. All occurrences of a given edge are in the same block.

For a given edge $e = \{k, \ell\}$, there exist $j_1 < j_2 < \cdots < j_{n(e)}$, using our first convention of ordered particles, such that

$$B_e = (\{x_{2j_1-1}, x_{2j_1}\}, \{x_{2j_2-1}, x_{2j_2}\}, \ldots, \{x_{2j_{n(e)}-1}, x_{2j_{n(e)}}\}).$$

In particular, half of the $x$'s are $k$ and other half are $\ell$. From this notation, we define

$$\mathcal{A}(B_e) = \mathbb{1}_{n(e) \text{ is even}} \prod_{p=1}^{\frac{n(e)}{2}} \langle A_{j_{2p-1}} A_{j_{2p}} \rangle.$$

We now define the set of perfect matchings which stabilize the block $B_e$ but also keep the edge structure, in other words $\mathcal{E}\text{-Stab}(B_e)$ is the set of perfect matchings $\Pi$ over the set $\{2j_1 - 1, 2j_1, \ldots, 2j_{n(e)} - 1, 2j_{n(e)}\}$, such that $\Pi \in \text{Stab}(B_e)$ and such that for all $p \in \{1, \ldots, n(e)\}$ there exists $q \in \{1, \ldots, n(e)\}$ with $q \neq p$ such that $\Pi(\{2j_p - 1, 2j_p\}) = \{2j_q - 1, 2j_q\}$.

Let $r$ be the number of distinct blocks of the configuration $\mathbf{x}$ and denote them by $B_{e_1}, \ldots, B_{e_r}$. We define the set $\mathcal{G}\text{-Stab}(\mathbf{x})$ to be the set of perfect matchings over the whole configuration of $2n$ particles such that the restriction of $\Pi$ to each block $B_{e_i}$ is in $\mathcal{E}\text{-Stab}(B_{e_i})$.

The first observation is that for a given edge $e = \{k, \ell\}$ we have

$$2^{\frac{n(e)}{2}\mathbb{1}_{k=\ell}} \sum_{\Pi\in\text{PM}[n(e)]} \prod_{i=1}^{n(e)} \langle A_{e^{(i)}} A_{e^{(\Pi(i))}} \rangle = \sum_{\Pi\in\mathcal{E}\text{-Stab}(B_e)} \mathcal{A}(\Pi \cdot B_e).$$

Note that the $2^{\frac{n(e)}{2}}$ is necessary in the case when $e = \{k, k\}$ for some $k$. Since the stabilization property is immediate, there are now two choices for each matching between edges. Thus we can write,

$$\begin{aligned}
F(\mathbf{x}, \mathbf{A}) &= \frac{1}{N^{n/2}\sqrt{\pi(\mathbf{x})}} \prod_{i=1}^{r} \sum_{\Pi\in\mathcal{E}\text{-Stab}(B_{e_i})} \mathcal{A}(\Pi \cdot B_{e_i}) \\
&= \frac{1}{N^{n/2}\sqrt{\pi(\mathbf{x})}} \sum_{\substack{\Pi_1\in\mathcal{E}\text{-Stab}(B_{e_1}) \\ \cdots \\ \Pi_r\in\mathcal{E}\text{-Stab}(B_{e_r})}} \prod_{i=1}^{r} \mathcal{A}(\Pi_i \cdot B_{e_i}) \\
&= \frac{1}{N^{n/2}} \sum_{\Pi\in\text{PM}[2n]} \left[\prod_{i=1}^{r} \mathcal{A}(\Pi_{|B_{e_i}} \cdot B_{e_i})\right] \frac{\mathbb{1}_{\Pi\in\mathcal{G}\text{-Stab}(\mathbf{x})}}{\sqrt{\pi(\mathbf{x})}}
\end{aligned}$$

where $\Pi_{|B_{e_i}}$ denote the restriction of $\Pi$ to $B_{e_i}$. Since $\mathcal{G}\text{-Stab}(\mathbf{x}) \subset \text{Stab}(\mathbf{x})$, we see that $F$ can be written as a linear combination of $\chi_\Pi$ and thus $F$ belongs to the kernel of $\mathscr{L}$. $\square$

### 3.2 Dyson vector flow analysis

The scheme of the analysis performed in this section is similar to [CES22b, CES22a]. Throughout this section we analyze the flow (3.3) on the event $\Omega_1 \cap \Omega_2$, with (cf. [CES22b, Eqs. (4.8), (4.29)])

$$\Omega_1 = \Omega_{1,\xi} := \Big\{ \sup_{0\leqslant t\leqslant T} \max_{i\in[N]} N^{2/3}\widehat{i}^{1/3}|\lambda_i(t) - \gamma_i(t)| \leqslant N^\xi \Big\} \tag{3.7}$$

where $\widehat{i} = \min(i, N-i+1)$, and

$$\begin{aligned}\Omega_2 = \Omega_{2,\xi,\omega} := &\bigcap_{\substack{\operatorname{Re} z_i\in[-3,3],\\ |\operatorname{Im} z_i|\geqslant N^{-1+\omega}}} \Bigg[ \Bigg\{ \sup_{0\leqslant t\leqslant T} \frac{1}{\sqrt{\rho_{i,t}}} \big|\langle G_t(z_1)\mathring{A}_1\rangle\big| \leqslant \frac{N^\xi \rho_1}{N\sqrt{|\operatorname{Im} z_1|}}\sqrt{\langle \mathring{A}_1^2\rangle} \Bigg\} \\ &\bigcap_{k=2}^{n} \Bigg\{ \sup_{0\leqslant t\leqslant T} \frac{1}{\sqrt{\rho_t^*}} \Big|\langle G_t(z_1)\mathring{A}_1 \dots G_t(z_k)\mathring{A}_k\rangle - \mathbb{1}_{k=2} m_t(z_1)m_t(z_2)\langle \mathring{A}_1\mathring{A}_2\rangle\Big| \\ &\leqslant \frac{N^{\xi+k/2-1}}{\sqrt{N\eta_*}} \prod_{i=1}^k \sqrt{\langle \mathring{A}_i^2\rangle} \Bigg\} \Bigg],\end{aligned} \tag{3.8}$$

with $G_t(z) := (W_t - z)^{-1}$, $m_t(z)$ denoting the Stieltjes transform of $\rho_t(x) = \frac{\sqrt{4(1+t^2)-x^2}}{2(1+t)\pi}$, $\eta_* := \min_i |\operatorname{Im} z_i|$, $\rho_{i,t} := |\operatorname{Im} m_t(z_i)|$, and $\rho_t^* := \max_i \rho_{i,t}$. Here $\gamma_i(t)$ denote the *quantiles* of $\rho_t(x)$, which are implicitly defined by

$$\int_{-\infty}^{\gamma_i(t)} \rho_t(x)\,\mathrm{d}x = \frac{i}{N}.$$

The fact that $\Omega_1$ holds with very high probability follows e.g. by [EKYY13, Theorem 7.6], while from [CES22b, Theorem 2.2] it follows that $\Omega_2$ holds with very high probability. Note that on $\Omega_1 \cap \Omega_2$ for $\langle A\rangle = 0$ it holds (see [CES22b, Eqs. (4.30)–(4.31)])

$$\big|\langle \mathbf{u}_i(t), A\mathbf{u}_j(t)\rangle\big|^2 \leqslant \frac{N^{2\omega}\langle A^2\rangle}{N[\rho_t(\gamma_i(t)) \wedge \rho_t(\gamma_j(t)) + N^{-1/3}]}. \tag{3.9}$$

We start by introducing an observable which is localized and is directed by a short-range approximation of the actual dynamics as introduced in [BY17, MY22].

**Definition 3.6.** *Let $\ell$, $K >$ be two $N$-dependent parameters, $\alpha \in (0,1)$ a given constant, and $\mathbf{y}$ a given configuration of $n$ particles such that $y_i \in [\![\alpha N, (1-\alpha)N]\!] =: \mathcal{I}_\alpha$, we define the short-range observable by*

$$\begin{cases} \partial_t g_t(\mathbf{x};\ell,K,\mathbf{y}) &= \mathscr{S} g_t(\mathbf{x};\ell,K,\mathbf{y}) \\ g_0(\mathbf{x};\ell,K,\mathbf{y}) &= \operatorname{Av}(\mathbf{x};K,\mathbf{y})\,(f_0(\mathbf{x}) - F(\mathbf{x})) \end{cases} \tag{3.10}$$

*where* $\operatorname{Av}(\mathbf{x};K,\mathbf{y})f(\mathbf{x}) := \frac{1}{K}\sum_{j=K}^{2K-1} f(\mathbf{x})\mathbb{1}_{\|\mathbf{x}-\mathbf{y}\|_1\leqslant j}$ *with* $\|\mathbf{x}-\mathbf{y}\|_1 = \sum_{i=1}^{2n}|x_i - y_i|$ *and*

$$\mathscr{S} = \sum_{\substack{|i-j|\leqslant\ell \\ i,j\in\mathcal{I}_\alpha,\, i\neq j}} c_{ij}\mathscr{L}_{ij}.$$

*In the rest of the paper we may omit the dependence of $g_t$ on the parameters $\ell$, $K$, and $\mathbf{y}$.*

We can quantify the difference between the full dynamics and the short-range approximation with the following lemma.

**Lemma 3.7** ([MY22, Proposition 5.7][CES22a, Lemma 4.4])**.** *Let $\alpha \in (0,1)$, $0 \leqslant s_1 \leqslant s_2 \leqslant s_1 + \frac{\ell}{N}$, and $f$ be a function on $\Lambda^n$. For any $\mathbf{x} \in \Lambda^n$ supported on $\mathcal{I}_\alpha$ we have*

$$|\mathscr{U}(s_1,s_2) - \mathscr{U}_{\mathscr{S}}(s_1,s_2;\ell)| \leqslant CN^{1+n\xi}\frac{s_2-s_1}{\ell}\|f\|_\infty$$

*where $\mathscr{U}$ and $\mathscr{U}_{\mathscr{S}}$ denotes the semigroup for the operator $\mathscr{L}$ and $\mathscr{S}$ respectively.*

We also define the following term

$$\Phi_{\mathbf{A}} = \prod_{i=1}^{n}\sqrt{\frac{\langle A_i^2\rangle}{N}} \quad \text{which gives} \quad |g_t(\mathbf{x})| \leqslant N^{n\xi}\Phi_{\mathbf{A}} \tag{3.11}$$

with very high probability since $\mathscr{S}$ is a contraction in $L^\infty$ and $|g_0(\mathbf{x})| \leqslant N^{n\xi}\Phi_{\mathbf{A}}$ on $\Omega_1 \cap \Omega_2$ from (3.9). The main goal of this section is to prove the following proposition.

**Proposition 3.8.** *Let $\eta > 0$, suppose that parameters are chosen so that*[4] *$N^{-1} \ll \eta \ll T \ll \ell/N \ll K/N$. On the event $\Omega_1 \cap \Omega_2$, it holds that*

$$\sum_{\mathbf{x}\in\Lambda^n}\pi(\mathbf{x})|g_T(\mathbf{x})|^2 \leqslant K^n\mathcal{E}_n^2, \quad \textit{with} \quad \mathcal{E}_n = N^{n\xi}\sqrt{\Phi_{\mathbf{A}}}\left(\frac{N^\varepsilon \ell}{K} + \frac{NT}{\ell} + \frac{1}{\sqrt{K}} + \frac{1}{\sqrt{N\eta}}\right) \tag{3.12}$$

*with $\pi(\mathbf{x})$ is the reversible measure for $\mathscr{L}$ and the bound holds with very high probability.*

The start of the proof is similar to that of [CES22a, Proposition 4.5]. To see the main difference between the two proofs, we split the proof of Proposition 3.8 into the following lemma and proposition. From now on, we often omit the dependence on $t$ of eigenvectors and eigenvalues.

**Lemma 3.9.** *On the event $\Omega_1 \cap \Omega_2$, it holds that*

$$\begin{aligned}
\partial_t\|g_t\|_2^2 &\leqslant -\frac{1}{\eta}\left(C_1(n) + \frac{1}{N\eta}\right)\|g_t\|_2^2 + C(n)\frac{N^\xi K^{n-1}}{\eta}\Phi_{\mathbf{A}} \\
&+ \frac{1}{\eta}\sum_{\mathbf{x}\in\Lambda^n}\sum_{\mathbf{i}}^{*}\sum_{\mathbf{a},\mathbf{b}}^{*}\mathrm{Av}(\mathbf{x})|g_t(\mathbf{x})|\Psi_{\mathbf{a},\mathbf{b}}^{\mathbf{i}}(\mathbf{x})\times \\
&\times\left(\sum_{\mathbf{j}}^{*}\left(\prod_{r=1}^{2n}a_{i_r,j_r}\right)\left(f_t(\mathbf{x}_{\mathbf{a},\mathbf{b}}^{\mathbf{ij}}) - F(\mathbf{x}_{\mathbf{a},\mathbf{b}}^{\mathbf{ij}})\right) + \Phi_{\mathbf{A}}\mathcal{O}\left(N^{n\xi}\left(\frac{N^\varepsilon\ell}{K} + \frac{NT}{\ell} + \frac{N\eta}{\ell}\right)\right)\right)
\end{aligned} \tag{3.13}$$

*where*

$$\mathbf{x}_{\mathbf{a},\mathbf{b}}^{\mathbf{ij}} := \mathbf{x} + \left(\prod_{r=1}^{2n}\delta_{x_{a_r}i_r}\delta_{x_{b_r}i_r}\right)\sum_{r=1}^{2n}(j_r - i_r)(\mathbf{e}_{a_r} + \mathbf{e}_{b_r})$$

*and*

$$\Psi_{\mathbf{a},\mathbf{b}}^{\mathbf{i}}(\mathbf{x}) := \prod_{r=1}^{2n}\delta_{x_{a_r}i_r}\delta_{x_{b_r}i_r}.$$

*Proof.* By [MY22, Lemma 4.16], the operator $\mathscr{L}$ is negative in $L^2$, hence we get

$$\begin{aligned}
\partial_t\|g_t\|_2^2 &= \sum_{|i-j|\leqslant\ell}c_{ij}(t)\langle g_t, \mathscr{L}_{ij}(t)g_t\rangle \\
&\leqslant \frac{1}{N\eta}\sum_{|j-i|\leqslant\ell}\sum_{\mathbf{x}\in\Lambda^n}\frac{\eta}{(\lambda_i-\lambda_j)^2+\eta^2}g_t(\mathbf{x})(\mathscr{M}_{ij} - \mathcal{E}_{ij})g_t(\mathbf{x}).
\end{aligned} \tag{3.14}$$

[4] where $u_N \ll v_N$ means that there exists an $\varepsilon > 0$ such that $u_N = N^{-\varepsilon}v_N$.

We first control the exchange operator $-\mathcal{E}_{ij}$,

$$\frac{1}{N\eta}\sum_{|j-i|\leqslant\ell}\sum_{\mathbf{x}\in\Lambda^n}\frac{\eta}{(\lambda_i-\lambda_j)^2+\eta^2}g_t(\mathbf{x})(-\mathcal{E}_{ij})g_t(\mathbf{x})\leqslant\frac{1}{N\eta^2}\|g_t\|_2^2, \tag{3.15}$$

where we used that

$$g_t(\mathbf{x})\mathcal{E}_{ij}g_t(\mathbf{x})\sim g_t(\mathbf{x})[g(s_{ab}^{ij}\mathbf{x})-g(\mathbf{x})]\leqslant|g(s_{ab}^{ij}\mathbf{x})|^2+|g_t(\mathbf{x})|^2,$$

and that only finitely many (dependent on $n$) $i,j,a,b$ contributes to the summation for each $\mathbf{x}$.

As in [CES22a], we replace the move operator $\mathcal{M}$ by

$$\mathcal{A}:=\frac{1}{\eta}\sum_{\mathbf{i},\mathbf{j}\in[N]^{2n}}^{*}\left(\prod_{r=1}^{2n}a_{i_r,j_r}^{\mathscr{S}}\right)\sum_{\mathbf{a},\mathbf{b}}^{*}\left(g_t(\mathbf{x}_{\mathbf{a},\mathbf{b}}^{\mathbf{ij}})-g_t(\mathbf{x})\right)$$

where $a_{i_r,j_r}:=\frac{\eta}{N[(\lambda_{i_r}-\lambda_{j_r})^2+\eta^2]}$, $a_{i_r,j_r}^{\mathscr{S}}=a_{i_r,j_r}$ for $|i_r-j_r|\leqslant\ell$ and equal to zero otherwise. Here $\sum^*$ denotes that the sum is over all distinct indices. By [CES22a, Lemma 4.6], we have

$$\frac{1}{N\eta}\sum_{|j-i|\leqslant\ell}\sum_{\mathbf{x}\in\Lambda^n}\frac{\eta}{(\lambda_i-\lambda_j)^2+\eta^2}g_t(\mathbf{x})\mathcal{M}_{ij}g_t(\mathbf{x})\leqslant C(n)\langle g_t,\mathcal{A}g_t\rangle\leqslant 0,$$

which together with (3.14) and (3.15), implies

$$\begin{aligned}\partial_t\|g_t\|_2^2&\leqslant\frac{C(n)}{N\eta}\sum_{|j-i|\leqslant\ell}\sum_{\mathbf{x}\in\Lambda^n}\frac{\eta}{(\lambda_i-\lambda_j)^2+\eta^2}g_t(\mathbf{x})\mathcal{A}_{ij}g_t(\mathbf{x})+\frac{1}{N\eta^2}\|g_t\|_2^2\\&=\frac{1}{\eta}\sum_{\mathbf{x}\in\Lambda^n}\sum_{\mathbf{i},\mathbf{j}\in[N]^{2n}}^{*}\left(\prod_{r=1}^{2n}a_{i_r,j_r}^{\mathscr{S}}\right)\sum_{\mathbf{a},\mathbf{b}}^{*}g_t(\mathbf{x})\big(g_t(\mathbf{x}_{\mathbf{a},\mathbf{b}}^{\mathbf{ij}})-g_t(\mathbf{x})\big)\Psi_{\mathbf{a},\mathbf{b}}^{\mathbf{i}}(\mathbf{x})+\frac{1}{N\eta^2}\|g_t\|_2^2.\end{aligned} \tag{3.16}$$

For the term with $|g_t(\mathbf{x})|^2$ we can use the same bound as in [CES22a, (4.34)] and write

$$\begin{aligned}-\sum_{\mathbf{x}\in\Lambda^n}|g_t(\mathbf{x})|^2\sum_{\mathbf{a},\mathbf{b}}\sum_{\mathbf{i},\mathbf{j}\in[N]^{2n}}^{*}\Psi_{\mathbf{a},\mathbf{b}}^{\mathbf{i}}(\mathbf{x})\prod_{r=1}^{2n}a_{i_r,j_r}^{\mathscr{S}}&\\ \leqslant -C(n)\sum_{\mathbf{x}\in\Lambda^n}|g_t(\mathbf{x})|^2\sum_{\mathbf{a},\mathbf{b}}\sum_{\mathbf{i}}\Psi_{\mathbf{a},\mathbf{b}}^{\mathbf{i}}(\mathbf{x})&\leqslant -C_1(n)\|g_t\|_2^2+C(n)N^{n\xi}\Phi_{\mathbf{A}}K^{n-1}\end{aligned}$$

which gives that

$$\begin{aligned}\partial_t\|g_t\|_2^2\leqslant&-\frac{1}{\eta}\left(C_1(n)+\frac{1}{N\eta}\right)\|g_t\|_2^2+C(n)\frac{N^{n\xi}K^{n-1}}{\eta}\Phi_{\mathbf{A}}\\&+\frac{1}{\eta}\sum_{\mathbf{x}\in\Lambda^n}\sum_{\mathbf{i},\mathbf{j}\in[N]^{2n}}^{*}\left(\prod_{r=1}^{2n}a_{i_r,j_r}^{\mathscr{S}}\right)\sum_{\mathbf{a},\mathbf{b}}^{*}g_t(\mathbf{x})g_t(\mathbf{x}_{\mathbf{a},\mathbf{b}}^{\mathbf{ij}})\Psi_{\mathbf{a},\mathbf{b}}^{\mathbf{i}}(\mathbf{x}).\end{aligned} \tag{3.17}$$

We can now use the fact that $F(\mathbf{x})$ is in the kernel of $\mathscr{L}$ by Lemma 3.5 and use similar bounds as [MY22, (5.89-91), (5.95-97)] and [CES22b] to obtain

$$\begin{aligned}\sum_{j}^{*}\left(\prod_{i=1}^{2n}a_{i_r,j_r}^{\mathscr{S}}\right)\sum_{\mathbf{a},\mathbf{b}}^{*}g_t(\mathbf{x}_{\mathbf{a},\mathbf{b}}^{\mathbf{ij}})=&\,\mathrm{Av}(\mathbf{x})\sum_{j}^{*}\left(\prod_{r=1}^{2n}a_{i_r,j_r}\right)\sum_{\mathbf{a},\mathbf{b}}^{*}\left(f_t(\mathbf{x}_{\mathbf{a},\mathbf{b}}^{\mathbf{ij}})-F(\mathbf{x}_{\mathbf{a},\mathbf{b}}^{\mathbf{ij}})\right)\\&+\mathcal{O}\left(\frac{N^{\varepsilon+n\xi}\ell}{K}\Phi_{\mathbf{A}}+\frac{N^{1+n\xi}T}{\ell}\Phi_A+\frac{N^{1+n\xi}\eta}{\ell}\Phi_{\mathbf{A}}\right)\end{aligned} \tag{3.18}$$

which gives the results using [CES22a, (4.51)]. □

We now need to analyze the term involving the configurations where jumps have occurred. The main difference with previous analysis from [CES22a] is that we do not have a sum over all perfect matchings of the configurations with the same matrix $A$ but we need to consider each colored perfect matching separately.

**Lemma 3.10.** *On the event $\Omega_1 \cap \Omega_2$, we have that*

$$\sum_{\mathbf{i}}^{*} \sum_{\mathbf{a},\mathbf{b}} \Psi^{\mathbf{i}}_{\mathbf{a},\mathbf{b}}(\mathbf{x}) \sum_{\mathbf{j}}^{*} \left( \prod_{r=1}^{2n} a_{i_r,j_r} \right) \left( f_t(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) - F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) \right) = \mathcal{O}\left( \frac{N^{n\varepsilon}}{\sqrt{N\eta}} \Phi_{\mathbf{A}} \right).$$

*Proof.* We start by using the graph convention from (3.6). Remember that we are going to make particles jump to sites where there are no particles, thus two vertices from each site are going to jump to a new empty site. There are several possible cases for our configuration after jumps $\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}$, we now fix an edge $e \in \mathcal{E}$ and consider the different cases both for our observable $f_t$ and for the ansatz observable $F$. For simplicity, we introduce the following notation

$$F_e(\mathcal{G}) = \sum_{\Pi \in \mathrm{PM}[n(e)]} \sum_{\{i,j\} \in \Pi} \langle A_{e^{(i)}} A_{e^{\Pi(i)}} \rangle$$

Note that since we $\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}$ has either 2 or 0 particles at each site, we see that $(n(i)-1)!! = 1$ and we can write

$$F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) = \frac{1}{N^{n/2}} \prod_{e=\{k,\ell\} \in \widetilde{\mathcal{E}}} 2^{n(e)\mathbb{1}_{k=\ell}} F_e(\mathcal{G}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}})$$

**Case 1:** $e = \{i,i\}$ For this case, we have three possible cases of jumps. We recall that at the site $i$, we have $n(e)$ occurrences of the edge $\{i,i\}$.

- If for at least one occurrence of the edge, both vertices are selected to move to the same new site $j_r$, we obtain a subleading term. Indeed, the resulting graph has only one occurrence of the edge $e' = \{j_r, j_r\}$. Thus, since it is impossible to perform a perfect matching on one vertex, we have $F_{e'}(\mathcal{G}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) = 0$ and thus $F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) = 0$. For the observable $f_t(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}})$, it means that we obtain a term of the form

$$\begin{aligned}\sum_{j_r=1}^{*} a_{i,j_r} \langle \mathbf{u}_{j_r}, A\mathbf{u}_{j_r} \rangle &= \sum_{j_r=1}^{*} \frac{\eta \langle \mathbf{u}_{j_r}, A\mathbf{u}_{j_r} \rangle}{N[(\lambda_{i_r} - \lambda_{j_r})^2 + \eta^2]} \\ &= \mathrm{Im}\langle G_t(z_i) A \rangle + \mathcal{O}\left( \frac{N^{\xi}\sqrt{\langle A^2 \rangle}}{N^{3/2}\eta} \right) = \mathcal{O}\left( \frac{N^{\xi}\sqrt{\langle A^2 \rangle}}{N\sqrt{\eta}} \right)\end{aligned}$$

  where we added in the second equality the finitely many missing terms in the sum and we used the matrix-valued local law (3.8) in the third one. So that we can bound, for this case of jumps,

$$\begin{aligned}&\sum_{\mathbf{j}}^{*} \left( \prod_{r=1}^{2n} a_{i_r,j_r} \right) \left( f_t(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) - F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) \right) \\ &\qquad = \mathcal{O}\left( \frac{N^{n\xi}}{N\sqrt{\eta}} \sqrt{\langle A_{i_r}^2 \rangle} \prod_{i \neq i_r} \sqrt{\frac{\langle A_i \rangle}{N}} \right) = \mathcal{O}\left( \frac{N^{n\xi}}{\sqrt{N\eta}} \Phi_{\mathbf{A}} \right).\end{aligned}$$


Figure 5: This jump structure makes each edge from the initial configuration jump together to an edge on a distinct sites.

- We can take vertices from a cycle of occurrences $e^{(m_1)}, \dots, e^{(m_p)}$. In the sense that each pair of vertices which jumps to the same site $j_r$ is taken from two different occurrences. Suppose that $k \geqslant 3$, this creates $p$ single edges and thus for the same reason as above $F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) = 0$. For $f_t(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}})$, we obtain a term of the form,

$$\sum_{j_1,\dots j_p=1}^{N} a_{i,j_1} \dots a_{i,j_p} \langle \mathbf{u}_{j_1}, A'_1 \mathbf{u}_{j_2} \rangle \langle \mathbf{u}_{j_2}, A'_2 \mathbf{u}_{j_3} \rangle \dots \langle \mathbf{u}_{j_p}, A'_p \mathbf{u}_{j_1} \rangle$$
$$= N^{1-p} \langle \operatorname{Im} G_t(z_i) A'_1 \dots \operatorname{Im} G_t(z_i) A'_p \rangle = \mathcal{O}\left( \frac{N^{\xi - p/2}}{\sqrt{N\eta}} \prod_{i=1}^{p} \sqrt{\langle (A'_i)^2 \rangle} \right).$$

  where we reindexed the matrices as $A'_i = A_{e^{(m_i)}}$. Thus, by trivially bounding the rest of the edges as for above, we obtain a term

$$\sum_{\mathbf{j}}^{*} \left( \prod_{r=1}^{2n} a_{i_r, j_r} \right) \left( f_t(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) - F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) \right) = \mathcal{O}\left( \frac{N^{(n-p+1)\xi}}{\sqrt{N\eta}} \Phi_{\mathbf{A}} \right).$$


Figure 6: These jumps take two vertices from different edges and place them on the same site $j_r$. We see that a cycle has been created in the new configuration of particles. If $k \geqslant 3$, as in the left figure ($k = 4$), the ansatz observable vanishes while if $k = 2$, as in the right, the ansatz observable is nonzero.

- Now, if we consider two occurrences of the edge $e = \{x_{2k-1}, x_{2k}\}$ and $e' = \{x_{2\ell-1}, x_{2\ell}\}$ such that $x_{2k-1} = x_{2k} = x_{2\ell-1} = x_{2\ell} = i$. If we choose the two particles $x_{2k-1}$ and $x_{2\ell}$ to jump to $j_r$ and $x_{2k}$ and $x_{2\ell-1}$ to jump to $j_{r'}$, the new graph has exactly two occurrences of the edge $e'' = \{j_r, j_{r'}\}$. Thus, $F_{e''}(\mathcal{G}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) = \langle A_k A_\ell \rangle$ and

$$\sum_{j_r, j_{r'}=1}^{N} a_{i,j_r} a_{i,j_{r'}} F_{e''}(\mathcal{G}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}}) = (\operatorname{Im} m_t(z_i))^2 \langle A_k A_\ell \rangle.$$

  For the observable, the same reasoning holds and we obtain

$$\sum_{j_r, j_{r'}=1}^{N} a_{i,j_r} a_{i,j_{r'}} \langle \mathbf{u}_{j_r}, A_k \mathbf{u}_{j_{r'}} \rangle \langle \mathbf{u}_{j_r}, A_\ell \mathbf{u}_{j_{r'}} \rangle$$
$$= (\operatorname{Im} m_t(z_i))^2 \langle A_k A_\ell \rangle + \mathcal{O}\left( \frac{N^{\xi} \sqrt{\langle A_k^2 \rangle \langle A_\ell^2 \rangle}}{\sqrt{N\eta}} \right).$$

**Case 2:** $e=\{i_1,i_2\}$ **with** $i_1\neq i_2$ In this case we have two subcases when choosing the jumping particles.

- Suppose that the particle at site $i_1$ and $i_2$ jump respectively to the sites $j_1$ and $j_2$. If at the site $i_1$, the second chosen particle is attached to a particle at a site $i_p$ with $i_p\neq i_2$. Then after the jumps, the edge $\{j_1,j_2\}$ will be a single edge and thus $F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{ab}})=0$. For the observable $f_t$, this creates a cycle of a certain order as in the second point of the Case 1 above and we get that, for some $p>0$,

$$\sum_{\mathbf{j}}^{*}\left(\prod_{r=1}^{2n}a_{i_r,j_r}\right)\left(f_t(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}})-F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}})\right)=\mathcal{O}\left(\frac{N^{(n-p+1)\xi}}{\sqrt{N\eta}}\Phi_{\mathbf{A}}\right).$$

- Suppose again that the particle at site $i_1$ and $i_2$ jump respectively to the sites $j_1$ and $j_2$. Suppose now that the other chosen particle at site $i_1$ belongs to an occurrence of the edge $\{i_1,i_2\}$ and that its attached vertex is also jumping to site $j_2$, then we are exactly in the same case as the third item of Case 1.

Thus, we see that if we are in the first or second item of Case 1 or the first item of Case 2 for at least one edge, we obtain that

$$\sum_{\mathbf{j}}^{*}\left(\prod_{r=1}^{2n}a_{i_r,j_r}\right)\left(f_t(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}})-F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}})\right)=\mathcal{O}\left(\frac{N^{n\xi}}{\sqrt{N\eta}}\Phi_{\mathbf{A}}\right).$$

If we perform for all particles in the configuration, jumps as in the third item of Case 1 or the second of Case 2 (if possible), we obtain an error of

$$\sum_{\mathbf{j}}^{*}\left(\prod_{r=1}^{2n}a_{i_r,j_r}\right)\left(f_t(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}})-F(\mathbf{x}^{\mathbf{ij}}_{\mathbf{a},\mathbf{b}})\right)=\mathcal{O}\left(\frac{N^{n\xi}}{\sqrt{N\eta}}\sqrt{\prod_{k=1}^{n}\frac{\langle A_k^2\rangle}{N}}\right)=\mathcal{O}\left(\frac{N^{n\xi}}{\sqrt{N\eta}}\Phi_{\mathbf{A}}\right)$$

which concludes the proof. □

We can now proof Proposition 3.8.

*Proof of Proposition 3.8.* The proof is now the same as the $L^2$ bound proved in [CES22a, Equation (4.50) and beyond] where the error $\mathcal{E}$ is slightly changed to take into account the different matrices from $\mathbf{A}$. □

The $L^2$ bound from Proposition 3.8 can now be converted into a $L^\infty$ bound using the same techniques as [CES22a, Section 4.4] based on the estimates from [MY22] as stated in the following corollary.

**Corollary 3.11.** *Let $\delta>0$. For any $n\in\mathbb{N}$, there exists a constant $c(n)$ such that for any $\varepsilon>0$ and $T\geqslant N^{-1+\varepsilon}$, we have*

$$\frac{1}{\Phi_{\mathbf{A}}}\sup_{\mathcal{G}}|f_{\boldsymbol{\lambda},T}(\mathcal{G},\mathbf{A})-F(\mathcal{G},\mathbf{A})|\leqslant C(n,\varepsilon,\delta)N^{-c(n)}$$

*where the supremum holds for $\mathcal{G}=(\mathcal{V},\mathcal{E})$ with $|\mathcal{V}|=2n$ and $\mathcal{V}\subset\mathcal{I}_\alpha^{2n}$ and the bounds hold with very high probability.*

*Proof.* We first introduce a new dynamics which is a contraction from $L^2\to L^\infty$. It coincides with the short range dynamics (3.10) up until a time $T_1$ and then becomes a full-range dynamics up until a time $T_2\geqslant 2T_1$. The full-range dynamics is given by

$$\mathscr{W}(t)=\sum_{i\neq j\in[\![1,N]\!]}\mathscr{W}_{ij}(t),\quad \mathscr{W}_{ij}(t)h(\mathbf{x})=c_{ij}^{\mathscr{W}}(t)\frac{n_j(\mathbf{x})+1}{n_i(\mathbf{x})-1}\sum_{a\neq b\in[\![1,2n]\!]}\left(h(\mathbf{x}^{ij}_{ab})-h(\mathbf{x})\right),$$

where we defined

$$c_{ij}^{\mathscr{W}}(t) = \begin{cases} c_{ij}(t) & \text{if } i,j \in \mathcal{I}_\alpha \text{ and } 1 \leqslant |i-j| \leqslant \ell, \\ \frac{N}{|i-j|^2} & \text{otherwise.} \end{cases}$$

Thus we define the new observable

$$\begin{cases} \partial_t q_t(\mathbf{x};\mathbf{y}) = \mathscr{S}(t) q_t(\mathbf{x};\mathbf{y}) & \text{for } 0 \leqslant t \leqslant T_1, \\ \partial_t q_t(\mathbf{x};\mathbf{y}) = \mathscr{W}(t) q_t(\mathbf{x};\mathbf{y}) & \text{for } T_1 \leqslant t \leqslant T_2, \\ q_0(\mathbf{x};\mathbf{y}) = \mathrm{Av}(\mathbf{x};K,\mathbf{y})(f_0(\mathbf{x}) - F(\mathbf{x})). \end{cases}$$

By Proposition 3.8 we now that

$$\sup_{\mathbf{y}\in\mathcal{I}_\alpha} \|q_{T_1}(\cdot;\mathbf{y})\|_2 \leqslant K^{n/2}\mathcal{E}_n. \tag{3.19}$$

By [CES22a, (4.64)], we have a bound on the $L^\infty$ norm at time $T_2$,

$$\sup_{\mathbf{y}} \|q_{T_2}(\cdot;\mathbf{y})\|_\infty \leqslant \frac{N^{n\xi}}{(N(T_2-T_1))^{n/2}} \sup_{\mathbf{y}} \|q_{T_1}(\cdot;\mathbf{y})\|_2 + N^{n\xi}\frac{K^n}{N^n}. \tag{3.20}$$

This bound used the fact that $\mathscr{W}$ satisfies a Poincaré inequality and thus has a contractivity property from $L^2$ to $L^\infty$ from [MY22, Proposition 6.29]. By combining (3.19) and (3.20), we obtain

$$\sup_{\mathbf{y}} \|q_{T_2}(\cdot,\mathbf{y})\|_\infty \leqslant N^{n\xi}\left(\frac{K}{NT_2}\right)^{n/2}\mathcal{E}_n. \tag{3.21}$$

To get our result on the original dynamics $f_t$, we can use Lemma 3.7 and [CES22a, Lemma 4.3] to get that

$$\sup_{\mathbf{y}} |q_{T_2}(\mathbf{y},\mathbf{y}) - (f_{T_2}(\mathbf{y}) - F(\mathbf{y}))| \leqslant CN^\varepsilon\left(\frac{\ell}{K} + \frac{NT_2}{\ell}\right). \tag{3.22}$$

We get the final result by combining (3.21) and (3.22) and keeping the conditions on the parameters from Proposition 3.8. □

We conclude this section with the proof of Theorem 2.2.

*Proof of Theorem 2.2.* Let $\widehat{\Omega}$ be the event on which Corollary 3.11 holds. Since $\widehat{\Omega}$ holds with very high probability and we have the trivial bounds coming from the normalization of eigenvectors on the complement, we obtain uniformly in the graph $\mathcal{G}$,

$$\frac{1}{\Phi_{\mathbf{A}}}\left|\mathbb{E}\left[\prod_{e=\{k,\ell\}\in\mathcal{E}}\langle\mathbf{u}_k(t), A_i\mathbf{u}_\ell(t)\rangle\right] - \prod_{e=\{k,\ell\}\in\widetilde{\mathcal{E}}} 2^{\frac{n(e)}{2}\mathbb{1}_{k=\ell}} \sum_{\pi\in\mathrm{PM}[n(e)]}\prod_{\{i,j\}\in\Pi}\langle A_{e^{(i)}}, A_{e^{(j)}}\rangle\right| \ll 1.$$

By Wick's theorem, we obtain a Gaussian field with the covariance structure given by Theorem 2.2 for matrices $W$ with a small Gaussian component of size $t \sim N^{-1+\varepsilon}$. Then, this Gaussian component can be easily removed by a standard Green's Function Comparison (GFT) argument detailed in Section A which gives

$$\frac{1}{\Phi_{\mathbf{A}}}\left|\mathbb{E}\left[\prod_{e=\{k,\ell\}\in\mathcal{E}}\langle\mathbf{u}_k(t), A_i\mathbf{u}_\ell(t)\rangle\right] - \mathbb{E}\left[\prod_{e=\{k,\ell\}\in\mathcal{E}}\langle\mathbf{u}_k(0), A_i\mathbf{u}_\ell(0)\rangle\right]\right| \ll 1 \tag{3.23}$$

and thus Theorem 2.2 for any Wigner matrix. □

# A Green's function comparison (GFT)

Consider the Ornstein–Uhlenbeck flow

$$\mathrm{d}\widehat{W}_t = -\frac{\widehat{W}_t}{2}\mathrm{d}t + \frac{\mathrm{d}\widehat{B}_t}{\sqrt{N}}, \qquad \widehat{W}_0 = W, \tag{A.1}$$

with $\widehat{B}_t$ being a matrix valued real symmetric or complex Hermitian Brownian motion, and $W$ being a Wigner matrix satisfying Assumption 2.1. Note that along the flow (A.1) the first two moments of $\widehat{W}_t$ are preserved. Note also that the dynamics is slightly different from (3.1) but it is chosen as an initial condition for (3.1) such that at time $T$, their distributions are the same, see [CES22b, (B.1)–(B.3)] for more details. We denote the resolvent of $\widehat{W}_t$ by $G_t(z) := (\widehat{W}_t - z)^{-1}$, with $z \in \mathbb{C} \setminus \mathbb{R}$.

In this section we show that for $t$ sufficiently small the distribution of the eigenvectors of $\widehat{W}_t$ is very close to their distribution at time $t = 0$. The next proposition shows that as long as $t \ll N^{1/2}$ the change in the distribution of eigenvectors is negligible. We remark that for the purpose of proving Theorem 2.2 we need to choose $t = N^{-1+\varepsilon}$, for some small fixed $\varepsilon > 0$, hence much smaller that the threshold $N^{-1/2}$. The fact that to prove (3.23) it is enough to prove Proposition A.1 for resolvents at a spectral parameter $\eta = N^{-1-\zeta}$ follows by level repulsion [KY13].

**Proposition A.1.** *Fix any $p \in \mathbb{N}, \zeta, \delta > 0$, and consider a smooth function $F : \mathbb{R}^p \to \mathbb{R}$. Then there exists a constant $C > 0$ such that*

$$\Big| \mathbb{E} F\left(N^{-(k-2)/2} C_k^{-1} \langle \operatorname{Im} G_t(z_1) A_1 \dots \operatorname{Im} G_t(z_k) A_k \rangle\right)_{k \in [p]} \\ - \mathbb{E} F\left(\operatorname{Im} G_t \to \operatorname{Im} G_0\right)_{k \in [p]} \Big| \lesssim t N^{1/2 + C\zeta}. \tag{A.2}$$

*Here $z_i = E_i + \mathrm{i}\eta$, with $\eta = N^{-1-\zeta}$ and $E_i \in [-2+\delta, 2-\delta]$, $A_1, A_2, \dots$ are Hermitian matrices with $\langle A_i \rangle = 0$, and*

$$C_k = C_k(A_1, \dots, A_k) := \prod_{i \in [k]} \langle A_i^2 \rangle^{1/2}.$$

*Proof.* The proof of this proposition is similar to [CES22b, Appendix B], we thus only present an outline of it. To keep the notation short we define

$$R_t := F\left(N^{-(k-2)/2} C_k^{-1} \langle \operatorname{Im} G_t(z_1) A_1 \dots \operatorname{Im} G_t(z_k) A_k \rangle\right)_{k \in [p]}.$$

Then, by Itô's formula, we compute

$$\mathbb{E}\frac{\mathrm{d}}{\mathrm{d}t} R_t = \mathbb{E}\left[ -\frac{1}{2}\sum_{\alpha} w_\alpha(t) \partial_\alpha R_t + \frac{1}{2}\sum_{\alpha,\beta} \kappa_t(\alpha,\beta) \partial_\alpha \partial_\beta R_t \right], \tag{A.3}$$

where $\alpha, \beta \in [N]^2$ are double indices, $w_\alpha(t)$ are the entries of $\widehat{W}_t$, and $\partial_\alpha := \partial_{w_\alpha}$. Here

$$\kappa_t(\alpha_1, \dots, \alpha_l) := \kappa(w_{\alpha_1}(t), \dots, w_{\alpha_l}(t))$$

denotes the joint cumulant of $w_{\alpha_1}(t), \dots, w_{\alpha_l}(t)$, with $l \in \mathbb{N}$.

Note that by (2.1) it follows $|\kappa_t(\alpha_1, \dots, \alpha_l)| \lesssim N^{-l/2}$ uniformly in $t \geqslant 0$. Performing a cumulant expansion in (A.3) we are left with

$$\mathbb{E}\frac{\mathrm{d}}{\mathrm{d}t} R_t = \mathbb{E}\sum_{l=3}^{L} \sum_{\alpha_1, \dots, \alpha_l} \kappa_t(\alpha_1, \dots, \alpha_l) \partial_{\alpha_1} \cdots \partial_{\alpha_l} R_t + \Omega(L), \tag{A.4}$$

with $\Omega(L)$ a negligible error.

Next, we note that by [CES22b, Corollary 2.4] it follows that

$$\left|\langle \boldsymbol{x}, G_t(z_1)A_1 \dots G_t(z_k)A_k G_t(z_{k+1})\boldsymbol{y}\rangle\right| \lesssim N^{C\zeta} N^{k/2} \prod_{i\in[k]} \langle A_i^2\rangle^{1/2}, \tag{A.5}$$

for some constant $C > 0$ possibly depending on $k$. We point out that in [CES22b, Corollary 2.4] the bound (A.5) is proven for $\operatorname{Im} z_i \geqslant N^{-1+\varepsilon}$, for any small $\varepsilon > 0$; the fact that it can be extended to $\operatorname{Im} z_i = \eta$, i.e. a bit below the level spacing $N^{-1}$, follows similarly to [CES21a, Appendix A]. We now show this considering a specific example, the general case is completely analogous and so omitted; we explain the details at the end of this section.

Finally, by (A.5) and $|(G)_{ab}| \leqslant N^{\zeta}$ it readily follows that for any $k \in [p]$ we have

$$\left|\partial_{\alpha_1} \cdots \partial_{\alpha_l} N^{-(k-2)/2} C_k^{-1} \langle \operatorname{Im} G_t(z_1)A_1 \dots \operatorname{Im} G_t(z_k)A_k\rangle\right| \leqslant N^{C\zeta}, \tag{A.6}$$

for some large fixed constant $C > 0$. We thus conclude that

$$\left|\mathbb{E}\frac{\mathrm{d}}{\mathrm{d}t} R_t\right| \lesssim \sum_{l=3}^{L} N^{-l/2+2} N^{C\zeta} = N^{1/2+C\zeta}. \tag{A.7}$$

Integrating in time we conclude (A.2).

We now give a brief proof of (A.5), for simplicity we do not distinguish the indices, i.e. we write

$$(GA)^l = G_t(z_1)A_1 \dots G_t(z_l)A_l.$$

Assume that $\operatorname{Im} z_l = \eta$ and $|\operatorname{Im} z_j| \geqslant \eta_* := N^{-1+\varepsilon}$ for $j \neq l$, then

$$\begin{aligned}\langle \boldsymbol{x}, (GA)^{l-1} G(z_l) A (GA)^{k-l} G\boldsymbol{y}\rangle &= \langle \boldsymbol{x}, (GA)^{l-1} G(E_l + \mathrm{i}\eta_*) A (GA)^{k-l} G\boldsymbol{y}\rangle \\ &\quad - \int_{\eta}^{\eta_*} \langle \boldsymbol{x}, (GA)^{l-1} G(E_{q_0} + \mathrm{i}\tau)^2 A (GA)^{k-l} G\boldsymbol{y}\rangle \,\mathrm{d}\tau.\end{aligned} \tag{A.8}$$

For the first term in the rhs. we use the local law from [CES22b, Corollary 2.4] obtaining a bound as in (A.5). For the second term we estimate (neglecting $\log N$–factors)

$$\begin{aligned}&\int_{\eta}^{\eta_*} \langle \boldsymbol{x}, (GA)^{l-1} G(E_l + \mathrm{i}\tau)^2 A (GA)^{k-l} G\boldsymbol{y}\rangle \,\mathrm{d}\tau \\ &\lesssim \int_{\eta}^{\eta_*} \frac{1}{\tau} \langle \boldsymbol{x}, (GA)^{l-1} \operatorname{Im} G(E_l + \mathrm{i}\tau) ((GA)^{l-1})^* \boldsymbol{x}\rangle^{1/2} \times \\ &\times \langle \boldsymbol{y}, (A(GA)^{k-l}G)^* \operatorname{Im} G(E_l + \mathrm{i}\tau) A (GA)^{k-l} G\boldsymbol{y}\rangle^{1/2} \,\mathrm{d}\tau \\ &\lesssim \frac{\eta_*}{\eta} \int_{\eta}^{\eta_*} \frac{1}{\tau} \langle \boldsymbol{x}, (GA)^{l-1} \operatorname{Im} G(E_l + \mathrm{i}\eta_*) ((GA)^{l-1})^* \boldsymbol{x}\rangle^{1/2} \times \\ &\times \langle \boldsymbol{y}, (A(GA)^{k-l}G)^* \operatorname{Im} G(E_l + \mathrm{i}\eta_*) A (GA)^{k-l} G\boldsymbol{y}\rangle^{1/2} \,\mathrm{d}\tau \\ &\lesssim \frac{\eta_*}{\eta} \sqrt{N^{l-1} N^{k-l+1} \prod_{i\in[k]} \langle A_i^2\rangle} \lesssim N^{\zeta} N^{k/2} \prod_{i\in[k]} \langle A_i^2\rangle^{1/2}.\end{aligned} \tag{A.9}$$

This proves (A.5) choosing $\varepsilon \leqslant \zeta$. We remark that in the second inequality we used the monotonicity of the map $\tau \mapsto \tau\langle \boldsymbol{v}, \operatorname{Im} G(E + \mathrm{i}\tau)\boldsymbol{w}\rangle$, and that in the third inequality we used the local law from [CES22b, Corollary 2.4].

□

## References


[ABLM18] M. Abert, N. Bergeron, and E. Le Masson, *Eigenfunctions and random waves in the Benjamini–Schramm limit*, arXiv preprint arXiv:1810.05601 (2018).

[AS19] N. Anantharaman and M. Sabri, *Quantum ergodicity on graphs: from spectral to spatial delocalization*, Ann. of Math. (2) **189** (2019), no. 3, 753–835. MR3961083

[Ben21] L. Benigni, *Fermionic eigenvector moment flow*, Probab. Theory Related Fields **179** (2021), no. 3-4, 733–775. MR4242625

[Ber77] M. V. Berry, *Regular and irregular semiclassical wavefunctions*, Journal of Physics A: Mathematical and General **10** (1977), no. 12, 2083. MR0489542

[BGS84] O. Bohigas, M.-J. Giannoni, and C. Schmit, *Characterization of chaotic quantum spectra and universality of level fluctuation laws*, Physical review letters **52** (1984), no. 1, 1. MR0730191

[BL22a] L. Benigni and P. Lopatto, *Fluctuations in local quantum unique ergodicity for generalized Wigner matrices*, Comm. Math. Phys. **391** (2022), no. 2, 401–454. MR4397177

[BL22b] L. Benigni and P. Lopatto, *Optimal delocalization for generalized wigner matrices*, Advances in Mathematics **396** (2022), 108109. MR4370471

[Bru89] M.-F. Bru, *Diffusions of perturbed principal component analysis*, Journal of multivariate analysis **29** (1989), no. 1, 127–136. MR0991060

[BY17] P. Bourgade and H.-T. Yau, *The eigenvector moment flow and local quantum unique ergodicity*, Comm. Math. Phys. **350** (2017), no. 1, 231–278. MR3606475

[BYY20] P. Bourgade, H.-T. Yau, and J. Yin, *Random band matrices in the delocalized phase I: Quantum unique ergodicity and universality*, Comm. Pure Appl. Math. **73** (2020), no. 7, 1526–1596. MR4156609

[CDV85] Y. Colin De Verdière, *Ergodicité et fonctions propres du laplacien*, Comm. Math. Phys. **102** (1985), no. 3, 497–502. MR0818831

[CES21a] G. Cipolloni, L. Erdős, and D. Schröder, *Edge universality for non-hermitian random matrices*, Probability Theory and Related Fields **179** (2021), no. 1, 1–28. MR4221653

[CES21b] G. Cipolloni, L. Erdős, and D. Schröder, *Eigenstate thermalization hypothesis for Wigner matrices*, Comm. Math. Phys. **388** (2021), no. 2, 1005–1048. MR4334253

[CES22a] G. Cipolloni, L. Erdős, and D. Schröder, *Normal fluctuation in quantum ergodicity for Wigner matrices*, Ann. Probab. **50** (2022), no. 3, 984–1012. MR4413210

[CES22b] G. Cipolloni, L. Erdős, and D. Schröder, *Rank-uniform local law for Wigner matrices*, Forum Math. Sigma **10** (2022), Paper No. e96. MR4502022

[Deu18] J. Deutsch, *Eigenstate thermalization hypothesis*, Rep. Prog. Phys. **81** (2018), no. 8, 082001. MR3829412

[EKYY13] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin, *The local semicircle law for a general class of random matrices*, Electronic Journal of Probability **18** (2013), no. 59. MR3068390

[ESY09a] L. Erdős, B. Schlein, and H.-T. Yau, *Local semicircle law and complete delocalization for Wigner random matrices*, Comm. Math. Phys. **287** (2009), no. 2, 641–655. MR2481753

[ESY09b] L. Erdős, B. Schlein, and H.-T. Yau, *Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices*, Ann. Probab. **37** (2009), no. 3, 815–852. MR2537522

[ESY11] L. Erdős, B. Schlein, and H.-T. Yau, *Universality of random matrices and local relaxation flow*, Invent. Math. **185** (2011), no. 1, 75–119. MR2810797

[Hol10] R. Holowinsky, *Sieving for mass equidistribution*, Ann. of Math. (2) (2010), 1499–1516. MR2680498

[HS10] R. Holowinsky and K. Soundararajan, *Mass equidistribution for Hecke eigenforms*, Ann. of Math. (2) (2010), 1517–1528. MR2680499

[Ing21] M. Ingremeau, *Local weak limits of Laplace eigenfunctions*, Tunisian Journal of Mathematics **3** (2021), no. 3, 481–515. MR4258632

[KY13] A. Knowles and J. Yin, *Eigenvector distribution of Wigner matrices*, Probab. Theory Related Fields **155** (2013), no. 3–4, 543–582. MR3034787

[Lin06] E. Lindenstrauss, *Invariant measures and arithmetic quantum unique ergodicity*, Ann. of Math. (2) (2006), 165–219. MR2195133

[MY22] J. Marcinek and H.-T. Yau, *High Dimensional Normality of Noisy Eigenvectors*, Comm. Math. Phys. **395** (2022), no. 3, 1007–1096. MR4496381

[PT56] C. E. Porter and R. G. Thomas, *Fluctuations of nuclear reaction widths*, Physical Review **104** (1956), no. 2, 483.

[RS94] Z. Rudnick and P. Sarnak, *The behaviour of eigenstates of arithmetic hyperbolic manifolds*, Comm. Math. Phys. **161** (1994), no. 1, 195–213. MR1266075

[Shn74] A. Shnirelman, *Ergodic properties of eigenfunctions*, Russian Math. Surveys **29** (1974), no. 6, 181–182. MR0402834

[TV12] T. Tao and V. Vu, *Random matrices: universal properties of eigenvectors*, Random Matrices Theory Appl. **1** (2012), no. 01, 1150001. MR2930379

[VW15] V. Vu and K. Wang, *Random weighted projections, random quadratic forms and random eigenvectors*, Random Structures Algorithms **47** (2015), no. 4, 792–821. MR3418916

[Wig57] E. Wigner, *Gatlinberg conference on nneutron physics by time-of-flight*, Report ORNL **2309** (1957), 55.

[Zel87] S. Zelditch, *Uniform distribution of eigenfunctions on compact hyperbolic surfaces*, Duke Math. J. **55** (1987), no. 4, 919–941. MR0916129