

SPECTRAL ALIGNMENT OF CORRELATED GAUSSIAN MATRICES

LUCA GANASSALI* ** AND

MARC LELARGE,* *Inria, DI/ENS, PSL Research University*LAURENT MASSOULIÉ,* *MSR–Inria Joint Center and Inria, DI/ENS, PSL Research University*

Abstract

In this paper we analyze a simple spectral method (*EIGI*) for the problem of matrix alignment, consisting in aligning their leading eigenvectors: given two matrices A and B , we compute two corresponding leading eigenvectors v_1 and v'_1 . The algorithm returns the permutation $\hat{\pi}$ such that the rank of coordinate $\hat{\pi}(i)$ in v_1 and that of coordinate i in v'_1 (up to the sign of v'_1) are the same.

We consider a model of weighted graphs where the adjacency matrix A belongs to the Gaussian orthogonal ensemble of size $N \times N$, and B is a noisy version of A where all nodes have been relabeled according to some planted permutation π ; that is, $B = \Pi^T(A + \sigma H)\Pi$, where Π is the permutation matrix associated with π and H is an independent copy of A . We show the following zero–one law: with high probability, under the condition $\sigma N^{7/6+\epsilon} \rightarrow 0$ for some $\epsilon > 0$, *EIGI* recovers all but a vanishing part of the underlying permutation π , whereas if $\sigma N^{7/6-\epsilon} \rightarrow \infty$, this method cannot recover more than $o(N)$ correct matches.

This result gives an understanding of the simplest and fastest spectral method for matrix alignment (or complete weighted graph alignment), and involves proof methods and techniques which could be of independent interest.

Keywords: Graph alignment; random matrices; spectral theory

2020 Mathematics Subject Classification: Primary 68W40; Secondary 60B20

1. Introduction

The graph alignment problem

Graph alignment (or graph matching, or network alignment) consists in recovering the underlying vertex correspondence between two correlated graphs, and hence can be viewed as the noisy version of the isomorphism problem. Many questions can be phrased as graph alignment problems. They are found in various fields, such as network privacy and data de-anonymization [15, 16], biology and protein–protein interaction networks [20], natural language processing [13], and pattern recognition in image processing [6].

For two graphs of size N with adjacency matrices A and B , the graph matching problem can be formalized as an optimization problem:

$$\arg \max_{P \in \mathcal{S}_N} \langle A, PBP^T \rangle, \quad (1.1)$$

Received 3 December 2019; revision received 24 May 2021.

* Postal address: Inria, 2 rue Simone Iff, 75012 Paris, France.

** Email address: luca.ganassali@inria.fr

© The Author(s), 2022. Published by Cambridge University Press on behalf of Applied Probability Trust.

where the maximum is taken over all $N \times N$ permutation matrices, and $\langle \cdot, \cdot \rangle$ is the canonical matrix inner product. Note that for each $P \in \mathcal{S}_N$, PBP^T is the matrix obtained from B by relabeling the nodes according to P^{-1} . This formulation is a special case of the well studied *quadratic assignment problem* [19], which is known to be NP-hard in the worst case, as well as some of its approximations [14]. A natural idea is then to study the average-case version of this problem, when A and B are random instances. Following a recent line of work [7, 9, 11], this paper focuses on the case where the signal lies in the weights of edges between all pairs of nodes.

Related work

Some general spectral methods for random graph alignment are introduced in [11], based on representation matrices and low-rank approximations. These methods are tested over synthetic graphs and real data; however, no precise theoretical guarantee—e.g. an error control of the inferred mapping depending on the signal-to-noise ratio—can be found for such techniques.

Most recently, a spectral method for matrix and graph alignment (*GRAMPA*) was proposed in [9, 10], which computes a similarity matrix which takes into account all pairs of eigenvalues (λ_i, μ_j) and eigenvectors (u_i, v_j) of matrices A and B . The authors study the regime in which the method exactly recovers the underlying vertex correspondence, meeting the state-of-the-art performances for alignment of Erdős–Ri graphs in polynomial time, and improving the performances among spectral methods for matrix alignment. This method can tolerate a noise σ up to $O(1/\log N)$ to recover the entire underlying vertex correspondence. Since the computation of all eigenvectors is required, the time complexity of *GRAMPA* is at least $O(N^3)$.

It is important to note that the signs of eigenvectors are ambiguous: in order to optimize the cost function in practice, it is necessary to test over all possible signs of eigenvectors. This additional complexity has no consequence when reducing A and B to rank-one matrices, but becomes costly when the reduction made is of rank $k \gg 1$. This combinatorial observation makes implementation and analysis of general rank-reduction methods (such as the ones proposed in [11]) more difficult. We therefore focus on the analysis of the rank-one reduction (*EIGI* hereafter) which is the simplest and most natural spectral alignment method, where only the leading eigenvectors of A and B are computed, with time complexity $O(N^2)$, which is significantly less than that of *GRAMPA*.

Gaussian weighted graph matching: model and method

As mentioned above, we focus on the case where the graphs are complete, weight-correlated. Matrices A and B are thus symmetric, with correlated entries. A natural model recently studied in [7, 9, 11] is as follows: A and H are two $N \times N$ independent normalized matrices of the Gaussian orthogonal ensemble (GOE), i.e. such that for all $1 \leq i \leq j \leq N$, $\{A_{i,j}\}_{i < j}$ are independent,

$$A_{i,j} = A_{j,i} \sim \begin{cases} \frac{1}{\sqrt{N}} \mathcal{N}(0, 1) & \text{if } i \neq j, \\ \frac{\sqrt{2}}{\sqrt{N}} \mathcal{N}(0, 1) & \text{if } i = j, \end{cases} \quad (1.2)$$

and H is an independent copy of A . We define $B = \Pi^T (A + \sigma H) \Pi$, where Π is the matrix of some permutation π —e.g. random uniform—of $\{1, \dots, N\}$ (i.e. such that $\Pi_{i,j} = 1$ if and only if $i = \pi(j)$), and $\sigma = \sigma(N)$ is the *noise parameter*.

Given two vectors $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ whose coordinates are all distinct, the permutation ρ which *aligns* x and y is the permutation such that for all $1 \leq i \leq n$, the rank (for the usual order) of $x_{\rho(i)}$ in x is the rank of y_i in y .

Algorithm 1. *EIGI* algorithm for matrix alignment

```

Compute  $v_1$ , a normalized leading eigenvector of  $A$ 
Compute  $v'_1$ , a normalized leading eigenvector of  $B$ 
Compute  $\Pi_+$ , the permutation aligning  $v_1$  and  $v'_1$ 
Compute  $\Pi_-$ , the permutation aligning  $v_1$  and  $-v'_1$ 
if  $\langle A, \Pi_+ B \Pi_+^T \rangle \geq \langle A, \Pi_- B \Pi_-^T \rangle$ 
  | return  $\Pi_+$ 
else
  | return  $\Pi_-$ 
end

```

Remark 1.1. Note that in our model, all the probability distributions are absolutely continuous with respect to Lebesgue measure; thus the eigenvectors of A and B all have almost surely pairwise distinct coordinates.

We recall that the aim is to infer the underlying permutation Π given the observation of A and B . We now introduce our simple spectral algorithm *EIGI*, derived from [11], which computes and aligns the leading eigenvectors v_1 and v'_1 of A and B . This very natural method can be thought of as the relaxation of the quadratic assignment problem formulation (1.1) when reducing A and B to rank-one matrices $\lambda_1 v_1 v_1^T$ and $\lambda'_1 v'_1 v'^T_1$. Indeed, as soon as v_1 and v'_1 have pairwise distinct coordinates, it is easy to see that

$$\arg \max_{P \in \mathcal{S}_N} \left\langle \lambda_1 v_1 v_1^T, P \lambda'_1 v'_1 v'^T_1 P^T \right\rangle = \arg \max_{P \in \mathcal{S}_N} \pm v_1^T P v'_1 = \rho,$$

where ρ is the aligning permutation of v_1 and $\pm v'_1$. Computing the two normalized leading eigenvectors (i.e. those corresponding to the highest eigenvalues) v_1 and v'_1 of A and B , the *EIGI* algorithm (Algorithm 1) returns the aligning permutation of v_1 and $\pm v'_1$. The method then decides which permutation to output according to the scores.

The aim of this paper is to find the regime in which *EIGI* achieves *almost exact recovery*, i.e. recovers all but a vanishing fraction of nodes of the planted truth Π .

2. Notation, main results, and proof scheme

In this section we introduce some notation that will be used throughout this paper; we also mention the main results and the proof scheme.

2.1 Notation

- Recall that A and H are two $N \times N$ matrices drawn under the model (1.2). The matrix B is defined as $\Pi^T (A + \sigma H) \Pi$, where Π is a uniform $N \times N$ permutation matrix and σ is the noise parameter, depending on N .
- In the following, (v_1, v_2, \dots, v_N) (resp. $(v'_1, v'_2, \dots, v'_N)$) denotes an orthonormal basis of eigenvectors of A (resp. of B) with respect to the (real) eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ of A (resp. $\lambda'_1 \geq \lambda'_2 \geq \dots \geq \lambda'_N$ of B). Throughout the paper, the sign of v'_1 is fixed so that $\langle \Pi v_1, v'_1 \rangle > 0$.

- Denote by $\|\cdot\|$ the Euclidean norm of \mathbb{R}^N . Let $\langle \cdot, \cdot \rangle$ denote the corresponding inner product.
- For any estimator $\hat{\Pi}$ of Π , define its overlap:

$$\mathcal{L}(\hat{\Pi}, \Pi) := \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\hat{\Pi}(i)=\Pi(i)}. \tag{2.1}$$

This metric is used to quantify the quality of a given estimator of Π .

- The equality $\stackrel{(d)}{=}$ will refer to equality in distribution. An event A_N is said to hold *with high probability* (w.h.p.) if $\mathbb{P}(A_N)$ converges to 1 when $N \rightarrow \infty$.
- For two random variables $u = u(N)$ and $v = v(N)$, we will use the notation $u = o_{\mathbb{P}}(v)$ if $\frac{u(N)}{v(N)} \xrightarrow{\mathbb{P}} 0$ when $N \rightarrow \infty$. We also use this notation when $X = X(N)$ and $Y = Y(N)$ are N -dimensional random vectors: $X = o_{\mathbb{P}}(Y)$ if $\frac{\|X(N)\|}{\|Y(N)\|} \xrightarrow{\mathbb{P}} 0$ when $N \rightarrow \infty$.
- Define

$$\mathcal{F} := \left\{ f : \mathbb{N} \rightarrow \mathbb{R}_+ \mid \forall t > 0, N^t f(N) \rightarrow \infty, \frac{f(N)}{N^t} \rightarrow 0 \right\}.$$

For two random variables $u = u(N)$ and $v = v(N)$, $u \asymp v$ refers to equivalence w.h.p. up to some sub-polynomial factor, meaning that there exists a function $f \in \mathcal{F}$ such that

$$\mathbb{P} \left(\frac{v(N)}{f(N)} \leq u(N) \leq f(N)v(N) \right) \rightarrow 1.$$

Throughout the paper, all limits are taken as $N \rightarrow \infty$, and the dependency on N will most of the time be elided, as an abuse of notation.

2.2 Main results, proof scheme

The result shown can be stated as follows: there exists a condition—a threshold—on σ and N under which the *EIG1* method enables us to recover Π almost exactly, in terms of the overlap \mathcal{L} defined in (2.1). Above this threshold, we show that *EIG1* cannot recover more than a vanishing part of Π .

Theorem 2.1. (Zero-one law for the method.) *For all N , let Π_N denote an arbitrary permutation of size N , and let $\hat{\Pi}_N$ be the estimator obtained with the algorithm *EIG1*, for A and B of the model (1.2), with permutation Π_N and noise parameter σ . We have the following zero-one law:*

- (i) *If there exists $\epsilon > 0$ such that $\sigma = o(N^{-7/6-\epsilon})$, then*

$$\mathcal{L}(\hat{\Pi}_N, \Pi_N) \xrightarrow{L^1} 1.$$

- (ii) *If there exists $\epsilon > 0$ such that $\sigma = \omega(N^{-7/6+\epsilon})$, then*

$$\mathcal{L}(\hat{\Pi}_N, \Pi_N) \xrightarrow{L^1} 0.$$

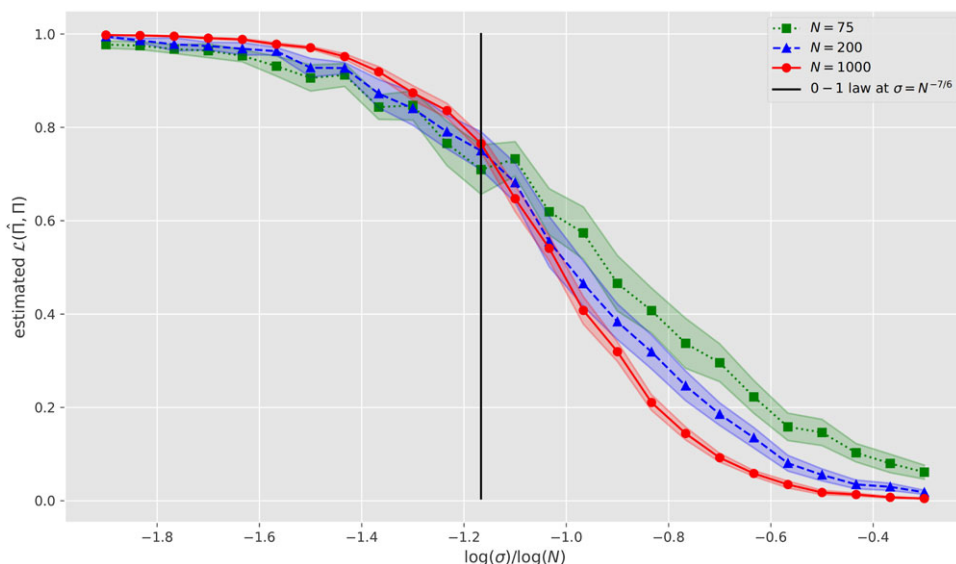


FIGURE 1. Estimated overlap $\mathcal{L}(\hat{\Pi}, \Pi)$ reached by *EIGI* in the model (1.2), for varying N and σ . (With 95% confidence intervals.)

The results of Theorem 2.1 are illustrated in Figure 1, which shows the zero–one law at $\sigma \asymp N^{-7/6}$. Note that the convergence to the step function appears to be slow.

Remark 2.1. We now underline that without loss of generality, we can assume that $\Pi = \text{Id}$, the identity mapping. Indeed, one can return to the general case by applying the transformations $A \rightarrow \Pi A \Pi^T$ and $H \rightarrow \Pi H \Pi^T$. From now on we will assume that $\Pi = \text{Id}$.

In order to prove this theorem, it is necessary to establish two intermediate results along the way, which may also be of independent interest. First, we study the behavior of v'_1 with respect to v_1 , showing that under some conditions on σ and N , the difference $v_1 - v'_1$ can be approximated by a renormalized Gaussian standard vector, multiplied by a variance term \mathbf{S} , where \mathbf{S} is a random variable whose behavior is well understood in terms of N and σ when $N \rightarrow \infty$. For this we work under the following assumption:

$$\exists \alpha > 0, \quad \sigma = o\left(N^{-1/2-\alpha}\right). \quad (2.2)$$

Proposition 2.1. Under the assumption (2.2), there exist a standard Gaussian vector $Z \sim \mathcal{N}(0, I_N)$, independent of v_1 , and a random variable $\mathbf{S} \asymp \sigma N^{1/6}$ such that

$$v'_1 = (1 + o_{\mathbb{P}}(1)) \left(v_1 + \mathbf{S} \frac{Z}{\|Z\|} \right).$$

Remark 2.2. The assumption (2.2) (or a tighter formulation) arises when studying the diffusion trajectories of eigenvalues and eigenvectors in random matrices, and corresponds to the microscopic regime in [2]. This assumption ensures that all eigenvalues of B are close enough to the eigenvalues of A . The comparison term is justified from random matrix theory: $N^{-1/2}$

is the typical amplitude of the spectral gaps $\sqrt{N}(\lambda_i - \lambda_{i+1})$ in the bulk—i.e. for i of order ηn with $\eta \in (0, 1)$ —which are the smaller ones.

Eigenvector diffusions in similar models (diffusion processes drawn with the scaling $\sigma = \sqrt{t}$) are studied in [2], where the main tool is the Dyson Brownian motion (see e.g. [3]) and its formulation for eigenvector trajectories, giving stochastic differential equations for the evolutions of $v'_{j(t)}$ with respect to vectors $v_i = v'_{i(0)}$. These equations lead to a system of stochastic differential equations for the overlaps $\langle v_i, v'_{j(t)} \rangle$, which is quite difficult to analyze rigorously. In this work we use a more elementary method to get an expansion of v'_1 around v_1 , for which this very condition (2.2) also appears.

Note that here, spectral gaps at the edge—i.e. for $i = O(1)$ —are of order $N^{-1/6}$, so the assumption (2.2) may not optimal for our study, and we expect Proposition 2.1 to hold up to $\sigma = o(N^{-1/6-\alpha})$. However, since the positive result of Theorem 2.1 holds in a much more restrictive regime (see the condition (i)), the condition (2.2) is enough for our purposes and allows a short and simple proof.

Proposition 2.1 suggests the study of v'_1 as a Gaussian perturbation of v_1 . The main question is now formulated as follows: what is the probability that the perturbation on v_1 has an impact on the overlap of the estimator $\hat{\Pi}$ from the *EIGI* method? To answer this question, we introduce a correlated Gaussian vectors model (referred to hereafter as the *toy model*) of parameters N and $s > 0$. In this model, we draw a standard Gaussian vector X of size N and $Y = X + sZ$ where Z is an independent copy of X . We will use the notation $(X, Y) \sim \mathcal{J}(N, s)$.

Define r_1 as the function that associates to any vector $T = (t_1, \dots, t_p)$ the rank of t_1 in T (for the usual decreasing order). For $(X, Y) \sim \mathcal{J}(N, s)$ we evaluate

$$p(N, s) := \mathbb{P}(r_1(X) = r_1(Y)).$$

Our second result shows that there is a zero–one law for the property of rank preservation in the toy model $\mathcal{J}(N, s)$.

Proposition 2.2. (Zero–one law for $p(N, s)$.) *In the correlated Gaussian vectors model we have the following:*

(i) *If $s = o(1/N)$ then*

$$p(N, s) \xrightarrow[N \rightarrow \infty]{} 1.$$

(ii) *If $s = \omega(1/N)$ then*

$$p(N, s) \xrightarrow[N \rightarrow \infty]{} 0.$$

These results are illustrated in Figure 2, which shows the zero–one law at $s \asymp N^{-1}$.

Organization of paper.

The Gaussian approximation of $v_1 - v'_1$ is established in Section 3 with the proof of Proposition 2.1. The toy model defined above is studied in Section 4, where Proposition 2.2 is established. Finally, we gather results of Propositions 2.1 and 2.2 in Section 5 to prove Theorem 2.1. Some additional proofs are deferred to Appendices A and B.

3. Behavior of the leading eigenvectors of correlated matrices

The main idea of this section is to find a first-order expansion of v'_1 around v_1 . Recall that we use the notation (v_1, v_2, \dots, v_N) for normalized eigenvectors of A , corresponding to the

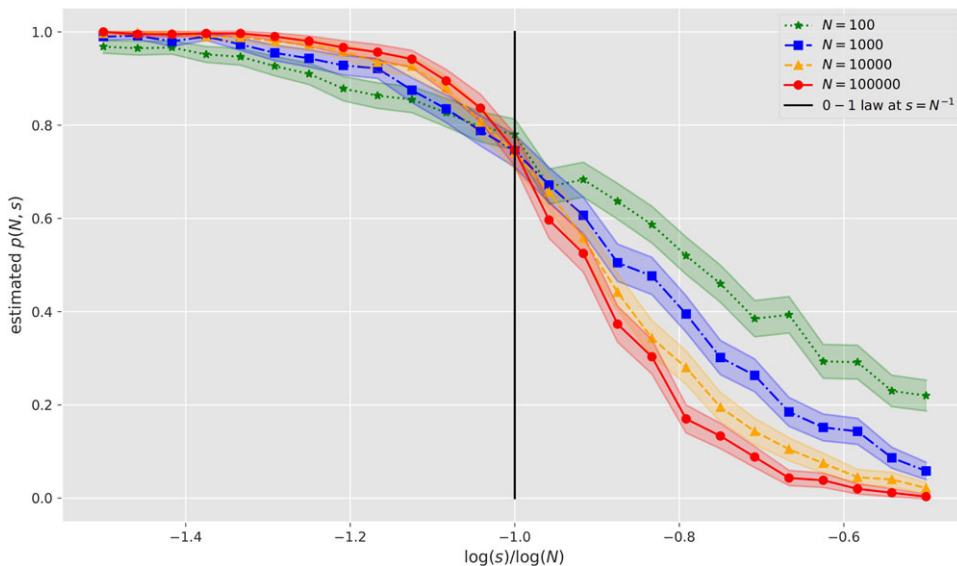


FIGURE 2. Estimated $p(N, s)$ in the toy model $\mathcal{J}(N, s)$. (With 95% confidence intervals.)

eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. Similarly, $(v'_1, v'_2, \dots, v'_N)$ and $\lambda'_1 \geq \lambda'_2 \geq \dots \geq \lambda'_N$ will refer to eigenvectors and eigenvalues of $B = A + \sigma H$. Since A and B are symmetric, all these eigenvalues are real and the vectors $\{v_i\}_i$ (resp. $\{v'_i\}_i$) are pairwise orthogonal. We also recall that v'_1 is taken such that $\langle v_1, v'_1 \rangle > 0$.

3.1 Computation of a leading eigenvector of B

Recall now that we are working under the assumption (2.2):

$$\exists \alpha > 0, \quad \sigma = o\left(N^{-1/2-\alpha}\right).$$

Let w' be a (non-normalized) eigenvector of B for the eigenvalue λ'_1 of the form

$$w' := \sum_{i=1}^N \theta_i v_i,$$

where we assume that $\theta_1 = 1$. Such an assumption can be made almost surely since any hyperplane of \mathbb{R}^N has a null Lebesgue measure in \mathbb{R}^N (see Remark 1.1).

The defining eigenvector equations projected on vectors v_i give

$$\begin{cases} \theta_1 = 1, \\ \forall i > 1, \theta_i = \frac{\sigma}{\lambda'_1 - \lambda_i} \sum_{j=1}^N \theta_j \langle H v_j, v_i \rangle, \\ \lambda'_1 - \lambda_1 = \sigma \sum_{j=1}^N \theta_j \langle H v_j, v_1 \rangle. \end{cases} \tag{3.1}$$

The strategy is then to approximately solve (3.1) with an iterative scheme, leading to the following expansion.

Proposition 3.1. *Under the assumption (2.2) one has the following:*

$$w' = v_1 + \sigma \sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle}{\lambda_1 - \lambda_i} v_i + o_{\mathbb{P}} \left(\sigma \sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle}{\lambda_1 - \lambda_i} v_i \right). \tag{3.2}$$

We refer to Appendix A.1 for the details regarding the definition of the mentioned iterative scheme, as well as a proof of Proposition 3.1. The proof uses the assumption (2.2) and builds upon some standard results on the distribution of eigenvalues in the GOE.

Remark 3.1. The above proposition could easily be extended to all eigenvectors of B , under the assumption (2.2). Based on the studies of the trajectories of the eigenvalues and eigenvectors in the Gaussian unitary ensemble (GUE) [2] and the GOE [1], since we are only interested here in the leading eigenvectors, we expect the conclusion of Proposition 3.1 to hold under the weaker assumption $\sigma N^{1/6+\alpha} \rightarrow 0$, for $N^{-1/6}$ is the typical spectral gap $\sqrt{N}(\lambda_1 - \lambda_2)$ at the edge. However, as explained before (see Remark 2.2), our analysis does not require this less restrictive assumption. We also know that the expansion (3.2) fails to hold as soon as $\sigma = \omega(N^{-1/6})$. A result proved by Chatterjee [5, Theorem 3.8] shows that the eigenvectors corresponding to the highest eigenvalues v_1 of A and v'_1 of $B = A + \sigma H$, when A and H are two independent matrices from the GUE, are delocalized (in the sense that $\langle v_1, v'_1 \rangle$ converges in probability to 0 as $N \rightarrow \infty$), when $\sigma = \omega(N^{-1/6})$.

3.2 Gaussian representation of $v'_1 - v_1$

We continue to work under the assumption (2.2). After renormalization, we have $v'_1 = \frac{w'}{\|w'\|}$. We are now able to study the behavior of the overlap $\langle v'_1, v_1 \rangle$:

$$\langle v'_1, v_1 \rangle = \left(1 + \sigma^2 (1 + o_{\mathbb{P}}(1)) \sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle^2}{(\lambda_1 - \lambda_i)^2} \right)^{-1/2}.$$

Hence

$$\langle v'_1, v_1 \rangle = 1 - \frac{\sigma^2}{2} \sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle^2}{(\lambda_1 - \lambda_i)^2} + o_{\mathbb{P}} \left(\sigma^2 \sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle^2}{(\lambda_1 - \lambda_i)^2} \right). \tag{3.3}$$

Let us give the heuristic to evaluate the first sum in the right-hand side of (3.3): since the GOE distribution is invariant by rotation (see e.g. [3]), the random variables $\langle H v_i, v_1 \rangle$ are zero-mean Gaussian, with variance $1/N$. Moreover, it is well known [3] that the eigenvalue gaps $\lambda_1 - \lambda_i$ are of order $N^{-1/6}$ when i is small, and $N^{-1/2}$ in the bulk (when i is typically of order N). These considerations lead to the following.

Lemma 3.1. *We have the following concentration:*

$$\sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle^2}{(\lambda_1 - \lambda_i)^2} \asymp N^{1/3}. \tag{3.4}$$

We refer to Appendix A.2 for a rigorous proof of this result. With this lemma, we are now able to give the first-order expansion of $\langle v'_1, v_1 \rangle$ with respect to σ :

$$\langle v'_1, v_1 \rangle = 1 - \frac{\sigma^2}{2} N^{1/3} + o_{\mathbb{P}} \left(\sigma^2 N^{1/3} \right).$$

Remark 3.2. The comparison between σ and $N^{1/6}$ made in [5] naturally reappears here, as $\sigma^2 N^{1/3}$ is the typical shift of v'_1 with respect to v_1 .

The intuition is that the scalar product $\langle v'_1, v_1 \rangle$ is sufficient to derive a Gaussian representation of v'_1 with respect to v_1 . We formalize this in the following lemma.

Lemma 3.2. Given v_1 , when writing the decomposition $w' = v_1 + w$, with

$$w := \sum_{i=2}^N \tilde{\theta}_i v_i,$$

the distribution of w is invariant by rotation in the orthogonal complement of v_1 . This implies in particular that given v_1 , $\|w\|$ and $\frac{w}{\|w\|}$ are independent, and that $\frac{w}{\|w\|}$ is uniformly distributed on \mathbb{S}^{N-2} , the unit sphere of v_1^\perp .

Proof of Lemma 3.2. We work conditionally on v_1 . Let O be an orthogonal transformation of the hyperplane v_1^\perp (such that $Ov_1 = v_1$). Since the GOE distribution is invariant by rotation and A and H are independent, $\tilde{B} := O^T A O + \sigma O^T H O$ has the same distribution as $B = A + \sigma H$.

Note that $Ow' = v_1 + Ow$ is an eigenvector of \tilde{B} for the eigenvalue λ_1 . Since the distribution of the matrix of eigenvectors (v_2, \dots, v_n) is the Haar measure on the orthogonal group $\mathcal{O}_{n-1}(v_1^\perp)$, denoted by $d\mathcal{H}$, the distribution of w is also invariant by rotation in the orthogonal complement of v_1 . Furthermore, for any bounded continuous functions f, g and $O \in \mathcal{O}_{n-1}(v_1^\perp)$,

$$\begin{aligned} \mathbb{E} \left[f(\|w\|) g \left(\frac{w}{\|w\|} \right) \right] &= \mathbb{E} \left[f(\|w\|) g \left(\frac{Ow}{\|Ow\|} \right) \right] \\ &= \mathbb{E} \left[f(\|w\|) \int_{\mathcal{O}_{n-1}(v_1^\perp)} d\mathcal{H}(O) g \left(\frac{Ow}{\|Ow\|} \right) \right] \\ &= \mathbb{E} \left[f(\|w\|) \int_{\mathbb{S}^{n-2}} \frac{g(u) du}{\text{Vol}(\mathbb{S}^{n-2})} \right] = \mathbb{E} [f(\|w\|)] \mathbb{E} \left[g \left(\frac{w}{\|w\|} \right) \right]. \end{aligned}$$

This completes the proof of Lemma 3.2. □

We can now prove the main result of this section, Proposition 2.1.

Proof of Proposition 2.1. Recall the decomposition $w' = v_1 + w$ with $w = \sum_{i=2}^N \theta_i v_i$. According to Lemma 3.2, conditioned to v_1 , $\frac{w}{\|w\|}$ is uniformly distributed on \mathbb{S}^{N-2} , the unit sphere of v_1^\perp . We now state a classical result about sampling uniform vectors on a sphere.

Lemma 3.3. Let E be p -dimensional Euclidean space, endowed with an orthogonal basis $\mathcal{B} = (e_1, \dots, e_p)$. Let u be a random vector uniformly distributed on the unit sphere \mathbb{S}^{p-1} of E . Then, in the basis \mathcal{B} , u has the same distribution as

$$\left(\frac{\xi_1}{\sqrt{\sum_{i=1}^p \xi_i^2}}, \dots, \frac{\xi_p}{\sqrt{\sum_{i=1}^p \xi_i^2}} \right),$$

where ξ_1, \dots, ξ_p are independent and identically distributed standard normal random variables.

We refer e.g. to [18, Lemma 10.1] for the proof of this result. In our context, this proves that the joint distribution of the coordinates w_2, \dots, w_n of w along v_2, \dots, v_n is always that of a normalized standard Gaussian vector (on \mathbb{R}^{N-1}). This joint probability does not depend on v_1 . Hence, there exist standard Gaussian independent variables Z_2, \dots, Z_N , independent from v_1 (and from $\|w\|$ by Lemma 3.2), such that

$$w' = v_1 + \frac{\|w\|}{\left(\sum_{i=2}^N Z_i^2\right)^{1/2}} \sum_{i=2}^N Z_i v_i.$$

Let Z_1 be another standard Gaussian variable, independent from everything else. Then

$$w' = \left(1 - \frac{\|w\|Z_1}{\left(\sum_{i=2}^N Z_i^2\right)^{1/2}}\right) v_1 + \frac{\|w\|}{\left(\sum_{i=2}^N Z_i^2\right)^{1/2}} \sum_{i=1}^N Z_i v_i.$$

Let $Z = \sum_{i=1}^N Z_i v_i$, which is a standard Gaussian vector. Since the distribution of Z is invariant by permutation of the $(Z_i)_{1 \leq i \leq N}$, Z and v_1 are independent. We have

$$\begin{aligned} v'_1 &= \frac{w'}{\|w'\|} = \frac{w'}{\sqrt{1 + \|w\|^2}} \\ &= \frac{1}{\sqrt{1 + \|w\|^2}} \left(1 - \frac{\|w\|Z_1}{\left(\sum_{i=2}^N Z_i^2\right)^{1/2}}\right) v_1 + \frac{\|w\|\|Z\|}{\sqrt{1 + \|w\|^2} \left(\sum_{i=2}^N Z_i^2\right)^{1/2}} \frac{Z}{\|Z\|}. \end{aligned}$$

Taking

$$\mathbf{S} = \frac{\|w\|\|Z\|}{\left(\sum_{i=2}^N Z_i^2\right)^{1/2} - \|w\|Z_1},$$

we get

$$v'_1 = \frac{1}{\sqrt{1 + \|w\|^2}} \left(1 - \frac{\|w\|Z_1}{\left(\sum_{i=2}^N Z_i^2\right)^{1/2}}\right) \left(v_1 + \mathbf{S} \frac{Z}{\|Z\|}\right). \tag{3.5}$$

Proposition 3.1 together with Lemma 3.1 yields

$$\|w\|^2 = \|w' - v_1\|^2 = (1 + o_{\mathbb{P}}(1)) \cdot \sigma^2 \sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle^2}{(\lambda_1 - \lambda_i)^2} \asymp \sigma^2 N^{1/3},$$

the last quantity being $o(1)$ under the assumption (2.2). With the previous computation, Equation (3.5) becomes

$$v'_1 = (1 + o_{\mathbb{P}}(1)) \left(v_1 + \mathbf{S} \frac{Z}{\|Z\|}\right),$$

with $\mathbf{S} = (1 + o_{\mathbb{P}}(1))\|w\| \asymp \sigma N^{1/6}$. □

4. Definition and analysis of a toy model

Now that we have established an expansion of v'_1 with respect to v_1 , our main question boils down to the study of the effect of a random Gaussian perturbation of a Gaussian vector in terms of the ranks of its coordinates: if these ranks are preserved, the permutation that aligns these two vectors will be $\hat{\Pi} = \Pi = \text{Id}$. Otherwise we want to understand the error between $\hat{\Pi}$ and $\Pi = \text{Id}$.

4.1 Definitions and notation

We refer to Section 2.1 for the definition of the toy model $\mathcal{J}(N, s)$. Recall that we want to compute, when $(X, Y) \sim \mathcal{J}(N, s)$, the probability

$$p(N, s) := \mathbb{P}(r_1(X) = r_1(Y)).$$

In this section, we denote by E the probability density function of a standard Gaussian variable, and by F its cumulative distribution function. Namely,

$$E(u) := \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \quad \text{and} \quad F(u) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-z^2/2} dz.$$

We hereafter elaborate on the link between this toy model and our first matrix model (1.2) in Section 3. Since v_1 is uniformly distributed on the unit sphere, we have the equality in distribution $v_1 = \frac{X}{\|X\|}$, where X is a standard Gaussian vector of size N , independent of Z by Proposition 2.1. We write

$$v_1 = \frac{X}{\|X\|},$$

$$v'_1 = (1 + o_{\mathbb{P}}(1)) \left(\frac{X}{\|X\|} + \mathbf{S} \frac{Z}{\|Z\|} \right).$$

Note that for all $\lambda > 0$, $r_1(\lambda T) = r_1(T)$; hence

$$r_1(v_1) = r_1(X), \quad r_1(v'_1) = r_1(X + \mathbf{s}Z), \tag{4.1}$$

where

$$\mathbf{s} = \frac{\mathbf{S}\|X\|}{\|Z\|} \asymp \sigma N^{1/6},$$

here we use the law of large numbers ($\|X\|/\|Z\| \rightarrow 1$ almost surely) as well as Proposition 2.1 in the last expansion. Equation (4.1) shows that this toy model is relevant for our initial problem, up to the fact that the noise term \mathbf{s} is random in the matrix model (though we know its order of magnitude to be $\asymp \sigma N^{1/6}$).

Remark 4.1. The intuition for the zero–one law for $p(N, s)$ is as follows. If we sort the N coordinates of X on the real axis, all coordinates being typically perturbed by a factor s , it seems natural to compare s with the typical gap between two coordinates of order $1/N$ to decide whether the rank of the first coordinate of X is preserved in Y .

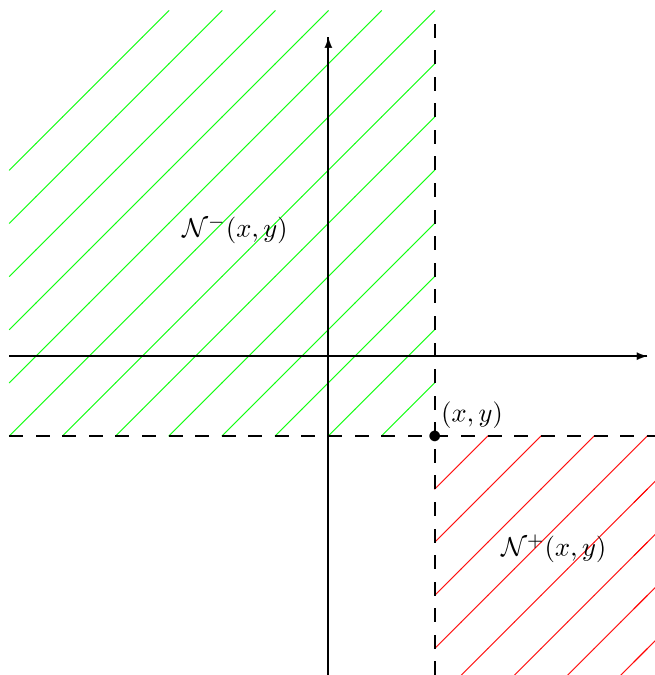


FIGURE 3. Areas corresponding to $\mathcal{N}^+(x, y)$ and $\mathcal{N}^-(x, y)$.

Let us show that this intuition is rigorously justified. For every pair (x, y) of real numbers, define

$$\mathcal{N}_{N,s}^+(x, y) := \#\{1 \leq i \leq N, X_i > x, Y_i < y\},$$

$$\mathcal{N}_{N,s}^-(x, y) := \#\{1 \leq i \leq N, X_i < x, Y_i > y\}.$$

In the following, we omit all dependencies on N and s , using the notation \mathcal{N}^+ and \mathcal{N}^- . The corresponding regions are shown in Figure 3. We will also need the following probabilities:

$$S^+(x, y) := \mathbb{P}(X_1 > x, Y_1 < y), \text{ and}$$

$$S^-(x, y) := \mathbb{P}(X_1 < x, Y_1 > y) = S^+(-x, -y).$$

In terms of distribution, the random vector

$$(\mathcal{N}^+(x, y), \mathcal{N}^-(x, y), N - 1 - \mathcal{N}^+(x, y) - \mathcal{N}^-(x, y))$$

follows a multinomial distribution of parameters

$$(N - 1, S^+(x, y), S^-(x, y), 1 - S^+(x, y) - S^-(x, y)).$$

In order to have $r_1(X) = r_1(Y)$, there must be the same number of points in the two domains in Figure 3, for $x = X_1$ and $y = Y_1$. We then have the following expression for $p(N, s)$:

$$\begin{aligned} p(N, s) &= \mathbb{E} \left[\mathbb{P}(\mathcal{N}^+(X_1, Y_1) = \mathcal{N}^-(X_1, Y_1)) \right] \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{P}(dx, dy) \mathbb{P}(\mathcal{N}^+(x, y) = \mathcal{N}^-(x, y)) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} E(x)E(z)\phi_{x,z}(N, s) dx dz, \end{aligned}$$

with

$$\phi_{x,z}(N, s) := \sum_{k=0}^{\lfloor (N-1)/2 \rfloor} \binom{N-1}{k} \binom{N-1-k}{k} (S_{x,z}^+)^k (S_{x,z}^-)^k (1 - S_{x,z}^+ - S_{x,z}^-)^{N-1-2k}, \quad (4.2)$$

using the notation $S_{x,z}^+ = S^+(x, x + sz)$ and $S_{x,z}^- = S^-(x, x + sz)$. A simple computation shows that

$$\begin{aligned} S^+(x, x + sz) &= \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \left(\int_{-\infty}^{z + \frac{x-u}{s}} \frac{1}{\sqrt{2\pi}} e^{-v^2/2} dv \right) du \\ &= \int_x^{+\infty} E(u) F \left(z - \frac{u-x}{s} \right) du, \end{aligned} \quad (4.3)$$

$$= s \int_0^{+\infty} E(x + vs) F(z - v) dv. \quad (4.4)$$

We have the classical integration result

$$\int_{-\infty}^z F(u) du = zF(z) + E(z). \quad (4.5)$$

From (4.3), (4.4), and (4.5) we derive the following easy lemma.

Lemma 4.1. *For all x and z ,*

$$\begin{aligned} S^+(x, x + sz) &\underset{s \rightarrow 0}{=} s [E(x) (zF(z) + E(z))] + o(s), \\ S^+(x, x + sz) &\underset{s \rightarrow \infty}{\rightarrow} F(z) (1 - F(x)), \\ S^-(x, x + sz) &\underset{s \rightarrow 0}{=} s [E(x) (-z + zF(z) + E(z))] + o(s), \\ S^-(x, x + sz) &\underset{s \rightarrow \infty}{\rightarrow} F(x) (1 - F(z)). \end{aligned}$$

Moreover, both $s \mapsto S^+(x, x + sz)$ and $s \mapsto S^-(x, x + sz)$ are increasing.

4.2. Zero-one law for $p(N, s)$

In this section we give a proof of Proposition 2.2.

Proof of Proposition 2.2. In the first case (i), if $s = o(1/N)$, we have the following inequality:

$$p(N, s) \geq \int_{\mathbb{R}} \int_{\mathbb{R}} dx dz E(x)E(z) \mathbb{P}(\mathcal{N}^+(x, x + sz) = \mathcal{N}^-(x, x + sz) = 0). \quad (4.6)$$

According to Lemma 4.1, for all $x, z \in \mathbb{R}$

$$\begin{aligned} \mathbb{P}(\mathcal{N}^+(x, x + sz) = \mathcal{N}^-(x, x + sz) = 0) &= (1 - S^+(x, x + sz) - S^-(x, x + sz))^{N-1} \\ &\sim \exp(-NsE(x)[z(2F(z) - 1) + 2E(z)]) \\ &\xrightarrow{N \rightarrow \infty} 1. \end{aligned}$$

By applying the dominated convergence theorem in (4.6), we conclude that $p(N, s) \rightarrow 1$.

In the second case (ii), if $sN \rightarrow \infty$, recall that

$$p(N, s) = \int_{\mathbb{R}} \int_{\mathbb{R}} dx dz E(x)E(z)\phi_{x,z}(N, s), \tag{4.7}$$

with $\phi_{x,z}$ defined in Equation (4.2). In the rest of the proof, we fix two real numbers x and z . Let

$$b(N, s, k) := \binom{N-1}{k} (S_{x,z}^+)^k (1 - S_{x,z}^+)^{N-1-k}$$

and

$$M(N, s) := \max_{0 \leq k \leq N-1} b(N, s, k).$$

Note that by Lemma 4.1, there exists $C = C(x, z) < 1$ such that for N large enough, $S_{x,z}^+ < C < 1$. Moreover, combining this lemma with the assumption (ii) gives that $NS_{x,z}^+ \rightarrow \infty$. It is also known that $M(N, s) = b(N, s, \lfloor NS_{x,z}^+ \rfloor)$, and a classical computation shows that in this case (see e.g. [4, Formula 1.5]),

$$\begin{aligned} M(N, s) &= \binom{N-1}{\lfloor NS_{x,z}^+ \rfloor} (S_{x,z}^+)^{\lfloor NS_{x,z}^+ \rfloor} (1 - S_{x,z}^+)^{N-1-\lfloor NS_{x,z}^+ \rfloor} \\ &\sim \frac{1}{\sqrt{2\pi Nt(1-t)}} t^{-(N-1)t} (1-t)^{-(N-1)(1-t)} (S_{x,z}^+)^{(N-1)t} (1 - S_{x,z}^+)^{(N-1)(1-t)} \\ &= (NS_{x,z}^+)^{-1/2} (1 + O(1)) \rightarrow 0, \end{aligned}$$

where

$$t := \frac{\lfloor NS_{x,z}^+ \rfloor}{N-1} \sim S_{x,z}^+.$$

Working with Equation (4.2), we obtain the following control:

$$\begin{aligned} \phi_{x,z}(N, s) &\leq M(N, s) \times \sum_{k=0}^{\lfloor (N-1)/2 \rfloor} \binom{N-1-k}{k} (S_{x,z}^-)^k \frac{(1 - S_{x,z}^+ - S_{x,z}^-)^{N-1-2k}}{(1 - S_{x,z}^+)^{N-1-k}} \\ &\stackrel{(a)}{=} M(N, s) \times \frac{(1 - S_{x,z}^+) \left(1 - \left(\frac{-S_{x,z}^+}{1 - S_{x,z}^-}\right)^N\right)}{1 + S_{x,z}^- - S_{x,z}^+} \\ &\stackrel{(b)}{=} M(N, s) \times O(1) \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

We used in (b) the facts that $S_{x,z}^+ + S_{x,z}^-$ is increasing in s , and that given x and z , for all $s > 0$, by Lemma 4.1,

$$S_{x,z}^+ + S_{x,z}^- < F(x)(1 - F(z)) + F(z)(1 - F(x)) < 1.$$

We used in (a) the following combinatorial result.

Lemma 4.2. For all $\alpha > 0$,

$$\sum_{k=0}^{\lfloor (N-1)/2 \rfloor} \binom{N-1-k}{k} \alpha^k = \frac{1}{\sqrt{1+4\alpha}} \left[\left(\frac{1+\sqrt{1+4\alpha}}{2} \right)^N - \left(\frac{1-\sqrt{1+4\alpha}}{2} \right)^N \right]. \quad (4.8)$$

We refer to Appendix B.1 for a proof of this result. To obtain (a) from Lemma 4.2, we apply (4.8) to

$$\alpha = \frac{S_{x,z}^- (1 - S_{x,z}^+)}{(1 - S_{x,z}^+ - S_{x,z}^-)^2},$$

with

$$\sqrt{1+4\alpha} = \frac{1 - S_{x,z}^+ + S_{x,z}^-}{1 - S_{x,z}^+ - S_{x,z}^-}.$$

Some straightforward simplifications then give the claimed result. The dominated convergence theorem in (4.7) shows that $p(N, s) \rightarrow 0$ and ends the proof. \square

Remark 4.2. The above computations also imply the existence of a non-degenerate limit of $p(N, s)$ in the critical case where $sN \rightarrow c > 0$: in this case, previous discussions as well as Lemma 4.1 show that the joint distribution of $(\mathcal{N}^+(x, x + sz), \mathcal{N}^-(x, x + sz))$ is asymptotically

$$\text{Poi}(c[E(x)(zF(z) + E(z))]) \otimes \text{Poi}(c[E(x)(-z + zF(z) + E(z))]).$$

Therefore, $p(N, s)$ has a non-degenerate limit given by

$$\int_{\mathbb{R}} \int_{\mathbb{R}} E(x)E(z) \cdot \mathbf{G}(c[E(x)(zF(z) + E(z))], c[E(x)(-z + zF(z) + E(z))]) dx dz, \quad (4.9)$$

where

$$\mathbf{G}(a, b) := \mathbb{P}(\text{Poi}(a) = \text{Poi}(b)) = e^{-(a+b)} \sum_{k \geq 0} \frac{a^k b^k}{(k!)^2}. \quad (4.10)$$

Note that $\mathbf{G}(a, b) \rightarrow 1$ when $a, b \rightarrow 0$, and $\mathbf{G}(a, b) \rightarrow 0$ when $a, b \rightarrow +\infty$.

5. Analysis of the EIGI method for matrix alignment

We now come back to our initial problem, which is the analysis of the EIGI method. Recall that for any estimator $\hat{\Pi}$ of Π , its overlap is defined as follows:

$$\mathcal{L}(\hat{\Pi}, \Pi) := \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\hat{\Pi}(i) = \Pi(i)}.$$

The aim of this section is to show how Propositions 2.1 and 2.2 can be assembled to prove the main result of our study, namely Theorem 2.1.

Proof of Theorem 2.1. In the first case (i), assuming $\sigma = o(N^{-7/6-\epsilon})$ for some $\epsilon > 0$, we have in particular that the condition (2.2) holds. Proposition 2.1 and Equation (4.1) in Section 4 enable us to identify v_1 and v'_1 with the following vectors:

$$v_1 \sim X, \quad v'_1 \sim X + \mathbf{s}Z, \tag{5.1}$$

where X and Z are two independent Gaussian vectors from the toy model, and where $\mathbf{s} \asymp \sigma N^{1/6}$ w.h.p. Recall that we work under the assumptions $\Pi = \text{Id}$ and $\langle v_1, v'_1 \rangle > 0$. In this case, we expect Π_+ to be very close to Id .

We will use the notation of Section 4 hereafter. Let us take $f \in \mathcal{F}$ such that w.h.p., $\sigma N^{1/6}f(N)^{-1} \leq \mathbf{s} \leq \sigma N^{1/6}f(N)$. For all $1 \leq i \leq N$, we have

$$\begin{aligned} \mathbb{P}(\Pi_+(i) = \Pi(i)) &= \mathbb{P}(\Pi_+(1) = \Pi(1)) \\ &= \mathbb{E} \left[\iint dx dz E(x)E(z)\phi_{x,z}(N, \mathbf{s}) \mathbf{1}_{\sigma N^{1/6}f(N)^{-1} \leq \mathbf{s} \leq \sigma N^{1/6}f(N)} \right] + o(1) \\ &= \iint dx dz E(x)E(z)\mathbb{E} \left[\phi_{x,z}(N, \mathbf{s}) \mathbf{1}_{\sigma N^{1/6}f(N)^{-1} \leq \mathbf{s} \leq \sigma N^{1/6}f(N)} \right] + o(1). \end{aligned}$$

When conditioning on the event \mathcal{A} where $\sigma N^{1/6}f(N)^{-1} \leq \mathbf{s} \leq \sigma N^{1/6}f(N)$, we know that $\mathbf{s}N \rightarrow 0$ by the condition (i), and for all x, z , we have $\mathbb{E}[\phi_{x,z}(N, \mathbf{s}) \mid \mathcal{A}] \rightarrow 1$ as shown in Section 4. Since \mathcal{A} occurs w.h.p. we have

$$\mathbb{E}[\phi_{x,z}(N, \mathbf{s}) \mathbf{1}_{\mathcal{A}}] \rightarrow 1,$$

which implies, together with the dominated convergence theorem, that

$$\mathbb{E}[\mathcal{L}(\Pi_+, \Pi)] \xrightarrow[N \rightarrow \infty]{} 1 \tag{5.2}$$

and thus

$$\mathcal{L}(\Pi_+, \Pi) \xrightarrow{L^1} 1.$$

We now check that w.h.p., Π_+ is preferred to Π_- in the *EIGI* method.

Lemma 5.1. *In the case (i), if $\langle v_1, v'_1 \rangle > 0$, we have w.h.p.*

$$\langle A, \Pi_+ B \Pi_+^T \rangle > \langle A, \Pi_- B \Pi_-^T \rangle;$$

in other words, the algorithm EIGI returns w.h.p. $\hat{\Pi} = \Pi_+$.

This lemma is proved in Appendix B.3 and implies, together with (5.2), that

$$\begin{aligned} \mathbb{E}[\mathcal{L}(\hat{\Pi}, \Pi)] &\geq \mathbb{E}[\mathcal{L}(\hat{\Pi}, \Pi)\mathbf{1}_{\hat{\Pi}=\Pi_+}] = \mathbb{E}[\mathcal{L}(\Pi_+, \Pi)\mathbf{1}_{\hat{\Pi}=\Pi_+}] \\ &= \mathbb{E}[\mathcal{L}(\Pi_+, \Pi)] - \mathbb{E}[\mathcal{L}(\Pi_+, \Pi)\mathbf{1}_{\hat{\Pi}=\Pi_-}] \\ &= 1 - o(1), \end{aligned}$$

and thus

$$\mathcal{L}(\hat{\Pi}, \Pi) \xrightarrow[N \rightarrow \infty]{L^1} 1. \tag{5.3}$$

In the second case (ii), if the condition (2.2) is verified, then the identification (5.1) still holds and the proof of the case (i) adapts well. However, if (2.2) is not verified, we can still make a link to the toy model studied in Section 4. We use a simple coupling argument: if $\sigma = \omega(N^{-1/2-\alpha})$ for some $\alpha \geq 0$, let us take $\sigma_1, \sigma_2 > 0$ such that

$$\sigma^2 = \sigma_1^2 + \sigma_2^2$$

and

$$N^{-7/6+\epsilon} \ll \sigma_1 \ll N^{-1/2-\alpha},$$

fixing for instance $\sigma_1 = N^{-1}$. We will use the notation \tilde{v}_1 , now viewed as the leading eigenvector of the matrix

$$\tilde{B} = A + \sigma_1 H + \sigma_2 \tilde{H},$$

where \tilde{H} is an independent copy of H . This has no consequence in terms of distribution: (A, \tilde{B}) is still drawn under the model (1.2). Let us denote by v'_1 the leading eigenvector of $B_1 = A + \sigma_1 H$, chosen so that $\langle v_1, v'_1 \rangle > 0$. It is clear that the condition (2.2) holds for σ_1 . We have the following result, based on the invariance by rotation of the GOE distribution.

Lemma 5.2. *We still have the following equality in distribution:*

$$(r_1(v_1), r_1(\tilde{v}_1)) \stackrel{(d)}{=} (r_1(X), r_1(X + \mathbf{s}Z)),$$

where X, Z are two standard Gaussian vectors from the toy model, with w.h.p.

$$\mathbf{s} \geq \mathbf{s}^1 \asymp \sigma_1 N^{1/6}.$$

We refer to Appendix B.2 for a proof. Since w.h.p. $\mathbf{s} \geq \mathbf{s}^1$ and $\mathbf{s}^1 N \asymp \sigma_1 N^{7/6} \rightarrow \infty$, for all $1 \leq i \leq N$ we have

$$\begin{aligned} \mathbb{P}(\Pi_+(i) = \Pi(i)) &= \mathbb{P}(\Pi_+(1) = \Pi(1)) \\ &= \mathbb{E} \left[\iint dx dz E(x) E(z) \phi_{x,z}(N, \mathbf{s}) \mathbf{1}_{\mathbf{s}N \rightarrow \infty} \right] + o(1) \\ &= \iint dx dz E(x) E(z) \mathbb{E} [\phi_{x,z}(N, \mathbf{s}) \mathbf{1}_{\mathbf{s}N \rightarrow \infty}] + o(1). \end{aligned}$$

With the same arguments as in the case (i), we show that $\phi_{x,z}(N, \mathbf{s}) \mathbf{1}_{\mathbf{s}N \rightarrow \infty} \xrightarrow{L^1} 0$, which implies

$$\mathbb{E} [\mathcal{L}(\Pi_+, \Pi)] \xrightarrow{N \rightarrow \infty} 0,$$

and hence $\mathcal{L}(\Pi_+, \Pi) \xrightarrow{N \rightarrow \infty} 0$. The last step is to verify that the overlap achieved by Π_- does not outperform that of Π_+ . We prove the following lemma in Appendix B.4.

Lemma 5.3. *In the case (ii), if $\langle v_1, v'_1 \rangle > 0$, we also have*

$$\mathcal{L}(\Pi_-, \Pi) \xrightarrow{N \rightarrow \infty} 0.$$

Lemma 5.3 then gives

$$\mathbb{E} [\mathcal{L}(\hat{\Pi}, \Pi)] \leq \mathbb{E} [\mathcal{L}(\Pi_+, \Pi)] + \mathbb{E} [\mathcal{L}(\Pi_-, \Pi)] \xrightarrow{N \rightarrow \infty} 0,$$

and thus

$$\mathcal{L}(\hat{\Pi}, \Pi) \xrightarrow[N \rightarrow \infty]{L^1} 0. \tag{5.4}$$

Of course, the convergences in (5.3) and (5.4) also hold in probability, by Markov’s inequality. \square

Remark 5.1. We conclude this work with a remark about the critical case in Theorem 2.1. In the light of Remark 4.2, if $\sigma \asymp N^{7/6}$, we expect $\mathcal{L}(\hat{\Pi}, \Pi)$ to have a non-degenerate limit \mathcal{L}_∞ in probability, whose behavior is random in $[0,1]$. Since our estimates are always made up to a sub-polynomial factor, it is nontrivial to establish such a result for Theorem 2.1 in the critical case without substantial additional work. However, we can make a guess about the expected result when $\sigma N^{7/6} \rightarrow c > 0$: in this case we conjecture that the expected value of \mathcal{L}_∞ is given by

$$\int_{\mathbb{R}} \int_{\mathbb{R}} E(x)E(z) \cdot \mathbf{G} (c [E(x) (zF(z) + E(z))], c [E(x) (-z + zF(z) + E(z))]) \, dx \, dz, \tag{5.5}$$

where \mathbf{G} is defined in (4.10) in Remark 4.2.

Appendix A. Additional proofs for Section 3

Throughout the proofs, all variables denoted by C_i with $i = 1, 2, \dots$ are unspecified, independent, positive constants.

A.1 Proof of Proposition 3.1

Proof of Proposition 3.1. Let us establish a first inequality: since the GOE distribution is invariant by rotation (see e.g. [3]), the random variables $\langle Hv_j, v_i \rangle$ are zero-mean Gaussian, with variance $1/N$ if $i \neq j$ and $2/N$ if $i = j$. Hence, w.h.p.

$$\sup_{1 \leq i, j \leq N} |\langle Hv_j, v_i \rangle| \leq C_1 \sqrt{\frac{\log N}{N}}. \tag{A.1}$$

We will use the following shorthand notation for $1 \leq i, j \leq N$:

$$m_{i,j} := \langle Hv_j, v_i \rangle.$$

The defining eigenvector equations projected on vectors v_i are written as follows:

$$\begin{cases} \theta_i &= \frac{\sigma}{\lambda'_1 - \lambda_i} \sum_{j=1}^N \theta_j m_{i,j}, \\ \lambda'_1 - \lambda_1 &= \sigma \sum_{j=1}^N \theta_j m_{1,j}. \end{cases} \tag{A.2}$$

In order to approximate the θ_i variables, we define the following iterative scheme:

$$\begin{cases} \theta_i^k &= \frac{\sigma}{\lambda_1^{k-1} - \lambda_i} \sum_{j=1}^N \theta_j^{k-1} m_{i,j}, \\ \lambda_1^k - \lambda_1 &= \sigma \sum_{j=1}^N \theta_j^{k-1} m_{1,j}, \end{cases} \tag{A.3}$$

with initial conditions $(\theta_i^0)_{2 \leq i \leq N} = 0$ and $\lambda_1^0 = \lambda_1$, and setting $\theta_1^k = 1$ for all k . For $k \geq 1$, define

$$\Delta_k := \sum_{i \geq 2} \left| \theta_i^k - \theta_i^{k-1} \right|,$$

and for $k \geq 0$,

$$S_k := \sum_{i \geq 1} |\theta_i^k|.$$

Recall that under the assumption (2.2), there exists $\alpha > 0$ such that $\sigma = o(N^{-1/2-\alpha})$. We define ϵ as follows:

$$\epsilon = \epsilon(N) = \sqrt{\sigma N^{1/2+\alpha}}.$$

The idea is to show that the sequence $\{\Delta_k\}_{k \geq 1}$ decreases geometrically with k at rate ϵ . More specifically, we show the following result.

Lemma A.1. *With the same notation and under the assumption (2.2) of Proposition 3.1, one has w.h.p.*

- (i) $\forall k \geq 1, \Delta_k \leq \Delta_1 \epsilon^{k-1}$;
- (ii) $\forall k \geq 0, \forall 2 \leq i \leq N, |\lambda_1^k - \lambda_i^k| \geq \frac{1}{2} |\lambda_1 - \lambda_i| (1 - \epsilon - \dots - \epsilon^{k-1})$;
- (iii) $\forall k \geq 0, S_k \leq 1 + (1 + \dots + \epsilon^{k-1}) \Delta_1$;
- (iv) $\sum_{i=2}^N |\theta_i - \theta_i^1|^2 = o\left(\sum_{i=2}^N |\theta_i^1|^2\right)$.

This lemma is proved in the next section. Equation (iv) of Lemma A.1 yields

$$\begin{aligned} w' &= v_1 + \sum_{i=2}^N \theta_i^1 v_i + \sum_{i=2}^N (\theta_i - \theta_i^1) v_i \\ &= v_1 + \sigma \sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle}{\lambda_1 - \lambda_i} v_i + o_{\mathbb{P}} \left(\sigma \sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle}{\lambda_1 - \lambda_i} v_i \right). \end{aligned}$$

□

Proof of Lemma A.1

Proof of Lemma A.1. In this proof we will use the same notation as defined in the proof of Proposition 3.1, and we make the assumption (2.2). We now state three technical lemmas controlling some statistics of eigenvalues in the GOE which will be useful hereafter.

Lemma A.2. *W.h.p., for all $\delta > 0$,*

$$\sum_{j=2}^N \frac{1}{\lambda_1 - \lambda_j} \leq O(N^{1+\delta}). \quad (\text{A.4})$$

Lemma A.3. *We have*

$$\sum_{j=2}^N \frac{1}{(\lambda_1 - \lambda_j)^2} \asymp N^{4/3}. \quad (\text{A.5})$$

Lemma A.4. *For any $C > 0$, w.h.p.*

$$\lambda_1 - \lambda_2 \geq N^{-2/3} (\log N)^{-C \log \log N}. \quad (\text{A.6})$$

Proofs of these three lemmas can be found in the next sections. We will work under the event (that occurs w.h.p.) on which the equations (A.4), (A.5), (A.6), (3.4), and (A.1) are satisfied. We show the following inequalities:

- (i) $\forall k \geq 1, \Delta_k \leq \Delta_1 \epsilon^{k-1}$;
- (ii) $\forall k \geq 0, \forall 2 \leq i \leq N, |\lambda_1^k - \lambda_i| \geq \frac{1}{2} |\lambda_1 - \lambda_i| (1 - \epsilon - \dots - \epsilon^{k-1})$;
- (iii) $\forall k \geq 0, S_k \leq 1 + (1 + \dots + \epsilon^{k-1}) \Delta_1$.

Recall that ϵ is given by

$$\epsilon = \epsilon(N) = \sqrt{\sigma N^{1/2+\alpha}}$$

We will denote by $f_i(N)$, with i an integer, functions as defined in Lemma A.3. All of the following inequalities will be valid for N large enough (uniformly in i and in k).

Step 1: Propagation of the first equation. Let $k \geq 3$. We work by induction, assuming that (i), (ii), and (iii) are verified up to $k - 1$. We have the following:

$$\begin{aligned} |\theta_i^k - \theta_i^{k-1}| &\leq \left| \frac{\sigma}{\lambda_1^{k-1} - \lambda_i} \sum_{j=2}^N (\theta_j^{k-1} - \theta_j^{k-2}) m_{i,j} \right| + \left| \frac{\sigma (\lambda_1^{k-2} - \lambda_1^{k-1})}{(\lambda_1^{k-1} - \lambda_i) (\lambda_1^{k-2} - \lambda_i)} \sum_{j=1}^N \theta_j^{k-2} m_{i,j} \right| \\ &\leq \frac{\sigma}{|\lambda_1^{k-1} - \lambda_i|} C_1 \sqrt{\frac{\log N}{N}} \Delta_{k-1} + \sigma C_1 \sqrt{\frac{\log N}{N}} S_{k-2} \frac{|\lambda_1^{k-2} - \lambda_1^{k-1}|}{|\lambda_1^{k-1} - \lambda_i| |\lambda_1^{k-2} - \lambda_i|} \\ &\stackrel{(a)}{\leq} \sigma \frac{3}{|\lambda_1 - \lambda_i|} C_1 \sqrt{\frac{\log N}{N}} \Delta_{k-1} + \sigma C_1 \sqrt{\frac{\log N}{N}} S_{k-2} \frac{9 |\lambda_1^{k-2} - \lambda_1^{k-1}|}{|\lambda_1 - \lambda_i|^2} \\ &\stackrel{(b)}{\leq} \sigma \frac{3}{|\lambda_1 - \lambda_i|} C_1 \sqrt{\frac{\log N}{N}} \Delta_{k-1} + \sigma C_1 \sqrt{\frac{\log N}{N}} 2 \frac{9 |\lambda_1^{k-2} - \lambda_1^{k-1}|}{|\lambda_1 - \lambda_i|^2}, \end{aligned}$$

where we applied (ii) to $k - 1, k - 2$ in (a) and (iii) to $k - 2$ in (b). Note that

$$|\lambda_1^{k-2} - \lambda_1^{k-1}| = \left| \sigma \sum_{j=1}^N (\theta_j^{k-2} - \theta_j^{k-3}) m_{i,j} \right| \leq \sigma C_1 \sqrt{\frac{\log N}{N}} \Delta_{k-2},$$

which yields the inequality

$$|\theta_i^k - \theta_i^{k-1}| \leq \frac{\sigma}{|\lambda_1 - \lambda_i|} f_1(N) N^{-1/2} \Delta_{k-1} + \frac{\sigma^2}{|\lambda_1 - \lambda_i|^2} f_2(N) N^{-1} \Delta_{k-2}.$$

We choose δ such that $0 < \delta < \alpha$ (where α is fixed by (2.2)), and we sum from $i = 2$ to N :

$$\begin{aligned} \Delta_k &\leq \sigma f_1(N) N^{1/2+\delta} \Delta_{k-1} + \sigma^2 f_3(N) N^{1/3} \Delta_{k-2} \\ &\stackrel{(a)}{\leq} o(\epsilon) \epsilon^{k-2} \Delta_1 + o(\epsilon^2) \epsilon^{k-3} \Delta_1 \\ &\leq \epsilon^{k-1} \Delta_1. \end{aligned}$$

Here we used $\sigma f_1(N) N^{1/2+\delta} = o(\epsilon), \sigma^2 f_3(N) N^{1/3} = o(\epsilon^2)$, and we applied (i) to $k - 1$ and $k - 2$ in (a).

Step 2: Propagation of the second equation. Let $k \geq 2$, and $0 < \delta < \alpha$. We work by induction, assuming that (i), (ii), and (iii) are verified up to $k - 1$. We have the following:

$$\begin{aligned} \left| \lambda_1^k - \lambda_1^{k-1} \right| &\leq \sigma f_1(N) N^{-1/2} \Delta_{k-1} \\ &\stackrel{(a)}{\leq} \sigma f_1(N) N^{-1/2} \epsilon^{k-2} \Delta_1 \\ &\leq N^{-2/3} (\log N)^{-C \log \log N} \epsilon^{k-2} \Delta_1 \\ &\leq \frac{\lambda_1 - \lambda_2}{2} \epsilon^{k-2} \Delta_1 \\ &\leq \frac{\lambda_1 - \lambda_i}{2} \epsilon^{k-2} \Delta_1. \end{aligned}$$

We applied (i) to $k - 1$ in (a). Note that

$$\Delta_1 = \sum_{j=2}^N \frac{\sigma}{\lambda_1 - \lambda_j} |m_{i,1}| \leq \sigma f_1(N) N^{1/2+\delta} \leq o(\epsilon).$$

Applying (ii) to $k - 1$, we get

$$\begin{aligned} \left| \lambda_1^k - \lambda_i \right| &\geq \left| \lambda_1 - \lambda_1^{k-1} \right| - \left| \lambda_1^k - \lambda_1^{k-1} \right| \\ &\geq \frac{\lambda_1 - \lambda_i}{2} \left(1 - \epsilon - \dots - \epsilon^{k-2} \right) - \frac{\lambda_1 - \lambda_i}{2} \epsilon^{k-1} \\ &\geq \frac{\lambda_1 - \lambda_i}{2} \left(1 - \epsilon - \dots - \epsilon^{k-1} \right). \end{aligned}$$

Step 3: Propagation of the third equation. Let $k \geq 1$. Here again, we work by induction, assuming that (i), (ii), and (iii) are verified up to $k - 1$. We have the following:

$$\begin{aligned} S_k &= 1 + \sum_{j=2}^N \left| \theta_j^k \right| \\ &\leq 1 + \Delta_k + S_{k-1} - 1 \\ &\stackrel{(a)}{\leq} \epsilon^{k-1} \Delta_1 + 1 + \left(1 + \dots + \epsilon^{k-2} \right) \Delta_1 \\ &\leq 1 + \left(1 + \epsilon + \dots + \epsilon^{k-1} \right) \Delta_1. \end{aligned}$$

We applied (i) to k and (iii) to $k - 1$ in (a).

Step 4: Proof of (i) for $k = 1, 2$, (ii) for $k = 0, 1$, and (iii) for $k = 0, 1$. The equation (i) for $k = 1$ is obvious. For $k = 2$, we have

$$\left| \theta_i^2 - \theta_i^1 \right| \leq \left| \frac{\sigma}{\lambda_1^1 - \lambda_i} \sum_{j=2}^N \left(\theta_j^1 - \theta_j^0 \right) m_{i,j} \right| + \left| \frac{\sigma (\lambda_1^0 - \lambda_1^1)}{(\lambda_1^1 - \lambda_i) (\lambda_1^0 - \lambda_i)} \sum_{j=1}^N \theta_j^0 m_{i,j} \right|.$$

We have

$$\begin{aligned} \left| \lambda_1^1 - \lambda_i \right| &\geq \left| \lambda_1 - \lambda_i \right| - \left| \lambda_1 - \lambda_1^1 \right| \geq \left| \lambda_1 - \lambda_i \right| - \sigma |m_{1,1}| \\ &\geq \left| \lambda_1 - \lambda_i \right| - \frac{1}{2} \left| \lambda_1 - \lambda_2 \right| \geq \frac{1}{2} \left| \lambda_1 - \lambda_i \right|, \end{aligned}$$

which shows (ii) for $k = 0, 1$. Thus, for $0 < \delta < \alpha$,

$$|\theta_i^2 - \theta_i^1| \leq \frac{2\sigma}{\lambda_1 - \lambda_i} C_1 \sqrt{\frac{\log N}{N}} \Delta_1 + \frac{4\sigma}{(\lambda_1 - \lambda_i)^2} \sigma |m_{1,1}| |m_{i,1}|,$$

and

$$\begin{aligned} \Delta_2 &\leq \sigma f_1(N) N^{1/2+\delta} \Delta_1 + 4\sigma |m_{1,1}| \sum_{i=2}^N \frac{\sigma |m_{i,1}|}{(\lambda_1 - \lambda_i)^2} \\ &\leq \sigma f_1(N) N^{1/2+\delta} \Delta_1 + 4\sigma f_4(N) N^{-1/2} N^{2/3} \sum_{i=2}^N \frac{\sigma |m_{i,1}|}{(\lambda_1 - \lambda_i)} \\ &\leq \sigma f_1(N) N^{1/2+\delta} \Delta_1 + 4\sigma f_4(N) N^{1/6} \Delta_1 \\ &\leq \epsilon \Delta_1. \end{aligned}$$

The proof of (iii) for $k = 0, 1$ is obvious.

Step 5: Proof of Equation (iv). Let $k \geq 2$ and $2 \leq i \leq N$. In the same way as in Step 1, we have

$$|\theta_i^k - \theta_i^{k-1}| \leq \frac{2\sigma C_1}{\lambda_1 - \lambda_i} \sqrt{\frac{\log N}{N}} \epsilon^{k-2} \Delta_1 + \frac{8\sigma^2 C_1^2}{(\lambda_1 - \lambda_i)^2} \frac{\log N}{N} \epsilon^{(k-3)+} \Delta_1.$$

In the right-hand term, the ratio of the second term to the first one is smaller than

$$\frac{4\sigma C_1}{\lambda_1 - \lambda_i} \sqrt{\frac{\log N}{N}} \epsilon^{-1} \leq \sigma N^{1/6} f(N) \epsilon^{-1} \leq \epsilon \rightarrow 0,$$

using Lemma A.4, with $f \in \mathcal{F}$. It follows that for N big enough (uniformly in k and i) one has

$$|\theta_i^k - \theta_i^{k-1}| \leq \frac{\sigma f(N)}{\lambda_1 - \lambda_i} N^{-1/2} \epsilon^{k-2} \Delta_1. \tag{A.7}$$

Equation (A.7) shows that the scheme (A.3) converges, and that the limits are indeed the solutions $\theta_1 = 1, \theta_2, \dots, \theta_N$ of the fixed-point equations. By a simple summation of (A.7) over $k \geq 2$, applying Lemma A.2 and the inequality (A.1), we have

$$|\theta_i - \theta_i^1| \leq \frac{2\sigma f(N)}{\lambda_1 - \lambda_i} N^{-1/2} \Delta_1 \leq \frac{2\sigma^2 f(N)}{\lambda_1 - \lambda_i} N^\delta,$$

where $\delta > 0$ is a positive quantity as in Lemma A.2, to be specified later. Using Lemma A.3 one has the following control:

$$\sum_{i=2}^N |\theta_i - \theta_i^1|^2 \leq 4\sigma^4 N^{2\delta} f(N) N^{4/3}.$$

Moreover, Lemma 3.1 shows that

$$\sum_{i=2}^N |\theta_i^1|^2 \asymp \sigma^2 N^{1/3} \geq g(N)^{-1} \sigma^2 N^{1/3},$$

where g is another function in \mathcal{F} . This yields

$$\sum_{i=2}^N \left| \theta_i - \theta_i^1 \right|^2 \leq \sum_{i=2}^N \left| \theta_i^1 \right|^2 4\sigma^2 N^{2\delta+1} f(N) g(N).$$

The proof is completed by taking $\delta = \alpha/2$ and applying (2.2). \square

Proof of Lemma A.4

Proof of Lemma A.4. This lemma provides a control of the spectral gap $\lambda_1 - \lambda_2$. Given a good rescaling (in $N^{2/3}$), the asymptotic joint law of the eigenvalues in the edge has been investigated in a great deal of research, for Gaussian ensembles and for more general Wigner matrices. The GOE case has mainly been studied by Tracy, Widom, and Forrester, among many others; in [12, 21], the convergence of the joint distribution of the first k eigenvalues towards a density distribution is established, as stated in the following proposition.

Proposition A.1. ([12, 21].) *For a given $k \geq 1$, and all s_1, \dots, s_k real numbers,*

$$\mathbb{P} \left(N^{2/3} (\lambda_1 - 2) \leq s_1, \dots, N^{2/3} (\lambda_k - 2) \leq s_k \right) \xrightarrow[N \rightarrow \infty]{} \mathcal{F}_{1,k}(s_1, \dots, s_k), \quad (\text{A.8})$$

where the $\mathcal{F}_{1,k}$ are continuous and can be expressed as solutions of nonlinear partial differential equations. Thus the rescaled spectral gap $N^{2/3} (\lambda_1 - \lambda_2)$ has a limit probability density law supported by \mathbb{R}_+ , which implies that

$$\mathbb{P} \left(N^{2/3} (\lambda_1 - \lambda_2) \geq (\log N)^{-C \log \log N} \right) \xrightarrow[N \rightarrow \infty]{} 1.$$

Of course, the choice of the function $N \mapsto (\log N)^{-C \log \log N}$ here is arbitrary, and the result is also true for any function tending to 0. \square

Proof of Lemma A.3

Proof of Lemma A.3. This result needs an understanding of the behavior of the spectral gaps of the matrix A , in the bulk and in the edges (left and right). The eigenvalues in the *edges* correspond to indices i such that $i = o(N)$ (left) or $i = N - o(N)$ (right). The eigenvalues in the *bulk* are the remaining eigenvalues. For this, we use a result on rigidity of eigenvalues, due to L. Erdős *et al.* [8], which consists in a control of the probability of the gap between the eigenvalues of A and the typical eigenvalues γ_j of the semicircle law, defined as follows:

$$\forall i \in \{1, \dots, n\}, \quad \frac{1}{2\pi} \int_{-2}^{\gamma_j} \sqrt{4 - x^2} dx = 1 - \frac{j}{N}. \quad (\text{A.9})$$

Proposition A.2. ([8].) *For some positive constants $C_5 > 0$ and $C_6 > 0$, for N large enough,*

$$\begin{aligned} \mathbb{P} \left(\exists j \in \{1, \dots, n\} \mid |\lambda_j - \gamma_j| \geq (\log N)^{C_5 \log \log N} (\min(j, N+1-j))^{-1/3} N^{-2/3} \right) \\ \leq C_5 \exp \left(- (\log N)^{C_6 \log \log N} \right). \end{aligned} \quad (\text{A.10})$$

Remark A.1. Another similar result that goes in the same direction for the GOE is already known: it has been shown by O'Rourke in [17] that the variables $\lambda_i - \gamma_i$ behave as Gaussian

variables when $N \rightarrow \infty$. However, the rigidity result in (A.10) obtained in [8] can be applied in more general models. This quantitative probabilistic statement was not previously known even for the GOE case.

Remark A.2. Let us note that one of the assumptions made in [8] is that variances of each column sum to 1, which is not directly the case in our model (1.2). Nevertheless, one may use (A.10) for the rescaled matrix $\tilde{A} := A \left(1 + \frac{1}{N}\right)^{-1/2}$, then easily check that there is a possible step back to A:

$$|\lambda_j - \gamma_j| \leq \left| \lambda_j \left(1 + \frac{1}{N}\right)^{-1/2} - \gamma_j \right| + N^{-1} + o(N^{-1}),$$

and $N^{-1} + o(N^{-1}) \leq 2 (\min(j, N + 1 - j))^{-1/3} N^{-2/3}$ for N big enough. Tolerating a slight increase of the constant C_5 , the result (A.10) is thus valid in the GOE.

Let us now compute an asymptotic expansion of γ_j in the right edge, which is for $j = o(N)$. Define

$$G(x) := \frac{1}{2\pi} \int_{-2}^x \sqrt{4 - t^2} dt = \frac{x\sqrt{4 - x^2} + 4 \arcsin(x/2)}{4\pi} + \frac{1}{2}, \tag{A.11}$$

for all $x \in [-2, 2]$. We have $\gamma_j = G^{-1}(1 - j/N) = -G^{-1}(j/N)$, observing that the integrand in (A.11) is an even function. We get the following expansion when $x \rightarrow -2$:

$$G(x) \underset{x \rightarrow -2}{=} \frac{2(x + 2)^{3/2}}{3\pi} + o\left((x + 2)^{3/2}\right),$$

which implies that

$$G^{-1}(y) \underset{y \rightarrow 0}{=} -2 + \left(\frac{3\pi y}{2}\right)^{2/3} + o\left(y^{2/3}\right),$$

and hence

$$\gamma_j \underset{j/N \rightarrow 0}{=} 2 - \left(\frac{3\pi j}{2N}\right)^{2/3} + o\left((j/N)^{2/3}\right). \tag{A.12}$$

Remark A.3. One can observe the coherence of this result that arises naturally in [17] as the expectation of the eigenvalues in the edge.

Let $\epsilon > 0$, to be specified later. To establish our result we will split the variables j into three sets:

$$A_1 := \left\{ 2 \leq j \leq (\log N)^{(C_5+1) \log \log N} \right\} \text{ (a small part of the right edge),}$$

$$A_2 := \left\{ (\log N)^{(C_5+1) \log \log N} < j \leq N^{1-\epsilon} \right\} \text{ (a larger part of the right edge),}$$

$$A_3 := \left\{ N^{1-\epsilon} < j \leq N \right\} \text{ (everything else).}$$

We show that the sum over A_1 is the major contribution in (A.5). The split in the right edge between A_1 and A_2 is driven by the error term of (A.10): this term is small compared to γ_j if and only if $(\log N)^{C_5 \log \log N} = o(j)$.

Step 1: Estimation of the sum over A_1 . According to (A.10) and Lemma (A.4), w.h.p.

$$N^{-4/3} (\log N)^{-C_6 \log \log N} \leq (\lambda_1 - \lambda_2)^2 \leq C_7 N^{-4/3} (\log N)^{C_6 \log \log N},$$

where C_6, C_7 are positive constants. Hence, w.h.p.

$$\begin{aligned} \frac{N^{4/3}}{C_7 (\log N)^{C_6 \log \log N}} &\leq \sum_{j \in A_1} \frac{1}{(\lambda_1 - \lambda_j)^2} \\ &\leq \sum_{j \in A_1} \frac{1}{(\lambda_1 - \lambda_2)^2} \\ &\leq N^{4/3} (\log N)^{(C_5 + C_6 + 1) \log \log N}. \end{aligned}$$

Step 2: Estimation of the sum over A_2 . Let us show that the sum over A_2 is asymptotically small compared to the sum over A_1 : using (A.10) and (A.13), we know that there exists $C_8 > 0$ such that for all $j \in A_2$, w.h.p.

$$\lambda_j = 2 - C_8 \left(\frac{j}{N}\right)^{2/3} + o\left((j/N)^{2/3}\right),$$

and we know furthermore (see e.g. [3]) that w.h.p.

$$\lambda_1 = 2 + o\left((j/N)^{2/3}\right) \quad \forall j \in A_2; \tag{A.13}$$

hence w.h.p.

$$\begin{aligned} \sum_{j \in A_2} \frac{1}{(\lambda_1 - \lambda_j)^2} &= N^{4/3} \sum_{j \in A_2} \frac{1}{C_9 j^{4/3} (1 + o(1))} \\ &= N^{4/3} (1 + o(1)) \sum_{j \in A_2} \frac{1}{C_9 j^{4/3}} = o\left(N^{4/3}\right), \end{aligned}$$

using in the last line the fact that the Riemann series $\sum j^{-4/3}$ converges.

Step 3: Estimation of the sum under A_3 . From the previous results (A.10), (A.13), and (A.14), assuming that $\epsilon < 1$, we get w.h.p.

$$\lambda_1 - \lambda_{N^{1-\epsilon}} = C_8 N^{-2\epsilon/3} + O\left(N^{-2\epsilon/3}\right),$$

which gives w.h.p. the following control:

$$\begin{aligned} \sum_{j \in A_3} \frac{1}{(\lambda_1 - \lambda_j)^2} &\leq (N - N^{1-\epsilon}) \frac{1}{(\lambda_1 - \lambda_{N^{1-\epsilon}})^2} \\ &= (N - N^{1-\epsilon}) \frac{N^{4\epsilon/3}}{C_9 (1 + o(1))} = O\left(N^{1+4\epsilon/3}\right) = o\left(N^{4/3}\right), \end{aligned}$$

as long as $\epsilon < 1/4$. Taking such an ϵ , these three controls complete the proof. \square

Proof of Lemma A.2

Proof of Lemma A.2. We follow the same steps as in the proof of Lemma A.3. Let us take $\delta > 0$. We split the j variables into three sets:

$$\begin{aligned} A_1 &:= \left\{ 2 \leq j \leq N^{1/3} \right\}, \\ A_2 &:= \left\{ N^{1/3} < j \leq N^{1-\delta} \right\}, \\ A_3 &:= \left\{ N^{1-\delta} < j \leq N \right\}. \end{aligned}$$

We use Lemma A.4 to obtain the following control: w.h.p.

$$\sum_{j \in A_1} \frac{1}{\lambda_1 - \lambda_j} \leq N^{1/3} N^{2/3} (\log N)^{C_5 \log \log N} = O(N^{1+\delta}).$$

Similarly, for A_2 ,

$$\begin{aligned} \sum_{j \in A_2} \frac{1}{\lambda_1 - \lambda_j} &\leq \sum_{j \in A_2} \frac{1}{o(N^{-2/3}) + C_8(j/N)^{2/3} + O((\log N)^{C_5 \log \log N} N^{-2/3} j^{-1/3})} \\ &= N^{2/3} \sum_{j \in A_2} \frac{1}{o(1) + C_8 j^{2/3}} \leq C_{10} N^{2/3} N^{(1-\delta)/3} \leq O(N^{1+\delta}). \end{aligned}$$

Finally, using the Cauchy–Schwarz inequality,

$$\sum_{j \in A_3} \frac{1}{\lambda_1 - \lambda_j} \leq \sqrt{N} \left(\sum_{j \in A_3} \frac{1}{(\lambda_1 - \lambda_j)^2} \right)^{1/2} \leq \sqrt{N} O(N^{1/2+2\delta/3}) = O(N^{1+\delta}).$$

□

A.2 Proof of Lemma 3.1

Proof of Lemma 3.1. We show that w.h.p.

$$\sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle^2}{(\lambda_1 - \lambda_i)^2} - \frac{1}{N} \sum_{i=2}^N \frac{1}{(\lambda_1 - \lambda_i)^2} = o \left(\frac{1}{N} \sum_{i=2}^N \frac{1}{(\lambda_1 - \lambda_i)^2} \right). \tag{A.14}$$

Let us recall that H is drawn according to the GOE; hence its law is invariant by rotation. This implies that the $\langle H v_i, v_1 \rangle$ are independent variables with variance $1/N$, independent of $\lambda_1, \dots, \lambda_N$. Define

$$M_N := \sum_{i=2}^N \frac{\langle H v_i, v_1 \rangle^2 - 1/N}{(\lambda_1 - \lambda_i)^2}.$$

Computing the second moment of M_N , we get

$$\mathbb{E} \left[M_N^2 \mid \lambda_1, \dots, \lambda_N \right] = \text{Var}(M_N \mid \lambda_1, \dots, \lambda_N) = \frac{1}{N^4} \sum_{i=2}^N \frac{2}{(\lambda_1 - \lambda_i)^4}.$$

Adapting the proof of Lemma A.3, following the same steps, one can also show that w.h.p.

$$\sum_{i=2}^N \frac{1}{(\lambda_1 - \lambda_i)^4} \asymp N^{8/3}. \quad (\text{A.15})$$

Let $\epsilon = \epsilon(N) > 0$ to be specified later. By Markov's inequality

$$\begin{aligned} \mathbb{P} \left(|M_N| \geq \frac{\epsilon}{N} \sum_{i=2}^N \frac{1}{(\lambda_1 - \lambda_i)^2} \middle| \lambda_1, \dots, \lambda_N \right) &\leq \frac{N^2 \mathbb{E} [M_N^2 | \lambda_1, \dots, \lambda_N]}{\left(\sum_{i=2}^N \frac{1}{(\lambda_1 - \lambda_i)^2} \right)^2} \\ &\asymp \frac{1}{\epsilon^2 N^2}, \end{aligned}$$

by Lemma A.3 and Equation (A.16). Taking e.g. $\epsilon(N) = N^{-1/2}$ concludes the proof. \square

Appendix B. Additional proofs for Sections 4 and 5

B.1 Proof of Lemma 4.2

Proof of Lemma 4.2. We fix $\alpha > 0$, and we want to prove

$$\sum_{k=0}^{\lfloor (N-1)/2 \rfloor} \binom{N-1-k}{k} \alpha^k = \frac{1}{\sqrt{1+4\alpha}} \left[\left(\frac{1+\sqrt{1+4\alpha}}{2} \right)^N - \left(\frac{1-\sqrt{1+4\alpha}}{2} \right)^N \right]. \quad (\text{B.1})$$

In the following, we define

$$\phi_+ := \frac{1 + \sqrt{1+4\alpha}}{2}, \quad \phi_- := \frac{1 - \sqrt{1+4\alpha}}{2},$$

and for all $N \geq 1$,

$$u_N = u_N(\alpha) := \sum_{k=0}^{\lfloor (N-1)/2 \rfloor} \binom{N-1-k}{k} \alpha^k.$$

We clearly have $u_N(\alpha) \leq (1+\alpha)^N$. For all $t > 0$ small enough (e.g. $t < 1/(1+\alpha)$), define

$$f(t) := \sum_{N=1}^{\infty} u_N t^N.$$

On the one hand,

$$\begin{aligned} \frac{t}{1-t-\alpha t^2} &= t \sum_{m=0}^{\infty} (t + \alpha t^2)^m = \sum_{m=0}^{\infty} \sum_{l=0}^m \binom{m}{l} \alpha^l t^{l+m+1} \\ &= \sum_{N=1}^{\infty} \left(\sum_{\substack{0 \leq l \leq m \\ l+m=N-1}} \binom{m}{l} \alpha^l \right) t^N = \sum_{N=1}^{\infty} u_N t^N = f(t). \end{aligned}$$

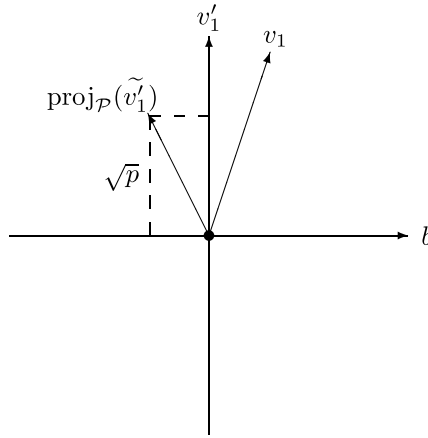


FIGURE 4. Orthogonal projection of \tilde{v}_1 on $\mathcal{P} := \text{span}(v'_1, v_1)$.

On the other hand,

$$\begin{aligned} \frac{t}{1-t-\alpha t^2} &= \frac{t}{(1-\phi_-t)(1-\phi_+t)} = \frac{1}{\phi_+ - \phi_-} \left(\frac{1}{1-\phi_+t} - \frac{1}{1-\phi_-t} \right) \\ &= \frac{1}{\sqrt{1+4\alpha}} \sum_{N=1}^{\infty} (\phi_+^N - \phi_-^N) t^N. \end{aligned}$$

This proves (B.1). □

B.2 Proof of Lemma 5.2

Proof of Lemma 5.2. Let us represent the situation in the plane spanned by v_1 and v'_1 , as shown in Figure 4.

Since \tilde{v}_1 is taken such that $\langle v_1, \tilde{v}_1 \rangle > 0$ and σ_1 satisfies (2.2), we have $\langle \tilde{v}_1, v'_1 \rangle > 0$ for N large enough by Proposition (2.1). Let $p := \langle \tilde{v}_1, v'_1 \rangle^2$ and $\tilde{w} := \tilde{v}_1 - \sqrt{p}v'_1 \in (v'_1)^\perp$. By invariance by rotation we obtain that

$$\frac{\tilde{w}}{\|\tilde{w}\|} = \frac{\tilde{w}}{\sqrt{1-p}}$$

is uniformly distributed on the unit sphere \mathbb{S}^{N-2} of $(v'_1)^\perp$, and independent of p , v_1 , and v'_1 . Hence

$$\langle b, \tilde{v}_1 \rangle = \langle b, \tilde{w} \rangle \stackrel{(d)}{=} \sqrt{1-p} \cdot \frac{\tilde{Z}_1}{\sqrt{\sum_{i=1}^{N-1} (\tilde{Z}_i)^2}},$$

where the \tilde{Z}_i are independent Gaussian standard variables, independent from everything else. From Section 3 we know that $1 - \langle v_1, v'_1 \rangle \asymp \sigma_1^2 N^{1/3}$ and thus $\langle v_1, b \rangle \asymp \sigma_1 N^{1/6}$. This yields, for

N large enough, w.h.p.

$$\begin{aligned} 0 < \langle \tilde{v}_1, v_1 \rangle &\leq \sqrt{p} \langle v_1, v'_1 \rangle + \sqrt{\frac{1-p}{N}} \tilde{Z}_1 \sigma_1 N^{1/6} f(N) \\ &\leq \sqrt{p} \langle v_1, v'_1 \rangle + \sqrt{1-p} N^{-4/3} g(N) \\ &\leq \max(\sqrt{p}, \sqrt{1-p}) \langle v_1, v'_1 \rangle \\ &\leq \langle v_1, v'_1 \rangle, \end{aligned}$$

where f and g are two functions as defined in Lemma A.3. From this point one can still make the link to the toy model, as done at the beginning of Section 4. By invariance by rotation, letting $t := \tilde{v}_1 - \langle \tilde{v}_1, v_1 \rangle v_1$, we know that $\|t\|$ and $\frac{t}{\|t\|}$ are independent, and that $\frac{t}{\|t\|}$ is uniformly distributed on the unit sphere in v_1^\perp . We have the following equality in distribution:

$$(r_1(v_1), r_1(\tilde{v}_1)) \stackrel{(d)}{=} (r_1(X), r_1(X + sZ)),$$

with w.h.p.

$$s \geq \mathbf{s}^1 = \frac{\|w\| \|X\|}{\left(\sum_{i=2}^N Z_i^2\right)^{1/2} \left(1 - \frac{\|w\| Z_1}{\left(\sum_{i=2}^N Z_i^2\right)^{1/2}}\right)} \asymp \sigma_1 N^{1/6},$$

where the X_i , Z_i , and w are defined in Section 4, for $\sigma = \sigma_1$. □

B.3 Proof of Lemma 5.1

Proof of Lemma 5.1. Recall that we work in the case (i) ($\sigma = o(N^{-7/6-\epsilon})$ for some $\epsilon > 0$), with $\langle v_1, v'_1 \rangle > 0$ and $\Pi = \text{Id}$. We want to show that w.h.p.

$$\langle A, \Pi_+ B \Pi_+^T \rangle > \langle A, \Pi_- B \Pi_-^T \rangle. \quad (\text{B.2})$$

Define

$$\mathcal{G} := \{i, \Pi_+(i) = \Pi(i) = i\}$$

and

$$\mathcal{A} := \left\{ \sigma N^{1/6} f(N)^{-1} \leq \mathbf{s} \leq \sigma N^{1/6} f(N) \right\},$$

with $f \in \mathcal{F}$ such that $\mathbb{P}(\mathcal{A}) \rightarrow 1$. For N large enough, on the event \mathcal{A} , we have $0 \leq \mathbf{s}N \leq N^{-\epsilon} f(N)$. Hence, repeating the proof of Proposition 2.2, we have

$$\begin{aligned} \phi_{x,z}(N, \mathbf{s}) &\geq \mathbb{P}(\mathcal{N}^+(x, x + sZ) = \mathcal{N}^-(x, x + sZ) = 0) \\ &\sim \exp(-\mathbf{s}NE(x)[z(2F(z) - 1) + 2E(z)]) = 1 - O(N^{-\epsilon} f(N)). \end{aligned}$$

Thus, with dominated convergence, for N large enough,

$$\mathbb{P}(\Pi_+(i) = \Pi(i) | \mathcal{A}) = \iint dx dz E(x) E(z) \mathbb{E}[\phi_{x,z}(N, \mathbf{s}) | \mathcal{A}] \geq 1 - O(N^{-\epsilon} f(N)). \quad (\text{B.3})$$

We use Markov's inequality with (B.3) to show that $\mathbb{P}(\#\mathcal{G} \leq N - N^{1-\epsilon/2} | \mathcal{A}) \leq O(N^{-\epsilon/2} f(N))$, and hence w.h.p.

$$\#\mathcal{G} \geq N - N^{1-\epsilon/2}. \quad (\text{B.4})$$

Splitting the sum

$$\langle A, \Pi_+ B \Pi_+^T \rangle = \sum_{i,j} A_{i,j} B_{\Pi_+(i), \Pi_+(j)} = \sum_{(i,j) \in \mathcal{G}^2} A_{i,j} B_{i,j} + \sum_{(i,j) \notin \mathcal{G}^2} A_{i,j} B_{\Pi_+(i), \Pi_+(j)},$$

one has, w.h.p.,

$$\begin{aligned} \langle A, \Pi_+ B \Pi_+^T \rangle &= \sum_{(i,j) \in \mathcal{G}^2} A_{i,j}^2 + \sum_{\substack{(i,j) \notin \mathcal{G}^2 \\ (\Pi_+(i), \Pi_+(j)) \neq (j,i)}} A_{i,j} A_{\Pi_+(i), \Pi_+(j)} \\ &\quad + \sum_{\substack{(i,j) \notin \mathcal{G}^2 \\ (\Pi_+(i), \Pi_+(j)) = (j,i)}} A_{i,j}^2 + \sigma \sum_{1 \leq i,j \leq N} A_{i,j} H_{\Pi_+(i), \Pi_+(j)} \\ &\geq C_1 \frac{(\#\mathcal{G})^2}{N} - C_2 \left(N^2 - (\#\mathcal{G})^2 \right) \frac{\log N}{N} - C_2 \sigma N^2 \frac{\log N}{N}. \end{aligned}$$

Here we applied the law of large numbers for the first sum, lower-bounded the third sum by zero, and used the classical inequality

$$\max_{i,j} \{A_{i,j}, H_{i,j}\} \leq C_2 \frac{\log N}{N}$$

(which holds w.h.p.) for the two others.

The inequality (B.4) and the condition (i) lead to, w.h.p.,

$$\langle A, \Pi_+ B \Pi_+^T \rangle \geq C_1 N - 2C_1 N^{1-\epsilon/2} - 2C_2 N^{1-\epsilon/2} \log N - C_2 N^{-1/6-\epsilon} \log N \geq C_3 N.$$

On the other hand, since by definition $\Pi_-(i) = \Pi_+(N + 1 - i)$, w.h.p.,

$$\begin{aligned} \langle A, \Pi_- B \Pi_-^T \rangle &= \sum_{(i,j) \in \mathcal{G}^2} A_{i,j} B_{N+1-i, N+1-j} + \sum_{(i,j) \notin \mathcal{G}^2} A_{i,j} B_{\Pi_-(i), \Pi_-(j)} \\ &\leq O(\log N) + \frac{(\#\mathcal{G})^2}{N} o(1) + C_2 \left(N^2 - (\#\mathcal{G})^2 \right) \frac{\log N}{N}. \end{aligned}$$

For the first sum, we used the law of large numbers: the variables $A_{i,j}$ and $B_{N+1-i, N+1-j}$ are independent in all cases but at most $N + 1$, and this part of the sum is bounded by $O(\log N)$. We used the same control on Gaussian variables as above.

This gives

$$\left(\langle A, \Pi_- B \Pi_-^T \rangle \right)_+ = o_{\mathbb{P}}(N),$$

where $(x)_+ := \max(0, x)$, which proves (B.2). □

B.4 Proof of Lemma 5.3

Proof of Lemma 5.3. Recall that we work in the case (ii) ($\sigma = \omega(N^{-7/6+\epsilon})$ for some $\epsilon > 0$), with $\langle v_1, v'_1 \rangle > 0$ and $\Pi = \text{Id}$. We want to show that the aligning permutation between v_1 and $-v'_1$ has a very bad overlap. Taking the pair $(X, -Y)$ where $(X, Y) \sim \mathcal{J}(N, s)$, one can adapt the proof of Proposition 2.2, with the new definitions

$$\begin{aligned} \widetilde{S}^+(x, y) &:= \mathbb{P}(X_1 > x, -Y_1 < -y) \quad \text{and} \\ \widetilde{S}^-(x, y) &:= \mathbb{P}(X_1 < x, -Y_1 > -y). \end{aligned}$$

The analysis is even easier since for all x, z , there exist two constants c, C such that

$$0 < c \leq \widetilde{S}^+(x, x + sz), \widetilde{S}^-(x, x + sz) \leq C < 1.$$

It is then easy to check that the proof of Proposition 2.2, case (ii), adapts well. \square

Funding Information

This work was partially supported by the French government under the management of the Agence Nationale de la Recherche, as part of the Investissements d'Avenir program, reference ANR19-P3IA-0001 (PRAIRIE 3IA Institute).

Competing Interests

There were no competing interests to declare which arose during the preparation or publication process for this article.

References

- [1] ALLEZ, R. AND BOUCHAUD, J.-P. (2014). Eigenvector dynamics under free addition. *Random Matrices Theory Appl.* **3**, 1450010.
- [2] ALLEZ, R., BUN, J. AND BOUCHAUD, J.-P. (2014). *The eigenvectors of Gaussian matrices with an external source*. Preprint. Available at <https://arxiv.org/abs/1412.7108>.
- [3] ANDERSON, G. W., GUIONNET, A. AND ZEITOUNI, O. (2009). *An Introduction to Random Matrices*. Cambridge University Press.
- [4] BOLLOBÁS, B. (2001). *Random Graphs*, 2nd edn. Cambridge University Press.
- [5] CHATTERJEE, S. (2014). *Superconcentration and Related Topics*. Springer, Cham.
- [6] CONTE, D., FOGGIA, P., VENTO, M. AND SANSONE, C. (2004). Thirty years of graph matching in pattern recognition. *Internat. J. Pattern Recogn. Artif. Intellig.* **18**, 265–298.
- [7] DING, J., MA, Z., WU, Y. AND XU, J. (2018). *Efficient random graph matching via degree profiles*. Preprint. Available at <https://arxiv.org/abs/1811.07821>.
- [8] ERDŐS, L., YAU, H.-T. AND YIN, J. (2010). *Rigidity of eigenvalues of generalized Wigner matrices*. Preprint. Available at <https://arxiv.org/abs/1007.4652>.
- [9] FAN, Z., MAO, C., WU, Y. AND XU, J. (2019). *Spectral graph matching and regularized quadratic relaxations I: the Gaussian model*. Preprint. Available at <https://arxiv.org/abs/1907.08880>.
- [10] FAN, Z., MAO, C., WU, Y. AND XU, J. (2019). *Spectral graph matching and regularized quadratic relaxations II: Erdős–Rényi graphs and universality*. Preprint. Available at <https://arxiv.org/abs/1907.08883>.
- [11] FEIZI, S. *et al.* (2017). *Spectral alignment of graphs*. Preprint. Available at <https://arxiv.org/abs/1602.04181>.
- [12] FORRESTER, P. (1993). The spectrum edge of random matrix ensembles. *Nuclear Phys. B* **402**, 709–728.
- [13] HAGHIGHI, A. D., NG, A. Y. AND MANNING, C. D. (2005). Robust textual inference via graph matching. In *Proc. Conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT '05)*, Association for Computational Linguistics, Stroudsburg, PA, pp. 387–394.
- [14] MAKARYCHEV, K., MANOKARAN, R. AND SVIRIDENKO, M. (2014). *Maximum quadratic assignment problem: reduction from maximum label cover and LP-based approximation algorithm*. Preprint. Available at <https://arxiv.org/abs/1403.7721>.
- [15] NARAYANAN, A. AND SHMATIKOV, V. (2008). Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy*, Institute of Electrical and Electronics Engineers, New York, pp. 111–125.
- [16] NARAYANAN, A. AND SHMATIKOV, V. (2009). De-anonymizing social networks. In *2009 IEEE Symposium on Security and Privacy*, Institute of Electrical and Electronics Engineers, New York, pp. 173–187.
- [17] O'ROURKE, S. (2010). Gaussian fluctuations of eigenvalues in Wigner random matrices. *J. Statist. Phys.* **138**, 1045–1066.
- [18] O'ROURKE, S., VU, V. AND WANG, K. (2016). *Eigenvectors of random matrices: a survey*. Preprint. Available at <https://arxiv.org/abs/1601.03678>.
- [19] PARDALOS, P., RENDL, F. AND WOLKOWICZ, H. (1994). The quadratic assignment problem: a survey and recent developments. In *Quadratic Assignment and Related Problems*, eds P. PARDALOS AND H. WOLKOWICZ, American Mathematical Society, Providence, RI, pp. 1–42.

- [20] SINGH, R., XU, J. AND BERGER, B. (2008). Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proc. Nat. Acad. Sci. USA* 105, 12763–12768.
- [21] TRACY, C. A. AND WIDOM, H. (1998). Correlation functions, cluster functions, and spacing distributions for random matrices. *J. Statist. Phys.* **92**, 809–835.