

# Uniform regular weighted graphs with large degree: Wigner's law, asymptotic freeness and graphons limit

Camille Male, Sandrine Péché

► **To cite this version:**

Camille Male, Sandrine Péché. Uniform regular weighted graphs with large degree: Wigner's law, asymptotic freeness and graphons limit. 2014. <hal-01458155>

**HAL Id: hal-01458155**

**<https://hal.inria.fr/hal-01458155>**

Submitted on 6 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Uniform regular weighted graphs with large degree: Wigner's law, asymptotic freeness and graphons limit

Camille Male, Sandrine Péché

ABSTRACT:

For each  $N \geq 1$ , let  $G_N$  be a simple random graph on the set of vertices  $[N] = \{1, 2, \dots, N\}$ , which is invariant by relabeling of the vertices. The asymptotic behavior as  $N$  goes to infinity of correlation functions:

$$\mathfrak{C}_N(T) = \mathbb{E} \left[ \prod_{(i,j) \in T} \left( \mathbb{1}_{\{\{i,j\} \in G_N\}} - \mathbb{P}(\{i,j\} \in G_N) \right) \right], \quad T \subset [N]^2 \text{ finite}$$

furnishes informations on the asymptotic spectral properties of the adjacency matrix  $A_N$  of  $G_N$ . Denote by  $d_N = N \times \mathbb{P}(\{i,j\} \in G_N)$  and assume  $d_N, N - d_N \xrightarrow{N \rightarrow \infty} \infty$ . If  $\mathfrak{C}_N(T) = \left(\frac{d_N}{N}\right)^{|T|} \times O(d_N^{-\frac{|T|}{2}})$  for any  $T$ , the standardized empirical eigenvalue distribution of  $A_N$  converges in expectation to the semicircular law and the matrix satisfies asymptotic freeness properties in the sense of free probability theory. We provide such estimates for uniform  $d_N$ -regular graphs  $G_{N,d_N}$ , under the additional assumption that  $|\frac{N}{2} - d_N - \eta\sqrt{d_N}| \xrightarrow{N \rightarrow \infty} \infty$  for some  $\eta > 0$ . Our method applies also for simple graphs whose edges are labelled by i.i.d. random variables.

## 1 Introduction and statement of results

The article is concerned with some statistical spectral properties of certain symmetric random matrices constructed from random graphs on  $N$  vertices. We are interested in the global asymptotic behavior when  $N$  is large of the (mean) empirical spectral distribution (e.s.d.) of such matrices  $M_N$  as well as in questions coming from free probability. Implicitly, whenever we consider a random matrix  $M_N$  or a random graph  $G_N$ , we actually mean a sequence  $(M_N)_{N \geq 1}$  (resp.  $(G_N)_{N \geq 1}$ ), where for each  $N \geq 1$ ,  $M_N$  is a random matrix of size  $N \times N$  (resp.  $G_N$  is a random graph) on the set of vertices  $[N] := \{1, \dots, N\}$ .

We can now define the model under study. Let  $G_N$  be an undirected random graph on  $[N]$ . To each edge  $\{i,j\} \in G_N$  is assigned the real weight  $\xi_{i,j} = \xi_{j,i}$ . We make the following hypotheses.

- Assumption 1.**
1.  $G_N$  is simple, i.e. has no loops and each edge is of multiplicity one.
  2. It is invariant in law by relabeling of its vertices, that is any weighted graph obtained by permuting the vertices of  $G_N$  has the same distribution.
  3. the  $\xi_{i,j}$ ,  $1 \leq i \leq j \leq N$ , are independent identically distributed (i.i.d.) random variables, not identically zero and which admit moments of any order (i.e.  $0 < \mathbb{E}[|\xi_{i,j}|^{2K}] < \infty$  for any  $K \geq 1$ ).

Most of our efforts in this article are dedicated to the uniform regular graph, that is when the random graph is chosen uniformly among the set of graphs in  $[N]$  with given degree  $d_N$ . We here assume that  $d_N$  goes to infinity with  $N$ . More generally our study extends to any random weighted graphs  $G_N$  satisfying the three above hypothesis.

The weighted adjacency matrix  $A_N = (A_N(i, j))_{i, j=1, \dots, N}$  of  $G_N$  is the symmetric matrix of size  $N$  defined by: for any  $i, j = 1, \dots, N$

$$A_N(i, j) = \begin{cases} \xi_{i, j} & \text{if } \{i, j\} \text{ is an edge of } G_N, \\ 0 & \text{otherwise.} \end{cases}$$

Denote by  $\lambda_1, \dots, \lambda_N$  the eigenvalues of a symmetric matrix  $A_N$ . Then its spectral measure  $\mu_{A_N}$  is the probability measure

$$\mu_{A_N} = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i}.$$

Note that  $A_N$  is invariant by conjugation by permutation matrices. The "standardized" version of the weighted adjacency matrix of  $G_N$  is the random matrix

$$M_N = \frac{A_N - m_1 \frac{d_N}{N-1} J_N}{\sqrt{d_N(m_2 - m_1^2 \frac{d_N}{N-1})}}, \quad (1.1)$$

where  $J_N$  is the matrix whose extra diagonal entries are ones and diagonal entries are zero,  $m_1 = \mathbb{E}[\xi_{i, j}]$ ,  $m_2 = \mathbb{E}[\xi_{i, j}^2]$ , and  $d_N$ , the mean degree of the graph, is defined by

$$d_N := \mathbb{E} \left[ \sum_{j \in [N] \setminus \{i\}} \mathbb{1}_{\{i, j\} \text{ is an edge of } G_N} \right], \quad \forall i \in [N]. \quad (1.2)$$

In the definition (1.2), there is no dependence on  $i$  since the graph is invariant by relabeling of its vertices. By standardized, we mean that  $M_N$  has centered entries and that for the polynomial  $P(x) = x^2$ , one has  $\mathbb{E} \mu_{M_N}(P) = \mathbb{E} \left[ \frac{1}{N} \text{Tr} M_N^2 \right] = 1$ .

Even in the case where the entries of  $M_N$  are independent, a variety of so-called "sparse" weighted random graphs (when  $d_N$  is bounded) exhibit spectral properties which differ from classical models of random matrix theory.

The Erdős-Rényi (E.R.) random graph  $G(N, \alpha_N)$  is the undirected weighted random graph for which any pair of vertices are connected with probability  $\alpha_N$ , independently of the other pairs of vertices. It fits the assumptions above with  $d_N = \alpha_N \times (N - 1)$ . When  $\alpha_N = \frac{d}{N}$  with  $d > 0$  fixed, the e.s.d. of the standardized adjacency matrix converges weakly and in moments to a symmetric measure. Since the measures  $\mu_{M_N}$  are random, it should be mentioned that this convergence is almost sure, in probability and in moments. Apart formulas for its moments, few is known about the limiting e.s.d. It always has unbounded support [8]. When the  $\xi_{i, j}$  are constant equal to one, it is known that it has a dense set of atoms if  $d \leq 1$  and it is conjectured that it has a density otherwise. The value of the mass at zero has been computed recently [3]. If one considers now an E.R. random graph with parameter  $\alpha_N$  converging to some constant  $\alpha \in ]0, 1[$ , then the standardized adjacency matrix is a special case of Wigner random matrices. Recall that a Wigner matrix is a random symmetric matrix  $X_N = (X(i, j))_{i, j=1, \dots, N}$  with independent, centered entries, such that the diagonal (resp. extra-diagonal) entries of  $\sqrt{N} X_N$  are distributed according to a measure  $\nu_0$  (resp.  $\nu$ ) that does not depend on  $N$ . Assume that  $\int x^{2+\varepsilon} d\nu$  and  $\int x^{2+\varepsilon} d\nu_0$  are finite. Then the e.s.d. of a Wigner matrix converges \*-weakly to the semicircular law  $\mu_{S, C}^{(a)}$  with radius  $2a$ , that is

$$\mu_{S, C}^{(a)}(dt) = \frac{1}{2\pi a^2} \sqrt{4a^2 - t^2} \mathbb{1}_{|t| \leq 2a}, \quad a^2 = \int t^2 \nu(dt).$$

The convergence for the standardized E.R. random weighted graph to the semicircular law actually holds as soon as  $\alpha_N = \frac{d_N}{N}$  where both  $d_N$  and  $N - d_N$  tend to infinity with  $N$ . The limitation that  $N - d_N$  goes to infinity is necessary. When  $\alpha_N = \frac{N-1-d}{N-1}$ ,  $d > 0$  fixed, the e.s.d. of  $G(N, \alpha_N)$  can be related to the sparse case where  $\alpha_N = \frac{d}{N-1}$ . To see that point, consider the complementary graph of  $G(N, \frac{N-1-d}{N-1})$ , for which two vertices are connected by an edge iff they

are not connected in  $G(N, \frac{N-1-d}{N-1})$ . Then the complementary graph is distributed as  $G(N, \frac{d}{N-1})$ . When the weights  $\xi_{i,j}$  are constant, the standardized adjacency matrix of  $G(N, \frac{N-1-d}{N-1})$  is distributed as the opposite of the standardized adjacency matrix of  $G(N, \frac{d}{N-1})$ .

The random regular graphs  $G_{N,d_N}$  have also been extensively studied in the literature ([12], [5],[4] [2]). Let  $N \geq 2$  and  $1 \leq d_N \leq N-1$  be given integers such that the product  $d_N \times N$  is even. Recall that  $G_{N,d_N}$  is a random graph chosen uniformly from the set of  $d_N$ -regular graphs. Let  $A_N$  be the adjacency matrix of the uniform random regular graph with degree  $d_N$  (unweighted, i.e.  $\xi_{i,j} = 1$  for any  $i, j$ ). When  $d$  is fixed independently of  $N$ , the e.s.d. of  $A_N$  converges weakly and in moments to the Mc Kay [12] law, i.e. the distribution with density

$$\mu_{\text{M.K.}(d)}(dx) = \frac{d\sqrt{4(d-1)-x^2}}{2\pi(d^2-x^2)} \mathbb{1}_{|x| \leq 2\sqrt{d-1}} dx.$$

If the graphs are weighted by random variables, it is still true that the e.s.d. converges but the distribution depends on the common distribution of the  $\xi_{i,j}$ .

Tran, Vu and Wang [14] consider the case where both  $d_N$  and  $N-d_N$  go to infinity with  $N$ . They show the weak convergence of the e.s.d. of the standardized matrix  $M_N$  of the unweighted  $d_N$  regular graph to the semicircular distribution  $\mu_{SC}^{(1)}$ . The convergence is almost sure and in probability. Moreover, a local version is proved: let  $I \subset (-2, 2)$  be an interval of length  $d_N^{-1/10} \ln^{1/5} d_N$ . Call  $N_I$  the number of eigenvalues falling in  $I$ . Then  $\mathbb{P}(|N_I - n \int_I \mu_{SC}^{(1)}(dx)| \geq \delta n \int_I \mu_{SC}^{(1)}(dx))$  decreases exponentially fast. The basic idea in [14] is that an E.R. graph with parameter  $\alpha_N$  is "quite often"  $d_N = (N-1) \times \alpha_N$ -regular. More precisely the authors compare the probability that it is indeed  $d_N$ -regular with the speed of concentration of linear statistics for the E.R. graph. This allows them to extend the asymptotic behavior of linear statistics of eigenvalues valid for E.R. graphs (based on random matrix theory results) to random regular graphs. Moreover Dumitriu and Pal [5] prove, under the hypothesis that  $d_N$  is smaller than any power of  $N$  ( $d_N = N^{o(1)}$ ), that the convergence in moments of  $\mu_{M_N}$  holds true. The local law is also improved therein. Their method is based on a local approximation of the  $d_N$  regular graph by the infinite  $d_N$ -regular tree.

A first application of our main results brings a slight contribution for that picture with new methods.

**Theorem 1.1** (Wigner's law). *Let  $M_{N,d_N}$  the matrix of a uniform  $d_N$ -regular weighted graph. Assume that  $d_N$ ,  $N-d_N$  and  $|\frac{N}{2} - d_N - \eta\sqrt{d_N}|$  tend to infinity for some constant  $\eta > 0$ . Then, the mean empirical spectral distribution of  $M_N$  converges in moments to the semicircular distribution, namely*

$$\forall P \text{ polynomial, } \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N P(\lambda_i) \right] \xrightarrow{N \rightarrow \infty} \int_{-2}^2 P(t) \frac{1}{2\pi} \sqrt{4-t^2} dt,$$

where  $\lambda_1, \dots, \lambda_N$  denote the eigenvalues of  $M_{N,d_N}$ .

We actually first obtain a criterion for the convergence to the semicircular law of a standardized matrix of a weighted graph  $G_N$  (see next section). Applying this criterion for the regular uniform graph is a large part of the problem. We proceed by combinatorial manipulations using intensively the so-called *switching-method* (see Section 5). The condition  $|\frac{N}{2} - d_N - \eta\sqrt{d_N}| \xrightarrow{N \rightarrow \infty} \infty$  is a limitation of our approach.

Our second application is an extension of this convergence in the context of free probability theory. Thanks to results from [10], [11], a minor modification of the mentioned criterion implies a phenomenon called "asymptotic freeness" (Definition 1.2 below). First, recall the notion of joint distribution of several random matrices  $\mathbf{A}_N = (A_N^{(j)})_{j \in J}$ , defined when the entries of each

matrix admit moments of any order. It is usually called the mean  $*$ -distribution (where the symbol  $*$  refers to the conjugate transpose) and defined as the linear map

$$\Phi_{\mathbf{A}_N} : P \mapsto \mathbb{E} \left[ \frac{1}{N} \text{Tr} P(\mathbf{A}_N) \right],$$

where the  $P$  are  $*$ -polynomials ( $P(\mathbf{A}_N)$  is a fixed finite complex linear combinations of words in the matrices  $\mathbf{A}_N^{(j)}$  and their transpose). Note that when the family consists only in one symmetric matrix, the  $*$ -distribution amounts just to restrict the e.s.d. to polynomials.

**Definition 1.2** (Asymptotic freeness). *Families  $\mathbf{A}_N^{(1)}, \dots, \mathbf{A}_N^{(n)}$  of random matrices are asymptotically free whenever,*

1. *the joint family  $\mathbf{A}_N = (\mathbf{A}_N^{(1)}, \dots, \mathbf{A}_N^{(n)})$  has a limiting  $*$ -distribution, that is  $\Phi_{\mathbf{A}_N}$  converges pointwise,*
2. *the limit of  $\Phi_{\mathbf{A}_N}$  is characterized in terms of the limiting  $*$ -distributions of each  $\mathbf{A}_N^{(j)}$ 's, thanks to the following formula. For  $n \geq 1$ , for any integers  $i_1 \neq i_2 \neq i_3 \neq \dots \neq i_n$  and any  $*$ -polynomials  $P_1, \dots, P_n$  such that  $\lim_{N \rightarrow \infty} \mathbb{E} \left[ \frac{1}{N} \text{Tr} P_j(\mathbf{A}_N^{(i_j)}) \right] = 0$  for any  $j = 1, \dots, n$ , one has*

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[ \frac{1}{N} \text{Tr} [P_1(\mathbf{A}_N^{(i_1)}) \dots P_n(\mathbf{A}_N^{(i_n)})] \right] = 0. \quad (1.3)$$

Freeness, which is defined as the relation (1.3), is considered as the analogue of classical independence in free probability. Voiculescu [15] and then Dikema [6] proved that independent Wigner matrices  $X_N^{(j)}, j \in J$ , are asymptotically free. Moreover, consider a family of deterministic matrices  $\mathbf{Y}_N$  converging in  $*$ -distribution. Assume the matrices of  $\mathbf{Y}_N$  are uniformly bounded in operator norm. Then  $X_N^{(j)}, j \in J$ , and  $\mathbf{Y}_N$  are asymptotically free (see [1]). The same holds for independent matrices  $M_N^{(j)}, j \in J$ , of weighted E.R. graphs with  $d_N^{(j)}$  and  $N - d_N^{(j)}$  going to infinity for any  $j \in J$  (by [10, Corollary 3.9] with no modification of the proof).

Our main result is the following.

**Theorem 1.3** (Asymptotic freeness). *Let  $M_N^{(j)}, j \in J$  be the adjacency matrices of independent uniform  $d_N^{(j)}$ -regular weighted graphs ( $d_N^{(j)}$  depends on  $j \in J$ ). Assume that there exists  $\eta > 0$  such that for any  $j \in J$ ,  $d_N^{(j)}, N - d_N^{(j)}$  and  $|\frac{N}{2} - d_N^{(j)} - \eta \sqrt{d_N^{(j)}}|$  tend to infinity. Moreover, let  $\mathbf{Y}_N$  be a family of deterministic matrices, uniformly bounded in operator norm as  $N$  goes to infinity and converging in  $*$ -distribution. Then the matrices  $M_N^{(j)}, j \in J$  and  $\mathbf{Y}_N$  are asymptotically free.*

To complete the comparison, let us mention that the sparse case does not behave exactly similarly to the E.R. case from the point of view of free probability. When  $\alpha_N^{(j)} \sim \frac{d^{(j)}}{N}$  for  $d^{(j)} > 0$  fixed for any  $j \in J$ , E.R. matrices  $M_N^{(j)}, j \in J$  are not asymptotically free [11]. A different notion of asymptotic freeness holds for  $M_N^{(j)}, j \in J$ , and  $\mathbf{Y}_N$ , called the asymptotic traffic-freeness [10]. We do not use so much this notion in this article. Just recall that the deterministic matrices  $\mathbf{Y}_N$  are required to satisfy stronger moment hypotheses. When  $d_N^{(j)} \xrightarrow{N \rightarrow \infty} d^{(j)} > 0$  for any  $j \in J$  and the  $\xi_{i,j}$  are constant equal to one, the adjacency matrices of uniform regular graphs  $M_N^{(j)}, j \in J$  are still asymptotically free [13, Section F]. Nevertheless, in general  $M_N^{(j)}, j \in J$  is not asymptotically free from deterministic matrices  $\mathbf{Y}_N$  converging in  $*$ -distribution. Under the assumption that  $\mathbf{Y}_N$  converges in distribution of traffics,  $M_N^{(j)}, j \in J$  and  $\mathbf{Y}_N$  are weakly asymptotically free in the sense of [10].

As a byproduct of our method, which does not concern the spectral properties of the adjacency matrices  $M_N$ , we obtain an estimate related to the geometry of uniform regular graphs with large degree. We follow the terminology introduced by Lovász [9]. A random (unweighted)

graph  $G_N$  is said to converge in distribution of "graphons" if and only if for any fixed finite graph  $T$

$$\lim_{N \rightarrow \infty} P(T \subset G_N) \text{ exists.}$$

Note that for an E.R. graph with parameter  $\alpha_N = \frac{d_N}{N}$ , one has  $P(T \subset G(N, \alpha_N)) = \left(\frac{d_N}{N}\right)^n$ . This implies the convergence in distribution of "graphons" for such random graphs.

Consider a sequence  $d_N$  for which there exists a constant  $\eta > 0$  such that  $d_N, N - d_N$  and  $|\frac{N}{2} - d_N - \eta\sqrt{d_N}|$  going to infinity as  $N \rightarrow \infty$ .

**Proposition 1.4** (Graphons limit estimate). *Let  $G_{N,d_N}$  be a uniform regular graph where  $d_N$  is a sequence as above. Then one has that*

$$\mathbb{P}(T \subset G_{N,d_N}) = \left(\frac{d_N}{N}\right)^n \left(1 + o(d_N)^{-\frac{1}{2}}\right). \quad (1.4)$$

In particular, it implies the convergence in the sense of graphons of  $G_{N,d_N}$  to the same limit as the one of the Erdős-Rényi random graph  $G(N, \frac{d_N}{N})$  when  $\left(\frac{d_N}{N}\right)$  admits a limit in  $]0, 1[\setminus \{\frac{1}{2}\}$ .

The roadmap of our proof is given in next section.

## 2 Main arguments: the correlation functions

Hereafter, we consider a random matrix  $M_N$  as in (1.1), constructed on a random graph  $G_N$  satisfying Assumption 1:  $G_N$  is a simple random graph invariant by relabeling of its vertices, weighted by i.i.d. random variables whose common distribution does not depend on  $N$ . The sequence  $d_N$  is the mean degree of the graph defined in (1.2).

Let  $T$  be a given simple graph with  $n$  edges  $e_1, \dots, e_n$ . Use integers to name the vertices of  $T$  and let  $N$  be larger than the maximum of these integers. Denote by  $G_N(e_j)$  the indicator function of the event " $e_j$  belong to  $G_N$ ". Note that  $\mathbb{E}[G_N(e_j)] = \frac{d_N}{N} =: \alpha_N$  for any  $j$  and  $\mathbb{E}\left[\prod_{j=1}^n G_N(e_j)\right] = \mathbb{P}(T \subset G_N)$ . Define the function

$$\mathfrak{C}_N(T) := \mathbb{E}\left[\prod_{i=1}^n (G_N(e_i) - \alpha_N)\right]. \quad (2.1)$$

Heuristically,  $\mathfrak{C}_N(T)$  measure some correlations between the edges of  $G_N$ .

**Proposition 2.1.** *Assume that  $d_N$  tends to infinity, and that either  $N - d_N$  tends also to infinity or that the random variables  $\xi_{i,j}$  are not constant. Moreover, assume that for any simple graph  $T$  with  $n$  edges, one has  $\mathfrak{C}_N(T) = \left(\frac{d_N}{N}\right)^n \times \varepsilon_N(T)$  where*

$$\varepsilon_N(T) \rightarrow 0 \text{ when } n = 2 \text{ and is } o\left(d_N^{1-\frac{n}{2}}\right) \text{ for } n \geq 3.$$

*Then, the mean empirical spectral distribution of  $M_N$  converges in moments to the semicircular distribution.*

Proposition 2.1 is proved in Section 3.

**Proposition 2.2.** *Let  $M_N^{(j)}, j \in J$  be independent matrices constructed as  $M_N$  in Proposition 2.1. Assume again that  $d_N$  tends to infinity, and that either  $N - d_N$  tends also to infinity or that the random variables  $\xi_{i,j}$  are not constant. Moreover, assume that for any simple graph  $T$  with  $n$  edges, one has  $\mathfrak{C}_N(T) = \left(\frac{d_N}{N}\right)^n \times \varepsilon_N(T)$  where*

$$\varepsilon_N(T) \rightarrow 0 \text{ if } n = 2 \text{ and is } O\left(d_N^{\frac{1}{2}-\frac{n}{2}}\right) \text{ for } n \geq 3.$$

*Let  $\mathbf{Y}_N$  be a family of deterministic matrices, uniformly bounded in operator norm as  $N$  goes to infinity and converging in \*-distribution. Then the  $M_N^{(j)}, j \in J$  and  $\mathbf{Y}_N$  are asymptotically free (Definition 1.2).*

The proof is given in Section 4. Propositions 2.1 and 2.2 obviously apply to the Erdős-Rényi graph for which  $\mathfrak{C}_N(T) = 0$  for any  $T$ . We also show these two theorems hold true for the uniform  $d_N$ -regular graph under slight assumptions below.

Let  $N \geq 2$  and  $1 \leq d_N \leq N - 1$  be given integers such that the product  $d_N \times N$  is even.

**Theorem 2.3.** *Assume that  $d_N, N - d_N$  and  $|\frac{N}{2} - d_N - \eta\sqrt{d_N}|$  go to infinity as  $N \rightarrow \infty$  for some constant  $\eta$ . Then, for the uniform  $d_N$ -regular graph  $G_{N,d_N}$  and for any finite graph  $T$  with  $n$  edges,*

$$\mathfrak{C}_N(T) = \left(\frac{d_N}{N}\right)^n \times O(d_N^{-\frac{n}{2}}).$$

In particular, Theorem 1.1 and Theorem 1.3 hold true.

Theorem 2.3 is proved in Section 5. As Theorem 2.3 readily implies Proposition 1.4, we here expose the proof.

*Proof of Proposition 1.4.* We recall that we assume Theorem 2.3 holds true.

Expanding the product in the definition of  $\mathfrak{C}_N(T)$  over the subgraphs  $\tilde{T}$  of  $T$ , we get

$$\mathfrak{C}_N(T) = \sum_{\tilde{T} \subset T} \left(-\frac{d_N}{N}\right)^{n-|\tilde{T}|} \mathbb{P}(\tilde{T} \subset G_N),$$

where  $|\tilde{T}|$  denotes the number of edges of  $\tilde{T}$ . Note that Theorem 2.3 states that

$$\mathfrak{C}_N(\tilde{T}) = \left(\frac{d_N}{N}\right)^{|\tilde{T}|} \times O(d_N^{-\frac{|\tilde{T}|}{2}}).$$

We reason by induction on the number of edges of  $T$ . If  $|T| = 1$ , the statement is obvious. Assume now that  $\mathbb{P}(\tilde{T} \subset G_N) = \left(\frac{d_N}{N}\right)^{|\tilde{T}|} \times (1 + \tilde{\varepsilon}_N(\tilde{T}))$  where  $\tilde{\varepsilon}_N(\tilde{T}) = O(d_N^{-1})$  for any  $\tilde{T}$  with at most  $n - 1$  edges. With this notation, we have

$$\mathbb{P}(T \subset G_N) = \left(\frac{d_N}{N}\right)^n \times \left(1 - \sum_{\substack{\tilde{T} \subset T \\ \tilde{T} \neq T}} (-1)^{n-|\tilde{T}|} \tilde{\varepsilon}_N(\tilde{T}) + O(d_N^{-\frac{n}{2}})\right).$$

This gives the desired estimate. □

### 3 Proof of Proposition 2.1

Our method uses the so-called injective trace formalism and extended moment method developed in [10]. To explain this method, we need to define more graph theory notations. Hereafter, we consider finite non oriented graphs  $T = (V, E)$ , independent of  $N$ . Such graphs  $T$  are to be seen as test functions that are evaluated at  $G_N$  or at the associated matrix  $M_N$ . Loops and multiple edges are allowed for  $T$ . We then denote by  $\overline{E}$  the collection of pairwise distinct edges of  $T$ . Then  $|\overline{E}|$  counts the number of edges without multiplicity. A tree is a graph which is connected and has no cycle nor multiple edge. A fat tree is a graph  $T$  such that  $\overline{T} := (V, \overline{E})$  is a tree. A double tree is a fat tree in which the multiplicity of each edge is 2. We call a simple edge of a graph an edge of multiplicity one which is not a loop. Recall the notation  $[N] = \{1, \dots, N\}$ .

We here make the entries of the matrices  $M_N$  explicit by writing  $M_N = (M_N(i, j))_{i, j=1, \dots, N}$ . For any finite graph  $T = (V, E)$  we set the function  $\text{Tr}^0$  (called the injective trace)

$$\text{Tr}_N^0[T(M_N)] = \sum_{\substack{\phi: V \rightarrow [N] \\ \text{injective}}} \prod_{\{v, w\} \in E} M_N(\phi(v), \phi(w)). \quad (3.1)$$

Let  $T_k$  be the graph consisting in a simple cycle with  $k$  vertices  $\{1, \dots, k\}$ , with edges  $\{1, 2\}, \dots, \{k-1, k\}, \{k, 1\}$ . Given a partition  $\pi$  of  $\{1, \dots, k\}$ , let  $T_k^\pi$  be the graph (with possibly multiple edges and loops) obtained by identifying vertices in a same block. Formally, its set of vertices is  $\pi$  and its multi-set of edges  $E$  are given as follows: there is one edge between two blocks  $V_i$  and  $V_j$  of  $\pi$  for each  $n$  in  $\{1, \dots, k\}$  such that  $n \in V_i$  and  $n+1 \in V_j$  (with notation modulo  $k$ ). Then for any finite connected graph, for any  $k \geq 1$  we have

$$\mathrm{Tr} M_N^k = \sum_{\pi \in \mathcal{P}(k)} \mathrm{Tr}^0 [T_k^\pi(M_N)], \quad (3.2)$$

where  $\mathcal{P}(k)$  is the set of partitions of  $[k]$ . The formula is obtained by classifying in the sum  $\mathrm{Tr} M_N^k = \sum_{i_1, \dots, i_k} M_N(i_1, i_2) \dots M_N(i_k, i_1)$  which indices are equal or not, see [10, Section 3.1].

We say that a graph of the form  $T = T_k^\pi$  for certain  $k$  and  $\pi$  is cyclic. A cyclic graph is a finite, connected graph such that there exists a cycle visiting each edge once, taking account into the multiplicity of the edges. By (3.2), the convergence of the expectation of  $\frac{1}{N} \mathrm{Tr}^0$  for any cyclic graph implies the convergence in moments of the e.s.d. of  $M_N$ . Our proof of Wigner's law (Proposition 2.1) is based on the following lemma.

**Lemma 3.1.** *Denote  $\tau_N^0 = \mathbb{E}[\frac{1}{N} \mathrm{Tr}^0]$ . For any cyclic graph  $T$*

$$\tau_N^0 [T(M_N)] \xrightarrow{N \rightarrow \infty} \begin{cases} 1 & \text{if } T \text{ is a double tree} \\ 0 & \text{otherwise.} \end{cases} \quad (3.3)$$

*Proof of Proposition 2.1.* We here assume that Lemma 3.1 holds true. One then gets from (3.2) and (3.3) that

$$\mathbb{E} \mathrm{Tr} M_N^k = \sum_{\pi \in \mathcal{P}(k): T_k^\pi \text{ is a double tree}} 1.$$

We may recall that the  $k$ -th moment of the semicircular law, given by the  $k$ -th Catalan number, is the number of oriented rooted trees with  $\frac{k}{2}$  edges (see [7]). As all oriented rooted trees can be obtained from a partition  $\pi$  in a unique way, this finishes the proof.  $\square$

We split the proof of Lemma 3.1 into the cases where  $T$  is a double tree or not, and state below Lemmas 3.3 and 3.4 accordingly. Let  $T = (V, E)$  be a cyclic graph. Denote its edges by  $e_1 = (v_1, w_1), \dots, e_n = (v_n, w_n)$  (with possible repetition due to multiplicity) and set

$$\delta_N^0 [T(M_N)] = \mathbb{E} \left[ \prod_{i=1}^n M_N(\phi(v_i), \phi(w_i)) \right],$$

where  $\phi$  is any injection  $V \rightarrow [N]$ . Since the matrix is invariant by conjugation by permutation matrices, this quantity does not depend on  $\phi$ . Then, by the definition of  $\mathrm{Tr}^0$ , namely Formula (3.1), we have

$$\tau_N^0 [T(M_N)] = \frac{1}{N} \frac{N!}{(N - |V|)!} \delta_N^0 [T(M_N)].$$

Denote  $\alpha_N = (\frac{d_N}{N})$  and recall that  $M_N = \frac{A_N - m_1 \alpha_N J_N}{\sqrt{d_N(m_2 - m_1^2 \alpha_N)}}$  with the notations in (1.1). Then, using the multi-linearity of  $\delta_N^0$  (with respect to the edges of the graph), for any finite graph  $T$  we get

$$\tau_N^0 [T(M_N)] = N^{|V|-1} d_N^{-\frac{|E|}{2}} \times \delta_N^0 \left[ T \left( \frac{A_N - m_1 \alpha_N J_N}{\sqrt{m_2 - m_1^2 \alpha_N}} \right) \right] \times (1 + o(N^{-1})). \quad (3.4)$$

For a given graph  $T$ , we denote by  $s$  the number of simple edges of  $T$ . We can then write that

$$N^{|V|-1} d_N^{-\frac{|E|}{2}} = \underbrace{N^{|V|-1-|\bar{E}|} d_N^{|\bar{E}| - \frac{|E|}{2} - \frac{s}{2}}}_{K_N} \times d_N^{\frac{s}{2}} \alpha_N^{-|\bar{E}|}. \quad (3.5)$$

We now appeal to the following classical lemma of graph theory (see [7, Lemma 1.1]).



**Lemma 3.2.** *For a finite graph with  $\mathcal{V}$  vertices,  $\mathcal{E}$  edges and  $\mathcal{C}$  connected components, then  $\mathcal{E} + \mathcal{C} - \mathcal{V}$  is the number of cycles of the graph. This is the maximal number of edges we can suppress without disconnecting a component of the graph. In particular,*

$$\mathcal{V} - \mathcal{E} - \mathcal{C} \leq 0$$

with equality if and only if the graph is a forest, i.e. its components are trees.

We first apply Lemma 3.2 to the graph  $\bar{T}$  obtained by forgetting the multiplicity of the edges of  $T$  so that

$$\mathcal{V} = |V|, \mathcal{E} = |\bar{E}| \text{ and } \mathcal{C} = 1.$$

Hence,  $K_N = 1$  if and only if  $T$  is a fat tree whose edges are of multiplicity one or two. When  $T$  is a double tree,  $K_N = 1$  and  $s = 0$ . Hence to prove the convergence  $\tau_N^0[T(M_N)] \xrightarrow{N \rightarrow \infty} 1$  for double trees  $T$ , it then suffices to prove the

**Lemma 3.3.** *For any double tree  $T$*

$$\delta_N^0 \left[ T \left( \frac{A_N - m_1 \alpha_N J_N}{\sqrt{m_2 - m_1^2 \alpha_N}} \right) \right] = \left( \frac{d_N}{N} \right)^{|\bar{E}|} \times (1 + o(1)). \quad (3.6)$$

*Proof of Lemma 3.3.* Let  $T = (V, E)$  be a double tree. Denote  $e_1, \dots, e_n$  the edges of  $T$  without multiplicity (so that each edge is repeated twice in  $T$ ). We want to prove that

$$\alpha_N^{-|\bar{E}|} \times \delta_N^0 \left[ T \left( \frac{A_N - m_1 \alpha_N J_N}{\sqrt{m_2 - m_1^2 \alpha_N}} \right) \right] = \alpha_N^{-n} (m_2 - m_1^2 \alpha_N)^{-n} \times \mathbb{E} \left[ \prod_{i=1}^n (A_N(e_i) - m_1 \alpha_N)^2 \right]$$

converges to one. By the proof of Proposition 1.4 one has  $\mathbb{E} \left[ \prod_{i=1}^n G_N(e_i) \right] = \mathbb{P}(T \subset G_N) = \alpha_N^n (1 + o(1)) = \prod_{i=1}^n \mathbb{E} [G_N(e_i)] (1 + o(1))$  (we only use  $\mathfrak{C}_N(T) = \alpha^n (1 + o(1))$  for any  $T$ ). Denote by  $\mathbb{E}[\cdot | \xi]$  conditional expectation with respect to the algebra spanned by the i.i.d. random variables  $\xi_{i,j}, 1 \leq i < j \leq N$ . Denote  $\xi(e) = \xi_{i,j}$  for  $e = \{i, j\}$ . One has

$$\begin{aligned} & \mathbb{E} \left[ \prod_{i=1}^n (A_N(e_i) - m_1 \alpha_N)^2 \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ \prod_{i=1}^n (\xi_{e_i} G_N(e_i) - m_1 \alpha_N)^2 \mid \xi \right] \right] \\ &= \mathbb{E} \left[ \prod_{i=1}^n \mathbb{E} \left[ \xi_{e_i}^2 G_N(e_i) - 2m_1 \alpha_N \xi_{e_i} G_N(e_i) + m_1^2 \alpha_N^2 \mid \xi \right] \times (1 + o(1)) \right] \\ &= \prod_{i=1}^n (m_2 \alpha_N - m_1^2 \alpha_N^2) \times (1 + o(1)) \\ &= \alpha_N^n (m_2 - m_1^2 \alpha_N)^n \times (1 + o(1)) \end{aligned}$$

as desired.  $\square$

Let us now consider a cyclic graph  $T$  which is not a double tree. Since  $T$  is cyclic, it is not possible that  $T$  has an edge of order one or three and is simultaneously a fat tree. So either  $T$  has an edge of multiplicity at least four, or  $\bar{T}$  admits a cycle. Hence for a cyclic graph which is not a double tree, in all cases one has that

$$K_N = O(d_N^{-1})$$

(since  $d_N \leq N - 1$ ). By Formula (3.5) and Lemma 3.2, the fact that  $\tau_N^0[T(M_N)] \xrightarrow{N \rightarrow \infty} 0$  when  $T$  is not a double tree follows from the

**Lemma 3.4.** *For any cyclic graph  $T$  which is not a double tree, with  $n$  edges and  $s$  simple edges,*

$$\delta_N^0 \left[ T \left( \frac{A_N - m_1 \alpha_N J_N}{\sqrt{m_2 - m_1^2 \alpha_N}} \right) \right] = \alpha_N^{|\bar{E}|} \times o(d_N^{1-\frac{s}{2}}). \quad (3.7)$$

*Proof of Lemma 3.4.* By multi linearity,

$$\delta_N^0 \left[ T \left( \frac{A_N - m_1 \alpha_N J_N}{\sqrt{m_2 - m_1^2 \alpha_N}} \right) \right] = (m_2 - m_1^2 \alpha_N)^{-\frac{s}{2}} \delta_N^0 [T(A_N - m_1 \alpha_N J_N)].$$

We claim that, without loss of generality, one can assume that  $m_2 - m_1^2 \alpha_N$  is bounded away from zero. Indeed, remark first that by the Cauchy Schwarz inequality, one has  $m_1^2 \leq m_2$  with equality if and only if the random variables  $\xi_{i,j}$  are constant. Hence, if the random variables are not constant, since  $\alpha_N \leq 1$ , then  $m_2 - m_1^2 \alpha_N$  is bounded away from zero. If the random variables are constant, one can assume this constant is one (this does not change the distribution of  $M_N$ ). Moreover, recall we considered in the introduction the complementary graph (for which edges belong to the graph if and only if edges do not belong to the initial graph). For unweighted graphs,  $-M_N$  is distributed as the matrix associated to the complementary graph and we can assume  $d_N \leq \frac{N}{2}$ . Finally  $m_2 - m_1^2 \alpha_N = 1 - \alpha_N$  is always bounded away from zero.

Hence it is sufficient to prove

$$\delta_N^0 [T(A_N - m_1 \alpha_N J_N)] = \alpha_N^{|\bar{E}|} \times o(d_N^{1-\frac{s}{2}}), \quad (3.8)$$

where  $s$  is the number of simple edges of  $T$ . Choose an enumeration of the edges of  $T : E = \{e_1, \dots, e_n\}$ , with possible repetitions. Then one has that

$$\delta_N^0 [T(A_N - m_1 \alpha_N J_N)] = \mathbb{E} \left[ \prod_{i=1}^n (A_N(e_i) - m_1 \alpha_N) \right], \quad (3.9)$$

where  $A_N(e_i)$  stands for the random variables  $A_N(v_i, w_i)$  whenever  $e_i = \{v_i, w_i\}$ .

If  $T$  is a simple graph, then

$$\mathfrak{C}_N(T) := \mathbb{E} \left[ \prod_{i=1}^n (G_N(e_i) - \alpha_N) \right] = \alpha_N^n \times \varepsilon_N(T), \quad (3.10)$$

where  $G_N(e_i) = \mathbb{1}_{e_i \text{ belongs to } G_N}$  and  $\varepsilon_N(T) = o(d_N^{1-\frac{n}{2}})$ . Equation (3.10) gives the good estimates (3.8) if  $T$  is a simple graph. Indeed, in that case, since the  $\xi_{i,j}$  are i.i.d., we obtain  $\delta_N^0 [T(A_N - m_1 \alpha_N J_N)] = m_1 \times \mathfrak{C}_N(T) = o(d_N^{1-\frac{n}{2}})$ . We then need to extend this fact for non simple graphs. We shall use the following lemma.

**Lemma 3.5.** *Let  $T$  be a simple graph with edges  $e_1, \dots, e_n$ . Then, for  $k = 0, \dots, n$ ,*

$$\mathbb{E} \left[ \prod_{i=1}^k G_N(e_i) \prod_{i=k+1}^n (G_N(e_i) - \alpha_N) \right] = \alpha_N^n (\delta_{n,k} + \varepsilon_N^{(k)}(T)), \quad (3.11)$$

where  $\varepsilon_N^{(k)}(T)$  tends to zero and  $\varepsilon_N^{(k)}(T) = o(d_N^{1-\frac{n-k}{2}})$ . Recall that the Kronecker symbol  $\delta_{n,k}$  is one if  $n = k$  and zero otherwise.

*Proof.* Note that  $\varepsilon_N^{(0)}(T) = \varepsilon_N(T)$ , where  $\varepsilon_N(T)$  is defined in (3.10), and  $\mathbb{P}(T \subset G_N) = \alpha_N^n \times (1 + \varepsilon_N^{(n)}(T))$ . Writing  $G_N(e_k) = (G_N(e_k) - \alpha_N) + \alpha_N$  we obtain  $\varepsilon_N^{(k)}(T) = \varepsilon_N^{(k-1)}(T) + \varepsilon_N^{(k-1)}(T \setminus e_k)$ . By a simple recursion,  $\varepsilon_N^{(k)}(T)$  is a finite linear combination of  $\varepsilon_N(\tilde{T})$  for subgraphs  $\tilde{T}$  of  $T$  with at least  $n - k$  edges. Hence the result follows.  $\square$

We can now prove (3.8), using Formula (3.9) and Lemma 3.5. Let  $T$  be a simple graph with edges  $e_1, \dots, e_n$ . Consider integers  $p_j, j = 1, 2, \dots, n$  such that  $p_j \geq 2$  for  $j = 1, \dots, k$  and  $p_j = 1$  for  $j = k + 1, \dots, n$ . Let prove the estimate for the (non simple) graph obtained from  $T$  by setting  $p_j$  for the multiplicity of  $e_j$ . One has

$$\begin{aligned}
& \mathbb{E} \left[ \prod_{i=1}^n (A_N(e_i) - m_1 \alpha_N)^{p_j} \right] \tag{3.12} \\
&= \sum_{\substack{\ell_1 \in [p_1] \\ \dots, \ell_k \in [p_k]}} \binom{p_1}{\ell_1} \dots \binom{p_k}{\ell_k} (-m_1 \alpha_N)^{\ell_1 + \dots + \ell_k} \mathbb{E} \left[ \prod_{i=0}^k A_N(e_i)^{p_i - \ell_i} \times \prod_{i=k+1}^n (A_N(e_i) - m_1 \alpha_N) \right] \\
&= \sum_{\substack{\ell_1 \in [p_1] \\ \dots, \ell_k \in [p_k]}} \binom{p_1}{\ell_1} \dots \binom{p_k}{\ell_k} (-m_1 \alpha_N)^{\ell_1 + \dots + \ell_k} \prod_{i=1}^k \mathbb{E} [\xi^{p_i - \ell_i} m_1^{n-k}] \\
&\quad \times \mathbb{E} \left[ \prod_{i=0}^k G_N(e_i)^{p_i - \ell_i} \times \prod_{i=k+1}^n (G_N(e_i) - \alpha_N) \right],
\end{aligned}$$

where  $\xi$  is distributed as the  $\xi_{i,j}$ . Note that  $G_N(e_i)^{m-\ell} = G_N(e_i)$  for any  $\ell \neq m$ . Let fix integers  $\ell_j \in [p_j]$  and denote by  $C$  is the number of  $\ell_i$ 's for which  $p_i - \ell_i$  is nonzero. Then, by Lemma 3.5, the associated expectation on the left hand side is of the form  $\alpha_N^{C+(n-k)} \varepsilon_N^{(C)}(\tilde{T})$  for a graph  $\tilde{T}$  with  $C + (n - k)$  edges. For each  $\ell_i = p_i$  we take an  $\alpha_N$  from the  $\ell_1 + \dots + \ell_k$ , and, since  $\alpha_N \leq 1$ , we get

$$\mathbb{E} \left[ \prod_{i=1}^n (A_N(e_i) - m_1 \alpha_N)^{p_j} \right] = \alpha_N^n \times o(d_N^{1 - \frac{n-k}{2}}).$$

This gives the estimate (3.7) and proves that  $\tau_N^0[T(M_N)] \xrightarrow{N \rightarrow \infty} 0$  when  $T$  is not a double tree.  $\square$

Hence we have proved Lemmas 3.3 and 3.4, which gives Lemma 3.1 and finishes the proof of Proposition 2.1.

## 4 Proof of Proposition 2.2

The proof of the Asymptotic freeness needs only slight modifications of the previous part, thanks to results from [10] and [11]. We use the notations of Lemma 3.1. To establish Proposition 2.2, we use the following Lemma.

**Lemma 4.1.** *Let  $\mathbf{M}_N = (M_N^{(j)})_{j=1, \dots, n}$  be a family of independent random matrices. Assume that each matrix is invariant in law by conjugation by permutation matrices, and that for any  $j = 1, \dots, n$*

1. *for any cyclic graph  $T$ ,  $\tau_N^0[T(M_N^{(j)})] \xrightarrow{N \rightarrow \infty} \mathbb{1}_{(T \text{ is a double tree})}$ ,*
2. *for any cyclic graphs  $T_1, \dots, T_k$ ,*

$$\mathbb{E} \left[ \prod_{i=1}^k \frac{1}{N} \text{Tr}^0 [T_i(M_N^{(j)})] \right] - \prod_{i=1}^k \tau_N^0 [T_i(M_N^{(j)})] \xrightarrow{N \rightarrow \infty} 0.$$

3. *for any (possibly non-cyclic) graphs  $T_1, \dots, T_k$ ,*

$$\mathbb{E} \left[ \prod_{i=1}^k \frac{1}{N} \text{Tr}^0 [T_i(M_N^{(j)})] \right] = O(1).$$

Let  $\mathbf{Y}_N$  be a family of deterministic matrices converging in  $*$ -distribution, whose matrices are uniformly bounded in operator norm. Then  $M_N^{(1)}, \dots, M_N^{(n)}, \mathbf{Y}_N$  are asymptotically free.

*Proof.* The proof is identical to the proof of [11, Corollary 3.9]. We recall briefly the idea. By [11, Theorem 2.3], the convergence of  $\tau_N^0[T(M_N^{(j)})]$  for cyclic graphs  $T$  and the two above assumptions implies that  $M_N^{(1)}, \dots, M_N^{(n)}, \mathbf{Y}_N$  satisfies a weak version of asymptotic freeness, under slighter stronger conditions on  $\mathbf{Y}_N$ . In [10], it is proved that when the limit of  $\tau_N^0[T(M_N)]$  is the indicator of double trees as in (3.3), then this weak asymptotic freeness is actually the Voiculescu's asymptotic freeness. A tightness argument implies that  $M_N^{(1)}, \dots, M_N^{(n)}, \mathbf{Y}_N$  are asymptotically free in the sense of Voiculescu, with the above assumptions on  $\mathbf{Y}_N$ .  $\square$

We now show that the assumptions of Lemma 4.1 are satisfied. First Lemma 3.1 implies that the first assumption holds true. Let us now show how to modify the proof of the previous Section to prove the second assumption of Lemma 4.1. Denote the cyclic graphs by  $T_i = (V_i, E_i)$  for any  $i = 1, \dots, k$ . Call  $\mathcal{P}(V_i)_{i=1, \dots, k}$  the set of partitions of  $V_1 \sqcup \dots \sqcup V_k$  whose blocks contain at most one element of each  $V_i$ . It is a matter of fact that

$$\prod_{i=1}^k \text{Tr}^0[T_i(M_N)] = \sum_{\sigma \in \mathcal{P}(V_i)_{i=1, \dots, k}} \text{Tr}^0[T_\sigma(M_N)], \quad (4.1)$$

where  $T_\sigma$  is the graph obtained from the disjoint union of the  $T_i$ 's by identifying vertices in a same block of  $\sigma$ . We get this formula by looking, for the injective maps in the injective traces, how their images intersect. Note that now  $T_\sigma$  is possibly disconnected, but each of its connected component are cyclic.

Now, with no modification of the proof of (3.4) in the previous part, we have: for any graph  $T$  (which is considered as a  $T_\sigma$ ) with  $c$  connected components and  $s$  simple edges,

$$\mathbb{E} \left[ \frac{1}{N^c} \text{Tr}^0[T(M_N)] \right] = \tilde{K}_N \times d_N^{\frac{s}{2}} \alpha_N^{-|\bar{E}|} \delta_N^0 \left[ T \left( \frac{A_N - m_1 \alpha_N J_N}{\sqrt{1 - \alpha_N}} \right) \right],$$

where  $\tilde{K}_N = N^{|V| - c - |\bar{E}|} d_N^{|\bar{E}| - \frac{|E|}{2} - \frac{s}{2}}$ . By Lemma 3.2, the quantity  $\tilde{K}_N$  is one if  $T$  is a fat forest of double trees. Otherwise, since each component is cyclic,  $\tilde{K}_N = O(d_N^{-1})$  with the same reasoning as in the previous part. The rest of the reasoning is as in the previous Section with no modification. Thus, in (4.1) the only term which is non negligible corresponds to  $\sigma$  being the trivial partition, and then

$$\mathbb{E} \left[ \prod_{i=1}^k \frac{1}{N} \text{Tr}^0[T_i(M_N^{(j)})] \right] \xrightarrow{N \rightarrow \infty} \mathbb{1}_{(\text{all the } T_i \text{'s are double trees)}}.$$

Let us now indicate where we modify this proof in order to get the third assumption of Lemma 4.1. Remove in the above computations the fact that the  $T_i$  are cyclic graphs. Hence the only detail that changes is now  $\tilde{K}_N$  is  $O(\sqrt{d_N}^{-1})$  if  $T_\sigma$  is not a fat forest of double trees. This yields the result.

## 5 Proof of Proposition 2.3

*Step 1: Preliminaries.* Let  $G_N (= G_{N, d_N})$  be a uniform  $d_N$ -regular graphs on  $[N] := \{1, \dots, N\}$ . Denote  $\alpha_N = \frac{d_N}{N-1}$  and fix for the rest of the Section a simple graph  $T = (V, E)$  with pairwise distinct edges  $e_1, e_2, \dots, e_n$ . Let give an estimate for  $\varepsilon_N(T)$  defined below

$$\mathfrak{C}_N(T) (= \mathfrak{C}_{N, d_N}(T)) = \mathbb{E} \left[ \prod_{i=1}^n (G_N(e_i) - \alpha_N) \right] =: \alpha_N^n \times \varepsilon_N(T). \quad (5.1)$$

Note that by considering the complementary graph  $\mathfrak{C}_{N, d_N}(T) = (-1)^n \mathfrak{C}_{N, N-1-d_N}(T)$ , so we can assume  $\alpha_N \leq \frac{1}{2}$  without loss of generality.

We use the combinatorial method, to expand the expectation with respect to the uniform measure on the set of all  $d_N$ -regular graphs. We give a first expression (5.3) of the correlation function  $\mathfrak{C}_N(T)$  and then explain the general idea of the proof. Denote by  $\mathcal{G}_N$  the set of all simple  $d_N$ -regular graphs on  $[N]$ . We partition the graphs  $G \in \mathcal{G}_N$  by considering the set of edges of  $T$  that belong to  $G$ . Given a subset  $J \subset [n] := \{1, \dots, n\}$ , we denote by  $\mathcal{G}_N(J)$  the set of graphs  $G \in \mathcal{G}_N$  such that  $e_j \subset G$  iff  $j \in J$ . For any  $j = 1, \dots, n$ , if  $e_j \subset G$  then  $G(e_j) - \alpha_N = \frac{N-1-d_N}{N-1}$  and otherwise  $G(e_j) - \alpha_N = \frac{-d_N}{N-1}$ . Hence we can write

$$\mathfrak{C}_N(T) = \sum_{J \subset [n]} (-1)^{n-|J|} \frac{\#\mathcal{G}_N(J) \times (N-1-d_N)^{|J|} d_N^{n-|J|}}{\#\mathcal{G}_N(J) \times (N-1)^n}. \quad (5.2)$$

Remark that the quantity  $N-1-d_N$  and  $d_N$  respectively count the number of non-neighbors and neighbors  $w'$  of a vertex  $v$  in  $G_N$ . For each  $j = 1, \dots, n$ , disregarding coincidences, we write

$$e_j = \{v_j, w_j\}$$

where  $v_j, w_j$  are (fixed) vertices. Using this notation, we need to be able to distinguish the vertices  $v_j$  and  $w_j$ , e.g. by orienting the edges  $e_j$ . We now use the notations

$$v \sim_J w \Leftrightarrow (v \sim w \text{ if } j \in J \text{ and } v \not\sim w \text{ if } j \notin J),$$

$$v \not\sim_J w \Leftrightarrow (v \not\sim w \text{ if } j \in J \text{ and } v \sim w \text{ if } j \notin J).$$

We use the notation  $\mathcal{G}_N(J) \cup (w'_j | w'_j \not\sim_J v_j)_{j \in [n]}$  for the set of graphs with labeled vertices  $(G, w'_j, j \in [n])$ , where

1.  $G \in \mathcal{G}_N(J)$ , i.e. it is a  $d_N$  regular graph and  $v_j \sim_J w_j$  for any  $j = 1, \dots, n$  (the vertices  $v_j$  and  $w_j$  are fixed),
2.  $w'_j \in [N]$  is a vertex, which varies in the enumeration, such that  $w'_j \not\sim_J v_j$ .

Starting with the graph  $T$ , we can write  $\mathfrak{C}_N(T)$  as the signed sum

$$\mathfrak{C}_N(T) = \sum_{J \subset [n]} (-1)^{n-|J|} \frac{\mathcal{G}_N(J) \cup (w'_j | w'_j \sim_J v_j)_{j \in [n]}}{\#\mathcal{G}_N \times (N-1)^n}. \quad (5.3)$$

Indeed for any  $j \in J$  (resp.  $j \notin J$ ) we choose a vertex which is not (resp. is) a neighbor of  $v_j$ . The basic idea is as follows: fix some subset  $J$ . Consider now an index  $j \in J$  so that  $(v_j, w_j) \in G$  but  $(v_j, w'_j) \notin G$ . Roughly speaking, switching these two edges and completing into a  $d_N$ -regular graph will result into almost the same number of graphs but with  $J \rightarrow J \setminus \{j\}$ . Therefore some cancellations will arise. However we need to be very precise if we want to stack up the different error terms for each  $j \in J$ . The plan is to pursue further this construction for each index  $i \in J$  and build hexagons as in Figure 1 below. There,  $T$  is the square with four edges  $e_1 e_2 e_3 e_4$  and  $J = \{1, 2\}$ . We anticipate in the figure below some notations defined later. We use a symmetry of the uniform regular graphs described here after to catch the simplifications in the sum (5.3).

*The switching method:* Let  $v^{(1)}, w^{(1)}, \dots, v^{(K)}, w^{(K)}$  be  $2K$  distinct vertices. It is convenient to define the notation  $w^{(K+1)} := w^{(1)}$  and  $v^{(K+1)} := v^{(1)}$ . We also define the  $2K$ -gone  $P$  obtained by drawing the edges  $\{v^{(k)}, w^{(k)}\}$  and  $\{w^{(k)}, v^{(k+1)}\}$ , for all  $k = 1, \dots, K$ . Consider two disjoint sets of edges  $\mathcal{E}$  and  $\mathcal{F}$  that do not contains edges of the  $2K$ -gone  $P$ . Then, the number of  $d_N$ -regular graphs  $G$  such that

$$v^{(k)} \sim w^{(k)}, w^{(k)} \not\sim v^{(k+1)} \quad \forall k = 1, \dots, K, \text{ and } e \subset G \quad \forall e \in \mathcal{E}, f \not\subset G \quad \forall f \in \mathcal{F} \quad (5.4)$$

is equal to the number of the number of  $d_N$ -regular graphs  $G$  such that

$$v^{(k)} \not\sim w^{(k)}, w^{(k)} \sim v^{(k+1)} \quad \forall k = 1, \dots, K, \text{ and } e \subset G \quad \forall e \in \mathcal{E}, f \not\subset G \quad \forall f \in \mathcal{F} \quad (5.5)$$

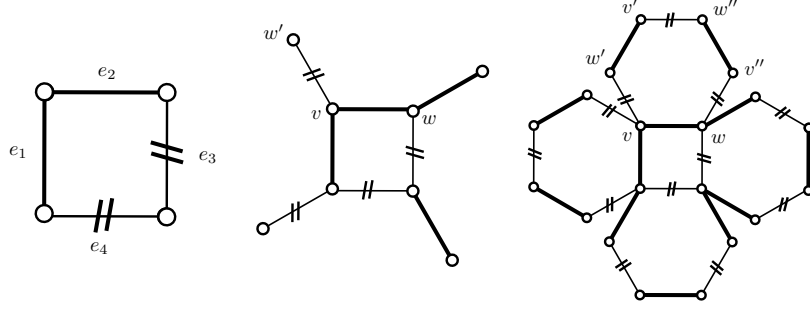


Figure 1: Construction of the hexagons (the plan).

Indeed, denote by  $\mathcal{G}^{(1)}$  and  $\mathcal{G}^{(2)}$  the sets of regular graphs defined by (5.4) and (5.5) respectively. Consider a graph from  $\mathcal{G}^{(1)}$ . By "switching" its edges, we mean removing the edges  $\{v^{(k)}, w^{(k)}\}$ 's and adding the  $\{w^{(k)}, v^{(k+1)}\}$ 's. This switching maps bijectively the graph to an element from  $\mathcal{G}^{(2)}$ . The important fact is that we do not modify the degree of any vertices when applying this switching. Moreover, as one can easily verify, one can add additional assumptions on the sets  $\mathcal{G}^{(1)}$  and  $\mathcal{G}^{(2)}$  that are preserved by switching. We may use this property later.

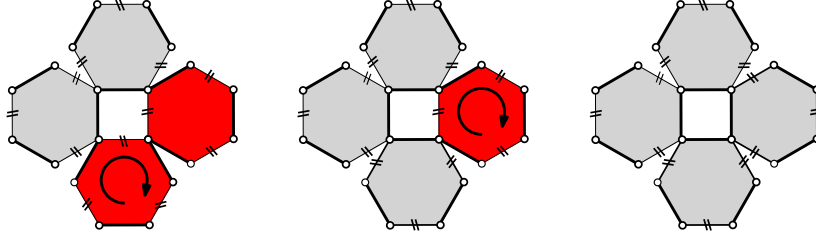


Figure 2: Switching of the hexagons.

We intend to obtain simplifications using successive switchings for  $j = 1, \dots, n$  and using the formula

$$\sum_{J \subset [n]} (-1)^{n-|J|} \prod_{j \in J} \left( \frac{d_N}{N} \right) (1 + O_j(d_N^{-\frac{1}{2}})) = \left( \frac{d_N}{N} \right)^n \times O(d_N^{-\frac{n}{2}}),$$

valid when the  $O_j(\cdot)$  depends on  $j \in [n]$  but not in  $J$ .

*Step 2: Construction of hexagons.* We consider now the graphs of  $\mathcal{G}_N(J) \cup (w'_j | w'_j \not\sim_J v_j)_{j \in [n]}$ , to which we add labels by  $2n$  more vertices  $v'_j$  and  $v''_j$ ,  $j = 1, \dots, n$ :

$$\mathcal{A}_N(J) = \mathcal{G}_N(J) \cup (w'_j | w'_j \not\sim_J v_j)_{j \in [n]} \cup (v'_j | v'_j \sim_J w'_j)_{j \in [n]} \cup (v''_j | v''_j \not\sim_J w_j)_{j \in [n]}.$$

In other words we consider the set of labeled graphs  $(G, w'_j, v'_j, v''_j, j \in [n])$ , where

A1:  $G \in \mathcal{G}_N$  is a  $d_N$ -regular graph on  $[N]$  and  $w'_j, v'_j, v''_j \in [N]$  are vertices.

A2: Four edges of each hexagon at each  $j$  are constructed:

$$v_j \sim_J w_j, w'_j \not\sim_J v_j, v'_j \sim_J w'_j, v''_j \not\sim_J w_j.$$

Note that the labels  $v_j, w_j$  are determined by  $T$ . On the contrary, the  $w'_j, v'_j, v''_j$  vary in the above enumeration. Choosing  $v'_j$  and  $v''_j$ , we have a total of  $(N - 1 - d_N)^n d_N^n$  possibilities (independently of the value of  $J$ ).

Hence, one has

$$\mathfrak{C}_N(T) := \sum_{J \subset [n]} (-1)^{n-|J|} \frac{\#\mathcal{A}_N(J)}{\#\mathcal{G}_N \times d_N^n (N - 1 - d_N)^n (N - 1)^n}. \quad (5.6)$$

In this definition, some of the vertices  $w'_j, v'_j$  or  $v''_j$  can be equal and/or coincide with vertices  $v_i, w_i$  for  $i \in [n]$ . This may be a problem for using the switching method later. Given a subset  $\tilde{J}$  of  $[n]$ , we denote by  $\mathcal{A}_N(J, \tilde{J})$  the set of  $(G, w'_j, v'_j, v''_j, j \in [n]) \in \mathcal{A}_N(J)$  such that:

- A3:  $j \in \tilde{J}$  if and only if the vertices  $v_j, w_j, w'_j, v'_j, v''_j$  are pairwise distinct and no vertex among  $w'_j, v'_j, v''_j$  belongs to  $\{v_i, w_i, w'_i, v'_i, v''_i\}_{i \neq j}$ .

In other words, the set  $\tilde{J}$  denotes the set of indices for which the vertices  $w'_j, v'_j, v''_j$  arise only once as labels (and do not arise as vertices of  $T$ ). We now fix a set  $\tilde{J} \subset [n]$  and finish the construction of the hexagons for indices in  $\tilde{J}$  only. The contribution of the indices in  $[n] \setminus \tilde{J}$  will be considered at the very end of our reasoning. Without loss of generality, we assume that  $\tilde{J} = \{1, \dots, \tilde{n}\}$  for some  $0 \leq \tilde{n} \leq n$ . Note that the choice of  $J \subset [n]$  can then be decomposed as a choice for a subset of  $[\tilde{n}]$  and of a subset of  $[n] \setminus [\tilde{n}]$ . We fix the second choice for the moment: given  $J \subset [\tilde{n}]$ , we set

$$\mathcal{A}_N(J, \tilde{n}) := \cup_{J_2 \subset [n] \setminus [\tilde{n}]} \mathcal{A}_N(J \cup J_2, \tilde{J})$$

and prove the estimate for

$$\mathfrak{C}_N(T, \tilde{n}) := \sum_{J \subset [\tilde{n}]} (-1)^{\tilde{n}-|J|} \frac{\#\mathcal{A}_N(J, \tilde{n})}{d_N^n (N-1-d_N)^n (N-1)^n \#G_N}. \quad (5.7)$$

For each graph of  $\mathcal{A}_N(J, \tilde{n})$  and each  $j \in [\tilde{n}]$ , we want to add the last vertex  $w''_j$  to close the associated hexagon (we do not add these vertices for indices  $j \notin [\tilde{n}]$ ). This is not always possible, given a labelled graph in  $\mathcal{A}_N(J, \tilde{n})$  to add such a vertex since it may happen that  $v'_j$  and  $v''_j$  have  $d_N$  neighbors in common.

Let  $\hat{J} \subset [\tilde{n}]$ . This partition represent the indices  $j = 1, \dots, \tilde{n}$  for which it will be possible to chose a vertex which is the neighbor of one but not both from  $v'_j$  and  $v''_j$ . We only create hexagons for  $j \in \hat{J}$ .

For each  $j \in [\tilde{n}]$ , we consider integers  $0 \leq L_j, M_j \leq 5n$  and  $\ell_j \in \{0, \dots, d_N - \max(L_j, M_j)\}$ . These integers will count in the following the number of non allowed choices at each step, when choosing vertices  $w''_j, j \in \hat{J}$ . The order of the steps will mater, since the number of allowed choices evolves along the construction. In short we denote  $\mathbf{L} = (L_j)_{j \in [\tilde{n}]}$ ,  $\mathbf{M} = (M_j)_{j \in [\tilde{n}]}$  and  $\boldsymbol{\ell} = (\ell_j)_{j \in [\tilde{n}]}$ . Let  $\mathcal{B}_N(J, \tilde{n}, \hat{J}, \mathbf{L}, \mathbf{M}, \boldsymbol{\ell})$  be the set of all  $(G, w'_j, v'_j, v''_j, w''_j, j \in [n], i \in \hat{J})$  such that

- B1:  $(G, w'_j, v'_j, v''_j, j \in [n]) \in \mathcal{A}_N(J, \tilde{n})$ ,  $w''_j \in [N]$  are vertices for any  $j \in \hat{J}$ .
- B2: The two remaining edges of each hexagon are chosen:  $\forall j \in \hat{J}, w''_j \not\sim_J v'_j, w''_j \sim_J v''_j$ ,
- B3: The construction is injective:  $\forall j \in \hat{J}$ , the  $w''_j$  are pairwise distinct and do not belong to  $\{v_i, w_i, w'_i, v'_i, v''_i\}_{i \in [n]}$ .
- B4: Some parameters are fixed and finite:  $\forall j \leq \tilde{n}, v'_j$  has exactly  $L_j$  neighbors in  $\{v_i, w_i, v'_i, w'_i, w''_i\}_{i \in [n]} \sqcup \{w''_i\}_{i < j, i \in \hat{J}} \setminus \{w'_j\}$ .
- B5:  $\forall j \leq \tilde{n}, v''_j$  has exactly  $M_j$  neighbors in  $\{v_i, w_i, v'_i, w'_i, w''_i\}_{i \in [n]} \sqcup \{w''_i\}_{i < j, i \in \hat{J}} \setminus \{w_j\}$ .
- B6: Some parameters are fixed and possibly large:  $\forall j \leq \tilde{n}, v'_j$  and  $v''_j$  have exactly  $\ell_j$  neighbors in common out of  $\{v_i, w_i, v'_i, w'_i, w''_i\}_{i \in [n]} \sqcup \{w''_i\}_{i < j, i \in \hat{J}}$ .
- B7: With boundary terms:  $\forall j \leq \tilde{n}$  such that  $j \notin \hat{J}, \ell_j = d_N - \max(L_j, M_j)$ .

See Figure 3 for a scheme of this labelled graph.

We now give a formula for  $\#\mathcal{A}_N(J, \tilde{n})$  in terms of the  $\#\mathcal{B}_N(J, \tilde{n}, \hat{J}, \mathbf{L}, \mathbf{M}, \boldsymbol{\ell})$ 's. Start with a graph  $G$  in  $\mathcal{A}_N(J, \tilde{n})$ , for which the vertices  $w'_j, v'_j$  and  $v''_j$  are chosen. Assume  $1 \in \hat{J}$ , so that we need to add the vertex  $w''_1$ . We partition  $\mathcal{A}_N(J, \tilde{n})$  into subsets according to the value of the constants  $L_1, M_1$  and  $\ell_1$  given by B4, B5 and B6. If  $1 \in J$ , we chose  $w''_1$  as a neighbor of  $v''_1$  (a priori  $d_N$  choices), with the restrictions that  $w''_1$  is not in the vertices previously considered

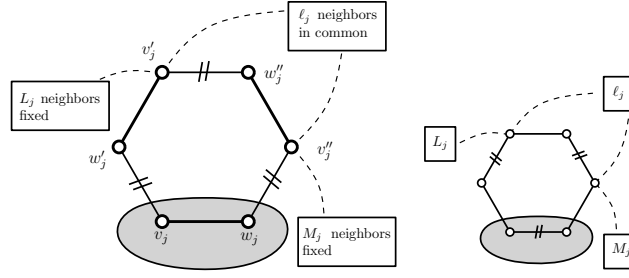


Figure 3: Construction of the hexagons: detail at  $j \in \hat{J}$  of an element of  $\mathcal{B}_N(J, \tilde{n}, \hat{J}, \mathbf{L}, \mathbf{M}, \ell)$ . This left picture represent the case where  $j \in J$ , the right one when  $j \notin J$ .

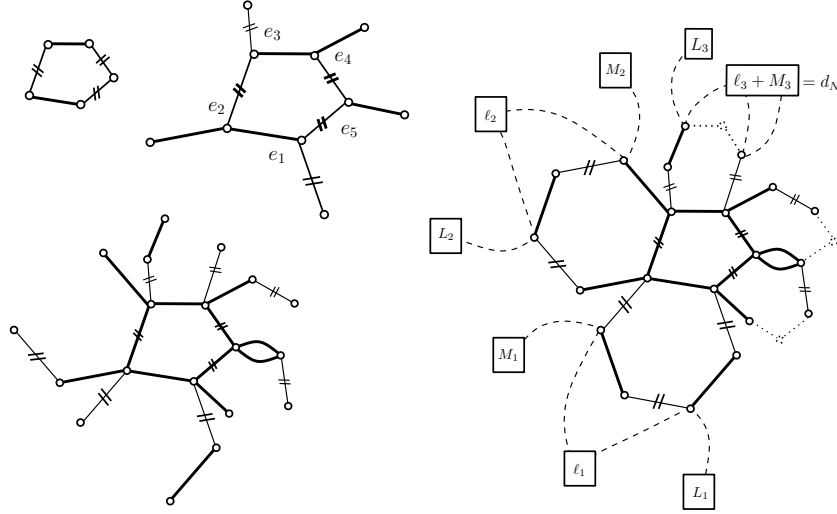


Figure 4: Construction of the hexagons: an element of  $\mathcal{B}_N(J, \tilde{n}, \hat{J}, \mathbf{L}, \mathbf{M}, \ell)$ ,  $n = 5$ ,  $J = \{1, 3\}$ ,  $\tilde{n} = 3$ ,  $\hat{J} = \{1, 2\}$ .

(remove  $M_1$  choices) and is not a neighbor of  $v'_1$  (remove  $\ell_1$  choices). If  $1 \notin J$ , we chose similarly  $v''_1$ , now among the neighbors of  $v'_1$ , with the same restrictions, then replacing the constant  $M_1$  by  $L_1$ . There are either  $(d_N - M_1 - \ell_1)$  or  $(d_N - L_1 - \ell_1)$  possible choices depending on the two above cases ( $1 \in J$  or not). Now, we use the same procedure in order to choose the vertex  $w''_2$  in  $2 \in \hat{J}$ , decomposing again our set of graphs thanks to  $L_2, M_2$  and  $\ell_2$ . If at some point  $j \notin \hat{J}$  we do not choose  $w''_j$  and look at  $j + 1$  continue this process up to  $\tilde{n}$  times. This gives

$$\#\mathcal{A}_N(J, \tilde{n}) = \sum_{\hat{J} \subset [\tilde{n}]} \sum_{\mathbf{L}, \mathbf{M}} \sum_{\ell} \frac{\#\mathcal{B}_N(J, \tilde{n}, \hat{J}, \mathbf{L}, \mathbf{M}, \ell)}{\kappa_N(J)}. \quad (5.8)$$

where

$$\kappa_N(J) = \prod_{j \in \hat{J} \cap J} (d_N - L_j - \ell_j) \times \prod_{j \in \hat{J} \setminus J} (d_N - M_j - \ell_j).$$

We assume without loss of generality  $\hat{J} = \{1, \dots, \hat{n}\}$  for  $1 \leq \hat{n} \leq \tilde{n}$ . As before, we fix the choice of the indices of  $J$  that are not in  $[\hat{n}]$  for the moment. Given  $J \subset [\hat{n}]$ , we set now

$$\mathcal{B}_N(J, \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \ell) := \cup_{\hat{J} \subset [\hat{n}] \setminus J} \mathcal{B}_N(J \cup \hat{J}, \tilde{n}, \hat{J}, \mathbf{L}, \mathbf{M}, \ell)$$



(we do not care about of the configuration at index  $\hat{n} < j < \tilde{n}$ ) and define

$$\mathfrak{C}_N(T, \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}) := \sum_{\ell_j, j \in \hat{J}} \sum_{J \subset [\hat{n}]} (-1)^{\hat{n}-|J|} \frac{\#\mathcal{B}_N(J, \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \boldsymbol{\ell})}{\kappa_N(J) \times \#\mathcal{G}_N \times d_N^n (N-1-d_N)^n (N-1)^n} \quad (5.9)$$

where the first sum is over all  $\ell_j = j, \dots, d_N - \max(L_j, M_j) - 1$  for  $j \leq \hat{n}$  (for  $\hat{n} + 1 \leq j < \tilde{n}$ , there is the restriction  $\ell_j = d_N - \max(L_j, M_j)$ ). Since there is a finite number of choices for  $\hat{n}, \mathbf{L}$  and  $\mathbf{M}$ , it is sufficient to prove estimate (5.1) for that quantity.

*Step 3: first switchings.* The switching method rotates the hexagon formed by the vertices  $\{v_j, w_j, v'_j, w'_j, v_j'', w_j''\}$  by one unit (edge) clockwise. Applied for each  $j \in \hat{J}$ , it turns a graph not containing  $e_j = (v_j, w_j)$  into a graph that contains  $e_j$  (changing the rightmost into the leftmost configurations in Figure 3). It is important to note that the constants  $\mathbf{L}, \mathbf{M}, \boldsymbol{\ell}$  remain the same after switching. This follows from the definition of the  $L_j, M_j, \ell_j$ 's. It follows that the cardinal of  $\mathcal{B}_N(J, \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \boldsymbol{\ell})$  does actually not depend on  $J$  and is equal to the one of

$$\mathcal{B}_{0,N}(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \boldsymbol{\ell}) := \mathcal{B}_N([\hat{n}], \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \boldsymbol{\ell})$$

for which each  $j = 1, \dots, \tilde{n}$  is in the leftmost configuration on Figure 3. Moreover,

$$\sum_{J \subset [\hat{n}]} (-1)^{\hat{n}-|J|} \frac{1}{\kappa_N(J)} = \prod_{j=1}^{\hat{n}} \left( \frac{1}{d_N - L_j - \ell_j} - \frac{1}{d_N - M_j - \ell_j} \right) = \prod_{j \in [\hat{n}]} O\left((d_N - \ell_j)^{-2}\right),$$

which yields

$$|\mathfrak{C}_N(T, \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M})| \leq \sum_{\ell_1, \dots, \ell_{\hat{n}}} \prod_{j=1}^{\hat{n}} O\left((d_N - \ell_j)^{-2}\right) \frac{\#\mathcal{B}_{0,N}(J, \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \boldsymbol{\ell})}{d_N^n (N-1-d_N)^n (N-1)^n \#\mathcal{G}_N}. \quad (5.10)$$

The conclusion of our construction so far is that we have gained a factor  $(d_N - \ell_j)^{-1}$  for each  $j \in [\hat{n}]$ , compared with the crude estimate where we forget the condition that  $w_j'' \not\sim v_j'$  and count the number of choices for  $w_j''$  (see Step 4 below). We first show that we get the desired factor  $d_N^{-\frac{1}{2}}$  provided  $\ell_j$  is not too close to  $d_N$ . We fix a constant  $0 < \eta' < \eta$ , where  $\eta$  has been chosen in the assumption of Theorem 2.3. We thus first consider the case where  $\ell_j \leq k_N := d_N - \eta' \sqrt{d_N}$ . In a second time, we somehow prove that when two given vertices have a large number of neighbors in common we get an exponentially small factor (the same argument applies for indices  $\hat{n} < j < \tilde{n}$ ).

We fix  $J \subset [\hat{n}]$  and the values of  $M_j, L_j, \ell_j$  for  $j \leq \tilde{n}$ . We first show how to suppress all the constructions for  $j \leq \hat{n}$  in Steps 5 and 6 below, proving the

**Lemma 5.1.**

$$\mathfrak{C}_N(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}) \leq O(d_N^{-\frac{\hat{n}}{2}}) \times \#\mathcal{C}_N(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}),$$

where  $\mathcal{C}_N(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M})$  is the set of labelled graphs  $(G, w'_j, v'_j, v''_j, \hat{n} < j \leq n)$  as in  $\mathcal{B}_N(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M})$  but where do not perform the constructions of labels for indices  $j \leq \hat{n}$ .

In steps 6 and 7 respectively, we investigate the cases where  $\hat{n} < j < \tilde{n}$  and  $\tilde{n} < j \leq n$  respectively, proving

**Lemma 5.2.**

$$\frac{\mathcal{C}_N(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \boldsymbol{\ell})}{(d_N(N-1-d_N)(N-1))^{n-\hat{n}}} = \sum_{J \subset [\tilde{n}] \setminus [\hat{n}]} O(d_N^{-\frac{(n-\hat{n})}{2}}) \times \frac{\#\mathcal{G}_N(J) \times \left(\frac{d_N}{N}\right)^{|J|}}{\#\mathcal{G}_N}$$

Assume Lemmas 5.1 and 5.2 hold true. Then, using the trivial bound  $\#\mathcal{G}_N(J)/\#\mathcal{G}_N \leq \mathbb{P}(T_J \subset G_N)$  where  $T_J$  is the graph formed by the edges  $e_j, j \in J$  of  $T$ , we have

$$|\mathfrak{C}_N(T)| \leq \sum_{\tilde{T} \subset T} \left(\frac{d_N}{N}\right)^n \varepsilon_N(\tilde{T}) \times \mathbb{P}(\tilde{T} \subset G_N), \quad \varepsilon_N(\tilde{T}) = O(d_N^{-\frac{n}{2}}). \quad (5.11)$$

with the sum bears over the subgraphs  $\tilde{T}$  of  $T$  with exactly  $n$  edges. With the same computation as in the proof of [ref], we deduce that  $P(\tilde{T} \subset G_N) = \alpha_N^n (1 + o(d_N^{-\frac{1}{2}}))$  for any  $\tilde{T}$ . As a feed-back, (5.11) yields  $\mathfrak{C}_N(T) = \left(\frac{d_N}{N}\right)^n \times O(d_N^{-\frac{n}{2}})$ , which is (5.1).

*Step 4: analysis for  $\ell$  smaller than  $k_N$ .* For simplicity, we denote  $\mathcal{B}_N(J, \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \ell)$  by  $\mathcal{B}_N(\cdot)$  and  $\ell_{\hat{n}}, M_{\hat{n}}, v_{\hat{n}} \dots$  by  $\ell, L, v, \dots$  in the computations below. Recall that  $v''$  is a neighbor of  $w'$  or (exclusively) of  $w''$ . Using the notation  $\mathcal{B}_N(\cdot) \setminus (w'')$  to mean that we forget the construction of  $w''$  in  $\mathcal{B}(\cdot)$ , we then have

$$\sum_{\ell=0}^{k_N} \frac{\#\mathcal{B}_N(\cdot)}{(d_N - \ell)^2} \leq \sum_{\ell=0}^{k_N} \frac{\#\mathcal{B}_N(\cdot) \setminus (w'')}{d_N - \ell}.$$

Similarly, with the notation  $\mathcal{B}_N(\cdot) \setminus (v'', \ell, M, L)$  meaning that we also forget the constraints B4, B5 and B6 for  $j = \tilde{n}$ , the following inequality

$$\sum_{\ell=0}^{k_N} \frac{\#\mathcal{B}_N(\cdot)}{(d_N - \ell)^2} \leq C \frac{\#\mathcal{B}_N(\cdot) \setminus (w'', \ell, M, L)}{d_N - k_N},$$

is valid for a constant  $C$  that bounds the number of choices for the  $M$  and  $L$ . Forgetting now the construction of the labels  $v', v''$  and  $w'$  yields

$$\sum_{\ell=0}^{k_N} \frac{\#\mathcal{B}_N(\cdot)}{(d_N - \ell)^2 d_N (N - d_N) (N - 1)} \leq C \times \frac{\#\mathcal{B}_N(\cdot) \setminus (\tilde{n})}{d_N - k_N}, \quad (5.12)$$

where the notation  $\#\mathcal{B}_N(\cdot) \setminus (\tilde{n})$  precisely means that we do not perform any choice of labels for the index  $j = \tilde{n}$ . We now choose  $k_N = d_N - \eta' \sqrt{d_N}$ . As a consequence, we get the final estimate of this step

$$\sum_{\ell=0}^{k_N} \frac{\#\mathcal{B}_N(\cdot)}{(d_N - \ell)^2 d_N (N - d_N) (N - 1)} \leq \frac{C}{\eta'} \times \frac{\#\mathcal{B}_N(\cdot) \setminus (\tilde{n})}{\sqrt{d_N}}. \quad (5.13)$$

*Step 5: analysis for  $\ell$  larger than  $k_N = d_N - \eta' \sqrt{d_N}$ .* We here show

**Lemma 5.3.** *There exists  $0 < \rho < 1$  such that*

$$\sum_{\ell \geq k_N} \frac{\#\mathcal{B}_N(J, \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \ell)}{(d_N - \ell)^2 d_N (N - d_N) (N - 1)} \leq \rho^{(\eta - \eta') \sqrt{d_N}} \#\mathcal{B}(\cdot) \setminus (\tilde{n}), \quad (5.14)$$

where  $\eta$  is the constant of Theorem 2.3.

To this aim, we use again the switching method. We now use the notation  $\mathcal{B}(\cdot, \ell)$  for  $\mathcal{B}_N(J, \tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}, \ell)$  to make apparent the dependance in  $\ell$ . In order to compare  $\mathcal{B}(\cdot, \ell)$  with  $\mathcal{B}(\cdot, \ell - 1)$  for  $\ell$  bigger than a quantity  $m_N := d_N - \eta \sqrt{d_N} < k_N$ , we introduce a square whose vertices are  $w'$  and new labeled vertices  $s, s'$  and  $t$ . Consider a graph  $G$  in  $\mathcal{B}(\cdot, \ell)$ . There are at least  $\ell - 6n$  possible choices of a vertex  $s$ , which is a common neighbor of  $w'$  and  $w''$  and which is out of the set  $\{v_i, w_i, v'_i, w'_i, w''_i\}_{i \in [n]} \sqcup \{v''_i\}_{i < j}$ . Hence we deduce that

$$\#\mathcal{B}_N(\cdot, \ell) \leq \frac{1}{\ell - 6n} \#\mathcal{B}_N(\cdot, \ell) \cup \{s \mid s \sim w' \text{ and } s \sim w''\}$$

where the symbol  $\mathcal{B}_N(\cdot, \ell) \cup (s \mid s \sim w' \text{ and } s \sim w'')$  represents the fact that  $s$  has been chosen according to the above constraints. This notation will be used hereafter while we introduce other vertices.

We then choose a second vertex  $s'$ , among those which are not connected to  $w'$  neither  $w''$ . Let us count the number of possible choices. There exist  $N - d_N$  vertices which are not neighbor of  $w'$ . Moreover,  $w''$  has  $d_N - \ell$  neighbors which are not neighbors of  $w'$ . So the number of vertices which are neither a neighbor of  $w'$  or  $w''$  is at least  $N - 2d_N + \ell$ . But  $\ell$  is larger than  $m_N$ . Then this number is at least  $N - d_N - \eta\sqrt{d_N}$ . We then choose  $s'$  which is neither a neighbor of  $w'$  or  $w''$ , different from  $s$  and out of the set  $\{v_i, w_i, v'_i, w'_i, w''_i\}_{i \in [n]} \sqcup \{v''_i\}_{i < j}$ . At last, we chose a neighbor  $t$  of  $s'$ , again out of the other vertices previously constructed. We then get

$$\#\mathcal{B}_N(\cdot, \ell) = \frac{\#\mathcal{B}_N(\cdot, \ell) \cup (s \mid s \sim w', s \sim w'') \cup (s' \mid s' \not\sim w', s' \not\sim w'') \cup (t \mid t \sim s')}{(\ell - c)(N - d_N - \eta\sqrt{d_N} - c)(d_n - c)} \quad (5.15)$$

where  $c$  depends only on  $n$ . See Figure 5 for a scheme of the set in the right hand side above.

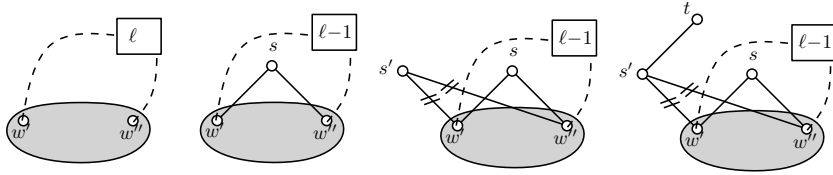


Figure 5: Notations for Formula (5.16). Left: a scheme of the set in the right hand side of (5.15). Right: the two different possibilities for the bond  $t \sim s$ .

From the dichotomy that either  $t$  and  $s$  are or not connected by an edge, we obtain that

$$\begin{aligned} & \#\mathcal{B}_N(\cdot, \ell) \cup (s \mid s \sim w', s \sim w'') \cup (s' \mid s' \not\sim w', s' \not\sim w'') \cup (t \mid t \sim s') \\ &= \#\mathcal{B}_N(\cdot, \ell) \cup (s \mid s \sim w', s \sim w'') \cup (s' \mid s' \not\sim w', s' \not\sim w'') \cup (t \mid t \sim s', t \not\sim s) \\ & \quad + \#\mathcal{B}_N(\cdot, \ell) \cup (s \mid s \sim w', s \sim w'') \cup (s' \mid s' \not\sim w', s' \not\sim w'') \cup (t \mid t \sim s', t \sim s) \end{aligned} \quad (5.16)$$

By the switching method, for the first quantity of the right hand side in 5.16, we can exchange the symbols  $\sim$  and  $\not\sim$  for the square formed by the sequence of vertices  $(w', s', t, s)$  as long as we decrease  $\ell$  by one:

$$\begin{aligned} & \#\mathcal{B}_N(\cdot, \ell) \cup (s \mid s \sim w', s \sim w'') \cup (s' \mid s' \not\sim w', s' \not\sim w'') \cup (t \mid t \sim s', t \not\sim s) \\ &= \#\mathcal{B}_N(\cdot, \ell - 1) \cup (s \mid s \not\sim w', s \sim w'') \cup (s' \mid s' \sim w', s' \not\sim w'') \cup (t \mid t \not\sim s', t \sim s). \end{aligned}$$

It is important to note that we do change the value of  $\ell$  exactly by one while switching, since we assumed that  $s'$  is not a neighbor of  $w''$ .

The quantity in the right hand side above is obviously bounded by the same quantity where we ignore the condition that  $s' \not\sim t$ . Then the number of choices for  $t$  is at most  $d_N$ , the number of choices of  $s'$  is at most  $d_N - \ell$  ( $s'$  is now neighbor of  $w'$  but not of  $w''$ ). Counting the number of possible choices for  $s$ , which is a neighbor of  $w''$  but not of  $w'$  yields

$$\begin{aligned} & \#\mathcal{B}_N(\cdot, \ell - 1) \cup (s \mid s \not\sim w', s \sim w'') \cup (s' \mid s' \sim w', s' \not\sim w'') \cup (t \mid t \not\sim s', t \sim s) \\ & \leq \#\mathcal{B}_N(\cdot, \ell - 1) \cup (s \mid s \not\sim w', s \sim w'') \cup (s' \mid s' \sim w', s' \not\sim w'') \cup (t \mid t \sim s) \\ & \leq d_N(d_N - \ell - 1) \#\mathcal{B}_N(\cdot, \ell - 1) \cup (s \text{ s.t. } s \not\sim w', s \sim w'') \\ & \leq d_N(d_N - \ell - 1)^2 \#\mathcal{B}_N(\cdot, \ell - 1), \end{aligned} \quad (5.17)$$

This computation is summed up in Figure 6.

Let now estimate the second quantity in the right hand side of (5.16). Using the trivial

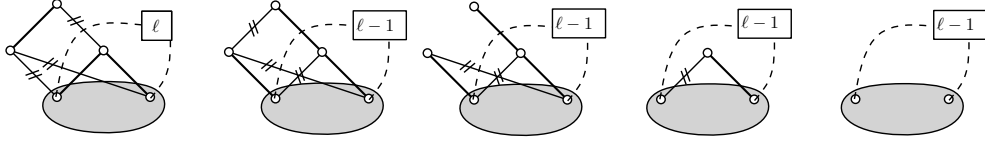


Figure 6: A scheme of the proof of Estimate (5.17).

bound where we forget that  $w' \not\sim s'$  and counting the choices for  $s'$ ,  $t'$  and then  $s$  yields

$$\begin{aligned}
& \#\mathcal{B}_N(\cdot, \ell) \cup (s \mid s \sim w', s \sim w'') \cup (s' \mid s' \not\sim w', s' \not\sim w'') \cup (t \mid, t \sim s', t \sim s) \\
& \leq \#\mathcal{B}_N(\cdot, \ell) \cup (s \mid s \sim w', s \sim w'') \cup (t \mid, t \sim s) \cup (s' \mid s' \sim t) \\
& \leq d_N \times \#\mathcal{B}_N(\cdot, \ell) \cup (s \mid s \sim w', s \sim w'') \cup (t \mid, t \sim s', t \sim s) \\
& \leq d_N^2 \ell \#\mathcal{B}(\cdot, \ell).
\end{aligned} \tag{5.18}$$

This computation is summed up in Figure 7.

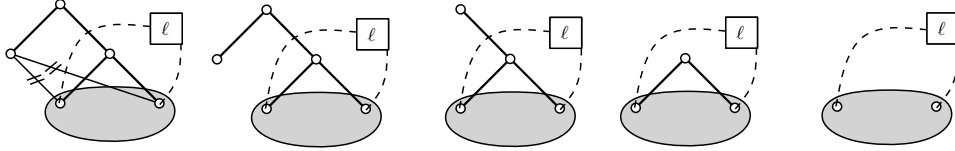


Figure 7: A scheme of the proof of Estimate (5.18).

To sum up, by (5.16), (5.17) and (5.18), there exists a constant  $c$  such that

$$(N - d_N - \eta\sqrt{d_N} - c)(d_N - c) \#\mathcal{B}_N(\cdot, \ell) \leq \frac{d_N(d_N - \ell)^2 \#\mathcal{B}_N(\cdot, \ell - 1) + d_N^2 \ell \#\mathcal{B}_N(\cdot, \ell)}{\ell - c},$$

or equivalently

$$\#\mathcal{B}_N(\cdot, \ell) \leq O\left(\frac{1}{N - 2d_N - \eta\sqrt{d_N}}\right) \frac{(d_N - \ell)^2}{\ell} \#\mathcal{B}_N(\cdot, \ell - 1).$$

But  $\ell \geq d_N - \eta\sqrt{d_N}$  yields  $\frac{(d_N - \ell)^2}{\ell} \leq \frac{\eta^2 d_N}{d_N - B\sqrt{d_N}} \leq \eta^2$ . Hence, thanks to the technical condition  $N - 2d_N - \eta\sqrt{d_N} \xrightarrow{N \rightarrow \infty} \infty$ , for  $N$  large enough there exists  $\rho < 1$  such that

$$\#\mathcal{B}_N(\cdot, \ell) \leq \rho \#\mathcal{B}_N(\cdot, \ell - 1) \leq \rho^{\ell - m_N} \#\mathcal{B}(\cdot, m_N)$$

where we recall that  $m_N = d_N - \eta\sqrt{d_N}$ . With the same reasoning as in step 3, we bound  $\mathcal{B}(\cdot, m_N)$  from above by successively: forgetting the condition  $f_j'' \notin G_N$ , counting the number of choices for  $v''$ , forgetting the conditions given by  $L, M$  and  $m_N$ , and counting the number of choices for  $w', v'$  and  $w''$ . This yields

$$\#\mathcal{B}(\cdot, m_N) \leq (d_N - m_N)d_N(N - d_N)^2 \#\mathcal{B}(\cdot) \setminus \{\tilde{n}\},$$

where we remind that  $\mathcal{B}(\cdot) \setminus \{\tilde{n}\}$  means we have forget all constructions for the index  $j = \tilde{n}$  in the construction of  $\mathcal{B}(\cdot)$ .

Hence, we obtain

$$\sum_{\ell=k_N}^{d_N - \max(M, L)} \frac{\#\mathcal{B}_N(\cdot)}{(d_N - \ell)^2 d_N (N - d_N) (N - 1)} \leq \rho^{k_N - m_N} \#\mathcal{B}_N(\cdot) \setminus \{\tilde{n}\} = \rho^{(\eta - \eta')\sqrt{d_N}} \#\mathcal{B}_N(\cdot) \setminus \{\tilde{n}\}.$$

This is the announced estimate (5.14).

The conclusion of Steps 4 and 5 is that Lemma 5.1 is true.

*Step 6: Indices  $\hat{n} < j \leq \tilde{n}$ .* By Step 4, it is clear that

$$\#\mathcal{C}_N(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}) \leq \sqrt{d_N}^{-\frac{(\tilde{n}-\hat{n})}{2}} \#\mathcal{D}_N(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M}),$$

where  $\mathcal{D}_N(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M})$  is the set of labelled graphs  $(G, w'_j, v'_j, v''_j, \tilde{n} < j \leq n)$  such that all the constructions in  $\mathcal{C}_N(\tilde{n}, \hat{n}, \mathbf{L}, \mathbf{M})$  for  $\hat{n} < j \leq \tilde{n}$  are forgotten.

*Step 7: Indices  $\hat{n} < j \leq \tilde{n}$ .* For each  $j = \tilde{n} + 1, \dots, n$ , there exists (at least) a vertex  $x_j$  among  $v'_j, w'_j$  and  $w''_j$  which coincides with (at least) another vertex in  $\{v_i, w_i\}_{i \in n} \cup \{v'_i, w'_i, w''_i\}_{i > \tilde{n}}$ . Assume that  $j \in J$ . If we forget the constraint A3, they would be  $d_N$  choices for  $w'_j$  and  $N - 1 - d_N$  choices for  $v'_j$  and  $w'_j$ . If instead  $j \notin J$ , these choices are respectively  $N - 1 - d_N$  and  $d_N$  choices.

As a consequence, if  $x_j$  is chosen in the fixed vertices  $v_i, w_i, i = 1, \dots, n$ , it is clear that we gain at least  $d_N^{-1}$  from that term. Assume now that  $x_j$  arises as a labeled vertex associated to  $m_j - 1$  other vertices among  $v'_i, w'_i, w''_i, i > \tilde{n}$ . In other word,  $m_j$  is the multiplicity of  $x_j$  as a labelled vertex. Without the constraint A3, the number of choices for  $x_j$  and these  $m_j - 1$  other vertices is of the form  $d_N^{p_j} \times (N - d_N)^{q_j}$  for  $p_j + q_j = m_j$ . Instead, it is less than  $O(d_N)$  if  $q_j \geq 1$  and less than  $O(N)$  otherwise. We gain for each copy of the copies of  $x_j$  at least the  $m_j$ -th root of its contribution  $d_N$  or  $N - d_N$ . Hence we gain at least the expected term  $\sqrt{d_N}$  for each index  $j$ .

## References

- [1] G. Anderson, A. Guionnet, O. Zeitouni An introduction to random matrices. Cambridge Studies in Advanced Mathematics, 118. Cambridge University Press, Cambridge, 2010. xiv+492 pp. ISBN: 978-0-521-19452-5
- [2] G. Ben Arous, K. Dang On fluctuations of eigenvalues of random permutation matrices arXiv:1106.2108 (2011) , To appear in *Ann. Inst. Henri Poincaré*.
- [3] C. Bordenave, M. Lelarge, J. Salez The rank of diluted random graphs *Annals of Probability*, Vol. 39 (2011), No. 3, 1097-1121
- [4] I. Dumitriu, T. Johnson, S. Pal, E. Paquette, Functional limit theorems for random regular graphs. *Probab. Theory Related Fields*, Vol. 156 (2013), no. 3-4, 921-975.
- [5] I. Dumitriu, S. Pal, Sparse regular random graphs: spectral density and eigenvectors. *Annals of Probability*, 40 (2012), no. 5, 2197-2235.
- [6] K. Dykema On certain free product factors via an extended matrix model. *J. Funct. Anal.* 112 (1993), no. 1, 31-60
- [7] A. Guionnet Large random matrices: lectures on macroscopic asymptotics. Lecture Notes in Mathematics. Lectures from the 36th Probability Summer School held in Saint-Flour, 2006. Springer-Verlag, Berlin (2009), xii+294.
- [8] A. Khorunzhy, Sparse random matrices: spectral edge and statistics of rooted trees. *Adv. in Appl. Probab.* 33 (2001), no. 1, 124-140.
- [9] L. Lovász Very large graphs. Current developments in mathematics 2008 (2009), 67-128.
- [10] C. Male, The distribution of traffics and their free product. arXiv:1111.4662v4, preprint.
- [11] C. Male, The limiting distributions of large heavy Wigner and arbitrary random matrices. arXiv:1209.2366, preprint

- [12] B. McKay, The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.* 40 (1981), 203–216.
- [13] B. Mohar, W. Woess, A Survey on Spectra of Infinite Graphs. *Bull. London Math. Soc.* 21 (1989), no. 3, 209-234.
- [14] L. V. Tran, V. H. Vu, K. Wang, Sparse Random Graphs: Eigenvalues and Eigenvectors *Random Structures and Algorithms*, Vol. 42 (2013), no. 1, p 110–134
- [15] D. Voiculescu, Operations on certain non-commutative operator-valued random variables. *Astérisque* 232 (1995), 243-275.