Weak convergence of U-statistics on a row-column exchangeable matrix

TÂM LE MINH[1]

[1] *Université Paris-Saclay, AgroParisTech, INRAE, UMR MIA-Paris, 75005, Paris, France,*
E-mail: tam.le-minh@inrae.fr

U-statistics are used to estimate a population parameter by averaging a function on a subsample over all the subsamples of the population. In this paper, the population we are interested in is formed by the entries of a row-column exchangeable matrix. We consider U-statistics derived from functions of quadruplets, i.e. submatrices of size 2×2. We prove a weak convergence result for these U-statistics in the general case and we establish a Central Limit Theorem when the matrix is also dissociated. We shed further light on these results using the Aldous-Hoover representation theorem for row-column exchangeable random variables. Finally, to illustrate these results, we give examples of hypothesis testing for bipartite networks.

MSC2020 subject classifications: Primary 60F05; secondary 60G09; 62G20
Keywords: Central Limit Theorem; hypothesis testing; row-column exchangeability; U-statistics

1. Introduction

U-statistics form a large class of statistics with powerful properties. They are built as the average of a given function on a subsample of a population, called kernel, applied to all the subsamples taken from this population. The population usually consists of i.i.d. individuals. In this case, Hoeffding (1948) gives a Central Limit Theorem (CLT), which ensures their asymptotic normality. For non-i.i.d. cases, similar results (Nandi and Sen, 1963) exist when the population is exchangeable, i.e. when the joint distribution of a subsample only depends on its size or equivalently for any finite permutation σ,

$$(Y_1, Y_2, ...) \overset{\mathcal{D}}{=} (Y_{\sigma(1)}, Y_{\sigma(2)}, ...).$$

In these cases, it is convenient to view the kernel h taken on each subsample of size k as a random variable indexed by k-tuple, e.g. $X_{\boldsymbol{i}} = X_{i_1 i_2 \dots i_k} := h(Y_{i_1}, Y_{i_2}, ..., Y_{i_k})$ and the U-statistic is therefore a sum of random variables. If the population is exchangeable, then the array X is not necessarily exchangeable, but it is jointly exchangeable, i.e. for any sequence of k-tuples $(\boldsymbol{i}, \boldsymbol{j}, ...)$ and for any finite permutation σ,

$$(X_{\boldsymbol{i}}, X_{\boldsymbol{j}}, ...) \overset{\mathcal{D}}{=} (X_{\sigma(i_1)\sigma(i_2)...\sigma(i_k)}, X_{\sigma(j_1)\sigma(j_2)...\sigma(j_k)}, ...).$$

A CLT exists for sums of jointly exchangeable variables, which has been proven by Eagleson and Weber (1978).

In our paper, the sample consists of the entries of a matrix Y of size $m \times n$, the rows and columns of which are separately exchangeable (row-column exchangeable, RCE), i.e. denoting $\mathbb{S}_n$ the symmetric group of order n, for any $\Phi = (\sigma_1, \sigma_2) \in \mathbb{S}_m \times \mathbb{S}_n$,

$$\Phi Y \overset{\mathcal{D}}{=} Y,$$

where $\Phi Y := (Y_{\sigma_1(i)\sigma_2(j)})_{1\leq i,j<\infty}$. We consider U-statistics based on submatrices of size 2×2, that we call quadruplets

$$Y_{\{i_1,i_2;j_1,j_2\}} := \begin{pmatrix} Y_{i_1j_1} & Y_{i_1j_2} \\ Y_{i_2j_1} & Y_{i_2j_2} \end{pmatrix}.$$

Their kernels are real functions h such that for any matrix Y, $h(Y_{\{1,2;1,2\}}) = h(Y_{\{2,1;1,2\}}) = h(Y_{\{1,2;2,1\}})$. Applied to a matrix of size $m\times n$, a quadruplet U-statistic is then defined by

$$U^h_{m,n} = \binom{m}{2}^{-1}\binom{n}{2}^{-1} \sum_{\substack{1\leq i_1<i_2\leq m\\ 1\leq j_1<j_2\leq n}} h(Y_{\{i_1,i_2;j_1,j_2\}}),$$

where $\binom{m}{2}$ is the number of 2-combinations from m elements. We can denote $X_{[i_1,i_2;j_1,j_2]} := h(Y_{\{i_1,i_2;j_1,j_2\}})$. However, contrarily to the case where Y is fully exchangeable, X is not jointly exchangeable in this case. Our aim is to establish a weak convergence theorem for these U-statistics using the martingale approach used by Eagleson and Weber (1978).

We apply our results to network analysis. The matrix Y can be seen as a weighted bipartite network, where the rows and the columns represent individuals of two different types, and the interactions can only happen between individuals of two different types. Each entry Y_{ij} represents the intensity of the interaction between the individuals i (of type 1) and j (of type 2). As an example, we consider two versions of the Weighted Bipartite Expected Degree Distribution (WBEDD) model, which is a weighted, bipartite and exchangeable extension of the Expected Degree Sequence model (Chung and Lu, 2002; Ouadah, Latouche and Robin, 2021). For binary graphs, the degree of a node is the number of edges that stem from it. For weighted graphs, the equivalent notion is the sum of the weights of these edges. It is sometimes called node strength (Barrat et al., 2004), but we will simply refer to it as node weight. The WBEDD model draws the node weights from two distributions, characterised by real functions f and g. The expected edge weights Y_{ij} are then proportional to the expected weights of the involved nodes. The model can be written as

$$\xi_i, \eta_j \overset{iid}{\sim} \mathcal{U}[0,1]$$
$$Y_{ij} \mid \xi_i, \eta_j \sim \mathcal{L}(\lambda f(\xi_i) g(\eta_j)).$$

where $\mathcal{L}$ is a family of probability distributions over positive real numbers such that the expectation of $\mathcal{L}(\mu)$ is μ and f and g are normalized by the condition $\int f = \int g = 1$. Consequently, λ is the mean intensity of the network. The two versions of the WBEDD are:

Version 1 λ is constant,
Version 2 λ is a positive random variable.

We explain the implications of the two versions and how our results apply to both of them. Then we suggest a framework to design statistical tests on these models using our CLT and we discuss how one can extend it.

Our results are presented and proven in Section 2. In addition to the RCE case, we prove that if the matrix is also dissociated, i.e. if any of its submatrices with disjoint indexing sets are independent, then we obtain a CLT. Section 3 gives examples of application of this CLT to hypothesis testing on networks.

2. Main result

2.1. Asymptotic framework

Our results apply in an asymptotic framework where the numbers of rows and columns of Y grow at the same rate, i.e. $m/(m+n) \to c$ and at each step, only one row or one column is added to the matrix Y. Now, we build a sequence of dimensions $(m_N, n_N)_{N\geq 1}$ that satisfies these conditions.

***Definition* 2.1** (Sequences of dimensions)**.** Let c be an irrational number such that $0 < c < 1$. For all $N \in \mathbb{N}$, we define $m_N = 2 + \lfloor c(N+1) \rfloor$ and $n_N = 2 + \lfloor (1-c)(N+1) \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function.

Proposition 2.2. *m_N and n_N satisfy:*

1. $\frac{m_N}{m_N+n_N} \xrightarrow[N\to\infty]{} c$,
2. $m_N + n_N = 4 + N$*, for all* $N \in \mathbb{N}$.

Corollary 2.3. *At each iteration $N \in \mathbb{N}^*$, one and only one of these two propositions is true:*

1. $m_N = m_{N-1} + 1$ *and* $n_N = n_{N-1}$,
2. $n_N = n_{N-1} + 1$ *and* $m_N = m_{N-1}$.

Such sequences m_N and n_N satisfy the desired growth conditions (proof given in Appendix A). We define the sequence of U-statistics as $U_N^h := U_{m_N,n_N}^h$.

2.2. Theorems

We establish the following results on the asymptotic behaviour of U-statistics over RCE matrices.

Theorem 2.4 (Main theorem)**.** *Let Y be a RCE matrix. Let h be a quadruplet kernel such that $\mathbb{E}[h(Y_{\{1,2;1,2\}})^2] < \infty$. Let $\mathcal{F}_N = \sigma\big((U_{kl}^h, k \geq m_N, l \geq n_N)\big)$ and $\mathcal{F}_\infty := \bigcap_{N=1}^\infty \mathcal{F}_N$. Set $U_\infty^h = \mathbb{E}[h(Y_{\{1,2;1,2\}})|\mathcal{F}_\infty]$. Then*

$$\sqrt{N}(U_N^h - U_\infty^h) \xrightarrow[N\to\infty]{\mathcal{D}} W,$$

where W is a random variable with characteristic function $\phi(t) = \mathbb{E}[\exp(-\frac{1}{2}t^2 V)]$, where

$$V = \frac{4}{c}\mathrm{Cov}\big(h(Y_{\{1,2;1,2\}}), h(Y_{\{1,3;3,4\}})\big|\mathcal{F}_\infty\big) + \frac{4}{1-c}\mathrm{Cov}\big(h(Y_{\{1,2;1,2\}}), h(Y_{\{3,4;1,3\}})\big|\mathcal{F}_\infty\big).$$

Theorem 2.4 states that the limit distribution of $\sqrt{N}(U_N^h - U_\infty^h)$ is a mixture of Gaussians, but we see that if V is constant, then it is a simple Gaussian. Next we identify a class of models where the limiting distribution of $\sqrt{N}(U_N^h - U_\infty^h)$ is a simple Gaussian.

***Definition* 2.5.** Y is a dissociated matrix if and only if $(Y_{ij})_{1\leq i\leq m, 1\leq j\leq n}$ is independent of $(Y_{ij})_{i>m, j>n}$, for all m and n.

In other words, Y is dissociated if submatrices that are not sharing any row or column are independent. Now we claim the following extension to Theorem 2.4 for dissociated RCE matrices.

Theorem 2.6. *In addition to the hypotheses of Theorem 2.4, if Y is dissociated, then U^h_∞ and V are constant and*

$$\sqrt{N}(U^h_N - U^h_\infty) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0,V),$$

More precisely,

1. $U^h_\infty = \mathbb{E}[h(Y_{\{1,2;1,2\}})]$,
2. $V = \frac{4}{c}\mathrm{Cov}\big(h(Y_{\{1,2;1,2\}}), h(Y_{\{1,3;3,4\}})\big) + \frac{4}{1-c}\mathrm{Cov}\big(h(Y_{\{1,2;1,2\}}), h(Y_{\{3,4;1,3\}})\big)$.

Now we shall explain this result in the light of the Aldous-Hoover representation theorem. Theorem 1.4 of Aldous (1981) states that for any RCE matrix Y, there exists a real function f such that if we denote $Y^*_{ij} = f(\alpha, \xi_i, \eta_j, \zeta_{ij})$, for $1 \le i, j < \infty$, where the α, ξ_i, η_j and ζ_{ij} are i.i.d. random variables with uniform distribution over $[0,1]$, then

$$Y \overset{\mathcal{D}}{=} Y^*.$$

It is possible to identify the role of each of the random variables involved in the representation theorem. We notice that each Y_{ij} is determined by α, ξ_i, η_j and ζ_{ij}. ζ_{ij} is entry-specific while ξ_i is shared by all the entries involving the row i and η_j by the ones involving the column j. Therefore, the ξ_i and η_j represent the contribution of each individual of type 1 and type 2 of the network, i.e. each row and column of the matrix. These contributions are i.i.d., which makes the network exchangeable. Finally, α is global to the whole network and shared by all entries.

Proposition 3.3 of Aldous (1981) states that if Y is dissociated, then Y^* can be written without α, i.e. it is of the form $Y^*_{ij} = f(\xi_i, \eta_j, \zeta_{ij})$, for $1 \le i, j < \infty$. In this case, because the ξ_i, η_j and ζ_{ij} are i.i.d., averaging with the U-statistic over an increasing number of nodes nullifies the contribution of each individual interaction (ζ_{ij}) and node (ξ_i and η_j). In the general case, i.e. when Y is not dissociated, then conditionally on α, Y is dissociated. It is easy to see that the mixture of Gaussians from Theorem 2.4 results from this conditioning.

We can also state with ease that Theorem 2.4 can be applied to matrices Y generated by the two versions of the WBEDD model. Theorem 2.6 only applies to Version 1, where the matrix is dissociated. Indeed, we see that in both models conditionally on λ, the expected mean of the interactions of any submatrix is λ. Therefore any 2 submatrices are independent if λ is constant. We could also have noticed that λ is determined by the α from the representation theorem of Aldous-Hoover.

In practice, dissociated exchangeable random graph models are widely spread. Notably, a RCE model is dissociated if and only if it can be written as a W-graph (or graphon), i.e. it is defined by a distribution $\mathcal{W}$ depending on two parameters in $[0,1]$ such that for $1 \le i, j < \infty$:

$$\xi_i, \eta_j \overset{i.i.d.}{\sim} \mathcal{U}[0,1]$$
$$Y_{ij} \mid \xi_i, \eta_j \sim \mathcal{W}(\xi_i, \eta_j)$$

In this definition, it is easy to recognize the variables from the representation theorem of Aldous-Hoover. We simply identify the ξ_i and η_j, then it suffices to take $\phi^{-1}_{\xi_i,\eta_j}$ the inverse distribution function of $\mathcal{W}(\xi_i, \eta_j)$ to see that defining the dissociated RCE matrix Y^* such that $Y^*_{ij} = f(\xi_i, \eta_j, \zeta_{ij}) := \phi^{-1}_{\xi_i,\eta_j}(\zeta_{ij})$ fulfills $Y^* \overset{\mathcal{D}}{=} Y$. It is also straightforward to remark that unlike Version 2, Version 1 of the WBEDD model can be written as a W-graph model, setting $\mathcal{W}(\xi_i, \eta_j) := \mathcal{L}(\lambda f(\xi_i) g(\eta_j))$.

2.3. Proof of Theorem 2.4

To prove Theorem 2.4, we adapt the proof of Eagleson and Weber (1978) establishing the asymptotic normality of sums of backward martingale differences. The definition of a backward martingale is reminded in Appendix B.

Theorem 2.7 (Eagleson and Weber, 1978). *Let $(M_n, \mathcal{F}_n)_{n\geq 1}$ be a square-integrable reverse martingale, V a $\mathcal{F}$-measurable, a.s. finite, positive random variable. Denote $M_\infty := \mathbb{E}[M_1|\mathcal{F}_\infty]$ where $\mathcal{F}_\infty := \bigcap_{n=1}^\infty \mathcal{F}_n$. Set $Z_{nk} := \sqrt{n}(M_k - M_{k+1})$. If:*

1. *$\sum_{k=n}^\infty \mathbb{E}[Z_{nk}^2|\mathcal{F}_{k+1}] \xrightarrow[n\to\infty]{\mathbb{P}} V$ (asymptotic variances),*
2. *for all $\epsilon > 0$, $\sum_{k=n}^\infty \mathbb{E}[Z_{nk}^2 \mathbb{1}_{\{|Z_{nk}|>\epsilon\}}|\mathcal{F}_{k+1}] \xrightarrow[n\to\infty]{\mathbb{P}} 0$ (conditional Lindeberg condition),*

then $\sum_{k=n}^\infty Z_{nk} = \sqrt{n}(M_n - M_\infty) \xrightarrow[n\to\infty]{\mathcal{D}} W$, where W is a random variable with characteristic function $\phi(t) = \mathbb{E}[\exp(-\frac{1}{2}t^2V)]$.

Proof of Theorem 2.4. The three steps to apply Theorem 2.7 to $(M_N)_{N\geq 1} = (U_N^h)_{N\geq 1}$ are to show that it is a backward martingale for a well chosen filtration and that it fulfills conditions 1 and 2. The expression of V is made explicit along the way. More precisely,

1. first, defining $\mathcal{F}_N = \sigma\big((U_{kl}^h, k \geq m_N, l \geq n_N)\big)$, Proposition C.1 states that $(U_N^h, \mathcal{F}_N)_{N\geq 1}$ is indeed a square-integrable reverse martingale ;
2. then, Proposition D.1 implies that $\sum_{k=n}^\infty \mathbb{E}[Z_{NK}^2|\mathcal{F}_{K+1}]$ does converge to a random variable V with the desired expression ;
3. finally, the conditional Lindeberg condition is ensured by Proposition E.1, since from it, we deduce that for all $\epsilon > 0$, $\sum_{K=N}^\infty \mathbb{E}[Z_{NK}^2 \mathbb{1}_{\{|Z_{NK}|>\epsilon\}}|\mathcal{F}_{K+1}] \xrightarrow[N\to\infty]{\mathbb{P}} 0$.

Hence Theorem 2.7 can be applied to U_N^h and we obtain that $\sqrt{N}(U_N^h - U_\infty^h) \xrightarrow[N\to\infty]{\mathcal{D}} W$, where W is a random variable with characteristic function $\phi(t) = \mathbb{E}[\exp(-\frac{1}{2}t^2V)]$ with V specified by Proposition D.1. The proofs of Propositions C.1, D.1 and E.1 are provided in Appendices C, D. and E respectively. □

2.4. Proof of Theorem 2.6

The proof of Theorem 2.6 relies on a Hewitt-Savage type zero-one law for events that are permutable in our row-column setup. Therefore, it is useful to define first what a row-column permutable event is. We remind the Aldous-Hoover representation theorem for dissociated RCE matrices as stated earlier: if Y is a dissociated RCE matrix, then its distribution can be written with $(\xi_i)_{1\leq i<m_N}$, $(\eta_j)_{1\leq j<n_N}$ and $(\zeta_{ij})_{1\leq i\leq m_N, 1\leq j\leq n_N}$ arrays of i.i.d. random variables.

Then let us consider such arrays of i.i.d. random variables $(\xi_i)_{1\leq i<m_N}$, $(\eta_j)_{1\leq j<n_N}$ and $(\zeta_{ij})_{1\leq i\leq m_N, 1\leq j\leq n_N}$. If we were to consider events depending only on them, there is no loss of generality in using the product probability space $(\Omega_N, \mathcal{A}_N, \mathbb{P}_N)$, where

$$\Omega_N = \Big\{(\omega^\xi, \omega^\eta, \omega^\zeta) : \omega^\xi \in \mathbb{R}^{m_N}, \omega^\eta \in \mathbb{R}^{n_N}, \omega^\zeta \in \mathbb{R}^{m_N n_N}\Big\} = \mathbb{R}^{m_N + n_N + m_N n_N},$$

$$\mathcal{A}_N = \mathcal{B}(\mathbb{R})^{m_N + n_N + m_N n_N},$$

$$\mathbb{P}_N = \mu^{m_N + n_N + m_N n_N}.$$

We then define the action of a row-column permutation on an element of Ω_N.

***Definition* 2.8.** Let $\Phi = (\sigma_1, \sigma_2) \in \mathbb{S}_{m_N} \times \mathbb{S}_{n_N}$. The action of Φ on $\omega \in \Omega_N$ is defined by

$$\Phi\omega = \left(\sigma_1\omega^\xi, \sigma_2\omega^\eta, (\sigma_1, \sigma_2)\omega^\zeta\right)$$

where $\sigma_1\omega^\xi = (\omega^\xi_{\sigma_1(i)})_{1\leq i<m_N}$, $\sigma_2\omega^\eta = (\omega^\eta_{\sigma_2(j)})_{1\leq j<n_N}$ and $(\sigma_1, \sigma_2)\omega^\zeta = (\omega^\zeta_{\sigma_1(i)\sigma_2(j)})_{1\leq i<m_N, 1\leq j<n_N}$

***Definition* 2.9.** Let $A \in \mathcal{A}_N$. A is invariant by the action of $\mathbb{S}_{m_N} \times \mathbb{S}_{n_N}$ if and only if for all $\Phi \in \mathbb{S}_{m_N} \times \mathbb{S}_{n_N}$, $\Phi^{-1}A = A$, i.e.

$$\{\omega : \Phi\omega \in A\} = \{\omega : \omega \in A\}\,.$$

Notation. In this section, we denote by $\mathcal{E}_N$ the collection of events of $\mathcal{A}_N$ that are invariant by row-column permutations of size $m_N \times n_N$, i.e. $\Phi \in \mathbb{S}_{m_N} \times \mathbb{S}_{n_N}$. We denote $\mathcal{E}_\infty := \bigcap_{n=1}^\infty \mathcal{E}_N$, which is the collection of events that are invariant by permutations of size $m_N \times n_N$, for all N.

The following theorem is an extension of the Hewitt-Savage zero-one law to the row-column setup.

Theorem 2.10. *For all $A \in \mathcal{E}_\infty$, $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$.*

The proof of Theorem 2.10 is given in Appendix F. Now we use this result to derive Theorem 2.6 from Theorem 2.4.

Proof of Theorem 2.6. In this proof, we specify the matrices over which the U-statistics are taken, i.e. we denote $U^h_{k,l}(Y)$ instead of $U^h_{k,l}$ the U-statistic of size $k \times l$ with kernel h taken on Y. We denote also $\mathcal{F}_N(Y) = \sigma\big((U^h_{kl}(Y), k \geq m_N, l \geq n_N)\big)$ which are sets of events depending on Y.

Since Y is RCE and dissociated, Proposition 3.3 of Aldous (1981) allows us to consider a real function f such that for $1 \leq i, j < \infty$, $Y^*_{ij} = f(\xi_i, \eta_j, \zeta_{ij})$ and $Y^* \overset{\mathcal{D}}{=} Y$, where ξ_i, η_j and ζ_{ij}, for $1 \leq i, j < \infty$ are i.i.d. random variables with uniform distribution on $[0, 1]$. Therefore we can consider these random variables, the product spaces $(\Omega_N, \mathcal{A}_N, \mathbb{P}_N)$ and the sets $\mathcal{E}_N$ of invariant events defined earlier.

But $\mathcal{F}_N(Y^*) = \sigma\big((U^h_{kl}(Y^*), k \geq m_N, l \geq n_N)\big) \subset \sigma(U_N(Y^*), \xi_i, \eta_j, \zeta_{ij}, i > m_N, j > n_N)$, so for all N, $\mathcal{F}_N(Y^*) \subset \mathcal{E}_N$. It follows that $\mathcal{F}_\infty(Y^*) \subset \mathcal{E}_\infty$, so $U_\infty(Y^*)$ is $\mathcal{F}_\infty(Y^*)$-measurable. Theorem 2.10 states that all the events in $\mathcal{E}_\infty$ happen with probability 0 or 1, so it ensures that $U_\infty(Y^*) = \mathbb{E}[h(Y^*_{\{1,2;1,2\}})|\mathcal{F}_\infty(Y^*)] = \mathbb{E}[h(Y^*_{\{1,2;1,2\}})]$ is constant. Moreover, since the distribution of $U^h_N(Y)$ is the same as this of $U^h_N(Y^*)$, we can conclude that $U_\infty(Y) = \mathbb{E}[h(Y_{\{1,2;1,2\}})|\mathcal{F}_\infty(Y)] = \mathbb{E}[h(Y_{\{1,2;1,2\}})]$.

Likewise, we deduce that $\mathbb{E}[h(Y_{\{1,2;1,2\}})h(Y_{\{1,3;3,4\}})|\mathcal{F}_\infty(Y)] = \mathbb{E}[h(Y_{\{1,2;1,2\}})h(Y_{\{1,3;3,4\}})]$ and $\mathbb{E}[h(Y_{\{1,2;1,2\}})h(Y_{\{3,4;1,3\}})|\mathcal{F}_\infty(Y)] = \mathbb{E}[h(Y_{\{1,2;1,2\}})h(Y_{\{3,4;1,3\}})]$ which gives the desired result for V. Thus we conclude that W of Theorem 2.4 follows a Gaussian distribution of variance V. $\square$

3. Applications

In this section, we illustrate how to build statistical tests on RCE networks using our result. Indeed, U-statistics can be used to build unbiased estimators. The advantage of taking quadruplets is to define functions over several interactions of the same row or column. This allows us to extract information on the row and column distribution. Theorem 2.6 then guarantees an asymptotic normality result, where the only unknown is V, which has to be estimated then plugged in with Slutsky's Theorem.

Now through different examples, we will show how one might use different kernels to estimate all the needed quantities to design tests on the Version 1 of the WBEDD (with constant density), to which Theorem 2.6 applies.

3.1. Heterogeneity in the row degrees of a network

Remember that the function f (resp. g) of the WBEDD model defines the expected weight distribution of the row (resp. column) nodes. For all $k > 0$, we denote $F_k = \int_0^1 f^k(u)du$ (resp. $G_k = \int_0^1 g^k(v)dv$). Consider that we are interested in the distribution of the row degrees only. We know that $F_1 = 1$, but we see that $F_2 = \int_0^1 f^2(u)du$ quantifies the heterogeneity in the row degrees. Indeed, if f is constant, i.e. $f \equiv 1$ and $F_2 = 1$, then the row degrees are homogeneous. Besides, the higher F_2, the more unbalanced their distribution. More specifically, a large value of F_2 indicates a strong distinction between generalist (with high degree) and specialists (with low degree) nodes. Then in order to evaluate the homogeneity of the rows of a network, it makes sense to test the following hypotheses : $\mathcal{H}_0 : f \equiv 1$ vs. $\mathcal{H}_1 : f \not\equiv 1$ using an estimator of F_2.

F_2 can be estimated with the U-statistic based on the quadruplet kernel $h_1(Y_{[i_1,i_2;j_1,j_2]}) = \frac{1}{2}(Y_{i_1j_1}Y_{i_1j_2} + Y_{i_2j_1}Y_{i_2j_2})$. We see that $\mathbb{E}[U_N^{h_1}] = \mathbb{E}[h_1(Y_{\{i_1,i_2;j_1,j_2\}})] = \lambda^2 F_2$. So Theorem 2.6 and the derivation of V gives the following result :

$$\sqrt{\frac{N}{V}}(U_N^{h_1} - \lambda^2 F_2) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0,1),$$

where $V = \lambda^4 c^{-1}(F_4 - F_2^2) + 4\lambda^4(1-c)^{-1}F_2^2(G_2 - 1)$.

We use the kernel $h_2(Y_{\{i_1,i_2;j_1,j_2\}}) = \frac{1}{4}(Y_{i_1j_1} + Y_{i_1j_2} + Y_{i_2j_1} + Y_{i_2j_2})$ to construct $U_N^{h_2}$, a consistent estimator of λ. It follows from Slutsky's theorem that

$$\sqrt{\frac{N}{V}}(U_N^{h_2})^2\left(\frac{U_N^{h_1}}{(U_N^{h_2})^2} - F_2\right) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0,1). \tag{1}$$

Under $\mathcal{H}_0$, $F_2 = F_4 = 1$, so $V = 4\lambda^4(1-c)^{-1}(G_2 - 1)$. Then, to estimate V, we consider the U-statistic based on the kernel $h_3(Y_{[i_1,i_2;j_1,j_2]}) = \frac{1}{2}(Y_{i_1j_1}Y_{i_2j_1} + Y_{i_1j_2}Y_{i_2j_2})$, which is a consistent estimator for $\lambda^2 G_2$. Thus, V can be consistently estimated by

$$\hat{V}_N = \frac{4}{1-c}(U_N^{h_2})^4\left[\frac{U_N^{h_3}}{(U_N^{h_2})^2} - 1\right].$$

Finally, a further application of Slutsky's theorem implies that under $\mathcal{H}_0$,

$$\sqrt{\frac{N}{\hat{V}_N}}(U_N^{h_2})^2\left(\frac{U_N^{h_1}}{(U_N^{h_2})^2} - 1\right) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0,1).$$

This result allows us to define an asymptotic test for $\mathcal{H}_0 : F_2 = 1$.

3.2. Network comparison

The previous example shows how to build a test for a single network. In fact, it is easy to extend this framework to network comparison, provided the networks are independent. Indeed, say network Y^A and Y^B are independent. Then for any quadruplet kernel h, the U-statistics $U_N^h(Y^A)$ and $U_N^h(Y^B)$ computed on each network are also independent. Therefore, if we were to compare the row degree unbalance of two networks, we can opt for a test of the type $\mathcal{H}_0 : F_2^A = F_2^B$ vs. $\mathcal{H}_1 : F_2^A \neq F_2^B$. One can simply notice that $U_N^h(Y^A) - U_N^h(Y^B)$ is still asymptotically normal, with $\mathbb{E}[U_N^h(Y^A) - U_N^h(Y^B)] = F_2^A - F_2^B$ and it is easy to find the asymptotic variance V as $\mathbb{V}[U_N^h(Y^A) - U_N^h(Y^B)] = \mathbb{V}[U_N^h(Y^A)] + \mathbb{V}[U_N^h(Y^B)]$.

3.3. Further remarks and leads

We have showcased an example of application of our result to statistical test design. One interesting feature of the kernels used is that they are simple to compute. Indeed, if we denote $Y_N := (Y_{ij})_{1 \leq i \leq m_N, 1 \leq j \leq n_N}$, one can write the U-statistics used in the previous example as

$$\begin{aligned}
U_N^{h_1} &= \frac{1}{n_N m_N (m_N - 1)} \left[|Y_N^T Y_N|_1 - \mathrm{Tr}(Y_N^T Y_N) \right], \\
U_N^{h_2} &= \frac{1}{n_N m_N} |Y_N|_1, \\
U_N^{h_3} &= \frac{1}{n_N (n_N - 1) m_N} \left[|Y_N Y_N^T|_1 - \mathrm{Tr}(Y_N Y_N^T) \right],
\end{aligned}$$

where Tr is the trace operator. We see that these U-statistics can be computed using only simple operations on matrices, which are optimized in most computing software.

However, one can define more elaborate kernels to test further hypotheses on other models. The only conditions on the model are that it should be RCE and dissociated, i.e. it can be written as a bipartite W-graph model. For example, given the W-graph model $Y_{ij} \mid \xi_i, \eta_j \sim \mathcal{P}(\lambda w(\xi_i, \eta_j))$ with $\int \int w = 1$, one could have tested if it is of product form, i.e. if $f(u) = \int w(u,v) dv$ and $g(v) = \int w(u,v) du$, w can be written as $w(u,v) = f(u) g(v)$ (as in the WBEDD model). An appropriated kernel for this test would be

$$\begin{aligned}
h(Y_{\{i_1,i_2;j_1,j_2\}}) =& \frac{1}{4} Y_{i_1 j_1} Y_{i_2 j_2} (Y_{i_1 j_1} + Y_{i_2 j_2} - Y_{i_1 j_2} - Y_{i_1 j_2} - 2) \\
&+ \frac{1}{4} Y_{i_1 j_2} Y_{i_1 j_2} (Y_{i_1 j_2} + Y_{i_1 j_2} - Y_{i_1 j_1} - Y_{i_2 j_2} - 2)
\end{aligned}$$

as $\mathbb{E}[h(Y_{\{i_1,i_2;j_1,j_2\}})] = \int \int w(u,v)(w(u,v) - f(u)g(v)) du dv$ and should be equal to 0 if the hypothesis is true.

The counts of bipartite motifs of size 2×2 can be expressed as quadruplet U-statistics and can be integrated in our framework. If Y is a binary matrix, then one can count the diagonal motifs using a kernel and obtain statistical guarantees. For example, motif 5 of Figure 7 in Ouadah, Latouche and Robin (2021) can be counted with the kernel

$$\begin{aligned}
h(Y_{\{i_1,i_2;j_1,j_2\}}) =& Y_{i_1 j_1} Y_{i_1 j_2} Y_{i_2 j_1} (1 - Y_{i_2 j_2}) + Y_{i_1 j_1} Y_{i_1 j_2} Y_{i_2 j_2} (1 - Y_{i_2 j_1}) \\
&+ Y_{i_1 j_1} Y_{i_2 j_1} Y_{i_2 j_2} (1 - Y_{i_1 j_2}) + Y_{i_1 j_2} Y_{i_2 j_1} Y_{i_2 j_2} (1 - Y_{i_1 j_1}).
\end{aligned}$$

It is legitimate to wonder if one can extend our framework to U-statistics over submatrices of size different from 2×2, for example $Y_{\{i_1,\dots,i_p;j_1,\dots,j_q\}}$ of size $p \times q$. If this can be done, then our framework can be used to count motifs of larger size. Also, one could have used formula (1) of the row heterogeneity example to derive an asymptotic confidence interval for F_2 (we do not necessarily have $F_2 = F_4 = 1$). Instead of using quadruplet kernels, we notice that one could have estimated the term $\lambda^4 F_4$ appearing in V with a kernel over submatrices of size 1×4 such as $h(Y_{\{i_1;j_1,j_2,j_3,j_4\}}) = Y_{i_1j_1}Y_{i_1j_2}Y_{i_1j_3}Y_{i_1j_4}$ and $\mathbb{E}[h(Y_{\{i_1;j_1,j_2,j_3,j_4\}})] = \lambda^4 F_4$. The possibility for an extension is discussed in the next section.

4. Discussion

We do not claim that the chosen kernels and the derived U-statistics necessarily lead to the most powerful tests. We have seen that one might combine several U-statistics to find a consistent estimator for V. Especially, this might make the convergence of $\hat{V}_N$ slow, especially when these U-statistics are correlated and there might exist more optimal kernels to build this test. In conclusion, these U-statistics based on quadruplets might not be theoretically the most efficient estimators, but more importantly, they are simple and easy to compute in practice.

It is possible to extend our theorem to U-statistics over submatrices of size different from 2×2, for example $Y_{\{i_1,\dots,i_p;j_1,\dots,j_q\}}$ of size $p \times q$. In this case, for some kernel h on these submatrices,

$$U_N^h = \left[\binom{m_N}{p}\binom{n_N}{q}\right]^{-1} \sum_{1\le i_1<\dots<i_p\le m_N} \sum_{1\le j_1<\dots<j_q\le n_N} h(Y_{\{i_1,\dots,i_p;j_1,\dots,j_q\}}),$$

would also be asymptotically normal. All the steps of our proof can be adapted to U-statistics of larger subgraphs. These U-statistics are indeed backward martingales and the equivalent of Proposition D.1 and Proposition E.1 require more calculus. As a consequence, the asymptotic variance also has a different expression. On the one hand, such an extension would allow more flexibility in the choice of the kernel, hence the ability to build more complex estimators. On the other hand, in practice, the computation of such U-statistics may also be more complex and computationally demanding, whereas simple functions on quadruplets can easily be expressed with matrix operations.

Further studies might be carried to investigate the rate of convergence of $\sqrt{N}(U_N^h - U_\infty^h)$ to its limiting distribution. A possible direction is the derivation of a Berry-Esseen-type bound. For specific applications, the computation of this rate through numerical simulation is also possible.

Appendix A: Properties of m_N and n_N

In this appendix, we provide the proofs for Proposition 2.2 and further properties of the sequences m_N and n_N defined as $m_N = 2 + \lfloor c(N+1) \rfloor$ and $n_N = 2 + \lfloor (1-c)(N+1) \rfloor$ for all $N \ge 1$, where c is an irrational number (Definition 2.1).

Proof of Proposition 2.2. The second result stems from the fact that

$$m_N + n_N = 4 + \lfloor c(N+1) \rfloor + \lfloor (1-c)(N+1) \rfloor = 4 + \lfloor c(N+1) \rfloor + \lfloor -c(N+1) \rfloor + N + 1$$

and $\lfloor c(N+1) \rfloor + \lfloor -c(N+1) \rfloor = -1$ because $c(N+1)$ is not an integer since c is irrational. Then, the first result simply follows as

$$\frac{m_N}{m_N + n_N} = \frac{\lfloor c(N+1) \rfloor + 2}{N+4} \underset{N}{\sim} \frac{c(N+1)+2}{N+4} \underset{N}{\sim} \frac{cN}{N},$$

where $\underset{N}{\sim}$ denotes the asymptotic equivalence when N grows to infinity, i.e. $a_N \underset{N}{\sim} b_N$ if and only if $a_N/b_N \underset{N\to\infty}{\longrightarrow} 1$. □

Proof of Corollary 2.3. As m_N and n_N are non decreasing, the corollary is a direct consequence of $m_N + n_N = 4 + N$, because then $m_{N+1} + n_{N+1} = 4 + N + 1 = m_N + n_N + 1$. □

***Definition* A.1.** We define $\mathcal{B}_c$ and $\mathcal{B}_{1-c}$ two complementary subsets of $\mathbb{N}^*$ as

$$\mathcal{B}_c = \{N \in \mathbb{N}^* : m_N = m_{N-1} + 1\} \text{ and } \mathcal{B}_{1-c} = \{N \in \mathbb{N}^* : n_N = n_{N-1} + 1\}.$$

Proposition A.2. *Set $\kappa_c(m) := \left\lfloor \frac{m-2}{c} \right\rfloor$ and $\kappa_{1-c}(n) := \left\lfloor \frac{n-2}{1-c} \right\rfloor$. If $N \in \mathcal{B}_c$, then $N = \kappa_c(m_N)$. Similarly, if $N \in \mathcal{B}_{1-c}$, then $N = \kappa_{1-c}(n_N)$.*

Proof. Remember that c is an irrational number, so if $N \in \mathcal{B}_c$, then

$$cN + 2 < \lfloor cN \rfloor + 3 = m_{N-1} + 1 = m_N = \lfloor c(N+1) \rfloor + 2 < c(N+1) + 2,$$

which means that $\frac{m_N - 2}{c} - 1 < N < \frac{m_N - 2}{c}$, thus $N = \left\lfloor \frac{m_N - 2}{c} \right\rfloor$. □

Appendix B: Backward martingales

In this appendix, we recall the definition of backward martingales and their convergence theorem.

***Definition* B.1.** Let $\mathcal{F} = (\mathcal{F}_n)_{n\geq 1}$ be a decreasing filtration and $M = (M_n)_{n\geq 1}$ a sequence of integrable random variables adapted to $\mathcal{F}$. $(M_n, \mathcal{F}_n)_{n\geq 1}$ is a backward martingale if and only if for all $n \geq 1$, $\mathbb{E}[M_n | \mathcal{F}_{n+1}] = M_{n+1}$.

Theorem B.2. *Let $(M_n, \mathcal{F}_n)_{n\geq 1}$ be a backward martingale. Then, $(M_n)_{n\geq 1}$ is uniformly integrable, and, denoting $M_\infty = \mathbb{E}[M_1 | \mathcal{F}_\infty]$ where $\mathcal{F}_\infty = \bigcap_{n=1}^{\infty} \mathcal{F}_n$, we have*

$$M_n \xrightarrow[n\to\infty]{a.s., L_1} M_\infty.$$

Furthermore, if $(M_n)_{n\geq 1}$ is square-integrable, then $M_n \xrightarrow[n\to\infty]{L_2} M_\infty$.

Appendix C: Square-integrable backward martingale

In this appendix, we prove Proposition C.1, which states that U_N^h is a square-integrable backward martingale.

Proposition C.1. *Let Y be a RCE matrix. Let h be a quadruplet kernel such that $\mathbb{E}[h(Y_{\{1,2;1,2\}})^2] < \infty$. Let $\mathcal{F}_N = \sigma\big((U_{kl}^h, k \geq m_N, l \geq n_N)\big)$ and $\mathcal{F}_\infty = \bigcap_{N=1}^{\infty} \mathcal{F}_N$. Set $U_\infty^h := \mathbb{E}[h(Y_{\{1,2;1,2\}}) | \mathcal{F}_\infty]$. Then $(U_N^h, \mathcal{F}_N)_{N\geq 1}$ is a square-integrable backward martingale and $U_N^h \xrightarrow[N\to\infty]{a.s., L_2} U_\infty^h = \mathbb{E}[h(Y_{\{1,2;1,2\}}) | \mathcal{F}_\infty]$.*

The proof relies on the following lemma.

Lemma C.2. *For all $1 \leq i_1 < i_2 \leq m_N$ and $1 \leq j_1 < j_2 \leq n_N$, $\mathbb{E}[h(Y_{\{i_1,i_2;j_1,j_2\}})|\mathcal{F}_N] = \mathbb{E}[h(Y_{\{1,2;1,2\}})|\mathcal{F}_N]$.*

Proof. In the proof of this lemma, we specify the matrices over which the U-statistics are taken, i.e. we denote $U^h_{k,l}(Y)$ instead of $U^h_{k,l}$ the U-statistic of kernel h and of size $k \times l$ computed on Y.

By construction, for all $k \geq m_N, l \geq n_N$, for all matrix permutation $\Phi \in \mathbb{S}_{m_N} \times \mathbb{S}_{n_N}$ (only changing the first m_N rows and n_N columns), we have $U^h_{k,l}(\Phi Y) = U^h_{k,l}(Y)$. Moreover, since Y is RCE, we also have $\Phi Y \overset{\mathcal{D}}{=} Y$. Therefore,

$$\Phi Y|(U^h_{k,l}(Y), k \geq m_N, l \geq n_N) \overset{\mathcal{D}}{=} Y|(U^h_{k,l}(Y), k \geq m_N, l \geq n_N).$$

That means that conditionally on $\mathcal{F}_N$, the first m_N rows and n_N columns of Y are exchangeable and the result to prove follows from this. □

Proof of Proposition C.1. First, we remark that as $\mathbb{E}[h(Y_{\{1,2;1,2\}})^2] < \infty$, then for all N, $\mathbb{E}[(U^h_N)^2] < \infty$. Thus, the $(U^h_N)_{N\geq 1}$ are square-integrable. Second, $\mathcal{F} = (\mathcal{F}_N)_{N\geq 1}$ is a decreasing filtration and for all N, U^h_N is $\mathcal{F}_N$-measurable.

Now using lemma C.2, we have for all $K \leq N$,

$$\begin{aligned}
\mathbb{E}[U^h_K|\mathcal{F}_N] &= \binom{m_K}{2}^{-2}\binom{n_K}{2}^{-2} \sum_{\substack{1\leq i_1<i_2\leq m_K \\ 1\leq j_1<j_2\leq n_K}} \mathbb{E}[h(Y_{\{i_1,i_2;j_1,j_2\}})|\mathcal{F}_N] \\
&= \binom{m_K}{2}^{-2}\binom{n_K}{2}^{-2} \sum_{\substack{1\leq i_1<i_2\leq m_K \\ 1\leq j_1<j_2\leq n_K}} \mathbb{E}[h(Y_{\{1,2;1,2\}})|\mathcal{F}_N] \\
&= \mathbb{E}[h(Y_{\{1,2;1,2\}})|\mathcal{F}_N],
\end{aligned}$$

In particular, $\mathbb{E}[U^h_{N-1}|\mathcal{F}_N] = \mathbb{E}[U^h_N|\mathcal{F}_N] = U^h_N$, which concludes the proof that $(U^h_N, \mathcal{F}_N)_{N\geq 1}$ is a square-integrable backward martingale. Finally, Theorem B.2 ensures that $U^h_N \xrightarrow[N\to\infty]{a.s., L_2} U^h_\infty$. □

Appendix D: Asymptotic variances

We prove Proposition D.1 which gives the convergence and an expression for the asymptotic variance. The proof involves some tedious calculations. Before that, we introduce some notations to make the proof of Proposition D.1 more readable.

Notation. In this appendix and in Appendix E, we denote

- $X_{[i_1,i_2;j_1,j_2]} := h(Y_{\{i_1,i_2;j_1,j_2\}})$,
- $Z_{NK} := \sqrt{N}(U_K - U_{K+1})$,
- $S_{NK} := \mathbb{E}[Z^2_{NK}|\mathcal{F}_{K+1}]$,
- $V_N := \sum_{K=N}^{\infty} S_{NK}$.

The exchangeability of Y implies that $\mathbb{E}[X_{[i_1,i_2;j_1,j_2]}X_{[i'_1,i'_2;j'_1,j'_2]}|\mathcal{F}_K]$ only depends on the numbers of rows and columns shared by both $[i_1,i_2;j_1,j_2]$ and $[i'_1,i'_2;j'_1,j'_2]$. For $0 \le p \le 2$ and $0 \le q \le 2$, we set

$$c_K^{(p,q)} := \mathbb{E}[X_{[i_1,i_2;j_1,j_2]}X_{[i'_1,i'_2;j'_1,j'_2]}|\mathcal{F}_K],$$

and

$$c_\infty^{(p,q)} := \mathbb{E}[X_{[i_1,i_2;j_1,j_2]}X_{[i'_1,i'_2;j'_1,j'_2]}|\mathcal{F}_\infty],$$

where they share p rows and q columns.

Proposition D.1. $V_N \xrightarrow[N\to\infty]{\mathbb{P}} V = 4c^{-1}(c_\infty^{(1,0)} - U_\infty^2) + 4(1-c)^{-1}(c_\infty^{(0,1)} - U_\infty^2)$.

The proof of Proposition D.1 will be based on the following five lemmas.

Lemma D.2. *If $K \in \mathcal{B}_c$, then*

$$Z_{N,K-1} = \sqrt{N}\frac{2}{m_K - 2}(U_K - \delta_K),$$

where

$$\delta_K = (m_K - 1)^{-1}\binom{n_K}{2}^{-1} \sum_{\substack{1\le i_1\le m_K-1\\ 1\le j_1<j_2\le n_K}} X_{[i_1,m_K;j_1,j_2]}.$$

Proof. Observe that

$$\sum_{\substack{1\le i_1<i_2\le m_K\\ 1\le j_1<j_2\le n_K}} X_{[i_1,i_2;j_1,j_2]} = \sum_{\substack{1\le i_1<i_2\le m_K-1\\ 1\le j_1<j_2\le n_K}} X_{[i_1,i_2;j_1,j_2]} + \sum_{\substack{1\le i_1\le m_K-1\\ 1\le j_1<j_2\le n_K}} X_{[i_1,m_K;j_1,j_2]}. \tag{2}$$

But if $K \in \mathcal{B}_c$ (see definition A.1), then $m_{K-1} = m_K - 1$ and $n_{K-1} = n_K$. Therefore, equation (2) is equivalent to

$$\binom{m_K}{2}\binom{n_K}{2}U_K = \binom{m_K-1}{2}\binom{n_K}{2}U_{K-1} + (m_K-1)\binom{n_K}{2}\delta_K,$$

so

$$U_{K-1} = \frac{1}{m_K - 2}\left(m_K U_K - 2\delta_K\right).$$

This concludes the proof since $Z_{N,K-1} = \sqrt{N}(U_{K-1} - U_K)$. □

We now calculate S_{NK} in the following lemmas.

Lemma D.3. *For all $0 \le p \le 2$ and $0 \le q \le 2$, $c_N^{(p,q)} \xrightarrow[N\to\infty]{a.s., L_1} c_\infty^{(p,q)}$.*

Proof. This follows from the fact that $(c_N^{(p,q)}, \mathcal{F}_N)_{N\ge1}$ is a backward martingale. □

Lemma D.4. *If $K \in \mathcal{B}_c$, then*

$$S_{N,K-1} = 4N\bigg(\frac{(n_K-2)(n_K-3)}{(m_K-1)(m_K-2)n_K(n_K-1)}c_K^{(1,0)} - \frac{1}{(m_K-2)^2}U_K^2 + \psi(K)\bigg),$$

where ψ does not depend on N and $\psi(K) = o(m_K^{-2})$.

Proof. Because of Lemma D.2 and the $\mathcal{F}_K$-measurability of U_K,

$$S_{N,K-1} = \frac{4N}{(m_K-2)^2}(\mathbb{E}[\delta_K^2|\mathcal{F}_K] + U_K^2 - 2U_K\mathbb{E}[\delta_K|\mathcal{F}_K]).$$

First, Lemma C.2 implies that

$$\mathbb{E}[\delta_K|\mathcal{F}_K] = U_K.$$

Then, we can calculate

$$\mathbb{E}[\delta_K^2|\mathcal{F}_K] = (m_K-1)^{-2}\binom{n_K}{2}^{-2} \sum_{\substack{1\le i_1\le m_K-1\\ 1\le j_1<j_2\le n_K}} \sum_{\substack{1\le i_1'\le m_K-1\\ 1\le j_1'<j_2'\le n_K}} \mathbb{E}[X_{[i_1,m_K;j_1,j_2]}X_{[i_1',m_K;j_1',j_2']}|\mathcal{F}_K].$$

Each term of the sum only depends on the number of rows and columns the quadruplets in $X_{[i_1,m_K;j_1,j_2]}$ and $X_{[i_1',m_K;j_1',j_2']}$ have in common. For example, if they share p rows and q columns, it is equal to $c_K^{(p,q)}$. So by breaking down the different cases for p and q, we may count the number of possibilities. For example, if $(p,q) = (1,2)$, then the number of possibilities is $(m_K-1)(m_K-2)\binom{n_K}{2}$. This gives

$$\begin{aligned}\mathbb{E}[\delta_K^2|\mathcal{F}_K] = (m_K-1)^{-1}\binom{n_K}{2}^{-1}\bigg\{&\frac{1}{2}(m_K-2)(n_K-2)(n_K-3)c_K^{(1,0)} + 2(m_K-2)(n_K-2)c_K^{(1,1)}\\ &+ (m_K-2)c_K^{(1,2)} + \frac{1}{2}(n_K-2)(n_K-3)c_K^{(2,0)} + 2(n_K-2)c_K^{(2,1)} + c_K^{(2,2)}\bigg\}.\end{aligned}$$

Finally, setting

$$\begin{aligned}\psi(K) := (m_K-1)^{-3}\binom{n_K}{2}^{-1}\bigg\{&2(m_K-2)(n_K-2)c_K^{(1,1)} + (m_K-2)c_K^{(1,2)}\\ &+ \frac{1}{2}(n_K-2)(n_K-3)c_K^{(2,0)} + 2(n_K-2)c_K^{(2,1)} + c_K^{(2,2)}\bigg\},\end{aligned}$$

we obtain the desired result, with $\psi(K) = o(m_K^{-2})$ since $\frac{m_K}{c} \underset{K}{\sim} \frac{n_K}{1-c} \underset{K}{\sim} K$. □

Remark. In the case where $K \in \mathcal{B}_{1-c}$, the equivalent formulas to those of Lemmas D.2 and D.4 are derived from similar proofs. If $K \in \mathcal{B}_{1-c}$, then

$$Z_{N,K-1} = \sqrt{N}\frac{2}{n_K-2}(U_K - \gamma_K),$$

where

$$\gamma_K = (n_K - 1)^{-1} \binom{m_K}{2}^{-1} \sum_{\substack{1 \le i_1 < i_2 \le m_K \\ 1 \le j_1 \le n_K - 1}} X_{[i_1, i_2; j_1, n_K]},$$

and

$$S_{N,K-1} = 4N\left(\frac{(m_K-2)(m_K-3)}{(n_K-1)(n_K-2)m_K(m_K-1)} c_K^{(0,1)} - \frac{1}{(n_K-2)^2} U_K^2 + \varphi(K)\right),$$

where φ does not depend on N and $\varphi(K) = o(n_K^{-2})$.

Lemma D.5. *Let $(R_n)_{n\ge1}$ be a sequence of random variables and $(\lambda_n)_{n\ge1}$ a sequence of real positive numbers. Set $C_n := n\sum_{k=n}^{\infty}\lambda_k R_k$. If*

- *$n\sum_{k=n}^{\infty}\lambda_k \xrightarrow[n\to\infty]{} 1$, and*
- *there exists a random variable R_∞ such that $R_n \xrightarrow[n\to\infty]{a.s.} R_\infty$,*

then $C_n \xrightarrow[n\to\infty]{a.s.} R_\infty$. Furthermore, if $R_n \xrightarrow[n\to\infty]{L_1} R_\infty$, then $C_n \xrightarrow[n\to\infty]{L_1} R_\infty$.

Proof. Notice that

$$\begin{aligned}
|C_n - R_\infty| &= \Big|n\sum_{k=n}^{\infty}\lambda_k R_k - R_\infty\Big| \\
&\le \Big|n\sum_{k=n}^{\infty}\lambda_k R_k - n\sum_{k=n}^{\infty}\lambda_k R_\infty\Big| + \Big|n\sum_{k=n}^{\infty}\lambda_k R_\infty - R_\infty\Big| \\
&\le \left(n\sum_{k=n}^{\infty}\lambda_k\right) \times \sup_{k\ge n}|R_k - R_\infty| + \Big|n\sum_{k=n}^{\infty}\lambda_k - 1\Big| \times |R_\infty|.
\end{aligned}$$

If $n\sum_{k=n}^{\infty}\lambda_k \xrightarrow[n\to\infty]{} 1$ and $R_n \xrightarrow[n\to\infty]{a.s.} R_\infty$, then for all ω fixed except a set of neglectable size, $C_n(\omega) \xrightarrow[n\to\infty]{} R_\infty(\omega)$, which gives the a.s. convergence. Now, consider also that

$$\begin{aligned}
\mathbb{E}\Big[|C_n - R_\infty|\Big] &\le n\sum_{k=n}^{\infty}\lambda_k \mathbb{E}\Big[|R_k - R_\infty|\Big] + \Big|n\sum_{k=n}^{\infty}\lambda_k - 1\Big|\mathbb{E}\Big[|R_\infty|\Big] \\
&\le \left(n\sum_{k=n}^{\infty}\lambda_k\right) \times \sup_{k\ge n}\mathbb{E}\Big[|R_k - R_\infty|\Big] + \Big|n\sum_{k=n}^{\infty}\lambda_k - 1\Big|\mathbb{E}\Big[|R_\infty|\Big].
\end{aligned}$$

So if $R_n \xrightarrow[n\to\infty]{L_1} R_\infty$, then $\mathbb{E}\big[|R_n - R_\infty|\big] \xrightarrow[n\to\infty]{L_1} 0$ and $\sup_{k\ge n}\mathbb{E}\big[|R_k - R_\infty|\big] \xrightarrow[n\to\infty]{L_1} 0$. Since $n\sum_{k=n}^{\infty}\lambda_k \xrightarrow[n\to\infty]{} 1$, the first term converges to 0, and the second term too because $\mathbb{E}\big[|R_\infty|\big] < \infty$. Finally, $\mathbb{E}\big[|C_n - R_\infty|\big] \xrightarrow[n\to\infty]{} 0$. $\square$

Lemma D.6. *Let $(Q_n)_{n\geq 1}$ be a sequence of random variables. Set $C_n := n\sum_{k=n}^{\infty} Q_k$. If there exists a random variable C_∞ such that $n^2 Q_n \xrightarrow[n\to\infty]{a.s.} C_\infty$, then $C_n \xrightarrow[n\to\infty]{a.s.} C_\infty$. Furthermore, if $n^2 Q_n \xrightarrow[n\to\infty]{L_1} C_\infty$, then $C_n \xrightarrow[n\to\infty]{L_1} C_\infty$.*

Proof. This is a direct application of Lemma D.5, where $R_n := n^2 Q_n$ and $\lambda_n := n^{-2}$, as $n\sum_{k=n}^{\infty} k^{-2} \xrightarrow[n\to\infty]{}$ 1. □

Proof of Proposition D.1. Recall that from Corollary 2.3, $\mathcal{B}_c$ and $\mathcal{B}_{1-c}$ form a partition of the set of the positive integers $\mathbb{N}^*$, so that we can write

$$V_N = V_N^{(c)} + V_N^{(1-c)},$$

where $V_N^{(c)} = \sum_{\substack{K=N+1\\K\in\mathcal{B}_c}}^{\infty} S_{N,K-1}$ and $V_N^{(1-c)} = \sum_{\substack{K=N+1\\K\in\mathcal{B}_{1-c}}}^{\infty} S_{N,K-1}$. Here, we only detail the computation of $V_N^{(c)}$, as one can proceed analogously with $V_N^{(1-c)}$.

In $V_N^{(c)}$, the sum is over the $K\in\mathcal{B}_c$. So, from Lemma D.4,

$$S_{N,K-1} = 4N\left(\frac{(n_K-2)(n_K-3)}{(m_K-1)(m_K-2)n_K(n_K-1)}c_K^{(1,0)} - \frac{1}{(m_K-2)^2}U_K^2 + \psi(K)\right).$$

Now we use Proposition A.2 to replace K with $\kappa_c(m_K) = \lfloor \frac{m_K-2}{c} \rfloor$ and

$$S_{N,\kappa_c(m_K)-1} = 4N\left(\frac{(\kappa_c(m_K)-m_K+2)(\kappa_c(m_K)-m_K+1)}{(m_K-1)(m_K-2)(\kappa_c(m_K)-m_K+4)(\kappa_c(m_K)-m_K+3)}c_{\kappa_c(m_K)}^{(1,0)}\right.$$
$$\left. - \frac{1}{(m_K-2)^2}U_{\kappa_c(m_K)}^2 + \psi(\kappa_c(m_K))\right).$$

Therefore, because for all $K\in\mathcal{B}_c$ we have $m_K = m_{K-1}+1$, we can then transform the sum over K into a sum over m and

$$V_N^{(c)} = \sum_{\substack{K=N+1\\K\in\mathcal{B}_c}}^{\infty} S_{N,K-1} = \sum_{m=m_{N+1}}^{\infty} S_{N,\kappa_c(m)-1} = N\sum_{m=m_{N+1}}^{\infty} R_m,$$

where $R_m := S_{N,\kappa_c(m)-1}/N$, i.e.

$$R_m = \frac{4(\kappa_c(m)-m+2)(\kappa_c(m)-m+1)}{(m-1)(m-2)(\kappa_c(m)-m+4)(\kappa_c(m)-m+3)}c_{\kappa_c(m)}^{(1,0)} - \frac{4}{(m-2)^2}U_{\kappa_c(m)}^2 + 4\psi(\kappa_c(m)).$$

But we notice that since $\psi(\kappa_c(m)) = o(m^{-2})$, then Lemma D.3 and Proposition C.1 give for all N,

$$m^2 R_m \xrightarrow[m\to\infty]{a.s.,L_1} 4(c_\infty^{(1,0)} - U_\infty^2).$$

And since $\frac{m_{N+1}}{N} \xrightarrow[N\to\infty]{} c$ from Proposition 2.2, we find with Lemma D.6 that

$$V_N^{(c)} = \frac{N}{m_{N+1}} \times m_{N+1}\sum_{m=m_{N+1}}^{\infty} R_m \xrightarrow[N\to\infty]{a.s.,L_1} \frac{4}{c}(c_\infty^{(1,0)} - U_\infty^2).$$

We can proceed likewise with $V_N^{(1-c)}$, where all the terms have $K \in \mathcal{B}_{1-c}$, to get

$$V_N^{(1-c)} \xrightarrow[N\to\infty]{a.s.,L_1} \frac{4}{1-c}(c_\infty^{(0,1)} - U_\infty^2),$$

which finally gives

$$V_N = V_N^{(c)} + V_N^{(1-c)} \xrightarrow[N\to\infty]{a.s.,L_1} V := \frac{4}{c}(c_\infty^{(1,0)} - U_\infty^2) + \frac{4}{1-c}(c_\infty^{(0,1)} - U_\infty^2).$$

□

Appendix E: Conditional Lindeberg condition

We verify the conditional Lindeberg condition as stated by Proposition E.1. We use the notations defined in Appendix D.

Proposition E.1. *Let $\epsilon > 0$. Then the conditional Lindeberg condition is satisfied :*

$$\sum_{K=N}^{\infty} \mathbb{E}\big[Z_{NK}^2 \mathbb{1}_{\{|Z_{NK}|>\epsilon\}} \big| \mathcal{F}_{K+1}\big] \xrightarrow[N\to\infty]{\mathbb{P}} 0$$

The proof relies on the four following lemmas.

Lemma E.2. *Let $(Q_n)_{n\geq 1}$ be a sequence of random variables. Set $C_n := n\sum_{k=n}^{\infty} Q_k$. If $n^2\mathbb{E}\big[|Q_n|\big] \xrightarrow[n\to\infty]{} 0$, then $C_n \xrightarrow[n\to\infty]{\mathbb{P}} 0$.*

Proof. Lemma D.6 and the triangular inequality give $\mathbb{E}\big[|C_n|\big] \leq n\sum_{k=n}^{\infty}\mathbb{E}\big[|Q_k|\big] \xrightarrow[n\to\infty]{} 0$. Let some $\epsilon > 0$, then Markov's inequality ensures that

$$\mathbb{P}(|C_n| > \epsilon) \leq \frac{\mathbb{E}\big[|C_n|\big]}{\epsilon} \xrightarrow[n\to\infty]{} 0.$$

□

Lemma E.3. *For sequences of random variables U_n and sets B_n, if $U_n \xrightarrow[n\to\infty]{L_2} U$ and $\mathbb{1}(B_n) \xrightarrow[n\to\infty]{\mathbb{P}} 0$, then $\mathbb{E}[U_n^2\mathbb{1}(B_n)] \xrightarrow[n\to\infty]{} 0$.*

Proof. Note that for all n, $a > 0$,

$$\begin{aligned}
\mathbb{E}\big[U_n^2\mathbb{1}(B_n)\big] &= \mathbb{E}\big[U_n^2\mathbb{1}(B_n)\mathbb{1}(U_n^2 > a)\big] + \mathbb{E}\big[U_n^2\mathbb{1}(B_n)\mathbb{1}(U_n^2 \leq a)\big] \\
&\leq \mathbb{E}\big[U_n^2\mathbb{1}(U_n^2 > a)\big] + \mathbb{E}\big[a\mathbb{1}(B_n)\big] \\
&\leq \mathbb{E}\big[U_n^2\mathbb{1}(U_n^2 > a)\big] + a\mathbb{P}(B_n)
\end{aligned}$$

Let $\epsilon > 0$. $U_n \xrightarrow[n\to\infty]{L_2} U$, so $(U_n^2)_{n\geq 1}$ is uniformly integrable and there exists $a > 0$ such that $\mathbb{E}[U_n^2 \mathbb{1}(U_n^2 > a)] \leq \sup_k \mathbb{E}[U_k^2 \mathbb{1}(U_k^2 > a)] \leq \frac{\epsilon}{2}$. Moreover, $\mathbb{1}(B_n) \xrightarrow[n\to\infty]{\mathbb{P}} 0$, which translates to $\mathbb{P}(B_n) \xrightarrow[n\to\infty]{} 0$ and there exists an integer n_0 such that for all $n > n_0$, $\mathbb{P}(B_n) \leq \frac{\epsilon}{2a}$. Choosing such a real number a, we can always find an integer n_0 such that for $n > n_0$, we have $\mathbb{E}[U_n^2 \mathbb{1}(B_n)] \leq \epsilon$. ☐

Lemma E.4. *For sequences of random variables M_n and sets B_n, if $(M_n)_{n\geq 1}$ is a backward martingale with respect to some filtration and $\mathbb{1}(B_n) \xrightarrow[n\to\infty]{\mathbb{P}} 0$, then $\mathbb{E}[M_n \mathbb{1}(B_n)] \xrightarrow[n\to\infty]{} 0$.*

Proof. We notice that from Theorem B.2, $(M_n)_{n\geq 1}$ is uniformly integrable, then the proof is similar to that of Lemma E.3. ☐

Lemma E.5. *Set $A_K := m_K^{-1}\binom{n_K}{2}^{-1} \sum_{\substack{2\leq i_2\leq m_K+1 \\ 1\leq j_1<j_2\leq n_K}} X_{[1,i_2;j_1,j_2]}$. If $K \in \mathcal{B}_c$, then $A_K \stackrel{\mathcal{D}}{=} \delta_K$,where δ_K is defined in Lemma D.2.*

Proof. Remember that if $K \in \mathcal{B}_c$ (see Definition A.1), then by symmetry of h, $\delta_K = (m_K - 1)^{-1}\binom{n_K}{2}^{-1} \sum\limits_{\substack{1\leq i_2\leq m_K-1 \\ 1\leq j_1<j_2\leq n_K}} X_{[m_K,i_2;j_1,j_2]}$. The exchangeability of Y says that all permutations on the rows and the columns of Y leave its distribution unchanged, hence for all $(\sigma_1,\sigma_2) \in \mathbb{S}_{m_K} \times \mathbb{S}_{n_K}$, we have

$$\delta_K \stackrel{\mathcal{L}}{=} (m_K - 1)^{-1}\binom{n_K}{2}^{-1} \sum_{\substack{1\leq i_2\leq m_K-1 \\ 1\leq j_1<j_2\leq n_K}} X_{[\sigma_1(m_K),\sigma_1(i_2);\sigma_2(j_1),\sigma_2(j_2)]}.$$

Consider σ_2 to be the identity and $\sigma_1 \in \mathbb{S}_{m_K}$ the permutation defined by :

- $\sigma_1(i) = i+1$ if $i < m_K$,
- $\sigma_1(m_K) = 1$,
- $\sigma_1(i) = i$ if $i > m_K$.

Then $A_K = (m_K - 1)^{-1}\binom{n_K}{2}^{-1} \sum\limits_{\substack{1\leq i_2\leq m_K-1 \\ 1\leq j_1<j_2\leq n_K}} X_{[\sigma_1(m_K),\sigma_1(i_2);\sigma_2(j_1),\sigma_2(j_2)]}$, hence $A_K \stackrel{\mathcal{L}}{=} \delta_K$. ☐

Proof of Proposition E.1. Similarly to the proof of the Proposition D.1, we can verify the conditional Lindeberg condition by decomposing the sum along with $K+1 \in \mathcal{B}_c$ and $K+1 \in \mathcal{B}_{1-c}$ (Corollary 2.3), so here we only consider $\sum_{\substack{K=N+1 \\ K\in\mathcal{B}_c}}^{\infty} \mathbb{E}\left[Z_{N,K-1}^2 \mathbb{1}_{\{|Z_{N,K-1}|>\epsilon\}} | \mathcal{F}_K\right]$.

Like previously, using Proposition A.2, we can transform the sum over K into a sum over m :

$$\sum_{\substack{K=N+1 \\ K\in\mathcal{B}_c}}^{\infty} \mathbb{E}\left[Z_{N,K-1}^2 \mathbb{1}_{\{|Z_{N,K-1}|>\epsilon\}} | \mathcal{F}_K\right] = \sum_{m=m_{N+1}}^{\infty} \mathbb{E}\left[Z_{N,\kappa_c(m)-1}^2 \mathbb{1}_{\{|Z_{N,\kappa_c(m)-1}|>\epsilon\}} | \mathcal{F}_{\kappa_c(m)}\right],$$

where $\kappa_c(m) = \lfloor \frac{m-2}{c} \rfloor$.

We remark that for $m \geq m_{N+1} = m_N + 1 > c(N+1) + 2$,

$$\begin{aligned}
\mathbb{1}_{\{|Z_{N,\kappa_c(m)-1}|>\epsilon\}} &\leq \mathbb{1}_{\left\{\frac{2\sqrt{N}}{m-2}|U_{\kappa_c(m)}-\delta_{\kappa_c(m)}|>\epsilon\right\}} \\
&\leq \mathbb{1}_{\left\{|U_{\kappa_c(m)}-\delta_{\kappa_c(m)}|>\frac{m-2}{2\sqrt{\frac{m-2}{c}}}\epsilon\right\}} \\
&\leq \mathbb{1}_{\left\{|U_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}} + \mathbb{1}_{\left\{|\delta_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}}.
\end{aligned}$$

So, using the identity $(U_{\kappa_c(m)} - \delta_{\kappa_c(m)})^2 \leq 2(U^2_{\kappa_c(m)} + \delta^2_{\kappa_c(m)})$, we get for $m \geq m_{N+1}$,

$$\begin{aligned}
&\mathbb{E}\big[Z^2_{N,\kappa_c(m)-1}\mathbb{1}_{\{|Z_{N,\kappa_c(m)-1}|>\epsilon\}}|\mathcal{F}_{\kappa_c(m)}\big] \\
&\leq \frac{8N}{(m-2)^2}\mathbb{E}\left[(U^2_{\kappa_c(m)} + \delta^2_{\kappa_c(m)})\left(\mathbb{1}_{\left\{|U_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}} + \mathbb{1}_{\left\{|\delta_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}}\right)\middle|\mathcal{F}_{\kappa_c(m)}\right].
\end{aligned}$$

This inequality and Lemma E.2 imply that a sufficient condition to have the conditional Lindeberg condition is

$$\mathbb{E}\left[(U^2_{\kappa_c(m)} + \delta^2_{\kappa_c(m)})\left(\mathbb{1}_{\left\{|U_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}} + \mathbb{1}_{\left\{|\delta_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}}\right)\right] \xrightarrow[m\to\infty]{} 0. \tag{3}$$

Next, we prove that this condition is satisfied.

First, note that

$$\mathbb{P}\left(|U_{\kappa_c(m)}| > \frac{\sqrt{c(m-2)}}{4}\epsilon\right) \leq \frac{4\mathbb{E}[|U_{\kappa_c(m)}|]}{\epsilon\sqrt{c(m-2)}} \xrightarrow[m\to\infty]{} 0$$

and

$$\mathbb{P}\left(|\delta_{\kappa_c(m)}| > \frac{\sqrt{c(m-2)}}{4}\epsilon\right) \leq \frac{4\mathbb{E}[|\delta_{\kappa_c(m)}|]}{\epsilon\sqrt{c(m-2)}} \xrightarrow[m\to\infty]{} 0.$$

Now, remember that from Proposition C.1, $U_K \xrightarrow[K\to\infty]{L_2} U_\infty$, therefore $U_{\kappa_c(m)} \xrightarrow[m\to\infty]{L_2} U_\infty$ and Lemma E.3 can be applied, which gives

$$\mathbb{E}\left[U^2_{\kappa_c(m)}\left(\mathbb{1}_{\left\{|U_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}} + \mathbb{1}_{\left\{|\delta_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}}\right)\right] \xrightarrow[m\to\infty]{} 0. \tag{4}$$

Likewise, we calculated $\mathbb{E}[\delta^2_K|\mathcal{F}_K]$ in the proof of Lemma D.4. The application of Lemma D.3 shows that $\mathbb{E}[\delta^2_{\kappa_c(m)}|\mathcal{F}_{\kappa_c(m)}]$ is a backward martingale. It follows from Lemma E.4 that

$$\mathbb{E}\left[\delta^2_{\kappa_c(m)}\mathbb{1}_{\left\{|U_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}}\right] = \mathbb{E}\left[\mathbb{E}[\delta^2_{\kappa_c(m)}|\mathcal{F}_{\kappa_c(m)}]\mathbb{1}_{\left\{|U_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}}\right] \xrightarrow[m\to\infty]{} 0. \tag{5}$$

Finally, applying Lemma E.5, we obtain

$$\mathbb{E}\left[\delta^2_{\kappa_c(m)}\mathbb{1}_{\left\{|\delta_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}}\right] = \mathbb{E}\left[A^2_{\kappa_c(m)}\mathbb{1}_{\left\{|A_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}}\right], \tag{6}$$

where $A_K = m_K^{-1}\binom{n_K}{2}^{-1}\sum_{\substack{2\le i_2\le m_K+1\\ 1\le j_1<j_2\le n_K}} X_{[1,i_2;j_1,j_2]}$. Using similar arguments as in the proof of Proposition C.1, it can be shown that A_K is a square integrable backward martingale with respect to the decreasing filtration $\mathcal{F}_K^A = \sigma(A_K, A_{K+1}, \ldots)$. Therefore, Theorem B.2 ensures that there exists A_∞ such that $A_K \xrightarrow[K\to\infty]{L_2} A_\infty$. This proves that $A_{\kappa_c(m)} \xrightarrow[m\to\infty]{L_2} A_\infty$, so applying Lemma E.3 again, we obtain

$$\mathbb{E}\left[A^2_{\kappa_c(m)}\mathbb{1}_{\left\{|A_{\kappa_c(m)}|>\frac{\sqrt{c(m-2)}}{4}\epsilon\right\}}\right] \xrightarrow[m\to\infty]{} 0. \tag{7}$$

Combining (4), (5), (6) and (7), we deduce that the sufficient condition (3) is satisfied, thus concluding the proof.

□

Appendix F: Hewitt-Savage theorem

Proof of Theorem 2.10. This proof adapts the steps taken by Feller (1971) and detailed by Durrett (2019) to our case. Let $A \in \mathcal{E}_\infty$.

First, let $\mathcal{A}_N = \sigma\big((\xi_i)_{1\le i\le m_N}, (\eta_j)_{1\le j\le n_N}, (\zeta_{ij})_{1\le i\le m_N, 1\le j\le n_N}\big)$, the σ-field generated by the random variables associated with the first m_N rows and n_N columns. Notice that $A \in \mathcal{A} := \bigcap_{n=1}^\infty \mathcal{A}_N$. Since $\mathcal{A}$ is the limit of $\mathcal{A}_N$, then for all $\epsilon > 0$, there exists a N and an associated set $A_N \in \mathcal{A}_N$ such that $\mathbb{P}(A - A\cap A_N) < \epsilon$ and $\mathbb{P}(A_N - A\cap A_N) < \epsilon$, so that $\mathbb{P}(A\Delta A_N) < 2\epsilon$, where Δ is the symmetric difference operator, i.e. $B\Delta C = (B-C)\cup(C-B)$. Therefore, we can pick a sequence of sets A_N such that $\mathbb{P}(A\Delta A_N) \longrightarrow 0$.

Next, we consider the row-column permutation $\Phi^{(N)} = (\sigma_1^{(N)}, \sigma_2^{(N)}) \in \mathbb{S}_{m_N}\times\mathbb{S}_{n_N}$ defined by

$$\sigma_1^{(N)}(i) = \begin{cases} i + m_N & \text{if } 1 \le i \le m_N, \\ i - m_N & \text{if } m_N + 1 \le i \le 2m_N, \\ i & \text{if } 2m_N + 1 \le i. \end{cases}$$

$$\sigma_2^{(N)}(j) = \begin{cases} j + n_N & \text{if } 1 \le j \le n_N, \\ j - n_N & \text{if } n_N + 1 \le j \le 2n_N, \\ j & \text{if } 2n_N + 1 \le j. \end{cases}$$

Since $A \in \mathcal{E}_\infty$, by the definition of $\mathcal{E}_\infty$, it follows that

$$\left\{\omega : \Phi^{(N)}\omega \in A\right\} = \{\omega : \omega \in A\} = A.$$

Using this, if we denote $A'_N := \left\{\omega : \Phi^{(N)}\omega \in A_N\right\}$, then we can write that

$$\left\{\omega : \Phi^{(N)}\omega \in A_N\Delta A\right\} = \left\{\omega : \omega \in A'_N\Delta A\right\} = A'_N\Delta A.$$

Furthermore, the $(U_i)_{1\le i<\infty}$, $(V_j)_{1\le j<\infty}$ and $(L_{ij})_{1\le i<\infty, 1\le j<\infty}$ are i.i.d., so

$$\mathbb{P}(A_N\Delta A) = \mathbb{P}(\omega : \omega \in A_N\Delta A) = \mathbb{P}(\omega : \Phi^{(N)}\omega \in A_N\Delta A).$$

and we conclude that $\mathbb{P}(A'_N\Delta A) = \mathbb{P}(A_N\Delta A) \longrightarrow 0$.

From this, we derive that $\mathbb{P}(A_N) \longrightarrow \mathbb{P}(A)$ and $\mathbb{P}(A'_N) \longrightarrow \mathbb{P}(A)$. We also remark that $\mathbb{P}(A_N \Delta A'_N) \leq \mathbb{P}(A_N \Delta A) + \mathbb{P}(A'_N \Delta A) \longrightarrow 0$, so $\mathbb{P}(A_N \cap A'_N) \longrightarrow \mathbb{P}(A)$.

But A_N and A'_N are independent, so we have $\mathbb{P}(A_N \cap A'_N) = \mathbb{P}(A_N)\mathbb{P}(A'_N) \longrightarrow \mathbb{P}(A)^2$, therefore $\mathbb{P}(A) = \mathbb{P}(A)^2$, which means that $\mathbb{P}(A) = 0$ or 1. $\square$

Acknowledgements

The author thanks Stéphane Robin (INRAE), Sophie Donnet (INRAE) and François Massol (CNRS) for many fruitful discussions and insights. This work was funded by a grant from Région Île-de-France and by the grant ANR-18-CE02-0010-01 of the French National Research Agency ANR (project EcoNet).

References

ALDOUS, D. J. (1981). Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis* **11** 581–598.

BARRAT, A., BARTHELEMY, M., PASTOR-SATORRAS, R. and VESPIGNANI, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences* **101** 3747–3752.

CHUNG, F. and LU, L. (2002). The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences* **99** 15879–15882.

DURRETT, R. (2019). *Probability: theory and examples* **49**. Cambridge university press.

EAGLESON, G. K. and WEBER, N. C. (1978). Limit theorems for weakly exchangeable arrays. In *Mathematical Proceedings of the Cambridge Philosophical Society* **84** 123–130. Cambridge University Press.

FELLER, W. (1971). *An Introduction to Probability theory and its application Vol II*. John Wiley and Sons.

HOEFFDING, W. (1948). A Class of Statistics with Asymptotically Normal Distribution. *The Annals of Mathematical Statistics* 293–325.

NANDI, H. and SEN, P. (1963). On the properties of U-statistics when the observations are not independent: Part two unbiased estimation of the parameters of a finite population. *Calcutta Statistical Association Bulletin* **12** 124–148.

OUADAH, S., LATOUCHE, P. and ROBIN, S. (2021). Motif-based tests for bipartite networks. *arXiv preprint arXiv:2101.11381*.