

ÉCOLE NORMALE SUPÉRIEURE

MÉMOIRE DE PREMIÈRE ANNÉE

---

Graphes aléatoires, graphons et tableaux  
échangeables

---

Nikola Sekulov, Louis Vanhalewyn

Encadré par Bastien Mallein

Juin 2023

# Table des matières

1	Introduction . . . . .	2
1.1	Bases de la théorie des graphes . . . . .	2
1.2	Les graphes aléatoires . . . . .	4
2	Echangeabilité . . . . .	8
2.1	Suites échangeables et le théorème de De Finetti . . . . .	8
2.2	Preliminaires . . . . .	9
2.3	Preuve du théorème de De Finetti . . . . .	11
2.4	Tableaux échangeables . . . . .	13
3	Une topologie pour les graphes infinis . . . . .	15
3.1	Convergence de grands graphes déterministes . . . . .	15
3.2	Convergence de graphes aléatoires . . . . .	18
4	Graphes aléatoires échangeables . . . . .	21
4.1	Représentations des limites de graphes et des graphes échangeables . .	25

# 1 Introduction

Le modèle du graphon (contraction de « *graph* » et de « *function* ») est un modèle permettant d'étudier le comportement typiques de limites de grands graphes. Il fut introduit pour la première fois en 2007 par Borgs, Chayes, Lovasz, Sos et Vesztergombi dans leur article [2], qui fait partie d'une série d'articles sur la convergence de grands graphes.

L'objectif de notre mémoire est de présenter la construction de ce modèle, en suivant la ligne directrice de l'article de Diaconis et Janson [4].

Dans un premier temps, on introduira des bases de théorie des graphes, nécessaires à la construction du graphon et d'une topologie sur les grands graphes, en particulier la notion de morphisme de graphes. Ces premières définitions nous permettront d'introduire le premier modèle historique de graphes aléatoires et d'en présenter rapidement les limites, qui justifient le développement d'une nouvelle théorie.

Dans un second temps, on présentera le concept d'échangeabilité, dont le théorème de De Finetti constitue un développement important. Une extension de ce théorème, le théorème d'Aldous-Hoover, sera appliquée par la suite sur les graphes aléatoires.

Puis, on développera une topologie pour les graphes infinis, qui nous permettra d'obtenir des caractérisations de convergence de suites de grands graphes, tout d'abord déterministes, puis aléatoires.

Enfin, on établira un théorème de représentations de graphes aléatoires à l'aide de graphons, que l'on démontrera grâce aux résultats d'Aldous-Hoover et à la topologie précédemment définie.

## 1.1 Bases de la théorie des graphes

**Définition 1.** Un **graphe**  $G$  est la donnée d'un ensemble  $V$  de **sommets**, et d'un ensemble  $E$  d'**arêtes** constitué de parties de  $V$  à deux éléments. On appelle la **taille** d'un graphe le cardinal de l'ensemble des sommets, noté  $v(G)$ .

Par la suite on ne s'intéressera qu'aux graphes dont aucun sommet n'est relié à lui-même par une boucle. On représente habituellement un graphe de la manière suivante : les sommets sont indiqués par une étiquette parfois inscrite dans un cercle, et sont reliés par un trait si et seulement une arête les lie dans le graphe. Un exemple de représentation est dessiné ci-dessous.

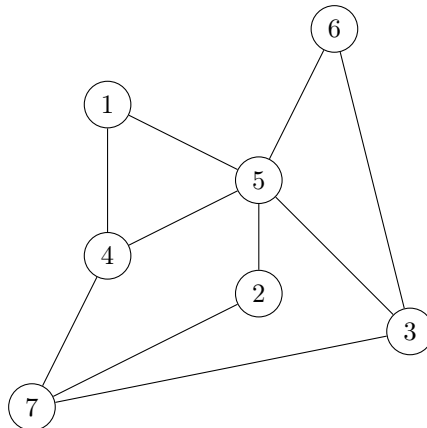


FIGURE 1 – Exemple d'une représentation d'un graphe  $G$

On note  $\mathcal{L}_n$  l'ensemble des graphes d'ensemble de sommets  $\{1, \dots, n\}$  donc de taille  $n$ , et  $\mathcal{L} = \bigcup_{n \geq 1} \mathcal{L}_n$  l'ensemble de tous les graphes.

**Définition 2.** Un **morphisme** de graphes est une application  $\varphi$  d'un graphe  $G = (V_G, E_G)$  vers un graphe  $H = (V_H, E_H)$ , telle que,

$$\forall \{x, y\} \in E_G, \quad \{\varphi(x), \varphi(y)\} \in E_H.$$

On dit que  $\varphi$  est un **isomorphisme**, s'il est un morphisme bijectif, dont la réciproque est un morphisme, c'est-à-dire que, pour tous  $(x, y) \in V^2$ ,

$$\{x, y\} \in E_G \iff \{\varphi(x), \varphi(y)\} \in E_H.$$

On dit que deux graphes sont **isomorphes** s'il existe un isomorphisme entre les deux. Cette relation est une relation d'équivalence, que l'on notera  $\simeq$ . On remarque en particulier que deux graphes sont isomorphes, si et seulement si l'un peut être obtenu en renommant les sommets du second.

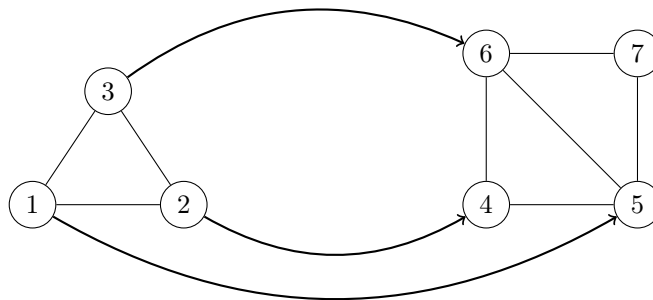


FIGURE 2 – Exemple de représentation de morphisme, d'un graphe triangle (à gauche) vers un graphe carré (à droite), une flèche représente l'image d'un sommet par le morphisme

On introduit alors,  $\mathcal{U}_n$  l'ensemble  $(\mathcal{L}_n)_{/\simeq}$  des classes d'équivalences de  $\mathcal{L}_n$  par la relation  $\simeq$ , ainsi que  $\mathcal{U} = (\mathcal{L}/\simeq) = \bigcup_{n \geq 1} \mathcal{U}_n$ . Autrement dit :  $\mathcal{U}$  est l'ensemble des graphes sans étiquettes.

Pour un graphe  $G$ , et un sous-ensemble  $(v_1, \dots, v_k)$  de sommets de  $G$ , on note  $G(v_1, \dots, v_k)$  le graphe de sommets  $\{1, \dots, k\}$ , et pour lequel  $i$  est relié à  $j$  si et seulement si  $v_i$  est relié à  $v_j$  dans  $G$ . On remarque, par ailleurs, que la définition autorise des répétitions dans la famille de sommets  $(v_1, \dots, v_k)$ .

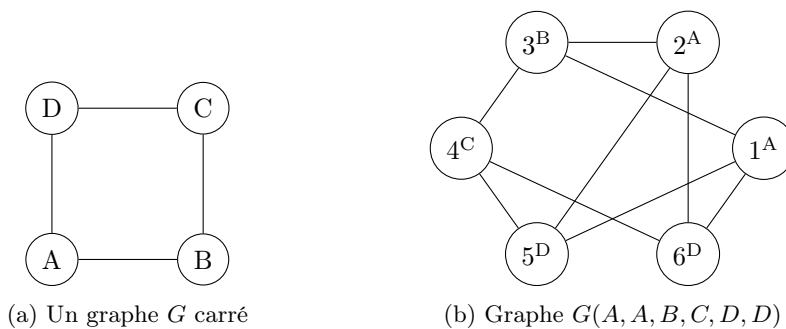


FIGURE 3 – Exemple de graphe  $G(v_1, \dots, v_k)$

Cette fonction nous permet de créer des graphes à partir d'« échantillons » de sommets. En choisissant cet échantillon de manière aléatoire, on peut donc obtenir des graphes aléatoires de  $k$  sommets (Borgs, Chayes, Lovász, et Sós [2] nomment ces graphes des «  $k$ -fold blow-up »). Ceux-ci vont nous permettre plus tard d'établir des limites de graphes rendant compte au mieux de propriétés locales typiques.

**Définition 3.** Soient  $G = (V, E)$  un graphe et  $k$  un entier non nul, on définit le  $k$ -sondage  $\mathbf{G}[k]$  comme étant le graphe aléatoire  $G(V_1, \dots, V_k)$ , où  $V_1, \dots, V_k$  sont des variables aléatoires indépendantes suivant une loi uniforme sur l'ensemble  $V$ .

De manière analogue on définit pour  $k \leq |V|$ ,  $\mathbf{G}'[k] = G(V_1, \dots, V_k)$ , où  $(V_1, \dots, V_k)$  est une famille de sommets choisie uniformément parmi les familles de sommets distincts.

Soient  $F = (V_F, E_F)$  et  $G = (V_G, E_G)$  deux graphes tels que  $V_F \subset V_G$ , on dit que  $F \subset G$ , si et seulement si  $E_F \subset E_G$ . En particulier, on s'intéressera à cette relation pour les graphes dont les ensembles de sommets sont inclus dans  $\mathbb{N}$ . On peut demander une condition plus forte à  $F$ , à savoir qu'il soit un **sous-graphe induit** de  $G$ , c'est-à-dire que deux sommets  $x, y \in V_F \subset V_G$  sont reliés dans  $F$  si et seulement s'ils sont reliés dans  $G$ .

**Définition 4.** Un graphe  $G = (V, E)$  peut se représenter sous forme de matrice. En numérotant chaque sommet de  $V$ ,  $v_1, \dots, v_n$  où  $n$  désigne le nombre de sommets de  $G$ , on définit la **matrice d'adjacence** du graphe  $G$  comme étant la matrice  $M$  dont les coefficients  $m_{ij}$  sont définis pour tous  $1 \leq i, j \leq n$ ,

$$m_{ij} = \begin{cases} 1 & \text{si } \{v_i, v_j\} \in E \\ 0 & \text{sinon} \end{cases}.$$

Un exemple de matrice d'adjacence est présenté en Figure 4.

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

FIGURE 4 – Matrice d'adjacence du graphe  $G$  de la figure 1

## 1.2 Les graphes aléatoires

On présente en premier lieu un modèle simple de graphe aléatoire, dont nous détaillerons les limites. On pourra se référer au livre de Mansuy [10] pour une introduction plus détaillée.

### Le graphe d'Erdős-Rényi

Un premier modèle de graphes aléatoires a été développé dans une série d'articles d'Erdős et Rényi de 1959 à 1968 (en particulier [5] le premier d'entre eux). Les graphes d'Erdős-Rényi constituent une première approche aux graphes aléatoires, et permettent d'établir un certain nombre de propriétés « typiques » des graphes. En particulier, ils ont permis à l'aide de la méthode probabiliste de démontrer des résultats généraux en théorie des graphes. On ne présente ici que la définition pour rapidement arriver à une limite de ce modèle pour l'étude de grands graphes.

**Définition 5.** On se donne un réel  $p$  compris entre 0 et 1, et un entier naturel  $n$  non nul, on va construire un graphe aléatoire d'ensemble de sommets  $\{1, \dots, n\}$  de la manière suivante. Pour chaque paire  $\{i, j\}$  d'entiers compris entre 1 et  $n$  distincts, l'arête  $\{i, j\}$  a une probabilité  $p$  d'apparaître dans le graphe. On obtient dès lors un graphe aléatoire  $G$  nommé **graphe d'Erdős-Rényi**, dont la loi est notée  $\mathcal{G}(n, p)$ .

**Remarque 1.** Ce graphe est parfois appelé *graphe binomial*, ce qui est justifié par le fait que le nombre d'arêtes d'un graphe d'Erdős-Rényi suit une loi binomiale de paramètres  $\binom{n}{2}, p$ .

### Le graphe de Rado

Dès lors, ce premier modèle construit, on peut s'intéresser à un modèle similaire, le **modèle de Rado**, où l'ensemble de sommets est l'ensemble  $\mathbb{N}^*$  des entiers naturels non

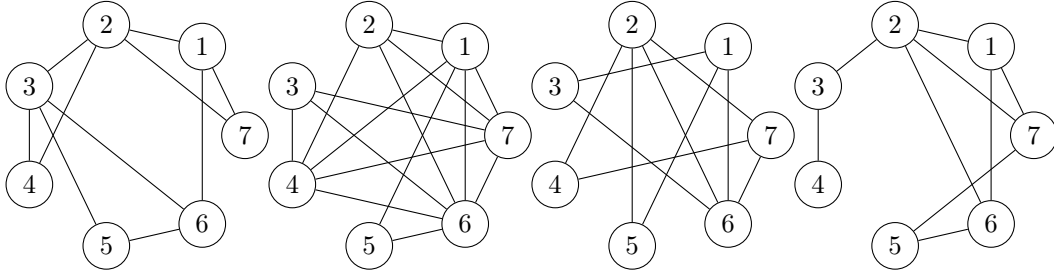


FIGURE 5 – Exemples de graphes d’Erdős-Rényi de loi  $\mathcal{G}(7, \frac{1}{2})$

nuls, et où les arêtes sont tirées une à une avec probabilité  $p$ . On note  $\mathcal{G}(\infty, p)$  la loi suivie par un tel graphe.

Ce modèle simple présente cependant des limites dans l’étude des propriétés locales de certains graphes, comme le montre le théorème suivant.

**Théorème 1.1.** Soient  $p, q \in ]0, 1[$ . Soient  $G \sim \mathcal{G}(\infty, p)$ ,  $H \sim \mathcal{G}(\infty, q)$ . Alors,  $G$  et  $H$  sont isomorphes presque sûrement.

On présente en annexe une preuve de ce théorème.

Ce théorème constitue un frein à ce premier modèle de graphes aléatoires, dans la mesure où deux graphes sont presque sûrement isomorphes (et dans un certain sens identiques) dès lors qu’ils présentent une certaine régularité. En particulier, cela empêche l’étude de propriétés locales de ces graphes infinis, et demande le développement d’un nouveau modèle.

### Modèle de graphe aléatoire généré par un graphon

Puisque être isomorphe s’avère trop restrictive pour les graphes aléatoires infinis on s’intéressera à les caractériser par une « limite d’échelle » de leurs matrices d’adjacence, le graphon. Soit  $\mathcal{W}$  l’ensemble de toutes les fonctions mesurables  $W : [0, 1]^2 \rightarrow [0, 1]$ . Soit  $\mathcal{W}_s$  le sous-ensemble des fonctions symétriques, i.e.  $W(a, b) = W(b, a)$ , appelées **graphons**. Pour chaque  $W \in \mathcal{W}_s$ , nous définissons un graphe aléatoire infini

$$\mathcal{G}(\infty, W) \in \mathcal{L}_\infty$$

par la construction suivante : nous commençons par choisir une suite  $U_1, U_2, \dots$  de variables aléatoires i.i.d. uniformément distribuées sur  $[0, 1]$ , puis, étant donné cette suite, pour chaque paire  $(i, j)$  avec  $i < j$ , nous ajoutons une arête  $\{i, j\}$  avec une probabilité  $W(X_i, X_j)$ , indépendamment pour toutes les paires  $(i, j)$  avec  $i < j$  (conditionnellement à  $\{X_i\}$ ). De plus, notons  $\mathcal{G}(n, W)$  la restriction  $\mathcal{G}(\infty, W)|_{[n]}$ , qui est obtenue par la même construction avec une suite finie  $U_1, \dots, U_n$ . Le modèle basé sur un graphon fixe  $W$  et sur  $n$  sommets est noté  $\mathcal{G}(n, W)$  par analogie avec le modèle des graphes aléatoires Erdős–Rényi.

Ce modèle est très riche et généralise les modèles de Erdős–Rényi et de Rado. En effet, on va montrer à la fin de cette mémoire qu’un graphe aléatoire infini à une loi invariante par re-étiquetage de ces sommets si et seulement s’il peut être défini en termes d’un graphon (éventuellement aléatoire) de cette manière. Il découle de cette construction et de la loi des grands nombres que, si  $W \neq 0$ , les modèles de graphes aléatoires échangeables sont presque sûrement denses. On donne quelques exemples concrets :

**Exemple 1.** L’exemple le plus simple d’un graphon est la fonction constante  $W(x, y) \equiv p$  pour un  $p \in [0, 1]$  fixé. On retrouve alors le modèle d’Erdős–Rényi. On a  $\mathcal{G}(n, W) \stackrel{\text{loi}}{=} \mathcal{G}(n, p)$ .

**Exemple 2.** Considérons les graphons constants par morceaux. En divisant le carré unité en blocs de  $k \times k$ , et attribuant à  $W$  la valeur  $p_{lm}$  sur le  $(l, m)$ <sup>ème</sup> bloc, on construit le modèle appelé de blocs stochastiques à  $k$  communautés (*stochastic block model* en anglais). On peut l’interpréter comme un modèle de graphes aléatoires composés de  $k$  blocs de graphes Erdős–Rényi distincts de paramètres  $p_{11}, \dots, p_{kk}$ , où chaque arête possible entre les blocs  $(l, l)$  et  $(m, m)$  est incluse indépendamment avec probabilité  $p_{lm}$ .

**Exemple 2.1.** Regardons plus en détails un cas spécifique important de graphons constants par morceaux pour se familiariser avec l'interaction entre les graphons et les graphes qu'ils génèrent. Le graphon biparti complet  $W$  est défini par  $W(x, y) = 1$  lorsque  $\min(x, y) < 1/2$  ou  $\max(x, y) > 1/2$ , et en posant  $W(x, y) = 0$  sinon. Alors la matrice d'adjacence de graphe aléatoire  $\mathcal{G}(n, W)$  ressemble à une matrice par blocs antidiagonale, avec deux blocs de uns de tailles aléatoires  $k$  et  $n - k$ . La seule source du hasard vient des tailles de ces blocs car le graphon ne prend que des valeurs dans  $\{0, 1\}$ . Dans ce cas,  $\mathcal{G}(n, W) = K_{k, n-k}$  p.s., où  $K_{n_1, n_2}$  est le graphe complet biparti i.e. son ensemble de sommets peut être divisé en deux sous-ensembles disjoints  $V = V_1 \cup V_2$  avec  $|V_1| = n_1$  et  $|V_2| = n_2$  tels que chaque arête ait une extrémité dans  $V_1$  et l'autre dans  $V_2$  et chaque sommet du  $V_1$  est relié à tous les sommets du  $V_2$ .

Par exemple, conditionnellement à l'évènement qu'on a tiré  $(U_1, U_2, U_3, U_4) \in [0, 1/4) \times [1/4, 1/2) \times [1/2, 3/4) \times [3/4, 1]$  la matrice d'adjacence de  $\mathcal{G}(4, W)$  est donnée par

$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

On peut déjà constater à l'aide de ce modèle simple qu'il existe différents graphons qui donnent lieu au même modèle. Par exemple, prenons le graphon  $\widehat{W} : (x, y) \mapsto W(\varphi(x), \varphi(y))$ , où

$$\varphi(x) = \begin{cases} x, & \text{si } x \in [0, 1/4) \\ x + 1, & \text{si } x \in [1/4, 1/2) \\ x - 1, & \text{si } x \in [1/2, 3/4) \\ x, & \text{si } x \in [3/4, 1] \end{cases},$$

de sorte que  $\widehat{W}$  ait une structure de damier de zéros et de uns (voir Figure 6). L'application

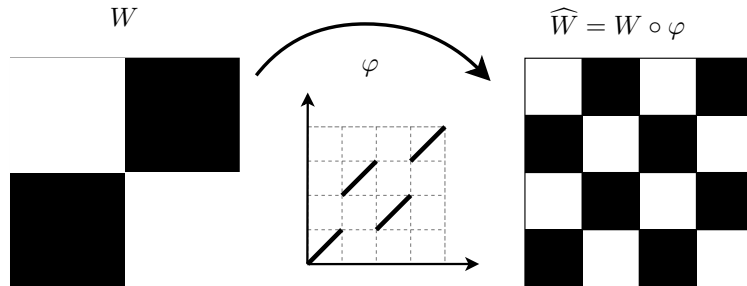


FIGURE 6 – Le graphon  $W$  et le graphon  $\widehat{W}$  donnent le même modèle de graphes aléatoires.

$\varphi$  est définie de manière que

$$\mathbb{P}(U_i \in \varphi^{-1}(A)) = \mathbb{P}(U_i \in A)$$

pour tout  $A \subset [0, 1]$  mesurable. On dit que  $\varphi$  est une **application préservant la mesure**.

Ce qui implique que  $\widehat{W}(U_i, U_j) \sim W(U_i, U_j)$  et donc les deux modèles sont identiques.

Reprenons l'exemple où on a conditionné par l'évènement  $(U_1, U_2, U_3, U_4) \in [0, 1/4) \times [1/4, 1/2) \times [1/2, 3/4) \times [3/4, 1]$ . Alors la matrice d'adjacence de  $\mathcal{G}(4, \widehat{W})$  est donnée par

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

On voit que même si ces deux matrices d'adjacence ne sont pas identiques elles représentent le même graphe non-étiqueté  $K_{2,2} \in \mathcal{U}_4$ . Cette propriété d'invariance par permutation des

lignes et colonnes est la motivation pour introduire la notion d'échangeabilité avec laquelle on s'occupera dans la section suivante.

## 2 Echangeabilité

Dans cette section, nous allons introduire la notion d'une suite de variables aléatoires échangeables ainsi que le théorème de De Finetti qui décrit le comportement des suites échangeables. Après avoir présenté la méthode de moments, nous allons pouvoir établir une démonstration simple de ce théorème. Les suites échangeables nous serviront de modèle jouet 1D pour ensuite introduire leur équivalent en 2D, les tableaux échangeables. À la fin de la section nous énonçons le théorème d'Aldous-Hoover qui décrit le comportement d'un tableau échangeable.

### 2.1 Suites échangeables et le théorème de De Finetti

Une suite de variables aléatoires est dite échangeable si la distribution jointe de la suite reste inchangée par toute permutation des indices. Le théorème de De Finetti caractérise toutes les suites échangeables à valeurs dans  $\{0, 1\}$  comme un « mélange » de suites de variables aléatoires de Bernoulli i.i.d.

**Définition 6.** Une suite de variables aléatoires  $\{X_i\}_{i \in \mathbb{N}}$  sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$  est dite échangeable si, pour tout  $N \in \mathbb{N}$  et toute permutation  $\sigma$  de  $\{1, 2, \dots, N\}$ , les distributions de  $(X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(N)})$  et de  $(X_1, X_2, \dots, X_N)$  sont identiques, c'est-à-dire si, pour toute famille d'ensembles  $A_1, A_2, \dots, A_N \in \mathcal{F}$ , on a

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2, \dots, X_N \in A_N) = \mathbb{P}(X_{\sigma(1)} \in A_1, X_{\sigma(2)} \in A_2, \dots, X_{\sigma(N)} \in A_N).$$

Remarquons que de manière équivalente,  $\{X_i\}_{i \in \mathbb{N}}$  est échangeable si ces lois fini-dimensionnelles sont préservées par les permutations sur  $\mathbb{N}$  à support finies, c'est-à-dire  $\{X_i\}_{i \in \mathbb{N}}$  et  $\{X_{\sigma(i)}\}_{i \in \mathbb{N}}$  ont les mêmes lois fini-dimensionnelles pour toutes permutations  $\sigma$  de  $\mathbb{N}$  qui vérifient  $\sigma(i) = i$  pour tout  $i$  suffisamment grand.

De Finetti a introduit les suites échangeables pour décrire une sorte de symétrie distributionnelle, une notion plus faible que celle d'indépendance. Par exemple, des variables indépendantes identiquement distribuées forment une suite échangeable. Cependant, la réciproque est fautive, c'est-à-dire qu'il existe des suites échangeables qui ne sont pas composées de variables aléatoires indépendantes, comme l'urne de Pólya (voir l'exemple 3).

Notons  $\pi_p$  la loi de Bernoulli sur  $\{0, 1\}$  donnée par  $\pi_p(1) = p$  et  $\pi_p(0) = 1 - p$ . Remarquons que sous forme analytique la fonction peut s'écrire  $\pi_p(x) = p^x(1 - p)^{1-x}$ . Si les variables aléatoires  $X_i$  sont indépendantes et telles que  $\mathbb{P}(X_i = 1) = \pi_p(1) = p$  et  $\mathbb{P}(X_i = 0) = \pi_p(0) = 1 - p$ , alors la fonction de masse de  $(X_1, \dots, X_N)$  est la mesure produit  $\bigotimes_{i=1}^N \pi_p$  sur  $\{0, 1\}^N$ . Avec ces définitions, nous énonçons le théorème de représentation de De Finetti publié en 1931 [3].

**Théorème 2.1.** Soit  $\{X_i\}_{i \in \mathbb{N}}$  une suite de variables aléatoires échangeables à valeurs dans  $\{0, 1\}$ . Alors, il existe une variable aléatoire  $\bar{X}_\infty : \Omega \rightarrow [0, 1]$  de distribution  $\mu$  telle que pour tous  $N$  et  $(x_1, \dots, x_N) \in \{0, 1\}^N$ ,

$$\mathbb{P}(X_1 = x_1, \dots, X_N = x_N) = \int_0^1 \prod_{i=1}^N \pi_p(x_i) \mu(dp) = \int_0^1 p^s (1 - p)^{N-s} \mu(dp),$$

où  $s = \sum_{i=1}^N x_i$ . De plus, on a

$$\bar{X}_n = \frac{1}{N} \sum_{i=1}^N X_i \xrightarrow[N \rightarrow \infty]{} \bar{X}_\infty \quad (2.1)$$

presque sûrement.

La mesure  $\mu$  est appelée la **mesure de De Finetti** associée à la suite  $\{X_i\}_{i \in \mathbb{N}}$  et le résultat de convergence 2.1 est connu sous le nom de **loi forte des grands nombres de De Finetti**, qui généralise la loi ordinaire des grands nombres pour les variables aléatoires i.i.d.

De manière intuitive, on peut comprendre ce résultat comme suit : une suite binaire échangeable est une suite de variables de Bernoulli i.i.d. de paramètre  $p$  aléatoire distribué selon une loi  $\mu$ .

Le théorème de De Finetti est fondamental pour le modèle bayésien en statistique et est utilisé dans la théorie des processus stochastiques. Il existe plusieurs façons de prouver ce théorème. Une façon est d'utiliser des martingales et la loi du 0-1 de Kolmogorov (cf. cours de processus stochastiques). Ici, nous présentons une preuve essentiellement autonome de Kirsch [8], qui repose sur la méthode des moments et permet de prouver la convergence étroite des mesures.

## 2.2 Préliminaires

Pour une mesure de probabilité  $\mu$  sur  $\mathbb{R}$ , nous définissons les moments d'ordre  $k$  par  $m_k(\mu) := \int x^k d\mu(x)$  chaque fois que l'intégrale existe (au sens où  $\int |x|^k d\mu(x) < \infty$ ). Par la suite, nous ne considérerons que des mesures à support compact de sorte que tous les moments existent et soient finis. Le lemme suivant nous dit que les moments déterminent la loi.

**Lemme 2.2.** Si  $\mu$  et  $\nu$  sont deux mesures de probabilité dont le support est contenu dans  $[a, b]$  et dont les moments sont égaux  $\forall k \in \mathbb{N}$ ,  $m_k(\mu) = m_k(\nu)$ , alors  $\mu = \nu$ .

*Démonstration.* On va démontrer que les moments d'une mesure déterminent sa transformée de Fourier et en conséquent la loi elle-même aussi. En effet,

$$\widehat{\mu}(\xi) := \int e^{i\xi x} \mu(dx) = \int \sum_{n=0}^{\infty} \frac{(i\xi)^n}{n!} x^n \mu(dx).$$

Or cette somme converge normalement sur un compact on peut échanger la somme et l'intégrale :

$$\widehat{\mu}(\xi) = \sum_{n=0}^{\infty} \frac{(i\xi)^n}{n!} \int x^n \mu(dx) = \sum_{n=0}^{\infty} \frac{(i\xi)^n}{n!} m_n(\mu).$$

Alors,  $\widehat{\mu} = \widehat{\nu}$  et donc  $\mu = \nu$ . □

Le lemme suivant est connu en théorie de probabilités comme la méthode des moments qui nous permettra d'établir une convergence de mesures à partir de convergence de ses moments.

**Lemme 2.3.** Soient  $\mu_n$  des mesures de probabilité dont le support est contenu dans un intervalle fixe  $[a, b]$  pour tout  $n \in \mathbb{N}$ . Si pour tout  $k$  les moments  $m_k(\mu_n)$  convergent vers une limite  $m_k$ , alors la suite  $\mu_n$  converge étroitement vers une mesure  $\mu$  à support dans  $[a, b]$ . En plus, les moments de  $\mu$  sont donnés par  $m_k(\mu) = m_k$ .

**Remarque 2.** Soient  $\mu_n$  et  $\mu$  des mesures de probabilité sur un espace polonais. La **convergence étroite** des mesures  $\mu_n$  vers  $\mu$  signifie que

$$\int f(x) d\mu_n(x) \xrightarrow{n \rightarrow \infty} \int f(x) d\mu(x)$$

pour toutes les fonctions bornées et continues  $f$ .

*Démonstration.* Notons  $H \subset \mathcal{C}([a, b])$  l'ensemble des polynômes restreints à l'intervalle  $[a, b]$ . Par le théorème d'approximation de Weierstrass, les polynômes sur  $[a, b]$  sont uniformément denses dans l'espace des fonctions continues sur  $[a, b]$ , c'est-à-dire que la fermeture de  $H$  (par rapport à la norme sup) contient  $\mathcal{C}([a, b])$ . Ainsi, l'intégrale  $\int f(x) d\mu(x)$  pour  $f$  continue peut être calculée à partir de la connaissance des moments de  $\mu$ . En effet, soit  $\varepsilon > 0$  alors il existe un polynôme  $p \in H$  telle que  $|f(x) - p(x)| < \frac{\varepsilon}{2(b-a)}$  pour tout  $x \in [a, b]$ . Alors, on a pour tout  $n \in \mathbb{N}$

$$\left| \int f d\mu_n - \int p d\mu_n \right| \leq \int |f - p| d\mu_n < \varepsilon/2.$$

Pour borner le terme  $\int p d\mu_n$ , on peut expliciter le polynôme  $p(x) = \sum_{k=0}^K a_k x^k$  avec  $K \in \mathbb{N}$  et  $a_0, \dots, a_K \in \mathbb{R}$ . On a,

$$\int p d\mu_n = \sum_{k=0}^K a_k m_k(\mu_n) \xrightarrow{n \rightarrow \infty} \sum_{k=0}^K a_k m_k(\mu)$$

donc pour tout  $n$  suffisamment grand

$$\left| \int p d\mu_n - \sum_{k=0}^K a_k m_k(\mu) \right| < \varepsilon/2.$$

Par l'inégalité triangulaire on obtient pour tout  $n$  suffisamment grand

$$\left| \int f d\mu_n - \sum_{k=0}^K a_k m_k(\mu) \right| < \varepsilon.$$

Nous avons que les intégrales  $\int f(x) d\mu_n(x)$  convergent pour tout  $f \in \mathcal{C}([a, b])$ . La limite est une forme linéaire positive. Ainsi, par le théorème de représentation de Riesz, les mesures de probabilité  $\mu_n$  convergent étroitement vers une unique mesure de Radon  $\mu$ .  $\square$

Avant de passer à la preuve du théorème étudions l'exemple emblématique de processus échangeable l'urne de Pólya pour qu'on puisse manipuler les notations introduites jusqu'ici.

**Exemple 3** (L'urne de Pólya). Supposons que nous ayons  $b$  billes bleues et  $r$  billes rouges dans une urne. Nous choisissons une bille au hasard et la remettons dans l'urne avec une bille en plus de la même couleur. Remarquons que dans un certain sens, le modèle de l'urne de Pólya est l'opposé du modèle d'échantillonnage sans remplacement, où chaque fois qu'une valeur particulière est observée, elle est moins probable d'être observée de nouveau, tandis que dans un modèle d'urne de Pólya, une valeur observée est plus susceptible d'être observée de nouveau. Considérons la variable aléatoire

$$X_i = \begin{cases} 1, & \text{si la } i\text{ème bille tirée est bleue} \\ 0, & \text{si elle est rouge.} \end{cases}$$

Les variables aléatoires  $X_i$  ne sont pas indépendantes par la remarque, mais elles sont échangeables. Par exemple,

$$\begin{aligned} \mathbb{P}(X_1 = 1, X_2 = 1, X_3 = 0) &= \frac{b}{b+r} \times \frac{b+1}{b+r+1} \times \frac{r}{b+r+2} \\ &= \frac{b}{b+r} \times \frac{r}{b+r+1} \times \frac{b+1}{b+r+2} = \mathbb{P}(X_1 = 1, X_2 = 0, X_3 = 1). \end{aligned}$$

Plus généralement, soit  $(x_1, \dots, x_n) \in \{0, 1\}^n$  et soit  $k = \sum_{i=1}^n x_i$  le nombre de fois qu'on tire une bille bleue pendant les  $n$  premières tirages. Alors,

$$\begin{aligned} \mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) &= \frac{\prod_{i=1}^k (b+i-1) \times \prod_{i=1}^{n-k} (r+i-1)}{\prod_{i=1}^n (r+b+i-1)} \\ &= \frac{(b+k-1)! (r+n-k-1)! (b+r-1)!}{(b-1)! (r-1)! (b+r+n-1)!}. \end{aligned}$$

Cette probabilité ne dépend pas de l'ordre d'observation des billes bleues et rouges et dépend uniquement du nombre total de billes rouges et du nombre total de billes bleues et donc la suite est échangeable. Pour identifier sa mesure de De Finetti  $\mu$ , examinons ses moments  $m_n(\mu) = \int x^n \mu(dx)$ . Pour tout  $n \in \mathbb{N}$  on a

$$m_n(\mu) = \mathbb{P}(X_1 = \dots = X_n = 1) = \frac{(b+n-1)! (b+r-1)!}{(b-1)! (b+r+n-1)!}.$$

On peut vérifier que les  $m_n(\mu)$  sont les moments de la distribution  $\text{Beta}(\alpha, \beta)$  de paramètres  $\alpha = b, \beta = r$ , i.e. la loi  $\mu$  est de densité

$$p_\mu(x) = \frac{(b+r)!}{a(b)!(r)!} x^{b-1} (1-x)^{r-1} \mathbf{1}_{0 \leq x \leq 1}.$$

Par la loi forte des grands nombres de De Finetti la proportion de billes bleues choisies pendant les  $n$  premiers tirages

$$\bar{X}_n = \frac{1}{n} (X_1 + \dots + X_n)$$

converge presque sûrement vers  $\bar{X}_\infty \sim \text{Beta}(b, r)$ . En plus, le théorème nous dit aussi que nous pouvons générer les  $X_i$  en choisissant d'abord  $p$  à partir de la distribution  $\text{Beta}(b, r)$ , puis en générant des  $X_i$  i.i.d. selon une loi de Bernoulli de paramètre  $p$ .

### 2.3 Preuve du théorème de De Finetti

Le théorème suivant est un substitut pour une loi faible des grands nombres.

**Théorème 2.4.** Soit  $\{X_i\}_{i \in \mathbb{N}}$  une suite de variables aléatoires échangeables à valeurs dans  $\{0, 1\}$ , alors la moyenne empirique  $\bar{X}_N := \frac{1}{N} \sum_{i=1}^N X_i$  converge en distribution vers une mesure de probabilité  $\mu$  qu'on appelle **la mesure de De Finetti** associée à  $\{X_i\}_{i \in \mathbb{N}}$ . De plus,  $\mu$  est concentrée sur  $[0, 1]$  et ses moments sont donnés par

$$m_k(\mu) = \mathbb{E}(X_1 \cdot X_2 \cdot \dots \cdot X_k).$$

Avant de passer à la démonstration, pour calculer l'espérance d'un produit des termes échangeables le lemme suivant nous sera utile.

**Lemme 2.5.** Soit  $\{X_i\}$  une suite échangeable de variables aléatoires à valeurs dans  $\{0, 1\}$ , soient  $i_1, \dots, i_k \in \mathbb{N}^*$  et

$$r := \#\{i_1, i_2, \dots, i_k\}$$

alors

$$\mathbb{E}(X_{i_1} \cdot X_{i_2} \cdot \dots \cdot X_{i_k}) = \mathbb{E}(X_1 \cdot X_2 \cdot \dots \cdot X_r)$$

*Démonstration du Lemme 2.5.* Puisque  $X_i \in \{0, 1\}$ , nous avons  $X_i^\ell = X_i$  pour tout  $\ell \in \mathbb{N}, \ell \geq 1$ . Ainsi, le produit de la partie gauche est en réalité un produit de  $r$  termes  $X_j$  différents, dont l'espérance est égale à la partie droite en raison de l'échangeabilité.  $\square$

*Démonstration du Théorème 2.4.* Pour exprimer les moments de  $\bar{X}_N$ , nous calculons :

$$\begin{aligned} \left( \sum_{i=1}^N X_i \right)^k &= \sum_{(i_1, \dots, i_k) \in \{1, \dots, N\}^k} X_{i_1} \cdot X_{i_2} \cdot \dots \cdot X_{i_k} \\ &= \sum_{r=1}^k \sum_{\substack{(i_1, \dots, i_k) \in \{1, \dots, N\}^k \\ \#\{i_1, i_2, \dots, i_k\} = r}} X_{i_1} \cdot X_{i_2} \cdot \dots \cdot X_{i_k}, \end{aligned}$$

où on a divisé la somme en fonction du nombre d'indices différents dans  $(i_1, \dots, i_k)$ . Ainsi, en prenant l'espérance on a

$$\begin{aligned} \mathbb{E} \left( \left( \frac{1}{N} \sum_{i=1}^N X_i \right)^k \right) &= \frac{1}{N^k} \sum_{\substack{i_1, \dots, i_k=1 \\ \#\{i_1, i_2, \dots, i_k\}=k}}^N \mathbb{E}(X_{i_1} \cdot X_{i_2} \cdot \dots \cdot X_{i_k}) \\ &+ \frac{1}{N^k} \sum_{\substack{i_1, \dots, i_k=1 \\ \#\{i_1, i_2, \dots, i_k\}<k}}^N \mathbb{E}(X_{i_1} \cdot X_{i_2} \cdot \dots \cdot X_{i_k}) \end{aligned} \quad (2.2)$$

On va montrer que le deuxième terme de 2.2 est négligeable quand  $N$  tend vers infini. Il existe au plus  $(k-1)^k N^{k-1}$  uplets d'indices  $(i_1, \dots, i_k)$  avec  $\#\{i_1, i_2, \dots, i_k\} < k$ . En effet, il y a  $N^{k-1}$  manières de choisir les indices possibles « candidats » pour  $(i_1, \dots, i_k)$ . Pour chacune des  $k$  positions du  $k$ -uplet, nous pouvons choisir l'un des  $k-1$  candidats, ce qui donne  $(k-1)^k$  possibilités. Cela couvre également les uplets avec moins de  $k-1$  indices différents car certains des candidats peuvent finalement ne pas apparaître dans l'uplet. Or  $\mathbb{E}(X_{i_1} \cdot X_{i_2} \cdot \dots \cdot X_{i_k}) \leq 1$  il s'ensuit que le deuxième terme dans (2.2) tend vers zéro. Donc

$$\mathbb{E} \left( \left( \frac{1}{N} \sum_{i=1}^N X_i \right)^k \right) \underset{N \rightarrow +\infty}{\sim} \frac{1}{N^k} \sum_{\substack{i_1, \dots, i_k=1 \\ \#\{i_1, i_2, \dots, i_k\}=k}}^N \mathbb{E}(X_{i_1} \cdot X_{i_2} \cdot \dots \cdot X_{i_k}),$$

ainsi, en utilisant Lemme 2.5 :

$$\begin{aligned} &= \frac{1}{N^k} \sum_{\substack{i_1, \dots, i_k=1 \\ \#\{i_1, i_2, \dots, i_k\}=k}}^N \mathbb{E}(X_1 \cdot X_2 \cdot \dots \cdot X_k) \\ &\sim \mathbb{E}(X_1 \cdot X_2 \cdot \dots \cdot X_k). \end{aligned}$$

Une application de la méthode de moments (Proposition 2.3) donne le résultat souhaité.  $\square$

Pour la preuve du Théorème 2.1, nous utiliserons le lemme simple suivant.

**Lemme 2.6.** Supposons que  $\{X_i\}_{i \in \mathbb{N}}$  est une suite échangeable à valeurs  $\{0, 1\}$ . Alors pour des  $i_1, i_2, \dots, i_k \in \mathbb{N}^*$  deux à deux distincts et  $x_1, \dots, x_k \in \{0, 1\}$  avec  $\sum x_i = m$  on a

$$\mathbb{P}(X_{i_1} = x_1, \dots, X_{i_k} = x_k) = \frac{1}{\binom{k}{m}} \mathbb{P} \left( \sum_{i=1}^k X_i = m \right)$$

*Démonstration.* Il existe  $\binom{k}{m}$  uplets  $x_1, \dots, x_k$  avec  $\sum x_i = m$ . En raison de l'échangeabilité, ils conduisent tous à la même probabilité :

$$\mathbb{P}(X_{i_1} = x_1, \dots, X_{i_k} = x_k) = \mathbb{P}(X_1 = 1, \dots, X_m = 1, X_{m+1} = 0, \dots, X_k = 0).$$

Et donc

$$\mathbb{P} \left( \sum_{i=1}^k X_i = m \right) = \binom{k}{m} \mathbb{P}(X_{i_1} = x_1, \dots, X_{i_k} = x_k)$$

$\square$

Avec ces résultats sous la main, on peut passer à la preuve du théorème de De Finetti.

*Démonstration de Théorème 2.1.* Soit  $\mu$  la mesure de De Finetti de  $X_i$  (voir le théorème 2.4) et définissons un processus  $\{Y_i\}_i$  à valeurs dans  $\{0, 1\}$  par

$$\mathbb{P}(Y_1 = y_1, \dots, Y_k = y_k) = \int \prod_{i=1}^k \pi_p(y_i) d\mu(p)$$

La suite  $Y_i$  est échangeable car cette probabilité ne dépend que de la somme  $\sum_{i=1}^k y_i$ . Nous allons démontrer que  $X_i$  et  $Y_i$  ont les mêmes lois fini-dimensionnelles.

Par le Lemme 2.6, il suffit de montrer que  $S_N := \sum_{i=1}^N X_i$  et  $T_N := \sum_{i=1}^N Y_i$  ont les mêmes distributions pour tout  $N$ , et pour cela, il suffit, d'après la Proposition 2.3, de montrer que leurs moments sont identiques. Par Théorème 2.4 on a

$$\mathbb{E}(X_1 \cdots X_r) = m_r(\mu) = \int p^r d\mu(p).$$

De l'autre coté par définition

$$\mathbb{E}(Y_1 \cdots Y_r) = \mathbb{P}(Y_1 = 1, \dots, Y_r = 1) = \int \prod_{i=1}^r \pi_p(1) d\mu(p) = \int p^r d\mu(p),$$

d'où on déduit que

$$\mathbb{E}(X_1 \cdots X_r) = \mathbb{E}(Y_1 \cdots Y_r). \quad (2.3)$$

On introduit la constante  $P_k^N(r)$  égale au nombre de  $k$ -uplets à valeurs dans  $[N] := \{1, \dots, N\}$  qui ont exactement  $r$  distinctes éléments i.e.

$$P_k^N(r) := \#\{(i_1, \dots, i_k) \in [N]^k : \#\{i_1, \dots, i_k\} = r\}$$

Par le Lemme 2.5 appliqué à la suite échangeable  $X_i$

$$\begin{aligned} \mathbb{E}((S_N)^k) &= \sum_{r=1}^k \sum_{\substack{i_1, \dots, i_k=1 \\ \#\{i_1, i_2, \dots, i_k\}=r}}^N \mathbb{E}(X_{i_1} \cdots X_{i_k}) \\ &= \sum_{r=1}^k P_k^N(r) \mathbb{E}(X_1 \cdots X_r). \end{aligned} \quad (2.4)$$

De même, par le Lemme 2.5 appliqué à la suite échangeable  $Y_i$  on a

$$\begin{aligned} \mathbb{E}((T_N)^k) &= \sum_{r=1}^k \sum_{\substack{i_1, \dots, i_k=1 \\ \#\{i_1, i_2, \dots, i_k\}=r}}^N \mathbb{E}(Y_{i_1} \cdots Y_{i_k}) \\ &= \sum_{r=1}^k P_k^N(r) \mathbb{E}(Y_1 \cdots Y_r). \end{aligned} \quad (2.5)$$

En utilisant (2.3) on conclut que  $\mathbb{E}((S_N)^k) = \mathbb{E}((T_N)^k)$ . □

## 2.4 Tableaux échangeables

Pour le traitement des graphes échangeables nous introduisons une notion d'échangeabilité qui sera satisfaite par la matrice d'adjacence. À savoir, l'invariance de sa loi sous permutation des sommets du graphe. Cela signifie que la loi de la matrice d'adjacence doit être préservée par permutation de ses lignes et colonnes, d'où la notion de tableaux échangeables. Elle peut être considérée comme un analogue bidimensionnel des suites échangeables.

**Définition 7.** Soient  $\{X_{ij}\}, 1 \leq i, j < \infty$ , des variables aléatoires binaires. Elles forment un **tableau** dit (conjointement) **échangeable** si

$$\mathbb{P}(X_{ij} = e_{ij}, 1 \leq i, j \leq n) = \mathbb{P}(X_{ij} = e_{\sigma(i)\sigma(j)}, 1 \leq i, j \leq n) \quad (2.6)$$

pour tout  $n$ , toute permutation  $\sigma \in \mathfrak{S}_n$  et tous  $e_{ij} \in \{0, 1\}$ .

De manière équivalente, le tableau  $\{X_{ij}\}$  est échangeable si le tableau  $\{X_{\sigma(i)\sigma(j)}\}$  a la même distribution que  $\{X_{ij}\}$  pour toute permutation  $\sigma$  de  $\mathbb{N}$  à support fini.

Pour motiver davantage cette définition dans notre étude des graphes aléatoires, nous remarquons qu'un sondage aléatoire d'un graphe est échangeable. En fait, considérons un graphe fini  $G$  et définissons maintenant  $\{X_{ij}\}, 1 \leq i, j < \infty$  comme la matrice d'adjacence de  $\mathbf{G}[\infty] = G(V_1, V_2, \dots)$  le  $\infty$ -sondage. Alors  $\{X_{ij}\}, 1 \leq i, j < \infty$  est un tableau échangeable. On va démontrer ce résultat dans Proposition 4.1 après avoir introduit la représentation des tableau échangeables.

Avec cette définition d'échangeabilité, une question naturelle est : existe-t-il une version bidimensionnelle du théorème de De Finetti caractérisant les tableaux échangeables ? On peut s'attendre à une version de (2.1) avec peut-être une double intégrale, mais le probabiliste David Aldous [1] et le logicien Douglas Hoover [6] ont montré que la réponse est plus complexe.

Pour aborder leurs résultat définissons un tableau binaire aléatoire  $\{X_{ij}\}$  comme suit : soient  $U_i, 1 \leq i < \infty$  v.a. indépendantes et uniformes sur  $[0, 1]$ . Prenons une fonction (déterministe)  $W : [0, 1]^2 \rightarrow [0, 1]$ . Posons  $X_{ij}$  égal à 1 ou 0 selon que la  $W(U_i, U_j)$ -pièce tombe pile ou face (i.e.  $X_{ij}$  est une variable Bernoulli de paramètre  $W(U_i, U_j)$ ). Remarquons que si  $W$  est un graphon on retrouve la matrice d'adjacence du graphe aléatoire infini  $\mathcal{G}(\infty, W)$  décrit dans l'introduction.

Notons  $P_W$  la loi de  $\{X_{ij}\}, 1 \leq i, j < \infty$ . Explicitement, soient  $i_1 j_1, \dots, i_m j_m \in \mathbb{N}^2$  des paires distinctes, on a

$$\mathbb{P}(X_{i_1 j_1} = \dots = X_{i_m j_m} = 1) = \mathbb{E} \prod_{r=1}^m W(U_{i_r}, U_{j_r}),$$

où l'espérance est prise sur les  $U_i, i \geq 1$  uniformes sur  $[0, 1]$  et indépendantes. Cette équation suffit pour déterminer la loi  $P_W$  parce que tout événement fini-dimensionnel  $x = \{x_{ij}\} \in \{0, 1\}^{\mathbb{N}^2}$  peut s'écrire comme une combinaison linéaire d'évènements sous la forme ci-dessus. On a donc,

$$P_W(x) = \mathbb{E} \prod_{(i,j) \in \mathbb{N}^2} W(U_i, U_j)^{x_{ij}} (1 - W(U_i, U_j))^{1-x_{ij}}.$$

Alors le tableau  $\{X_{ij}\}$  est échangeable en raison de l'échangeabilité des  $\{U_i\}$ . Le théorème d'Aldous-Hoover stipule que la loi de tout tableau binaire échangeable est un mélange de  $P_W$  construit de cette manière. Dans le cas qui nous intéresse, le théorème s'énonce :

**Théorème 2.7** (Aldous-Hoover). Soit  $X = \{X_{ij}\}, 1 \leq i, j < \infty$ , un tableau binaire échangeable avec  $X_{ii} = 0, 1 \leq i < \infty$ . Alors, il existe une fonction mesurable  $W_\bullet : [0, 1] \times [0, 1]^2 \rightarrow [0, 1]$  avec  $(\alpha, x, y) \mapsto W_\alpha(x, y)$  telle que

$$\mathbb{P}\{X \in A\} = \int_0^1 P_{W_\alpha}(A) d\alpha$$

Plus explicitement, on peut générer le tableau  $X$  comme suit : premièrement, on prend une réalisation de  $\alpha \sim \mathcal{U}[0, 1]$  et une échantillon  $U_i \stackrel{i.i.d.}{\sim} \mathcal{U}[0, 1]$ . Ensuite, on échantillonne

$$X_{ij} = X_{ji} \sim \mathcal{B}(W_\alpha(U_i, U_j)),$$

indépendemment pour tout  $i < j$ .

De manière analogue à l'intuition derrière le théorème de De Finetti, le théorème d'Aldous-Hoover nous dit que chaque tableau échangeable est comme l'exemple ci-dessus  $X_{ij} \sim W(U_i, U_j)$  mais au lieu d'avoir une fonction fixe  $W$  pour les paramètres de nos variables, cette fois  $W = W_\alpha$  est aléatoire.

Il est à noter que cette fonction  $W_\bullet$  n'est pas unique. Si  $\widehat{W}_\alpha(u, v) := W_\alpha(g_1(u), g_2(v))$  est obtenu à partir de  $W_\bullet$  par un changement préservant la mesure de chaque variable, le processus associé  $\{X_{ij}\}$  a la même loi que  $\{X_{ij}\}$  (voire Figure 6). Pourtant, il est connu que changements préservant la mesure sont la seule source de non-unicité (cf. Kallenberg [7]).

### 3 Une topologie pour les graphes infinis

On s'efforce ici de munir l'ensemble des graphes finis  $\mathcal{U}$  d'une structure d'espace métrique de clôture compacte. Dans un premier temps, on construira cette topologie sur l'ensemble des graphes déterministes, puis on cherchera à caractériser la convergence en loi de graphes aléatoires, lorsque le nombre de sommets tend vers l'infini.

#### 3.1 Convergence de grands graphes déterministes

Afin de permettre l'étude du comportement typique de grands graphes, il est nécessaire de s'intéresser à la notion de convergence de suite de graphes, et donc de munir l'espace des graphes d'une topologie. On s'appuie ici sur les articles de Lovász ([9] et [2]) et de Diaconis et Janson ([4]).

Pour construire cette topologie, on considère une famille dénombrables de fonctions  $t$  à valeurs dans  $[0, 1]$ , la convergence d'une suite dans l'espace produit caractérise la convergence d'une suite de graphes. On utilise ici la notion de  $k$ -sondage introduit dans la définition 3.

**Définition 8.** Soient  $F, G$  deux graphes de  $\mathcal{L}$ . Soit  $k$  le nombre de sommets de  $F$ . On pose,

$$t(F, G) = \mathbb{P}(F \subset \mathbf{G}[k]),$$

la **densité d'apparition** du graphe  $F$  dans le graphe  $G$ .

Cette quantité peut aussi s'exprimer en termes de morphismes de graphes : c'est la proportion de fonctions des sommets de  $F$  vers les sommets de  $G$  qui sont des morphismes. On retrouve dans l'article de Lovász [9] la définition suivante équivalente à la première,

$$t(F, G) = \frac{\text{hom}(F, G)}{v(G)^{v(F)}},$$

où  $\text{hom}(F, G)$  est le nombre de morphismes de  $F$  sur  $G$ .

On peut définir de la manière les fonctions analogues suivantes,

**Définition 9.** Soient  $F, G$  deux graphes de  $\mathcal{L}$ . Soit  $k$  le nombre de sommets de  $F$ , on pose, si  $k = v(F) \leq v(G)$ ,

$$t_{\text{inj}}(F, G) = \mathbb{P}(F \subset \mathbf{G}'[k]),$$

$$t_{\text{ind}}(F, G) = \mathbb{P}(F = \mathbf{G}'[k]),$$

et  $t_{\text{inj}}(F, G) = t_{\text{ind}}(F, G) = 0$  sinon.

Comme précédemment, on peut exprimer les fonctions  $t_{\text{inj}}$  et  $t_{\text{ind}}$  en terme de nombres de morphismes certains morphismes de  $F$  sur  $G$ . On note  $\text{inj}(F, G)$  le nombre de morphismes injectifs de  $F$  dans  $G$ , et  $\text{ind}(F, G)$  le nombre de morphismes de  $F$  sur  $G$ , tels que  $F$  est en bijection avec le sous-graphe induit par  $G$  sur l'image de  $F$ . On a alors,

$$t_{\text{inj}}(F, G) = \frac{\text{inj}(F, G)}{v(G)^{v(F)}}, \quad t_{\text{ind}}(F, G) = \frac{\text{ind}(F, G)}{v(G)^{v(F)}},$$

où  $v(G)^{v(F)} = v(G)(v(G) - 1) \dots (v(G) - v(F) + 1)$  désigne le nombre de fonctions injectives de  $F$  dans  $G$ .

Ces trois fonctions sont liés par les deux équations suivantes, présentées par Borgs, Chayes, Lovász, et Sós[2] par exemple, et que nous admettrons,

$$|t(F, G) - t_{\text{inj}}(F, G)| \leq \frac{v(F)^2}{2v(G)}, \quad (3.1)$$

$$t_{\text{inj}}(F, G) = \sum_{F' \in \mathcal{L}_k, F \subset F'} t_{\text{ind}}(F', G). \quad (3.2)$$

De l'équation 3.2 on déduit grâce au prince d'inclusion-exclusion que,

$$t_{\text{ind}}(F, G) = \sum_{F' \in \mathcal{L}_k, F' \supset F} (-1)^{e(F') - e(F)} t_{\text{inj}}(F', G), \quad (3.3)$$

où  $e(F)$  désigne le nombre d'arêtes d'un graphe  $F$ .

Avant de continuer, on remarque que la fonction  $t$  est invariante par isomorphisme. Plus formellement, on a la propriété suivante,

**Proposition 3.1.** Soient  $F, G \in \mathcal{L}$ , et soit  $\varphi$  un isomorphisme du graphe  $F$ , et  $\psi$  un isomorphisme du graphe  $G$ . Alors,

$$t(\varphi(F), \psi(G)) = t(F, G).$$

*Démonstration.* Soient  $F, G, \varphi, \psi$  comme dans l'énoncé. On remarque tout d'abord que  $\varphi$  induit une bijection entre les morphismes de  $F$  vers  $G$  et les morphismes de  $\varphi(F)$  vers  $G$ , en particulier  $\text{hom}(F, G) = \text{hom}(\varphi(F), G)$ . Il en va de même avec  $\psi$ , on en déduit donc que,  $\text{hom}(F, G) = \text{hom}(\varphi(F), \psi(G))$

$$t(F, G) = \frac{\text{hom}(F, G)}{v(F)v(G)} = \frac{\text{hom}(\varphi(F), \psi(G))}{v(\varphi(F))v(\psi(G))} = t(\varphi(F), \psi(G)).$$

□

Cette proposition montre que l'on peut aussi bien considérer  $t$  comme une fonction de  $\mathcal{U} \times \mathcal{U}$  dans  $[0, 1]$ , tout comme les fonctions  $t_{\text{inj}}$  et  $t_{\text{ind}}$ .

**Définition 10.** Pour un graphe  $G \in \mathcal{U}$ , on définit,

$$\tau(G) = (t(F, G))_{F \in \mathcal{U}} \in [0, 1]^{\mathcal{U}}.$$

L'espace  $[0, 1]^{\mathcal{U}}$  est un espace produit d'un nombre dénombrable d'espace  $[0, 1]$  muni de la topologie héritée de  $\mathbb{R}$ . Par ailleurs,  $[0, 1]$  est un espace compact métrique. D'après le théorème de Tykhonov,  $[0, 1]^{\mathcal{U}}$  est lui aussi compact. On peut le rendre métrique en considérant la distance suivante, si  $(F_n)_{n \in \mathbb{N}}$  est une énumération des graphes de  $\mathcal{U}$ ,

$$d(x, y) = \sum_{n \in \mathbb{N}} 2^{-n} |x_{F_n} - y_{F_n}|.$$

On admet ici que cette distance  $d$  est compatible avec la topologie produit.

Pour la suite on note  $\mathcal{U}^* \subset [0, 1]^{\mathcal{U}}$  l'image de  $\mathcal{U}$  par la fonction  $\tau$ . L'adhérence  $\overline{\mathcal{U}^*}$  de  $\mathcal{U}^*$  est un fermé dans un compact donc elle-même un compact. On s'intéressera particulièrement à cet ensemble qui est l'espace où évoluent les suites de graphes et leur limite.

La fonction  $\tau$  nous permet donc de munir l'ensemble des graphes d'une topologie. Elle n'est cependant pas injective. Comme nous le montre l'exemple suivant :

**Exemple 4.** Pour un entier  $n$  non nul, on note  $K_{n,n}$  la graphe biparti complet présenté en 2.1. Alors pour tout  $m, n$  entiers non nuls,  $\tau(K_{n,n}) = \tau(K_{m,m})$ . Pour démontrer cette assertion, montrons tout d'abord que les graphes  $\mathbf{K}_{n,n}[k]$  et  $\mathbf{K}_{m,m}[k]$  sont égaux en loi. On remarque que les graphes aléatoires  $\mathbf{K}_{n,n}[k]$  sont des graphes bipartis aléatoires, leur structure est donc caractérisée par le nombre d'éléments dans chaque composante. Le choix de chaque sommet, lors de la construction de  $\mathbf{K}_{n,n}[k]$ , parmi les deux composantes peut se modéliser par une variable de Bernoulli, la valeur 0 étant attribuée à l'une et 1 à l'autre. Alors le nombre de sommets dans chaque composante suit une loi binomiale de paramètre  $(k, \frac{1}{2})$ . Comme la loi ne dépend pas du nombre  $n$  de sommets des composantes, on en déduit que les lois de  $\mathbf{K}_{n,n}[k]$  et  $\mathbf{K}_{m,m}[k]$  sont identiques pour tout  $k$ . D'où l'on déduit que pour tout graphe  $F$ ,  $t(F, K_{n,n}) = t(F, K_{m,m})$ , et ainsi on prouve la non-injectivité de  $\tau$ .

Cela constitue un frein à l'identification de  $\mathcal{U}$  avec  $\overline{\mathcal{U}^*}$ , puisque deux graphes ayant la image par  $\tau$  seraient identifiés par cette topologie. Pour palier à cet écueil, on considère une fonction  $\tau^+$  basée sur la fonction  $\tau$ .

**Définition 11.** Pour un graphe  $G \in \mathcal{U}$ , on définit,

$$\tau^+(G) = (\tau(G), v(G)^{-1}),$$

où  $v(G)$  désigne le nombre de sommets du graphe  $G$ .

Cette fonction possède de meilleures propriétés que  $\tau$ , que l'on détaille notamment dans la propriété suivante.

**Proposition 3.2.**  $\tau^+$  est injective sur  $\mathcal{U}$

La démonstration repose sur les deux lemmes suivants.

**Lemme 3.3.** La famille  $(\mathbb{P}(F = \mathbf{G}[k]))_{F \in \mathcal{U}}$  est entièrement déterminée par  $\tau(G)$ .

*Démonstration.* On montre par récurrence décroissante sur  $e(F)$ , le nombre d'arêtes de  $F$ , que  $\mathbb{P}(F = \mathbf{G}[k])$  est déterminé par  $\tau$ . C'est le cas pour le graphe complet à  $k$  sommets. Soit maintenant  $F \in \mathcal{L}_k$ , avec  $e(F) < \binom{k}{2}$ . On remarque que,

$$\mathbb{P}(F \subset \mathbf{G}[k]) = \sum_{J \in \mathcal{L}_k, F \subset J} \mathbb{P}(\mathbf{G}[k] = J).$$

Donc  $\mathbb{P}(F = \mathbf{G}[k]) = \mathbb{P}(F \subset \mathbf{G}[k]) - \sum_{J \in \mathcal{L}_k, F \subset J} \mathbb{P}(\mathbf{G}[k] = J)$ . D'où l'on conclut par récurrence.  $\square$

**Lemme 3.4.** La famille  $(\mathbb{P}(F = \mathbf{G}'[k]))_{F \in \mathcal{U}}$  est entièrement déterminée par  $\tau(G)$ .

*Démonstration.* Pour tout morphisme de sous-graphe induit  $\varphi$  de  $F$  dans  $G$  (i.e. tel que  $\varphi(F)$  est un sous-graphe induit de  $G$ ), on introduit la relation d'équivalence  $\mathcal{R}$  suivante : pour deux sommets  $x, y$  de  $F$ ,  $x \mathcal{R} y$  si et seulement si  $\varphi(x) = \varphi(y)$ . On peut alors définir un graphe  $H = F/\mathcal{R}$ , dont chaque sommet est une classe d'équivalence, et pour lequel  $\{\bar{x}, \bar{y}\}$  est une arête si et seulement si  $\{\varphi(x), \varphi(y)\}$  est une arête dans  $G$  (ce qui est équivalent à  $\{x, y\}$  est une arête de  $F$ ).

Si  $\varphi$  n'est pas injectif, on observe d'une part que  $H$  possède strictement moins de sommets que  $F$ , et d'autre part que  $\varphi$  induit un morphisme de sous-graphe induit injectif de  $H$  dans  $G$ .

À l'aide de ces remarques on peut compter le nombre de morphismes de sous-graphe induit de  $F$  vers  $G$ , on note  $\text{HomInd}$  leur nombre.

- soit le morphisme est injectif, auquel cas, il y en a :  $\text{Ind}(F, G)$ ,
- sinon, il y en a autant que de graphes  $H$  issue d'une relation  $\mathcal{R}$  possible, fois le nombre de morphismes de sous-graphes induits injectifs de  $H$  dans  $G$ .

Au total,

$$\text{HomInd}(F, G) = \text{Ind}(F, G) + \sum_{\substack{H \in \mathcal{L} \\ v(H) < v(F)}} \text{HomInd}(H, F) \text{Ind}(H, G). \quad (3.4)$$

À normalisation près, les quantités  $(\text{HomInd}(H, G))_{H \in \mathcal{U}, v(H) < v(F)}$  (ainsi que  $\text{HomInd}(F, G)$ ) sont entièrement déterminées par  $\tau(G)$  d'après le lemme 3.3. En raisonnant par récurrence, on en déduit que les  $(\text{Ind}(H, G))_{v(H) < v(F)}$  sont eux aussi déterminés par  $\tau(G)$ . Ainsi 3.4 nous garantit que  $\text{Ind}(F, G)$  est entièrement déterminé par  $\tau(G)$ .  $\square$

Ces deux lemmes démontrés, on peut maintenant démontrer la proposition.

*Démonstration de la proposition 3.2.* Soient  $G$  et  $H$  deux graphes de même nombre de sommets. Supposons que  $\tau(G) = \tau(H)$ . D'après le lemme 3.4, pour tout  $F \in \mathcal{L}$ ,  $\mathbb{P}(F = \mathbf{G}'[k]) = \mathbb{P}(F = \mathbf{H}'[k])$ . En appliquant cela au graphe  $G$ , lui-même, et en remarquant que si  $F$  a autant de sommets que  $G$ ,  $\mathbb{P}(F = G[v(F)]) = \mathbb{P}(F = G)$  est égal à 1 si  $F = G$  et 0 sinon, on obtient donc que  $\mathbb{P}(G = H) = 1$ . Donc  $H = G$  dans  $\mathcal{U}$ .  $\square$

En particulier,  $\tau^+$  est une bijection de  $\mathcal{U}$  sur son image incluse dans  $[0, 1]^{\mathcal{U}^+} = [0, 1]^{\mathcal{U}} \times [0, 1]$ . On peut donc identifier  $\mathcal{U}$  à  $\tau^+(\mathcal{U})$ . L'ajout de la dernière composante  $v(G)^{-1}$  nous permet de caractériser plus simplement la convergence de suite de graphes dans son adhérence  $\bar{\mathcal{U}}$ . En effet, soit le nombre de sommets est constant à partir d'un certain rang, soit il tend vers l'infini.

Détaillons le cas où une suite  $(G_n)_{n \in \mathbb{N}}$  converge vers un graphe  $G$  de nombre de sommets fini. Il existe un rang  $N$  tel que pour tout  $n \geq N$ ,  $v(G_n) = v(G)$ , ainsi  $G_n \in L_{v(G)}$  à partir d'un certain rang.  $G_n$  prend donc un nombre fini de valeurs et converge, donc  $G_n$  est constant égal à  $G$  à partir d'un certain rang.

La projection  $\pi : [0, 1]^{\mathcal{U}^+} \rightarrow [0, 1]^{\mathcal{U}}$  établit un lien entre  $\tau^+(G)$  et  $\tau(G)$ , elle est continue pour les deux topologies que l'on a définies sur ces ensembles. Avec l'identification précédente, cette fonction peut être vue comme une fonction de  $\bar{\mathcal{U}}$  dans  $\bar{\mathcal{U}}^*$ .

Ceci nous indique qu'une suite de graphes converge dans  $\mathcal{U}$  si et seulement si son nombre de sommets converge vers un entier, et elle est alors constante à partir d'un certain rang. Ainsi l'ensemble  $\bar{\mathcal{U}}$  peut être divisé en deux ensembles disjoints  $\mathcal{U} \cup \bar{\mathcal{U}} \setminus \mathcal{U}$ , le second pouvant être identifié à la limite des suites dont le nombre de sommets diverge. On pose alors  $\mathcal{U}_\infty = \bar{\mathcal{U}} \setminus \mathcal{U}$ .

**Définition 12.** On nomme l'ensemble  $\mathcal{U}_\infty$  l'ensemble des **limites propres de graphes**, à savoir les éléments pouvant s'obtenir comme limite d'une suite de graphes dont la taille tend vers l'infini.

### 3.2 Convergence de graphes aléatoires

Nous avons pour le moment défini une notion de convergence de grands graphes déterministes, dans toute cette section, on s'intéressera particulièrement à une suite  $(G_n)_{n \in \mathbb{N}}$  de graphes aléatoires, dont la suite de variables aléatoires entières  $(v(G_n))_{n \in \mathbb{N}}$  tend vers l'infini en probabilité.

**Théorème 3.5.** Soit  $(G_n)_{n \in \mathbb{N}}$  une suite de graphes aléatoires non-étiquetés, tels que  $v(G_n) \rightarrow +\infty$  en probabilité. Alors les propositions suivantes sont équivalentes,

- (i)  $G_n \rightarrow \Gamma$  en loi, pour une certaine variable aléatoire  $\Gamma$  sur  $\bar{\mathcal{U}}$ ,
- (ii) Pour toute famille  $(F_1, \dots, F_m)$  de graphes déterministes, la suite de familles  $(t(F_1, G_n), \dots, t(F_m, G_n))_{n \in \mathbb{N}}$  converge en loi, pour la mesure produit,
- (iii) Pour tout graphe  $F \in \mathcal{U}$ , la suite de variables aléatoires  $(t(F, G_n))_{n \in \mathbb{N}}$  converge en loi,
- (iv) Pour tout graphe  $F \in \mathcal{U}$ , la suite des espérances  $(\mathbb{E}(t(F, G_n)))_{n \in \mathbb{N}}$  converge.

Par ailleurs, si l'une de ces propriétés est vraie, les limites sont respectivement  $(t(F_i, \Gamma))_{1 \leq i \leq m}$ ,  $t(F, \Gamma)$  et  $\mathbb{E}(t(F, \Gamma))$ . De plus,  $\Gamma \in \mathcal{U}_\infty$  presque sûrement.

*Démonstration.* (i)  $\Leftrightarrow$  (ii). Comme  $v(G_n) \rightarrow +\infty$ ,  $v(G_n)^{-1} \rightarrow 0$ , on ne s'intéresse donc qu'à la convergence en loi de la famille  $((t(F, G_n))_{F \in \mathcal{U}})_{n \in \mathbb{N}}$ , c'est à dire que pour toute fonction bornée Il s'agit de la définition de la convergence pour la topologie produit, à savoir qu'un élément converge si et seulement si chaque cylindre fini est convergent.

(ii)  $\Rightarrow$  (iii). Cette implication est immédiate.

(iii)  $\Rightarrow$  (iv). On applique ici le théorème de convergence dominée, en justifiant l'hypothèse de domination par le fait que  $t$  est bornée par 1.

(iv)  $\Rightarrow$  (ii). Le résultat repose sur l'observation suivante, pour une famille de graphes  $(F_1, \dots, F_m)$ , en considérant le graphe  $F$ , réunion disjointe des graphes de ces graphes,

$$t(F, G_n) = \mathbb{P}(F \subset \mathbf{G}_n[k]) = \prod_{i=1}^m \mathbb{P}(F_i \subset \mathbf{G}_n[k]) = \prod_{i=1}^m t(F_i, G_n), \quad (3.5)$$

que l'on peut justifier en termes de nombre morphismes de  $F$  vers  $G$ , en effet un tel morphisme peut-être construit en choisissant un morphisme pour chaque graphe  $F_i$  vers  $G$ , ainsi,  $\text{hom}(F, G) = \prod_{i=1}^m \text{hom}(F_i, G)$ , et on conclut en remarquant que  $v(G)^{v(F)} = \prod_{i=1}^m v(G)^{v(F_i)}$ . En particulier, pour tout graphe  $F$ , pour tout entier  $k$ ,  $\mathbb{E}(t(F, G_n)^k)$  converge. On fixe maintenant un entier  $m$ . D'après le théorème de Stone-Weierstrass, les fonctions polynomiales à

$m$  variables sont denses dans l'ensemble des fonctions continues de  $[0, 1]^m$  à valeurs réelles, munie de la norme infinie. Ainsi pour toute fonction  $\varphi$  continue de  $[0, 1]^m$  dans  $\mathbb{R}$ , il existe  $(P_k)_{k \in \mathbb{N}}$  convergeant uniformément vers  $\varphi$ , par linéarité de l'espérance et avec la remarque précédente,  $\lim_{n \rightarrow +\infty} \mathbb{E}(\varphi(t(F_1, G_n) \dots, t(F_m, G_n))) = \mathbb{E}(\varphi(t(F_1, G) \dots, t(F_m, G)))$ , en utilisant les polynômes précédents. Ce qui prouve le sens voulu.  $\square$

**Corollaire 3.6.** Soit  $\Gamma$  un élément aléatoire de  $\mathcal{U}_\infty = \overline{\mathcal{U}} \setminus \mathcal{U}$ , alors pour chaque famille  $F_1, \dots, F_m$  de graphes, non forcément distincts,

$$\mathbb{E} \left( \prod_{i=1}^m t(F_i, \Gamma) \right) = \mathbb{E} \left( t \left( \bigoplus_{i=1}^m F_i, \Gamma \right) \right),$$

où  $\bigoplus_{i=1}^m F_i$  désigne l'union disjointe des  $F_1, \dots, F_m$ .

La distribution de  $\Gamma$  est déterminée entièrement par la famille  $(\mathbb{E}(t_{\text{ind}}(F, \Gamma)))_{F \in \mathcal{U}}$ .

*Démonstration.* Comme  $\mathcal{U}$  est dense dans  $\overline{\mathcal{U}}$ , il existe une suite  $(G_n)_{n \in \mathbb{N}}$  convergeant presque sûrement vers  $\Gamma$ . En particulier  $G_n$  tend vers  $\Gamma$  en loi et  $v(G_n)$  tend vers l'infini en probabilités. D'après le théorème 3.5, et on déduit l'égalité voulue par 3.5, en passant à la limite.

Ce point nous montre que l'ensemble  $(\mathbb{E}(t(F, \Gamma)))_{F \in \mathcal{U}}$  détermine entièrement la loi de la distribution jointe  $(t(F_i, \Gamma))_{1 \leq i \leq m}$  et donc, d'après le théorème 3.5, la loi de  $\tau(\Gamma)$ , donc la loi de  $\Gamma$  par identification. Comme  $v(\Gamma) = \infty$ , 3.1 nous donne que  $t_{\text{inj}}(F, \Gamma) = t(F, \Gamma)$ , et 3.2 nous garantit que  $(\mathbb{E}(t_{\text{ind}}(F, \Gamma)))_{F \in \mathcal{U}}$  détermine la loi de  $(\mathbb{E}(t(F, \Gamma)))_{F \in \mathcal{U}}$ .  $\square$

On introduit maintenant l'ensemble  $\mathcal{L}_\infty$  des graphes infinis étiquetés par l'ensemble de sommets  $\mathbb{N}$ . On peut voir chacun de ses graphes comme un ensemble d'arêtes à savoir un sous-ensemble de  $\mathcal{P}_2(\mathbb{N})$ , et ainsi identifier  $\mathcal{L}_\infty$  à l'ensemble produit  $\{0, 1\}^{\mathcal{P}_2(\mathbb{N})}$ , que l'on peut munir la topologie produit. Ainsi  $\mathcal{L}_\infty$  est un espèce métrique compact.

On pourra plonger chaque ensemble  $\mathcal{L}_n$  dans  $\mathcal{L}_nfty$  en ajoutant un nombre infini de sommets isolés. Réciproquement pour un graphe  $H \in \mathcal{L}_\infty$  on pourra considérer le sous-graphe  $H|_{[n]} \in \mathcal{L}_n$  induit par  $\{1, \dots, n\}$  dans  $H$ .

De plus un graphe  $G$  (étiqueté ou non), peut être considéré comme un graphe étiqueté aléatoire  $\widehat{G}$  en lui attribuant un étiquetage aléatoire de ses sommets. On pourra par la suite plonger  $\widehat{G}$  dans  $\mathcal{L}_\infty$  comme suggéré précédemment.

Le théorème suivant présente une caractérisation de la convergence dans  $\mathcal{L}_\infty$  pour la topologie produit associée, à l'aide de la topologie de  $\mathcal{U}$ .

**Théorème 3.7.** Soit  $(G_n)_{n \in \mathbb{N}}$  une suite de graphes aléatoires à valeurs dans  $\mathcal{U}$ , tel que  $v(G_n)$  tend en probabilité vers l'infini. Les propositions suivantes sont équivalentes,

- (i)  $G_n \rightarrow \Gamma$  en loi, pour une certaine variable aléatoire  $\Gamma$  sur  $\overline{\mathcal{U}}$ ,
- (ii)  $\widehat{G}_n \rightarrow H$  en loi, pour une certaine variable aléatoire  $H$  sur  $\mathcal{L}_\infty$ .

Si l'une des deux assertions est vérifiée, alors pour tout  $F \in \mathcal{L}_k$ ,  $\mathbb{P}(H|_{[k]} = F) = \mathbb{E}(t_{\text{ind}}(F, \Gamma))$ , et  $\Gamma \in \mathcal{U}_\infty$  presque sûrement.

*Démonstration.* Soit  $G$  un graphe étiqueté, on considère  $\widehat{G}|_{[k]}$ , avec  $k \leq v(G)$ . Ce graphe aléatoire a la même loi que la variable aléatoire,  $\mathbf{G}'[k]$ . Ainsi pour tout  $F \in \mathcal{L}_k$ , si  $k \leq v(G)$ ,

$$\mathbb{P}(\widehat{G}|_{[k]} = F) = t_{\text{ind}}(F, G).$$

On obtient alors pour le graphe aléatoire  $G_n$ ,

$$\mathbb{E}(t_{\text{ind}}(F, G_n)) \leq \mathbb{P}(\widehat{G}_n|_{[k]} = F) \leq \mathbb{E}(t_{\text{ind}}(F, G_n)) + \mathbb{P}(v(G_n) < k)$$

Par hypothèse,  $\mathbb{P}(v(G_n) < k) \rightarrow 0$  lorsque  $n \rightarrow +\infty$ , de l'encadrement précédent et du théorème 3.5, que  $(G_n)_{n \in \mathbb{N}}$  converge en loi vers  $\Gamma$  dans  $\overline{\mathcal{U}}$  si et seulement si, pour tout  $k \geq 1$ , et tout  $F \in \mathcal{L}_k$ ,

$$\mathbb{P}(\widehat{G}|_{[k]} = F) \rightarrow \mathbb{E}(t_{\text{ind}}(F, G_n))$$

Comme  $\mathcal{L}_k$  est fini, on déduit de l'équation précédente que  $\widehat{G}_n \Big|_{[k]}$  converge en loi vers un certain graphe aléatoire  $H_k$  à valeurs dans  $\mathcal{L}_k$ , tel que  $\mathbb{P}(H_k = F) = \mathbb{E}(t_{\text{ind}}(F, \Gamma))$ , pour tout  $F \in \mathcal{L}_k$ . Comme  $\mathcal{L}_\infty$  est muni de la topologie produit, la convergence est équivalente à la convergence des cylindres finis, ce qui implique que  $\widehat{G}_n$  converge en loi vers une certaine variable aléatoire  $H$  à valeurs dans  $\mathcal{L}_\infty$ , tel que pour tout  $k$ ,  $H|_{[k]}$  soit égal en loi à  $H_k$ .

Réciproquement si  $\widehat{G}_n$  converge en loi vers  $H \in \mathcal{L}_\infty$ , alors  $\widehat{G} \Big|_{[k]}$  converge en loi vers  $H|_{[k]}$ , d'où l'on déduit de l'encadrement ci-dessus que, lorsque  $n \rightarrow +\infty$ , pour tout  $F \in \mathcal{L}_k$ ,

$$\mathbb{E}(t_{\text{ind}}(F, G_n)) = \mathbb{P}(\widehat{G} \Big|_{[k]} = F) + o(1) \rightarrow \mathbb{P}(H|_{[k]} = F).$$

On en déduit d'après le théorème 3.5 l'existence d'une variable aléatoire  $\Gamma \in \mathcal{U}_\infty$  telle que  $G_n$  converge en loi vers  $\Gamma$  et  $\mathbb{E}(t_{\text{ind}}(F, \Gamma)) = \mathbb{P}(H|_{[k]} = F)$ .  $\square$

## 4 Graphes aléatoires échangeables

Dans cette section, nous introduisons les graphes aléatoires échangeables et donnons une correspondance bijective entre les graphes aléatoires échangeables infinis et les distributions sur l'espace des limites propres des graphes 4.4, ce qui donne en particulier une correspondance bijective entre les limites propres des graphes et les points extrêmes de l'ensemble des distributions sur les graphes aléatoires échangeables 4.5. Ces résultats nous permettent d'explicitier le lien entre les limites propres des graphes, telles que développées par Lovasz et Szegedy [9] et présentées dans la section précédente, et la théorie des représentations d'Aldous et Hoover [1][7].

**Définition 13.** Un **graphe** aléatoire infini  $H \in \mathcal{L}_\infty$  est dit **échangeable** si sa distribution est invariante sous toute permutation à support fini de ses sommets.

Remarquons qu'il est équivalent de ne considérer que les permutations à support finies, c'est-à-dire les permutations  $\sigma$  de  $\mathbb{N}$  qui vérifient  $\sigma(i) = i$  pour tout  $i$  suffisamment grand, donc  $\sigma$  peut être considéré comme une permutation dans  $\mathfrak{S}_n$  pour un  $n$ .

De manière équivalente, si on note  $H = (\mathbb{N}, E \subseteq \mathbb{N}^2)$  et  $X_{ij} := \mathbf{1}_{\{i,j\} \in E}$  l'indicateur de l'existence de l'arête  $\{i, j\}$  dans  $H$ , alors le graphe  $H$  est échangeable si et seulement si le tableau  $\{X_{ij}\}$ ,  $1 \leq i, j < \infty$ , est un tableau échangeable comme défini dans Définition 7.

Avec la représentation des tableaux échangeables on peut montrer qu'en effet une sondage infini d'un graphe  $G$  donne lieu à un graphe échangeable.

**Proposition 4.1.** Soit  $G \in \mathcal{U}$  un graphe. Alors son  $\infty$ -sondage  $\mathbf{G}[\infty]$  est un graphe aléatoire infini échangeable.

*Démonstration.* Soit  $G = (V, E) \in \mathcal{U}_n$  avec sommets  $V = \{v_1, v_2, \dots, v_n\}$  et soit  $A_G$  sa matrice d'adjacence. Définissons la fonction étagée  $W_G : [0, 1]^2 \rightarrow [0, 1]$ , définie en partitionnant  $[0, 1]$  en  $n$  intervalles  $I_1, I_2, \dots, I_n$  de même longueur, i.e. tels que l'intérieur de  $I_j$  soit  $(\frac{j-1}{n}, \frac{j}{n})$ , et pour chaque  $(x, y) \in I_i \times I_j$ , on pose  $W_G(x, y) = (A_G)_{ij}$ , i.e.  $W_G(x, y) = 1$  si  $\{v_i, v_j\} \in E$  et  $W_G(x, y) = 0$  sinon. Cette fonction  $W_G$  est appelée le **graphon associé au graphe**  $G$ . Soient  $U_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{U}[0, 1]$  et  $V_i$  i.i.d. uniformes sur  $V$ . Soit  $X$  la matrice d'adjacence de  $\mathbf{G}[V_1, V_2, \dots]$ . Alors,

$$\mathbb{P}(X_{ij} = 1) = \mathbb{P}(\{V_i, V_j\} \in E) = \mathbb{E}W_G(U_i, U_j).$$

Par le Théorème 2.7 on conclut que la loi de  $\mathbf{G}[\infty]$  est  $P_{W_G}$  qui est bien échangeable.  $\square$

Le lemme suivant donne deux conditions techniques équivalentes à l'échangeabilité qui seront utiles.

**Lemme 4.2.** Soit  $H \in \mathcal{L}_\infty$  un graphe infini aléatoire. Alors les propositions suivantes sont équivalentes :

- (i)  $H$  est échangeable.
- (ii)  $H|_{[k]}$  a une distribution invariante sous toutes les permutations dans  $\mathfrak{S}_k$ , pour tout  $k \geq 1$ .
- (iii)  $\mathbb{P}(H|_{[k]} = F)$  ne dépend que de la classe d'isomorphisme de  $F$ , et peut donc être vu comme une fonction de la classe d'équivalence de  $F$  dans  $\mathcal{U}_k$ , pour chaque  $k \geq 1$ .

*Démonstration.* Soit  $\sigma \in \mathfrak{S}_n$ . Nous notons  $H \circ \sigma$  le graphe obtenu en renommant les  $n$  premiers sommets de  $H$  selon  $\sigma$ . Alors,  $H$  est échangeable si et seulement si  $H \circ \sigma \stackrel{\text{loi}}{=} H$  pour chaque  $\sigma \in \mathfrak{S}_n$  et  $n \in \mathbb{N}$ .

(i)  $\implies$  (ii). En prenant  $\sigma \in \mathfrak{S}_k$ , nous avons, grâce à l'échangeabilité,  $H \circ \sigma \stackrel{\text{loi}}{=} H$ , et en restreignant aux premiers  $k$  sommets, nous obtenons

$$H|_{[k]} \stackrel{\text{loi}}{=} H \circ \sigma|_{[k]} = H|_{[k]} \circ \sigma.$$

Par conséquent,  $H|_{[k]}$  possède une distribution invariante sous toutes les permutations de  $[k]$ .

(ii)  $\implies$  (i). Si  $\sigma$  est une permutation finie de  $\mathbb{N}$ , alors  $\sigma$  se restreint à une permutation de  $[k]$  pour chaque  $k$  suffisamment grand, et il en résulte que pour tous grands  $k$

$$H \circ \sigma|_{[k]} = H|_{[k]} \circ \sigma \stackrel{\text{loi}}{=} H|_{[k]},$$

ce qui implique  $H \circ \sigma \stackrel{\text{loi}}{=} H$ .

(ii)  $\implies$  (iii). Soient  $F$  et  $F' \in \mathcal{U}_k$  deux graphes isomorphes, dont on étiquette les sommets avec  $[k]$  d'une manière quelconque, ce qui est équivalent à l'existence d'une permutation  $\sigma \in \mathfrak{S}_k$  telle que  $F' \circ \sigma = F$ . On veut prouver que  $\mathbb{P}(H|_{[k]} = F') = \mathbb{P}(H|_{[k]} = F)$ . Comme  $\sigma$  est une permutation, nous avons

$$\mathbb{P}(H|_{[k]} = F') = \mathbb{P}(H|_{[k]} \circ \sigma = F)$$

mais grâce à l'invariance de  $H|_{[k]}$  sous ré-étiquetage :

$$\mathbb{P}(H|_{[k]} \circ \sigma = F) = \mathbb{P}(H|_{[k]} = F)$$

et nous concluons.

(iii)  $\implies$  (ii). Soit  $\sigma \in \mathfrak{S}_k$  et  $F \in \mathcal{L}_k$ . Pour montrer que la lois de  $H|_{[k]}$  est invariante sous ré-étiquetage on montre que

$$\mathbb{P}(H|_{[k]} \circ \sigma = F) = \mathbb{P}(H|_{[k]} = F).$$

Soit  $\sigma^{-1} \in \mathfrak{S}_k$  l'inverse de la permutation  $\sigma$ , alors par l'hypothèse (iii) nous avons :

$$\mathbb{P}(H|_{[k]} = F) = \mathbb{P}(H|_{[k]} = F \circ \sigma^{-1})$$

et comme

$$\mathbb{P}(H|_{[k]} = F \circ \sigma^{-1}) = \mathbb{P}(H|_{[k]} \circ \sigma = F)$$

nous concluons.  $\square$

**Théorème 4.3.** La limite  $H$  du Théorème 3.7 est échangeable c'est-à-dire que si une suite de graphes aléatoire  $G_n$  a une limite propre (dans  $\overline{\mathcal{U}}$ ) alors la suite aléatoirement étiquetée  $\widehat{G}_n$  converge (dans  $\mathcal{L}_\infty$ ) vers un graphe échangeable.

*Démonstration.*  $H$  satisfait le Lemme 4.2(iii) et alors  $H$  est échangeable. En effet, avec les notation du Théorème 3.7 on a  $\mathbb{P}(H|_{[k]} = F) = \mathbb{E}t_{\text{ind}}(F, \Gamma)$  pour tout  $F \in \mathcal{L}_k$  et, en particulier, cette probabilité ne dépend que de la classe d'isomorphisme de  $F$  par définition de  $t_{\text{ind}}$ .  $\square$

Une question naturelle maintenant serait de savoir si chaque loi de graphe infini échangeable correspond à une loi de limite propre de graphes. Le théorème principal dans l'article de Diaconis et Janson [4] affirme en effet qu'il existe une bijection entre les deux ensembles de distributions. Le Théorème 3.7 implique la connexion suivante entre les éléments aléatoires  $\Gamma \in \mathcal{U}_\infty$  et les graphes aléatoires  $H \in \mathcal{L}_\infty$  :

**Théorème 4.4.** Il existe une correspondance bijective entre les distributions des éléments aléatoires  $\Gamma \in \mathcal{U}_\infty$  et les distributions des graphes aléatoires infinis échangeables  $H \in \mathcal{L}_\infty$  donnée par

$$\mathbb{E}t_{\text{ind}}(F, \Gamma) = \mathbb{P}(H|_{[k]} = F) \tag{4.1}$$

pour tout  $k \geq 1$  et tout  $F \in \mathcal{L}_k$ , ou, de manière équivalente,

$$\mathbb{E}t(F, \Gamma) = \mathbb{P}(H \supset F) \tag{4.2}$$

pour tout  $F \in \mathcal{L}$ . De plus,  $H|_{[n]} \xrightarrow{\text{loi}} \Gamma$  dans  $\overline{\mathcal{U}}$  lorsque  $n \rightarrow \infty$ .

*Démonstration.* Remarquons d'abord que (4.1) et (4.2) sont équivalents. D'un côté puisque  $t(F, \Gamma) = t_{\text{inj}}(F, \Gamma)$  par (3.1) on peut écrire la relation (3.2) sous la forme

$$\mathbb{E}t(F, \Gamma) = \sum_{F' \in \mathcal{L}_k, F' \supset F} \mathbb{E}t_{\text{ind}}(F', \Gamma).$$

De l'autre côté, puisque  $H \supset F$  si et seulement si  $H|_{[k]} \supseteq F$  lorsque  $F \in \mathcal{L}_k$  on a

$$\mathbb{P}(H \supset F) = \sum_{F' \in \mathcal{L}_k, F' \supset F} \mathbb{P}(F' = H|_{[k]}).$$

Donc (4.1) implique (4.2). Pour l'autre direction on résonne de même en utilisant la relation

$$\mathbb{E}t_{\text{ind}}(F, \Gamma) = \sum_{F' \in \mathcal{L}_k, F' \supset F} (-1)^{e(F') - e(F)} \mathbb{E}t_{\text{inj}}(F', \Gamma)$$

obtenue par le principe d'inclusion-exclusion (3.3).

Supposons que  $\Gamma$  soit un élément aléatoire de  $\mathcal{U}_\infty \subset \bar{\mathcal{U}}$ . Puisque  $\mathcal{U}$  est dense dans  $\bar{\mathcal{U}}$ , il existe (comme dans la démonstration du Corollaire 3.6) des graphes aléatoires non étiquetés  $G_n$  tels que  $G_n \xrightarrow{\text{p.s.}} \Gamma$  dans  $\bar{\mathcal{U}}$ , et donc  $v(G_n) \xrightarrow{\text{p.s.}} \infty$  et  $G_n \xrightarrow{\text{loi}} \Gamma$ . Ainsi, les Théorèmes 3.7 et 4.3 montrent que  $\widehat{G}_n \xrightarrow{\text{loi}} H$  pour un certain graphe infini aléatoire échangeable  $H$  satisfaisant (4.1). De plus, (4.1) détermine la lois de  $H|_{[k]}$  pour tout  $k$ , et donc la lois de  $H$ .

Inversement, si  $H$  est un graphe aléatoire infini échangeable, soit  $G_n \stackrel{\text{def}}{=} H|_{[n]}$ . Par le Lemme 4.2(ii), la distribution de chaque  $G_n$  est invariante par permutations de sommets, donc si  $\widehat{G}_n$  est  $G_n$  avec un (re)étiquetage aléatoire, nous avons  $\widehat{G}_n \stackrel{\text{loi}}{=} G_n$ . Puisque  $G_n \xrightarrow{\text{loi}} H$  dans  $\mathcal{L}_\infty$  (par la topologie produit), on a  $\widehat{G}_n \xrightarrow{\text{loi}} H$  dans  $\mathcal{L}_\infty$ , donc le Théorème 3.7 s'applique et montre l'existence d'un  $\Gamma$  aléatoire dans  $\mathcal{U}_\infty$  tel que  $G_n \xrightarrow{\text{loi}} \Gamma$  et (4.1) est satisfaite. Enfin, (4.1) détermine la distribution de  $\Gamma$  selon le Corollaire 3.6.  $\square$

**Remarque 3.** Observons que,  $H|_{[n]}$  converge p.s. vers une variable aléatoire  $\Gamma \in \mathcal{U}_\infty$ , car

$$\{t_{\text{ind}}(F, H|_{[n]})\}_{n \geq v(F)}$$

est une martingale rétrograde pour tout graphe  $F$ .

**Corollaire 4.5.** Il existe une correspondance bijective entre les éléments  $\Gamma$  de  $\mathcal{U}_\infty$  et les points extrêmes de l'ensemble des distributions de graphes aléatoires infinis échangeables  $H \in \mathcal{L}_\infty$ . Cette correspondance est donnée par

$$t(F, \Gamma) = \mathbb{P}(H \supset F)$$

pour tout  $F \in \mathcal{L}$ . De plus,  $H|_{[n]} \xrightarrow{\text{p.s.}} \Gamma$  dans  $\bar{\mathcal{U}}$  lorsque  $n \rightarrow \infty$ .

*Démonstration.* Le Théorème 4.4 nous dit que l'ensemble des distributions sur  $\mathcal{U}_\infty$  et l'ensemble des lois échangeables sur  $\mathcal{L}_\infty$  sont en bijection. Alors en particulier les points extrêmes de ces deux ensembles sont aussi en bijection. Or les points extrêmes de l'ensemble des distributions sur  $\mathcal{U}_\infty$  sont les masses ponctuelles  $\delta_\Gamma$ , qui sont en correspondance bijective avec les éléments de  $\mathcal{U}_\infty$ .  $\square$

Le théorème suivant nous permet de caractériser ces distributions de points extrêmes des graphes infinis aléatoires échangeables.

**Théorème 4.6.** Soit  $H$  un graphe infini aléatoire échangeable. Alors les énoncés suivants sont équivalents :

- (i) La distribution de  $H$  est un point extrême de l'ensemble des distributions échangeables dans  $\mathcal{L}_\infty$ .

- (ii) Si  $F_1$  et  $F_2$  sont deux graphes (finis) avec des ensembles de sommets disjoints  $V(F_1)$ ,  $V(F_2) \subset \mathbb{N}$ , alors

$$\mathbb{P}(H \supset F_1 \cup F_2) = \mathbb{P}(H \supset F_1) \mathbb{P}(H \supset F_2).$$

- (iii) Les restrictions  $H|_{[k]}$  et  $H|_{[k+1, \infty)}$  sont indépendantes pour tout  $k$ .  
(iv) Soit  $\mathcal{F}_n$  la  $\sigma$ -algèbre engendrée par  $H|_{[n, \infty)}$ . Alors la  $\sigma$ -algèbre terminale  $\bigcap_{n=1}^{\infty} \mathcal{F}_n$  est triviale, c'est-à-dire qu'elle ne contient que des événements de probabilité 0 ou 1.

*Démonstration.* (i)  $\implies$  (ii). Selon le corollaire 4.5,  $H$  correspond à un certain  $\Gamma \in \mathcal{U}_{\infty}$  (non aléatoire) tel que

$$\mathbb{P}(H \supset F) = t(F, \Gamma) \quad (4.3)$$

pour tout  $F \in \mathcal{L}$ . Nous avons défini  $\mathcal{L}$  de telle manière qu'un graphe  $F \in \mathcal{L}$  est étiqueté de 1 à  $v(F)$ , mais les deux côtés de (4.3) sont invariants sous un étiquetage arbitraire de  $F$  par des entiers positifs; le côté gauche parce que  $H$  est échangeable et le côté droit parce que  $t(F, \Gamma)$  dépend uniquement de  $F$  en tant que graphe non étiqueté. Par conséquent, (4.3) est vérifié pour tout graphe fini  $F$  avec  $V(F) \subset \mathbb{N}$ .

De plus, étant donné que  $\Gamma$  est non aléatoire, le corollaire 3.6 donne  $t(F_1 \cup F_2, \Gamma) = t(F_1, \Gamma) t(F_2, \Gamma)$ . Par conséquent,

$$\mathbb{P}(H \supset F_1 \cup F_2) = t(F_1 \cup F_2, \Gamma) = t(F_1, \Gamma) t(F_2, \Gamma) = \mathbb{P}(H \supset F_1) \mathbb{P}(H \supset F_2).$$

(ii)  $\implies$  (iii). Par inclusion-exclusion, (ii) implique que si  $1 \leq k < l < \infty$ , alors pour tout graphes  $F_1$  et  $F_2$  avec  $V(F_1) = \{1, \dots, k\}$  et  $V(F_2) = \{k+1, \dots, k+l\}$ , les événements  $H|_{[k]} = F_1$  et  $H|_{\{k+1, \dots, l\}} = F_2$  sont indépendants. Ainsi,  $H|_{[k]}$  et  $H|_{\{k, \dots, l\}}$  sont indépendants pour chaque  $l > k$ , et le résultat s'ensuit.

(iii)  $\implies$  (iv). Supposons que  $A$  soit un événement dans la  $\sigma$ -algèbre terminale  $\bigcap_{n=1}^{\infty} \mathcal{F}_n$ . Soit  $\mathcal{F}_n^*$  la  $\sigma$ -algèbre engendrée par  $H|_{[n]}$ . D'après (iii),  $A$  est indépendant de  $\mathcal{F}_n^*$  pour chaque  $n$ , et donc de la  $\sigma$ -algèbre  $\mathcal{F}$  engendrée par  $\cup \mathcal{F}_n^*$ , qui est égale à la  $\sigma$ -algèbre  $\mathcal{F}_1$  engendrée par  $H$ . Cependant,  $A \in \mathcal{F}_1$ , donc  $A$  est indépendant de lui-même et donc  $\mathbb{P}(A) = 0$  ou 1.

(iv)  $\implies$  (i). Soit  $F \in \mathcal{L}_k$  pour un certain  $k$ , et soit  $F_n$  l'ensemble  $F$  dont tous les sommets ont été décalés de  $n$ . Considérons les deux indicateurs  $I = \mathbf{1}[H \supseteq F]$  et  $I_n = \mathbf{1}[H \supseteq F_n]$ . Puisque  $I_n$  est mesurable par rapport à  $\mathcal{F}_n$ ,

$$\mathbb{P}(H \supset F \cup F_n) = \mathbb{E}(II_n) = \mathbb{E}(\mathbb{E}(I | \mathcal{F}_n) I_n). \quad (4.4)$$

De plus,  $\mathbb{E}(I | \mathcal{F}_n)$ ,  $n = 1, 2, \dots$ , est une martingale rétrograde, et donc p.s.

$$\mathbb{E}(I | \mathcal{F}_n) \rightarrow \mathbb{E} \left( I \mid \bigcap_{n=1}^{\infty} \mathcal{F}_n \right) = \mathbb{E}I$$

en utilisant (iv). Ainsi,  $(\mathbb{E}(I | \mathcal{F}_n) - \mathbb{E}I) I_n \rightarrow 0$  p.s., et par convergence dominée

$$\mathbb{E}((\mathbb{E}(I | \mathcal{F}_n) - \mathbb{E}I) I_n) \rightarrow 0$$

En conséquence, (4.4) donne

$$\mathbb{P}(H \supset F \cup F_n) = \mathbb{E}I \mathbb{E}I_n + o(1) = \mathbb{P}(H \supset F) \mathbb{P}(H \supset F_n) + o(1).$$

De plus, puisque  $H$  est échangeable,  $\mathbb{P}(H \supset F \cup F_n)$  (pour  $n \geq v(F)$ ) et  $\mathbb{P}(H \supset F_n)$  ne dépendent pas de  $n$ , et nous obtenons lorsque  $n \rightarrow \infty$

$$\mathbb{P}(H \supset F \cup F_k) = \mathbb{P}(H \supset F)^2. \quad (4.5)$$

Soit  $\Gamma$  un élément aléatoire de  $\mathcal{U}_{\infty}$  correspondant à  $H$  comme dans le théorème 4.4 Par (4.4) et corollaire 3.6, (4.5) peut être écrit

$$\mathbb{E}t(F, \Gamma)^2 = (\mathbb{E}t(F, \Gamma))^2$$

Ainsi, la variable aléatoire  $t(F, \Gamma)$  a une variance de 0, elle est donc constante. Or cela est vrai pour tout  $F \in \mathcal{L}$ , il s'ensuit que  $\Gamma$  est constante p.s., c'est-à-dire que nous pouvons prendre  $\Gamma$  non aléatoire, et (i) découle du Corollaire 4.5.  $\square$

## 4.1 Représentations des limites de graphes et des graphes échangeables

D'un côté on a une caractérisation des tableaux échangeables et en particulier des graphes aléatoires infinis échangeables par le théorème de Aldous-Hoover (2.7). De l'autre côté on a une caractérisation des limites de graphes  $\mathcal{U}_\infty$  par les graphons. Dans cette sous-section on va expliciter le lien entre ces deux caractérisations.

Rappelons le modèle des graphes générés par les graphons. Soit  $W : [0, 1]^2 \rightarrow [0, 1]$  un graphon. Alors le graphe aléatoire infini

$$\mathcal{G}(\infty, W) \in \mathcal{L}_\infty$$

est construit de la manière suivante : nous commençons par choisir une séquence  $X_1, X_2, \dots$  de variables aléatoires i.i.d. uniformément distribuées sur  $[0, 1]$ , puis, étant donné cette séquence, pour chaque paire  $(i, j)$  avec  $i < j$ , nous ajoutons une arête  $\{i, j\}$  avec une probabilité  $W(X_i, X_j)$ , indépendamment pour toutes les paires  $(i, j)$  avec  $i < j$  (conditionnellement à  $\{X_i\}$ ).

Le graphe  $\mathcal{G}(\infty, W)$  est un graphe aléatoire infini échangeable de loi  $P_W$ , et le résultat d'Aldous et Hoover est que chaque graphe aléatoire infini échangeable est obtenu comme un mélange de tels  $\mathcal{G}(\infty, W)$ . En d'autres termes, tout graphe infini échangeable se met sous la forme  $\mathcal{G}(\infty, W)$  avec un  $W$  aléatoire.

En considérant à nouveau un  $W$  déterministe de  $\mathcal{W}_s$ , il est évident que  $\mathcal{G}(\infty, W)$  satisfait la condition (ii) de Théorème 4.6, et donc la loi de  $\mathcal{G}(\infty, W)$  est un point extrême de l'ensemble des distributions échangeables dans  $\mathcal{L}_\infty$ . Alors le Corollaire 4.5 montrent que  $\mathcal{G}(\infty, W)$  correspond à un élément  $\Gamma_W \in \mathcal{U}_\infty$ . De plus, selon le Théorème 4.4 et la Remarque 3,  $\mathcal{G}(n, W) \rightarrow \Gamma_W$  p.s. lorsque  $n \rightarrow \infty$ . En plus par (4.3) on a

$$t(F, \Gamma_W) = \mathbb{P}(F \subseteq \mathcal{G}(k, W))$$

pour tout  $F \in \mathcal{L}_k$ , et alors

$$t(F, \Gamma_W) = \int_{[0,1]^k} \prod_{ij \in \mathbb{E}(F)} W(x_i, x_j) dx_1 \dots dx_k. \quad (4.6)$$

Le résultat principal de Lovász et Szegedy [9] est que chaque élément de  $\mathcal{U}_\infty \cong \overline{\mathcal{U}^*}$  peut être obtenu comme  $\Gamma_W$  satisfaisant 4.6 pour certain graphon  $W \in \mathcal{W}_s$ .

Les théorèmes de représentation d'Aldous-Hoover 2.7 et de Lovász et Szegedy [9] sont liés par le Théorème 4.4 et le Corollaire 4.5 ci-dessus, et une caractérisation découle facilement de l'autre.

**Remarque 4.** Les représentations par  $W$  ne sont pas uniques. Par exemple, si  $\varphi : [0, 1] \rightarrow [0, 1]$  est une application préservant la mesure, alors  $W$  et  $\widehat{W} = W \circ \varphi$ , définie par  $\widehat{W}(x, y) = W(\varphi(x), \varphi(y)) := W(\varphi(x), \varphi(y))$ , définissent la même limite de graphe et le même tableau échangeable (en distribution).

Bien que cela soit en principe la seule source de non-unicité, les détails sont plus compliqués, principalement parce que l'application préservant la mesure  $\varphi$  n'a pas besoin d'être une bijection, et donc la relation  $\widehat{W} = W \circ \varphi$  n'est pas symétrique : elle peut être vraie sans qu'il existe une application préservant la mesure  $\varphi'$  telle que  $W = \widehat{W} \circ \varphi'$ . Pour un exemple en une dimension, considérez  $f(x) = x$  et  $f'(x) = \varphi(x) = 2x \bmod 1$  ; pour un exemple en deux dimensions, posons  $W(x, y) = f(x)f(y)$  et  $\widehat{W}(x, y) = f'(x)f'(y)$ .

Pour les tableaux échangeables Kallenberg a démontré [7] que le problème d'équivalence se réduit à l'existence des applications préservant la mesure  $\varphi, \varphi' : [0, 1] \rightarrow [0, 1]$  telles que  $W \circ \varphi = \widehat{W} \circ \varphi'$  p.p. sur  $[0, 1]^2$ , i.e.  $W(\varphi(x), \varphi(y)) = \widehat{W}(\varphi'(x), \varphi'(y))$  p.p.

**Remarque 5.** On peut considérer de manière plus générale une fonction mesurable symétrique  $W : \mathcal{S}^2 \rightarrow [0, 1]$  pour tout espace de probabilité  $(\mathcal{S}, \mu)$ , et définir  $\mathcal{G}(\infty, W)$  comme précédemment avec des variables aléatoires i.i.d.  $X_i$  dans  $\mathcal{S}$  avec une distribution  $\mu$ . Cela ne

fournit pas de nouveaux objets limites  $\mathcal{G}(\infty, W)$  ou  $\Gamma_W$ , mais ils peuvent parfois donner des représentations utiles.

Un cas intéressant est lorsque  $W$  est la matrice d'adjacence d'un graphe (fini)  $G$ , avec  $\mathcal{S} = V(G)$  et  $\mu$  la mesure uniforme sur  $\mathcal{S}$ ; nous laissons donc  $X_i$  être des sommets aléatoires i.i.d. de  $G$  et  $G(n, W)$  équivaut au graphe aléatoire  $G[n]$  qu'on a défini au début. Il découle de (4.6) que  $t(F, \Gamma_W) = t(F, G)$  pour chaque  $F \in \mathcal{U}$ , et ainsi  $\Gamma_W = G$  en tant qu'éléments de  $\overline{\mathcal{U}^*}$ .

# Appendice

## Preuve du théorème 1.1

On présente ici la preuve du théorème d'isomorphisme présentée dans l'introduction.

**Théorème.** Soient  $p, q \in ]0, 1[$ . Soient  $G \sim \mathcal{G}(\infty, p)$ ,  $H \sim \mathcal{G}(\infty, q)$ . Alors,  $G$  et  $H$  sont isomorphes presque sûrement.

*Démonstration.* On construit par récurrence un isomorphisme  $f$  entre  $G$  et  $H$ .

- On pose en premier lieu  $f(1) = 1$ .
- Soit  $k \in \mathbb{N}^*$ , on suppose avoir construit  $f$  sur l'ensemble  $\{1, \dots, k\}$  et sur l'ensemble  $f^{-1}(\{1, \dots, k\})$ , et que  $f$  induit un isomorphisme entre  $G(1, \dots, k)$  et  $H(1, \dots, k)$ . On cherche à prolonger la définition de  $f$  sur  $\{1, \dots, k+1\}$ , et à attribuer un sommet de  $H$  au sommet  $k+1$  de  $G$ . Seul le cas où  $k+1 \notin f^{-1}(\{1, \dots, k\})$  nous intéresse, alors plaçons-nous dans ce cas. Le sommet  $k+1$  de  $G$  est relié à certains sommets  $u_1, \dots, u_j$  de  $G$  parmi  $\{1, \dots, k\}$ . On cherche un sommet distinct de  $f^{-1}(\{1, \dots, k\})$  relié à  $f^{-1}(u_1), \dots, f^{-1}(u_j)$ , et non relié aux autres sommets de  $f^{-1}(\{1, \dots, k\})$ . La probabilité qu'un sommet vérifie cette propriété est  $p^j(1-p)^{k-j}$ , celle qu'aucun sommet de  $\{1, \dots, n\} \setminus \{f^{-1}(1), \dots, f^{-1}(k)\}$ , pour  $n$  quelconque, est  $(1 - p^j(1-p)^{k-j})^{n-k}$ , ce qui tend vers 0 quand  $n$  tend vers l'infini. En particulier, il existe presque sûrement un sommet dans  $H$  qui vérifie les conditions souhaitée, en sélectionnant ce sommet (ou plutôt le sommet de plus petit indice, qui convient) en temps qu'image de  $k$  par  $f$ , on prolonge la récurrence.

On peut effectuer le même travail, en cherchant un antécédant à  $k+1$ .

Cette récurrence nous permet d'affirmer qu'il existe un nombre dénombrable d'évènements presque sûrs décroissants pour l'inclusion, dont l'intersection est l'existence d'un isomorphisme  $f$  entre  $G$  et  $H$ , de probabilité 1.  $\square$

# Bibliographie

- [1] David J. ALDOUS. « Representations for partially exchangeable arrays of random variables ». In : *Journal of Multivariate Analysis* 11.4 (1981), p. 581-598.
- [2] C. BORGS et al. *Convergent Sequences of Dense Graphs I : Subgraph Frequencies, Metric Properties and Testing*. 2007.
- [3] B. DE FINETTI. « Funzione caratteristica di un fenomeno aleatorio ». In : *Atti della R. Accademia Nazionale dei Lincei, Ser. 6. Memorie, Classe di Scienze Fisiche, Matematiche e Naturali* 4 (1931), p. 251-299.
- [4] Persi DIACONIS et Svante JANSON. *Graph limits and exchangeable random graphs*. 2007.
- [5] Paul L. ERDOS et Alfréd RÉNYI. « On random graphs. I. » In : *Publicationes Mathematicae Debrecen* (1959).
- [6] D. HOOVER. *Relations on Probability Spaces and Arrays of Random Variables*. 1979.
- [7] Olav KALLENBERG. *Probabilistic symmetries and invariance principles*. Springer, 2005.
- [8] Werner KIRSCH. « An elementary proof of de Finetti's theorem ». In : *Statistics & Probability Letters* 151.C (2019), p. 84-88.
- [9] Laszlo LOVASZ et Balazs SZEGEDY. *Limits of dense graph sequences*. 2004.
- [10] Roger MANSUY. *Introduction aux graphes aléatoires (et à la méthode probabiliste)*. Calvage & Mounet, 2020.