

# Contribution à la modélisation des réseaux complexes

A. Lachgar, A. Achahbar

3 mai 2018

---

# Remerciements

gggggggggggggggggggg

---

# Abstract

Complex networks is a fundamental field of research that model and study artificial and natural networks in our real world. The discovery of common universal properties, almost to all real networks, such as small-world property and scale-free distribution, has revolutionized the way as these networks are studied, modeled and processed. A complex network consists of a large number of interacting entities, hence the emergence of properties at large-scale. The object of this thesis is to introduce some contributions to this domaine. In the beginning, we present the state of the art necessary to the readers, next we analyze the model of Barabási and Albert relying on a new model having the complement of the probability of the first one. After, we study uncorrelated scale-free networks, based on simple steps and assumptions. Then, we treat Newman and Watts' small-world model by relying on renormalization group transformation. Finally, we propose a method to predict the type of phase transitions in the percolation phenomenon in the vicinity of the critical point.

---

# Résumé

La science des réseaux complexes est un domaine de recherche fondamental qui permet de modéliser et étudier les réseaux artificiels et naturels dans notre monde réel. La découverte des propriétés universelles communes, quasiment à tous les réseaux réels, tel que la propriété petit-monde et la distribution sans-échelle, a révolutionné la façon d'étudier, modéliser et traiter ces réseaux. Un réseau est dit complexe s'il est constitué d'un grand nombre d'entités en interactions d'où l'émergence des propriétés à grande échelle. L'objectif de cette thèse est d'apporter quelques contributions à ce domaine. Au début, nous présentons l'état de l'art nécessaire aux lecteurs, puis nous analysons le modèle de Barabási et Albert en s'appuyant sur un nouveau modèle ayant le complément de la probabilité du précédent. Puis, nous étudions les réseaux sans-échelle non corrélés, en se basant sur des étapes et des hypothèses simples. Ensuite, nous traitons le modèle petit-monde de Newman et Watts en s'appuyant sur la transformation de groupe de renormalisation. Enfin, nous proposons une méthode permettant de prédire le type des transitions de phase dans le phénomène de percolation au voisinage du point critique.

---



# Table des matières

<b>Remerciements</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Résumé</b>	<b>vii</b>
<b>Table des figures</b>	<b>xiii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Notions et préliminaire</b>	<b>5</b>
1.1 Les réseaux complexes . . . . .	5
1.1.1 Définition . . . . .	5
1.1.2 Différents types des Réseaux complexes . . . . .	6
1.2 Les réseaux complexes, la physique moderne et l'unification . . . . .	10
1.3 La théorie des graphes . . . . .	12
1.3.1 Bref historique . . . . .	12
1.3.2 L'expérience de Milgram . . . . .	13
1.3.3 Représentation d'un graphe . . . . .	14
1.4 Caractéristiques des réseaux complexes . . . . .	16
1.4.1 Plus court chemin . . . . .	16
1.4.2 Coefficient de regroupement (Clustering) . . . . .	17
1.4.3 Distribution des degrés . . . . .	18
1.4.4 Degré de corrélation . . . . .	19
1.4.5 Mesures de la centralité . . . . .	21

1.5	Propriétés des réseaux réels . . . . .	22
1.5.1	La propriété petit-monde . . . . .	22
1.5.2	La distribution des degrés sans-échelle . . . . .	23
1.5.3	Structure communautaire . . . . .	24
1.6	Les modèles théoriques les plus connus . . . . .	25
<b>2</b>	<b>Les réseaux en croissance et l'attachement préférentielle</b>	<b>29</b>
2.1	Introduction . . . . .	29
2.2	Les processus dynamiques : théorie et simulation . . . . .	30
2.2.1	Introduction . . . . .	30
2.2.2	Équation maîtresse . . . . .	31
2.2.3	Modélisation et simulations numériques . . . . .	33
2.3	L'évolution dynamique et l'attachement préférentiel dans les réseaux réels . . .	34
2.3.1	L'évolution dynamique . . . . .	34
2.3.2	Le modèle sans-échelle de Barabási-Albert . . . . .	34
2.3.3	L'attachement préférentielle et le mécanisme "rich get richer" . . . . .	39
2.4	Attachement préférentiel sans l'effet "rich get richer" . . . . .	40
2.4.1	Le modèle . . . . .	40
2.4.2	Distribution des degrés en utilisant l'équation maîtresse . . . . .	40
2.4.3	Comparaison au niveau microscopique avec le modèle de BA . . . . .	43
2.5	conclusion . . . . .	48
<b>3</b>	<b>Structure des réseaux sans-échelle non corrélés : couches et plus court chemin</b>	<b>49</b>
3.1	Introduction . . . . .	49
3.2	Mesurer les lois de puissance . . . . .	50
3.3	Quelques notions mathématiques pour la loi de puissance . . . . .	51
3.3.1	Normalisation . . . . .	52
3.3.2	Moments . . . . .	52
3.3.3	Lois de puissance pour les variables discrètes . . . . .	53
3.3.4	Degré maximale . . . . .	53

3.3.5	Fonction génératrice . . . . .	55
3.4	Les réseaux sans-échelle non corrélés . . . . .	57
3.5	Les anciennes études sur les couches et le plus court chemin . . . . .	59
3.5.1	Les couches . . . . .	59
3.5.2	Plus court chemin . . . . .	63
3.6	Structure des couches dans les réseaux sans-échelle non corrélés . . . . .	63
3.6.1	L'étude théorique . . . . .	63
3.6.2	Comparaison avec les données réels . . . . .	70
3.7	Plus court chemin dans les réseaux sans-échelle non corrélés . . . . .	71
3.7.1	L'étude théorique . . . . .	71
3.8	conclusion . . . . .	72
<b>4</b>	<b>Émergence de la propriété petit-monde dans le modèle petit-monde</b>	<b>75</b>
4.1	Introduction . . . . .	75
4.1.1	Le modèle petit-monde de Watts-Strogatz . . . . .	76
4.2	La théorie de groupe de renormalisation et transition de phase . . . . .	79
4.3	Plus court chemin et transition de phase dans le modèle NW : ancienne contribution et la notre . . . . .	80
4.3.1	Ancienne contribution . . . . .	80
4.3.2	Notre contribution . . . . .	81
4.4	Structure du réseau NW : couche et plus court chemin . . . . .	82
4.4.1	Les couches aléatoires et régulières . . . . .	82
4.4.2	Plus court chemin . . . . .	92
4.5	Émergence spectaculaire vers la propriété petit-monde . . . . .	95
4.5.1	Invalidité de l'ancienne fonction universelle et la saturation des raccourcis	95
4.5.2	La nouvelle fonction universelle . . . . .	97
4.6	Conclusion . . . . .	100
<b>5</b>	<b>Une nouvelle approche pour prédire les transitions de phases dans le phénomène de percolation</b>	<b>101</b>
5.1	Introduction . . . . .	101

## TABLE DES MATIÈRES

---

5.2	Transition de phase . . . . .	102
5.3	Percolation . . . . .	103
5.4	Percolation explosive . . . . .	103
5.5	Percolation dans le réseau aléatoire d'Erdős-Rényi . . . . .	105
5.5.1	Percolation explosive dans le processus d'Achlioptas . . . . .	107
5.5.2	Percolation explosive dans d'autres modèles . . . . .	109
5.6	Percolation au voisinage du point critique : approche théorique et simulation .	111
5.6.1	Introduction . . . . .	111
5.6.2	Approche théorique . . . . .	113
5.7	Conclusion . . . . .	120
<b>Conclusion</b>		<b>123</b>
<b>Bibliographie</b>		<b>125</b>

# Table des figures

1.1	Représentation de quelques types de réseaux composés par des nœuds et des liens. . . . .	6
1.2	Nous avons ici un affichage hyperbolique de blogs en utilisant à la fois les ensembles de données WWE et ICWSM 2007 ( <a href="http://datamining.typepad.com/">http://datamining.typepad.com/</a> ). . . . .	7
1.3	Un réseau modulaire est illustré au moyen du protéome humain (données obtenues à partir de la base de données DIP : <a href="http://dip-doe-mbi.ucla.edu">http://dip-doe-mbi.ucla.edu</a> ). Les nœuds sont des protéines et des liens indiquent leur interaction physique (protéine-protéine). . . . .	8
1.4	Le graphe social derrière Facebook, graphe des relations d'amis de 500 million personnes, image par Paul Butley, 2010. . . . .	9
1.5	La structure de l'Internet au niveau des systèmes autonomes. Les nœuds de cette représentation sont des systèmes autonomes et les bords montrent les itinéraires empruntés par les données qui circulent entre eux. La photo a été créée par Hal Burch et Bill Cheswick en 2009. . . . .	10
1.6	Ponts de Königsberg, 1736 . . . . .	13
1.7	Exemple d'un graphe non-orienté et sa matrice d'adjacence $M_G$ avec 5 nœuds et 6 liens. . . . .	15
1.8	Exemple d'un graphe orienté et sa matrice d'adjacence $M_G$ avec 5 nœuds et 6 liens. . . . .	15
1.9	Exemple d'un graphe non-orienté et sa liste d'adjacence avec 5 nœuds et 6 liens. . . . .	16
1.10	Exemple de coefficient de regroupement. . . . .	18
1.11	Exemple des réseaux assortatif et disassortatif. . . . .	20

1.12	Quelques exemples des réseaux sans-échelle, (A) Graphe de collaboration d'acteur avec un nombre des nœuds $n = 212250$ et un degré moyen $\langle k \rangle = 28.78$ , (B) WWW, $n = 325729$ , $\langle k \rangle = 5.46$ , (C) Données du réseau électrique, $n = 4941$ et $\langle k \rangle = 2.69$ . Les lignes pointillées ont des pentes (A) $\gamma_{actor} = 2.3$ , (B) $\gamma_{www} = 2.1$ et (C) $\gamma_{electrique} = 4$ . . . . .	23
1.13	Exemples d'un réseau communautaire, on voit trois groupes où chacun est plus connecté par rapport au reste du réseau. . . . .	25
1.14	Simple réseau régulier de deux dimensions . . . . .	26
2.1	L'évolution de nombre d'hôtes sur Interent de 1994 à 2013, cette figure obtenues à partir de site : <a href="http://facesncups.com/inforev.html">http://facesncups.com/inforev.html</a> . . . . .	35
2.2	La distribution des degrés du réseau BA en échelle logarithmique, où les cercles sont les simulations numériques et la ligne continue représente la formule théorique, Eq. (2.9). La taille du réseau est $n = 10^6$ et $m = 2$ . . . . .	38
2.3	les résultats de la simulation (cercles) pour $n = 2 \cdot 10^6$ , $m = 2$ , $m_0 = 3$ , et la solution numérique ( ligne continue) pour Eq. (2.15). Dans l'encart, nous tracerons les mêmes données dans l'échelle log-linéaire. . . . .	43
2.4	(a) Évolution de $\langle k_s(t) \rangle$ dans le modèle BA, la ligne continue représente Eq. (2.19). (b) Évolution des fluctuations de $\langle k_s(t) \rangle$ , la ligne continue représente Eq. (2.23). Les cercles joints par des lignes pointillées dans les deux cas sont des données de simulation moyennées sur 20 réalisations pour $m = 2$ , $m_0 = 3$ . . .	45
2.5	(a) Évolution de $\langle k_s(t) \rangle$ dans notre modèle, la ligne continue représente Eq. (2.26). (b) Évolution des fluctuations de $\langle k_s(t) \rangle$ , la ligne continue représente la solution numérique de Eq. (2.26) et Eq. (2.28). Les cercles joints par des lignes pointillées dans les deux cas sont des données de simulation moyennées sur 20 réalisations pour $m = 2$ , $m_0 = 3$ . . . . .	47
3.1	Illustration de la loi de puissance pour une fonction, $P(k) = k^{-\gamma}$ , avec l'exposant $\gamma = 3$ . Dans a) l'échelle est logarithmique et dans b) l'échelle est linière. . . . .	50

3.2	Les corrélations des degrés dans le réseau aléatoire sans-échelle pour différentes valeurs de $m$ et de $\gamma$ . Le nombre de nœuds est $n = 10^4$ et le nombre de réalisations pour chaque simulations est 50. . . . .	58
3.3	Illustration du réseau construit. La taille des cercles pleins (nœuds) est le degré dépendant. Le degré maximum de couche $\ell$ est $K_\ell$ . . . . .	64
3.4	Nombre de nœuds dans chaque couche pour différentes valeurs de $\gamma$ . Les lignes pleines sont Eq. (3.39) et Eq. (3.41). Les symboles sont des simulations d'un réseau de taille $n = 4 \times 10^6$ et une moyenne de 200 réalisation pour chaque point. Les couleurs noir, rouge et bleu se réfèrent respectivement à $m = 8, 4$ et $2$ . . . . .	67
3.5	Proportion de cycles dans les couches par rapport à $m$ . De haut en bas, $\gamma$ est respectivement $4, 3, 2.6$ et $2.2$ . Nombre de nœuds $n = 4.10^6$ , le nombre de points pour chaque simulation est 200. . . . .	69
3.6	Comparaison entre les données empiriques (cercle) et notre équation théorique (ligne continue) : Les cercles représentent les couches du réseau d'acteurs Hollywoodien de 2283910 acteurs et de $\gamma = 2.13$ lorsque Kevin Bacon est au centre du réseau et la ligne continue représente l'Eq. (3.39) avec les mêmes valeurs $\gamma$ et $n$ du réseau Hollywoodien. Les résultats empiriques ont été extraits du site : " <a href="http://oracleofbacon.org/center.php">http://oracleofbacon.org/center.php</a> ". . . . .	70
3.7	La longueur moyenne du trajet en fonction du nombre de nœuds. Les valeurs de $\gamma$ de haut en bas sont respectivement $4, 3, 2.6$ et $2.2$ . Les lignes pleines correspondent à l'équation Eq. (3.46) et à l'équation Eq. (3.47). Chaque simulation est moyennée à plus de 200. . . . .	73
4.1	Un réseau petit-monde WS pour $n = 20$ , $k = 2$ sans raccourcis. . . . .	76
4.2	Représentation schématique de l'évolution du processus de "reconnexion" dans le modèle Watts-Strogatz. . . . .	77
4.3	Représentation de la variation de la longueur moyenne du chemin $L(p)$ et du coefficient de regroupement $C(p)$ avec la modification de la probabilité de reconnexion $p$ dans le modèle Watts-Strogatz. . . . .	78

4.4	Illustration des nœuds aléatoires (nœuds noirs) qui deviennent plus proches du nœud arbitraire $R$ , et nœuds réguliers (nœuds blanches) qui restent dans leurs distance initial par rapport au même nœud arbitraire $R$ , le réseau est de taille $n = 20$ , $k = 1$ et un seul raccourci. . . . .	83
4.5	La transformation GR d'un réseau (a) de $n = 20$ et $k = 2$ vers un réseau (b) de $\acute{n} = 10$ et $\acute{k} = 1$ . . . . .	87
4.6	Fraction des nœuds aléatoires $S_{al}/n$ (noir) et des nœuds réguliers $S_r/n$ (rouge) en fonction de nombre des raccourcis $x$ . Les lignes représentent les fonction $h(x)$ (noir) et $1 - h(x)$ (rouge). Les symboles représentent les simulations numériques d'un réseau de taille $n = 10^4$ , avec $k = 1$ (étoile et cercle) et $k = 5$ (carré et triangle). Chaque simulation est moyennée à 200 réalisations. . . . .	91
4.7	L'expression universelle $f(x) = \ell/\acute{n}$ en fonction de $x$ pour un réseau de taille $n = 10^6$ . La formule de Newman et watts Eq. (4.6) (ligne rouge), Eq. (4.35) (ligne vert), Eq. (4.34) (ligne noir) et les simulations numériques pour $k = 1$ (cercle) et $k = 5$ (carré). L'échelle est linéaire-log dans <b>a)</b> et log-log dans <b>b)</b> . Chaque simulation est moyennée à 100 réalisations. . . . .	95
4.8	Fluctuation $\sigma$ en fonction de $x$ d'un réseau de taille $n = 10^4$ pour différents valeurs de $k$ , $k = 1$ (étoile), $k = 2$ (carré) et $k = 5$ (cercle), le nombre de réalisations est 1000. . . . .	96
4.9	La valeur de $\Delta$ et $\rho$ en fonction de $y$ pour différentes valeurs de $n$ et $k$ : $n = 10^{-3}$ et $k = 1$ (triangle), $n = 10^{-3}$ et $k = 4$ (carré), $n = 10^{-4}$ et $k = 1$ (cercle), $n = 10^{-4}$ et $k = 4$ (étoile). Dans <b>a)</b> le nombre de réalisations pour chaque simulation est 1000 et dans <b>b)</b> le nombre de réalisations pour chaque simulation est 5000. . . . .	97
4.10	L'expression universelle de PCC en fonction de $y$ , par Eq. (4.39) (ligne continue) et par les simulations numériques pour différents valeurs de $n$ et $k$ , $n = 10^{-6}$ et $k = 20$ (étoile), $n = 10^{-5}$ et $k = 5$ (carré), $n = 10^{-4}$ et $k = 2$ (cercle), $n = 10^{-3}$ et $k = 1$ (triangle), nombre de réalisations pour chaque simulations est 100. En <b>a)</b> l' échelle est linéaire-log et en <b>b)</b> l'échelle est log-log. . . . .	98



5.1	Une démonstration de percolation sur une grille à deux dimensions pour différentes valeurs de $p$ . Sous le seuil de percolation le système est composé de petites grappes, après un certain point critique $p_c$ une grappe étendue et occupe la grille. . . . .	104
5.2	Les simulations numériques du seuil de percolation dans le modèle ER en fonction du degré moyen, pour un réseau de taille $n = 10^5$ . On observe que le point où la fraction de la composante géante, $S$ , émerge est à $\langle k \rangle = 1$ . . .	106
5.3	Illustration de la distribution des degrés d'un réseau aléatoire ER de degré moyen égale 6, les cercles représentent les simulations numériques et la ligne noire représente Eq. (5.2). . . . .	107
5.4	Comparaison entre la percolation aléatoire et PA. <b>(A)</b> Percolation ER classique, où les liens sont ajoutés au hasard dans le réseau. <b>(B)</b> Percolation PA, où à chaque étape deux liens sont en compétition pour être établis. <b>(C)</b> Paramètre d'ordre : la taille relative de la composante géante par rapport au nombre des liens ajoutés normalisés par la taille du système. . . . .	108
5.5	La distribution d'amas selon deux méthodes différentes pour choisir les liens dans le même réseau, où l'exposant de la distribution d'amas dans le réseau est $\gamma = 2.5$ , la première méthode est aléatoire comme le processus ER (cercles verts), la deuxième est par le processus Achlioptas [179] (carrés bleus). La ligne pointillée a pour but de montrer l'exposant $\gamma = 2.5$ . La taille du réseau est $n = 10^6$ , chaque simulation a une moyenne de $n$ réalisations. . . . .	113
5.6	Fraction de la composante géante $S$ en fonction de la densité des liens $\langle k \rangle$ pour différentes valeurs de $\gamma$ , les couleurs de gauche à droite représentent respectivement les simulations numériques pour $\gamma = 2, 2.5, 3, 4$ , les lignes noires représentent la solution numérique de l'Eq. (5.12). Le nombre des nœuds est $n = 10^5$ . . . . .	116
5.7	Estimation de point critique par les simulations numériques (cercles) et par Eq. (5.15) (ligne). Les valeurs de $\gamma$ de haut en bas sont respectivement 4, 3, 2.5, 2. Chaque simulation est moyennée à 500 réalisations. . . . .	117



# Introduction

Un réseau complexe est un grand nombre des nœuds liés entre eux selon des topologies de connexion spécifiques. Les réseaux complexes sont omniprésents dans le monde naturel et artificiel grâce aux développements technologiques des 60 dernières années. En effet, la plupart des réseaux réels peuvent être représentés par des réseaux complexes, on peut distinguer plusieurs types de réseaux dans les différents domaines, tels que le domaine social, technologique, biologique, physique, etc. Par exemple, un système qui se compose de différents types de molécules qui s'affectent les unes aux autres par des réactions enzymatiques [1], Internet relie un grand nombre de serveurs et d'ordinateurs dans le monde entier échangeant constamment des quantités gigantesques de paquets d'informations [2], le World Wide Web est un réseau virtuel de sites Web liés avec des hyperliens [3], et les réseaux alimentaires relient également, via des relations trophiques, un grand nombre d'espèces interdépendantes [4, 5]. Bien que l'existence de ces réseaux dans divers domaines était connue depuis longtemps, les physiciens n'ont commencé à s'y intéresser que depuis la découverte de certaines lois universelles communes à différents systèmes réels. Parmi ces lois les plus importantes on cite : la distribution de degrés libre-échelle<sup>1</sup>, la propriété **petit-monde** et la valeur élevée du **coefficient de regroupement** (Clustering). Afin de bien comprendre les comportements de ces réseaux et établir les lois qui les gouvernent, les physiciens, les mathématiciens et les informaticiens se consacrent à développer des modèles théoriques et des techniques permettant de découvrir et d'analyser les propriétés de ces réseaux qui sont presque partout. Il semble que le début du troisième millénaire va connaître une nouvelle révolution emportée par les principes des réseaux complexes.

---

1. Libre-échelle signifie qu'il n'y a pas une valeur de degré caractéristique dans le réseau, mais plutôt une distribution des degrés qui suit une loi de puissance.

Il est évident que la physique statistique est l'outil adéquat pour étudier les systèmes ayant un grand nombre d'éléments en interaction. En effet, elle a développé au cours du temps un ensemble de théories et d'outils mathématiques permettant, à partir des comportements microscopiques, de comprendre l'émergence des caractéristiques macroscopiques et d'étudier systématiquement la topologie de ces grands réseaux complexes. En outre la physique statistique est le cadre théorique le plus perfectionné pour étudier les problèmes de transitions de phases et les points critiques, ce qui est incontournable dans l'étude des systèmes complexes.

Ce travail de recherche entreprend quelques contributions à la modélisation des réseaux complexes, en abordant quelques problèmes concernant la dynamique de la croissance dans le modèle de Barabási-Albert (BA) [3], la structure des réseaux libre-échelle, l'émergence de la propriété petit-monde dans le modèle de Newman-Watts [6], et la caractérisation des transitions de phases dans le phénomène de percolation. De façon plus détaillée, cette thèse est constituée de cinq chapitres. Le premier commence par définir et mettre en évidence l'importance de notre cadre de recherche (réseaux complexes), et se termine par un état de l'art de certaines techniques et concepts fondamentaux pertinents de la physique statistique et de la théorie des graphes. Le deuxième chapitre commence par traiter la notion du réseau en croissance et de l'attachement préférentiel, puis propose un simple modèle en utilisant le complément de la probabilité utilisée dans le modèle BA, et finit par développer le calcul de la valeur moyenne du degré sélectionné à chaque instant et sa fluctuation dans notre modèle et celui de BA. Le but étant de faire une comparaison au niveau microscopique entre les deux et de vérifier si la distribution de degrés en loi de puissance persiste en l'absence du mécanisme "rich get richer". Le troisième chapitre a pour but de donner l'expression explicite de la structure des couches dans les réseaux libre-échelle non corrélés, puis celle du plus court chemin, en suggérant que le plus court chemin correspond à la distance à laquelle la couche est maximale. Ce chapitre établit également une comparaison avec les anciens travaux théoriques et avec les données réels. Le quatrième chapitre contient une étude détaillée sur le modèle introduit par Newman-Watts en utilisant la transformation de groupe de renormalisation en espace réel, les résultats obtenus sont très intéressants et nous donnent

une perspective plus clair sur l'émergence vers la propriété petit-monde dans ce modèle, on trouve à partir d'une prédiction purement théorique<sup>2</sup> qu'en fonction d'un certain paramètre il y a une émergence vers la phase petit-monde. Nous constatons également la disparition de la fonction universelle usuelle et l'apparition d'une nouvelle fonction universelle liée à ce paramètre. Nous en déduisons que cette émergence n'est pas ordinaire mais de nature spectaculaire. Le dernier chapitre est consacré au problème de transition de phase dans le phénomène de percolation, qui est un problème pas encore bien compris et le plus urgent à résoudre dans la physique des systèmes désordonnés. En admettant que le point critique suit toujours une lois de puissance nous avons établi une formule théorique qui permet de déterminer les conditions requises pour que la transition de phase sera de premier ou de deuxième ordre, indépendamment des mécanismes microscopiques de l'évolution du système. Autrement dit nous avons étudié le phénomène de percolation au voisinage du point critique en démontrant théoriquement les conditions nécessaires pour obtenir une transition de premier ordre ou de deuxième ordre.

Il convient de signaler que dans les quatre derniers chapitres nous avons développé plusieurs programmes en Fortran, basés principalement sur la méthode de Monte Carlo et sur les méthodes d'analyse numérique. Ces programmes servent à vérifier nos équations théoriques ou à confirmer les hypothèses énoncées. En outre, afin de pouvoir travailler sur des réseaux de très grande taille (plus de  $10^7$ ), et exécuter un maximum d'itérations, les programmes ont été amplement optimisés et perfectionnés afin qu'ils puissent fonctionner sur une machine ordinaire.

---

2. C'est-à-dire que cette idée est une conclusion qui revient directement de nos équations théoriques et pas par une inspiration intellectuelle ou empirique.



# Chapitre 1

## Notions et préliminaire

Dans le présent chapitre, nous présentons l'importance de notre domaine de recherche de façon très générale, puis on introduit des termes et des concepts qui seront utilisés tout au long de cette thèse : réseaux complexe, théorie des graphes, distribution des degrés, etc.

### 1.1 Les réseaux complexes

#### 1.1.1 Définition

Un **réseau** est tous simplement une collection de points réunis sous forme de paires, les points sont appelés nœuds ou sommets et les lignes sont appelées liens ou bords. Le mots **complexe** est en général le résultat de l'évolution décentralisée et non planifiée dans ces réseaux (Fig. 1.1). De nombreux objets d'intérêt dans les sciences physiques, biologiques et sociales peuvent être considérés comme des réseaux complexes, cette nouvelle façons de penser conduit souvent à de nouvelles et utiles idées. Alors un réseau est une représentation simplifiée qui réduit un système à une structure abstraite capturant uniquement les bases des modèles de connexion, les nœuds et les liens d'un réseau peuvent être accompagnés d' informations supplémentaires pour capturer plus de détails sur le système, même s'il y a un inconvénient de perte d'informations dans le processus de réduction d'un système complet à une représentation par réseau, mais il présente également des grands avantages.

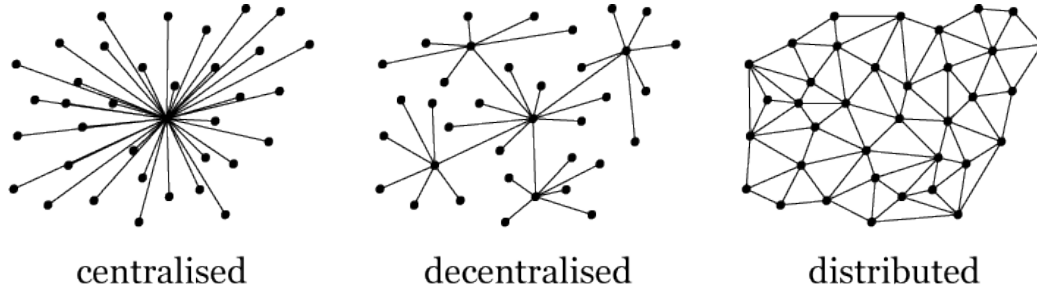


FIGURE 1.1 – Représentation de quelques types de réseaux composés par des nœuds et des liens.

### 1.1.2 Différents types des Réseaux complexes

L'étude des réseaux complexes a été inspirée par le désir de comprendre les différents systèmes réels, allant des réseaux de communications aux réseaux écologiques. Les bases de données empiriques disponibles pour l'étude couvrent plusieurs disciplines, en générale, elles se divisent en quatre grandes catégories : Réseaux Technologique, Réseaux Biologiques, Réseaux Sociaux et Réseaux d'Informations.

#### Réseaux Technologique

Les réseaux technologiques sont des réseaux artificiels, qui ont grandi au cours du siècle dernier et qui constituent une grande partie de notre société moderne, comme les réseaux électriques, réseaux téléphoniques, réseaux de transports, etc [7, 8, 9, 10]. L'Internet est parmi les exemples les plus connus et les plus largement étudiés des réseaux technologiques (voir Fig. 1.2), on peut le définir comme un réseau de données informatiques dans lequel les nœuds sont des ordinateurs et les liens sont des connexions de données physiques entre eux, tels que des câbles à fibres optiques ou des lignes téléphoniques [2, 11]. Il est curieux que, bien que l'Internet est un réseau artificiel, nous ne connaissons pas exactement sa structure, nos meilleures données actuelles sur sa structure proviennent d'études expérimentales.

Il existe un certain nombre d'excellentes raisons pratiques pour être intéressé à étudier la structure du réseau d'Internet, la fonction d'Internet consiste à transporter des données entre ordinateurs dans différentes parties du monde, ce qui se fait en divisant les données en pièces ou en paquets et en les transportant d'un nœud à l'autre sur le réseau jusqu'à ce qu'ils atteignent leur destination, sans aucun doute, la structure du réseau affectera la manière dont



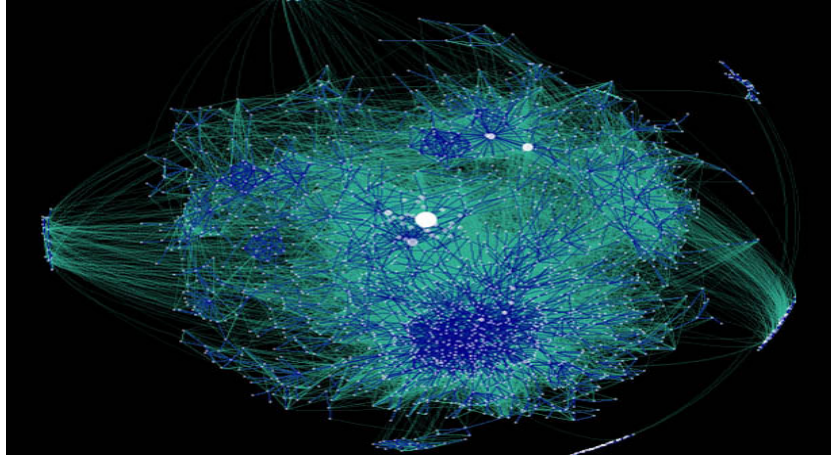


FIGURE 1.2 – Nous avons ici un affichage hyperbolique de blogs en utilisant à la fois les ensembles de données WWE et ICWSM 2007 (<http://datamining.typepad.com/>).

il accomplit efficacement cette fonction et si nous connaissons la structure du réseau, nous pouvons aborder de nombreuses questions et problèmes de pertinence pratique.

## Réseaux Biologiques

Une autre classe des réseaux les plus étudiées dans la littérature est celle des réseaux biologiques. Cette classe contient une grande variété de réseaux naturels. Le corps, qu'il soit humain ou animal, contient un grand nombre de réseaux, dont certains se produisent dans l'espace réel, tels que le système nerveux. Ces réseaux ont été étudiés depuis longtemps [12]. Une autre classe de réseau, ce sont les réseaux d'interactions gène-gène, protéines-gène et protéines-protéines [13], ainsi que les réseaux d'interactions entre les espèces dans les écosystèmes, comme la prédation ou la coopération.

La structure de ces réseaux se diffère selon chaque cas. Par exemple, les réseaux métaboliques sont des réseaux de protéines interagissant les uns avec les autres à l'intérieur de la cellule Fig. 1.3, il s'agit d'un réseau dirigé, car chaque protéine peut catalyser ou réprimer la création d'autres protéines, ce qui n'implique pas nécessairement le processus inverse. La structure à grande échelle des réseaux métaboliques a été étudiée pour de nombreuses espèces, il a été trouvé que ces réseaux ont une distribution des degrés sans-échelle [1]. De plus, il a été observé que le diamètre du réseau est très petit et presque indépendant de la taille du réseau, cette indépendance s'explique par que certain classe de réseaux sans-échelle est Ultra-small

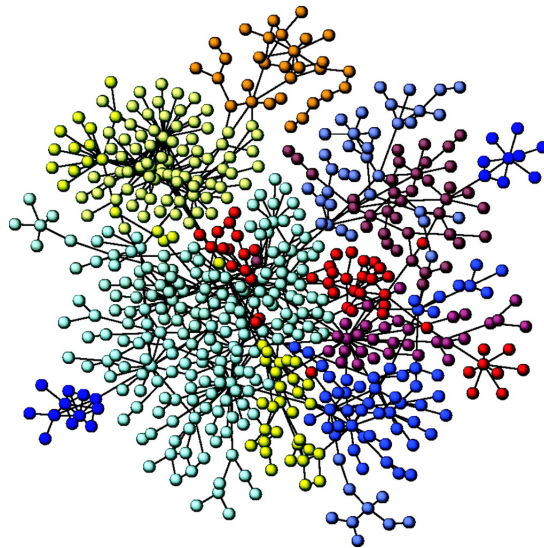


FIGURE 1.3 – Un réseau modulaire est illustré au moyen du protéome humain (données obtenues à partir de la base de données DIP : <http://dip-doe-mbi.ucla.edu>). Les nœuds sont des protéines et des liens indiquent leur interaction physique (protéine-protéine).

[14, 15, 16], ainsi que notre résultats du troisième chapitre renforcent cette explication et indiquent que la dépendance du diamètre avec la taille  $N$  est très faible.

Dans les réseaux génétiques, les nœuds représentent des gènes, et les liens sont dirigés et représentent l'influence d'un gène sur un autre, le réseau *E. coli*<sup>1</sup> est un des réseaux génétiques qui est bien étudiés dans la littérature [17].

## Réseaux Sociales

Un réseau social est un ensemble de nœuds, où ces nœuds se représente par des personnes (individus ou groupes sociaux) et les liens par une relation qui peut être de parenté, amitié, statut, etc [18], cette diversité des liens est une chose appréciée dans l'étude des réseaux sociaux, car il existe de nombreuses définitions possibles d'un tel réseau et la définition particulière que l'on utilisera dépendra des questions auxquelles on est intéressé à répondre. Tel que les réseaux d'amitiés entre les individus [19, 20], les relations d'affaires entre les entreprises [21]. La société offre une grande variété d'organisations de groupes possibles : les familles, les milieux de travail et d'amitié, les villages, les villes, les nations. La diffusion d'Internet a également conduit à la création de groupes virtuels, en direct sur le Web, comme

---

1. *Escherichia coli* (en abrégé *E. coli*) sont des bactéries présentes dans l'environnement, les aliments et les intestins des humains et des animaux. *E. coli* est un groupe important et diversifié de bactéries.

Facebook qui reliant presque tout notre monde entier (voir Fig. 1.4).



FIGURE 1.4 – Le graphe social derrière Facebook, graphe des relations d’amis de 500 million personnes, image par Paul Butley, 2010.

Bien avant que l’Internet commence à jouer un rôle important dans la vie de beaucoup de personnes, les sociologues et d’autres chercheurs des sciences humaines ont examiné la structure des groupes de personnes. Dans la plupart des cas, des groupes relativement petits ont été considérés, nécessairement parce que l’analyse de grands groupes n’était pas souvent possible.

Une contribution importante à l’analyse des réseaux sociaux est venue de Jacob Moreno qui a introduit des sociogrammes dans les années 1930. Un sociogramme peut être considéré comme une représentation graphique d’un réseau : les personnes sont représentées par des points (appelés nœuds) et leurs relations par des lignes reliant ces points (appelés liens).

L’analyse des réseaux sociaux a été importante pour le développement ultérieur de la théorie des graphes, par exemple en ce qui concerne l’introduction de mesures pour identifier l’importance des personnes ou des groupes. Par exemple, une personne ayant de nombreuses connexions avec d’autres personnes peut être considérée comme relativement importante. De même, une personne au centre d’un réseau semble être plus influente que quelqu’un au bord. Ce que la théorie des graphes nous fournit, ce sont les outils pour décrire formellement l’importance et l’influence des nœuds. En outre, en utilisant la théorie des graphes, nous pouvons facilement proposer des solutions alternatives pour décrire l’importance des personnes. L’existence de ces outils a également amélioré la précision des déclarations concernant le poste ou le rôle de ces personnes au sein d’une communauté.

## Réseaux d'informations

Les réseaux de citation et le World Wide Web (WWW), comme dans Fig. 1.5, sont un bon exemple des réseaux d'informations, car leur contenu d'informations étant stockée dans des nœuds, c'est pour cette raison que l'on utilise le terme réseau d'informations. Parfois on rencontre une certaine confusion à propos de réseau WWW et le réseau d'Internet. Dans la WWW, les nœuds sont les pages HTML, et les bords représentent les liens entre les pages, par contre dans l'Internet, les liens correspondent aux câbles physiques entre les ordinateurs. Alors le WWW est virtuelle et Internet est physique.

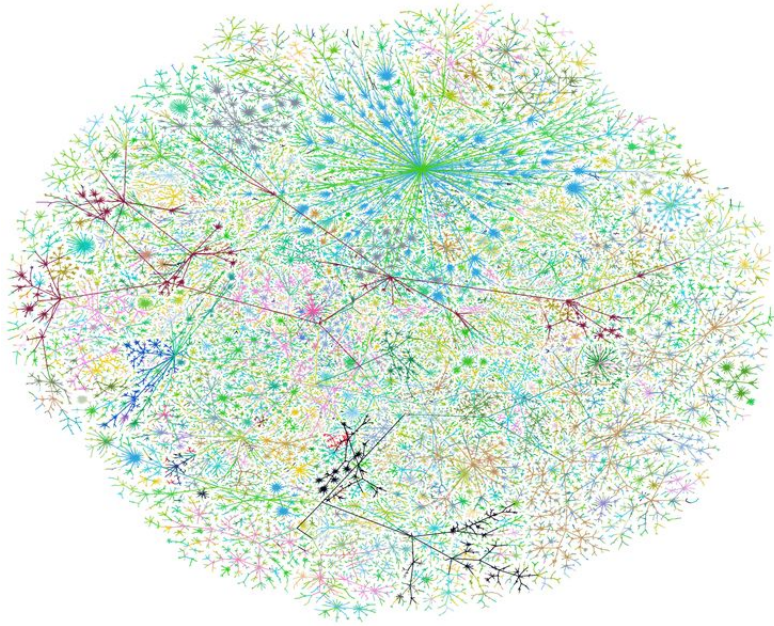


FIGURE 1.5 – La structure de l'Internet au niveau des systèmes autonomes. Les nœuds de cette représentation sont des systèmes autonomes et les bords montrent les itinéraires empruntés par les données qui circulent entre eux. La photo a été créée par Hal Burch et Bill Cheswick en 2009.

## 1.2 Les réseaux complexes, la physique moderne et l'unification

La physique explique les phénomènes de la nature en les réduisant à une interaction de lois fondamentales simples. Cette méthode plutôt réussie semble rencontrer certaines difficultés

lorsqu'il s'agit des systèmes complexes en général et des réseaux complexes en particulier. Dans ces derniers il reste peu clair s'il existe des lois universelles uniques expliquant une variété de similitudes structurelles et dynamiques trouvées dans de nombreux réseaux réels différents [22, 23, 24, 25]. En revanche, l'unification des lois universelles qui paraissaient jusqu'alors complètement séparées sont-elles d'une origine commune ? c'est ce que les physiciens théoriques rêvent de découvrir. Une telle **unification** va être, sans doute, un grand pas dans notre compréhension de la nature. En outre, l'existence de cette belle idée au cœur de l'unification montre le pouvoir mystérieux que les êtres humains peuvent découvrir derrière les apparences de la nature [26].

Le réseau causal représentant la structure à grande échelle de l'espace-temps dans notre univers accéléré est un graphe de loi de puissance avec un regroupement fort, similaire à de nombreux réseaux complexes tels que les réseaux Internet, sociaux ou biologiques [22]. Cette similitude structurelle est une conséquence de l'équivalence asymptotique entre la dynamique de croissance à grande échelle des réseaux complexes et des réseaux causaux. Par conséquent, un intérêt croissant est adressé à l'étude de la gravité quantique à partir de la théorie de l'information et de la perspective des réseaux complexes [27, 28]. Récemment, des relations intrigantes entre les propriétés des réseaux de communication quantique avec des topologies du réseau et la physique statistique ont été rapportées. Sur la base des concepts classiques de percolation [29], il a été montré que ces réseaux quantiques peuvent présenter une transition de phase de percolation d'enchevêtrement [30, 31]. L'avancement rapide de la technologie de l'information quantique a suscité un intérêt considérable pour les propriétés dynamiques des réseaux quantiques formés par les systèmes élémentaires, tels que les qubits, en raison de leur rôle privilégié dans la communication et le calcul quantiques [32, 33, 34].

Une autre domaine de recherche en relation avec les sciences des réseaux et notamment aux réseaux complexes. La science générale des réseaux et de ses diverses applications a une pertinence significative pour les praticiens de l'intelligence artificiel (IA). Par exemple, la compréhension de la structure d'Internet et du World Wide Web sera importante pour l'orientation de la direction de l'intelligent, l'équilibrage de charge, la recherche de toutes

sortes et le déploiement d'agents<sup>2</sup> intelligents qui assistent les utilisateurs dans leurs tâches réseau. La réflexion en réseau sera également fondamentale pour développer des algorithmes décentralisés efficaces pour les réseaux de calcul, de communication et de détection de plus en plus distribués et liés, ainsi que des méthodes de sécurité efficaces pour ces systèmes de plus en plus vulnérables. Ce sont tous des domaines dans lesquels la recherche sur l'IA et l'apprentissage automatique ont joué et joueront un rôle majeur [35, 36, 37].

## 1.3 La théorie des graphes

En terme générale, un réseau se décrit comme un graphe dont les nœuds (sommets) identifient les éléments du système et les liens de connexion (arêtes) représente la présence d'une relation ou d'une interaction entre ces éléments. Avec un tel niveau de généralité, il est facile de percevoir qu'un large éventail de systèmes peuvent être abordés dans le cadre de la théorie du graphe. Alors nous fournissons ici un bref historique et quelques notations de base nécessaires dans la théorie de graphe pour décrire les réseaux. Le cadre naturel pour une description mathématique rigoureuse des réseaux se trouve dans la théorie des graphes, mais il faut noter que la théorie des graphes constitue une branche des mathématiques vaste et compliquée et nous ne sommes pas dans l'état de fournir une présentation formelle et complète de celle-ci. Cependant notre but dans ce chapitre d'introduction est de fournir seulement quelques notions utiles pour décrire les réseaux dans la suite de cette thèse. Pour les lecteurs souhaitant aller plus en détail dans l'étude de la théorie des graphes pourraient consulter ces livres [38, 39].

### 1.3.1 Bref historique

En plus de la topologie, Euler est devenue le père de la théorie des graphes quand il a résolu, en 1736, un problème célèbre sous le nom *"le problème du pont de Königsberg"*, la question était de savoir s'il était possible de visiter les quatre quartiers de la ville séparés les uns des autres par un bras de rivière, en passant exactement une fois par chaque pont et en revenant à son point de départ (voir Fig. 1.6). Afin de trouver une solution à ce problème

---

2. Un agent est un terme important en IA qui désigne une entité capable d'interagir avec son environnement.



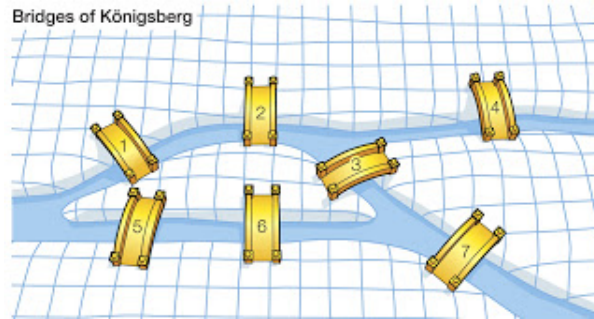


FIGURE 1.6 – Ponts de Königsberg, 1736

Euler a remplacé chaque zone de terrain par un point et chaque pont par une ligne joignant les points correspondants, produisant ainsi un *"graphe"*. Ainsi il a montré que le problème est insoluble et que le graphe de cette ville ne peut pas être parcouru d'une certaine manière.

En 1847 Kirchhoff a développé la théorie des arbres afin de résoudre le système d'équations linéaires simultanées qui donnent le courant dans chaque branche et autour de chaque circuit d'un réseau électrique. Bien qu'un physicien, il a pensé comme un mathématicien lorsqu'il a remplacé un réseau électrique par sa structure combinatoire correspondante constituée uniquement de points et de lignes sans indication du type d'élément électrique représenté par des lignes individuelles. Dans les années 1960, deux mathématiciens, Paul Erdős et Alfred Rényi (ER), ont introduit une nouvelle idée ingénieuse, ils ont combiné les concepts de la théorie des graphes avec les outils de la théorie des probabilités, cela permettant d'envisager des familles de graphes plutôt que des graphes spécifiques.

### 1.3.2 L'expérience de Milgram

En 1967, Stanley Milgram a effectué une expérience intéressante. Dans sa première expérience, Milgram a demandé à des personnes choisies au hasard au Nebraska d'envoyer des lettres à une personne cible éloignée à Boston. Les participants ne pouvaient passer que les lettres (à la main) aux connaissances personnelles qu'ils pensaient pouvoir atteindre la cible, soit directement, soit via un "ami d'un ami". Dans sa première expérience seules trois lettres ont finalement atteint leur destination. Mais dans les expériences ultérieures, Milgram a réussi à augmenter le taux de réussite à 95%, pour plus de détails (voir [40, 41]). La conclusion principale de cette expérience de Milgram était que la plupart des gens sur notre planète n'est

séparé que par six autres personnes en moyen. Cette idée de Milgram a été reprise encore une fois en 2001 par Duncan Watts et ses collègues en utilisant un message électronique qui devait être livré à des expéditeurs autour de monde, étonnamment, Watts a constaté que le nombre moyen d'intermédiaires était 6. Pour plus de détails et une analyse statistique beaucoup plus étendue des données par rapport à l'analyse de Milgram, voir Dodds et al [42].

### 1.3.3 Représentation d'un graphe

Comme dans toute abstraction mathématique, lorsque nous décrivons un système en tant que graphe, nous décidons de rejeter plusieurs des particularités particulières des phénomènes réels et de nous concentrer uniquement sur quelques caractéristiques d'intérêt. En particulier, un graphe est essentiellement un moyen de coder une relation, liens physiques, interactions, etc, entre les éléments d'un système. Les éléments du système identifient l'ensemble  $V$  (ensemble des nœuds) et les relations entre ceux de l'ensemble  $E$  (ensemble des liens). Le graphe indiqué comme  $G(V, E)$  peut être présenté en traçant les nœuds en tant que points et les bords comme lignes entre eux.

Parmi les façons de présentation basées sur cette définition on peut citer la matrice d'adjacence et la liste d'adjacence :

#### La matrice d'adjacence

Un graphe est représenté fréquemment par une matrice d'adjacence,  $M_{i,j}$ , dans laquelle chaque ligne et chaque colonne représente un nœud du graphe. Dans un graphe non-dirigé l'élément  $M_{i,j} = 1$  si un lien existe entre le  $i^{eme}$  et  $j^{eme}$  nœud sinon  $M_{i,j} = 0$  (voir Fig. 1.7), ainsi la matrice sera en général asymétrique et les éléments diagonaux ne sont pas nécessairement 0.

Un graphe orienté est un graphe où les bords sont dirigés, c'est-à-dire que chaque bord est une paire de nœuds ordonnée avec  $(i,j)$  désignant un bord (flèche) qui commence au nœud  $i$  et se termine au nœud  $j$  Fig. 1.8.



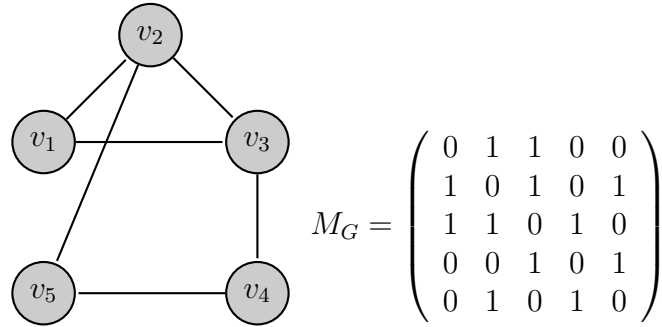


FIGURE 1.7 – Exemple d’un graphe non-orienté et sa matrice d’adjacence  $M_G$  avec 5 nœuds et 6 liens.

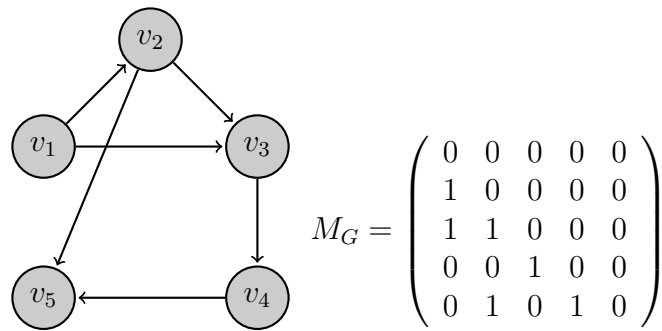


FIGURE 1.8 – Exemple d’un graphe orienté et sa matrice d’adjacence  $M_G$  avec 5 nœuds et 6 liens.

### La liste d’adjacence

Le format de liste d’adjacence est utile pour les graphes sans données associées aux nœuds ni aux bords et à des nœuds qui peuvent être représentés de manière significative sous la forme de chaînes, il se compose de lignes avec des étiquettes de nœud, la première étiquette dans une ligne est le nœud source, les autres étiquettes dans la ligne sont considérées comme des nœuds cibles et sont ajoutées au graphe avec un bord entre le nœud source et le nœud cible.

En règle générale, les matrices sont utilisées pour les tableaux denses et les listes d’adjacence pour les tableaux dispersés. La raison est que les matrices consomment moins d’espace pour les tableaux denses et les listes d’adjacence consomment moins d’espace pour les tableaux dispersés. Cependant, l’espace n’est qu’une considération, d’autres facteurs doivent également être pris en compte.

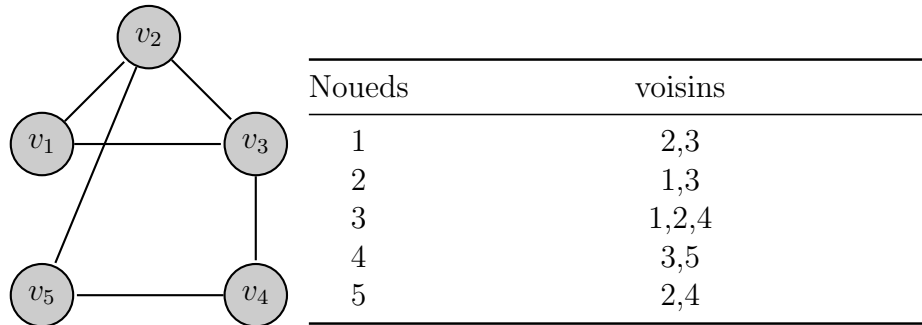


FIGURE 1.9 – Exemple d’un graphe non-orienté et sa liste d’adjacence avec 5 nœuds et 6 liens.

## 1.4 Caractéristiques des réseaux complexes

On croit que ne comprend pas encore les réseaux de manière appropriée. Par exemple, si les données (la matrice d’adjacence) d’un grand graphe sont présentes et que vous n’êtes pas autorisé à visualiser le réseau, il semble assez complexe de dire simplement quelles sont les propriétés du réseau. Alors l’étude de ces réseaux oblige à ajouter d’autres caractéristiques supplémentaires, telles que, le plus court chemin, Clustering, etc.

### 1.4.1 Plus court chemin

Le plus court chemin entre deux nœuds d’un graphe est défini comme étant la longueur du trajet le plus court parmi tous les trajets possibles. Une mesure statistique globale de la distance entre les nœuds peut alors être exprimée comme la valeur moyenne des chemins les plus courts entre tous les couples possibles de nœuds du réseau, mathématiquement on peut le définir par la forme suivante

$$l = \frac{1}{n(n-1)} \sum_{i,j} d_{i,j}, \quad (1.1)$$

avec  $d_{ij}$  est la distance la plus courte du nœud  $i$  au nœud  $j$  et  $n$  est le nombre total de nœuds dans le réseau. Sachant que la distance entre deux nœuds qui ne sont pas atteignables de l’un à l’autre est 0 et la distance d’un nœud à lui-même est également 0.

Dans de nombreux réseaux à grande échelle, la distance moyenne entre les nœuds est très faible par rapport à la taille du graphe, ce phénomène est connu sous le nom de la propriété *petit-monde*. Cette propriété a été popularisé dans le contexte sociologique où elle est parfois appelée *six degrés de séparation*[40].

L'importance de cette propriété consiste en son rôle important dans le transport et la communication au sein d'un réseau. Supposons qu'il soit nécessaire d'envoyer un paquet de données d'un ordinateur à un autre via Internet : la géodésique fournit un chemin optimal, car on pourrait obtenir un transfert rapide et conserver les ressources du système [43]. Pour une telle raison, les chemins les plus courts ont également joué un rôle important dans la caractérisation de la structure interne d'un graphe [44, 18, 45].

Les ensembles de données massifs du réseau s'accumulent à un rythme énorme dans des champs variés [46]. En utilisant la mesure moyenne du plus court chemin, les réseaux de petit-monde peuvent être considérés comme des systèmes à la fois efficaces à l'échelle mondiale et locale [47]. Nous utilisons souvent la longueur du plus court chemin comme mesure de l'efficacité du réseau, ce qui nous permet de donner une analyse quantitative précise de l'efficacité du flux d'information dans les réseaux. Le calcul du plus court chemin a également été utilisé pour estimer la précision des approximations analytiques de la dynamique sur les réseaux [48], en examinant l'apparition de la synchronisation [49] et en évaluant la résilience des réseaux de communication aux attaques et aux échecs [50].

### 1.4.2 Coefficient de regroupement (Clustering)

En plus de l'effet du petit-monde un haut niveau de regroupement s'en accompagne dans de nombreux réseaux sociaux et différents autres réseaux ont montré cette tendance également, notamment le réseau mondial [51], les réseaux de transport [52] et les réseaux métaboliques [53, 54]. Le concept de regroupement d'un graphe se réfère à la tendance observée dans de nombreux réseaux naturels à former des cliques au voisinage d'un nœud donné, cette propriété est appelée également la transitivité dans le contexte de la sociologie [44].

Le coefficient de regroupement peut être considéré comme la fraction de paires de nœuds avec un nœud commun ou équivalent comme la probabilité moyenne que deux nœuds voisins ont un nœud commun, c'est peut-être la manière la plus utile de définir le coefficient de regroupement. En notation mathématique :

$$C = \frac{3 \times (\text{Nombre de triangles})}{(\text{Nombre de triplets connectés})}. \quad (1.2)$$

Le facteur 3 dans le numérateur compense le fait que chaque triangle complet de trois nœuds contribue à trois triplets connectés, l'un centré sur chacun des trois nœuds et assure que  $0 \leq C \leq 1$ .

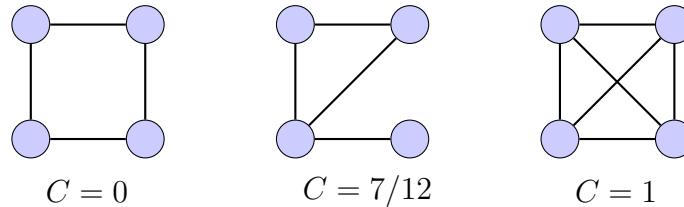


FIGURE 1.10 – Exemple de coefficient de regroupement.

Une autre définition du coefficient de regroupement, également largement utilisé, a été donnée par Watts et Strogatz [55], qui a proposé de définir une valeur locale

$$C_i = \frac{(\text{Nombre de triangles connectés au nœud } i)}{(\text{Nombre de triples centrés sur le nœud } i)}. \quad (1.3)$$

Dans le cas où le nœud a le degré 0 ou 1, nous mettons  $C_i = 0$ . Ensuite, le coefficient de regroupement pour l'ensemble du réseau est la moyenne

$$C = \frac{1}{n} \sum_i C_i. \quad (1.4)$$

Le coefficient de regroupement mesure la densité des triangles dans un réseau. Une généralisation évidente a été envisagée à propos de la densité des boucles plus longues que trois, boucles de longueur quatre et plus. Un certain nombre d'auteurs ont examiné ces coefficients de regroupement d'ordre supérieur [56, 57, 58, 59], bien qu'il n'y ait jusqu'à présent aucune théorie propre qui sépare les contributions indépendantes des différents ordres l'un de l'autre.

### 1.4.3 Distribution des degrés

La propriété la plus importante qui caractérise une structure de réseau est la distribution des degrés  $P(k)$ , définie comme la probabilité qu'un nœud choisi uniformément au hasard ait un degré  $k$  ou, de manière équivalente, la fraction de nœuds dans le graphe ait le degré  $k$ . Si le graphe est dirigé, le degré du nœud comporte deux composantes : le nombre de liens

sortants  $k^{out}$  (appelé "out-degree") et le nombre de liens entrants  $k^{in}$  (appelé "in-degree"). Le degré total est alors défini comme  $k = k^{out} + k^{in}$ .

Un réseau ordinaire a une séquence de degré simple parce que tous les nœuds ont le même nombre de bords, et donc une forme de distribution des degrés qui contient une seule pointe forte. En outre dans le cas d'un réseau complètement aléatoire, la séquence de degré obéit à la distribution de Poisson qui diminue exponentiellement, loin de la valeur moyenne  $\langle k \rangle$ . En raison de ce déclin exponentiel, la probabilité de trouver un nœud avec  $k$  bord devient négligeable pour  $k \gg \langle k \rangle$ . Au cours des dernières années, de nombreux résultats empiriques ont montré que pour la plupart des réseaux réels à grande échelle, la distribution des degrés s'écarte de manière significative de la distribution de Poisson.

En particulier, pour un certain nombre de réseaux, la distribution des degrés peut être mieux décrite par une loi de puissance de la forme  $P(k) \sim k^{-\gamma}$ . Cette distribution de la loi de puissance diminue progressivement et permet de créer quelques nœuds de très grande importance. Étant donné que ces lois de puissance sont libres de toute échelle caractéristique, un tel réseau est appelé un réseau sans-échelle.

#### 1.4.4 Degré de corrélation

Un grand nombre de réseaux réels sont corrélés en sens que la probabilité qu'un nœud de degré  $k$  soit connecté à un autre nœud de degré, disons  $k'$ , dépend de  $k$ . Dans ces cas, il est nécessaire d'introduire la probabilité conditionnelle  $P(k' \setminus k)$ , étant défini comme la probabilité qu'un lien d'un nœud de degré  $k$  soit connecté à un nœud de degré  $k'$  [60]. Bien que les corrélations de degré soient formellement caractérisées par  $P(k' \setminus k)$ , l'évaluation directe de la probabilité conditionnelle donne des résultats extrêmement bruyants pour la plupart des réseaux réels en raison de leur taille finie  $N$ . Ce problème peut être surmonté en définissant le degré moyen des voisins les plus proches d'un nœud  $i$  comme

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in N_i} k_j, \quad (1.5)$$

où la somme s'exécute sur les nœuds appartenant à  $N_i$  qui signifie l'ensemble des premiers voisins de  $i$ . on peut calculer le degré moyen des voisins les plus proches des nœuds avec le degré  $k$ , noté  $k_{nn}(k)$ , obtenant une expression qui intègre implicitement la dépendance de  $k$ . Une telle quantité peut, en effet, être exprimée en termes de probabilité conditionnelle comme

$$k_{nn}(k) = \sum_{k'} k' P(k'|k). \quad (1.6)$$

S'il n'y a pas de corrélation de degré, l'équation Eq. (1.6) donne  $k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$ , c'est-à-dire que  $k_{nn}(k)$  est indépendant de  $k$  [45].

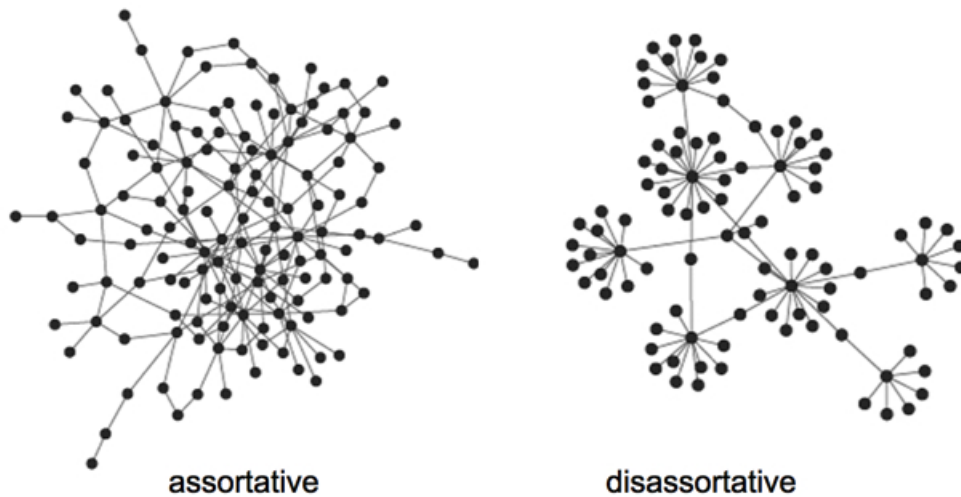


FIGURE 1.11 – Exemple des réseaux assortatif et disassortatif.

Selon l'équation Eq. (1.6) on peut distinguer deux différents types de réseaux, si les nœuds de haut degré dans un réseau s'associent préférentiellement avec d'autres nœuds à haut degré on dit que le réseau est assortatif, et s'ils préfèrent s'attacher à ceux à faible degré on dit que le réseau est disassortatif (voir Fig. 1.11). Les deux situations sont observées dans certains réseaux, mais le cas du réseau assortatif est particulièrement intéressant, car le degré est lui-même une propriété de la topologie des graphes, alors les corrélations de degré peuvent donner lieu à des effets de structure du réseau intéressants [61, 56].

### 1.4.5 Mesures de la centralité

L'identification des nœuds importants dans les réseaux est un problème intéressant qui a suscité beaucoup d'attention, surtout dans le contexte des réseaux de communication. Par exemple, la communication entre un groupe d'humains forme un réseau de communication [62]. Les sciences sociales à la fin des années 1940 ont développé des mesures théoriques pour détecter des nœuds importants dans les réseaux. Une classe importante de ces mesures est basée sur le concept de centralité [63, 64] qui tente intuitivement d'identifier les nœuds qui sont au centre de la communication au sein du réseau parmi tous les nœuds. Il existe deux types fondamentalement différents de mesures de centralité [65]. Le premier type évalue la centralité de chaque nœud dans un réseau et s'appelle des mesures de centralité des points où le mot "point" se réfère à un nœud ou un sommet. Le second type s'appelle les mesures de centralité des graphes car il attribue une valeur de centralité à l'ensemble du réseau. Parmi les paramètres de mesure de centralité on cite :

i) Centralité d'intermédierité (Betweenness) :

L'importance d'un nœud dans un réseau dépend de nombreux facteurs. Un site Web peut être important en raison de son contenu, d'un routeur pour sa capacité, d'un métabolite dû à sa fonction biochimique, etc. Bien sûr, toutes ces propriétés dépendent de la nature du réseau étudié et peuvent avoir très peu à faire avec la structure graphique du réseau. L'une des définitions les plus acceptées de centralité est basée sur les chemins de comptage traversant un nœud  $i$  dans le réseau, on compte le nombre de chemins de "routage" vers tous les autres nœuds passant par  $i$ , et ce nombre détermine la centralité  $i$ . La sélection la plus courante ne prend que les chemins les plus courts en tant que chemins de routage. Cela conduit à la définition suivante : la centralité de l'intermédierité d'un nœud  $i$  équivaut au nombre de chemins les plus courts entre toutes les paires de nœuds dans le réseau qui l'entoure, c'est-à-dire,

$$C_B(i) = \sum_{l \neq j} \frac{\sigma_{lj}(i)}{\sigma_{lj}}, \quad (1.7)$$

avec  $\sigma_{lj}$  indique le nombre des plus court chemins de  $l$  à  $j$ , et  $\sigma_{lj}(i)$  indique le nombre

des plus court chemins de  $l$  à  $j$  passant par  $i$ .

ii) La centralité de proximité :

La centralité de proximité tente de mesurer la proximité d'un nœud avec d'autres nœuds du réseau. Ceci se fait en termes de distance de communication mesurée par le nombre d'arêtes entre deux nœuds si connecté par le chemin le plus court.

$$C_C(i) = \frac{1}{\sum_{j=1}^n d_{ji}}, \quad (1.8)$$

$d_{ji}$  est le plus court chemin entre le nœud  $j$  et  $i$ .

En ce qui concerne la centralité de proximité, les gens se réfèrent généralement à sa forme normalisée qui représente la longueur moyenne des chemins les plus courts au lieu de leur somme. Il est généralement donné par la formule précédente multipliée par  $n - 1$ . Pour les larges réseaux de très grand nombre de nœuds on écrit :

$$C_C(i) = \frac{n}{\sum_{j=1}^n d_{ji}}. \quad (1.9)$$

## 1.5 Propriétés des réseaux réels

### 1.5.1 La propriété petit-monde

La propriété petit monde se réfère au fait que dans plusieurs réseaux, peut-être la plupart des réseaux à grande échelle, la distance moyenne entre les nœuds est très faible par rapport à la taille des graphes. La distance entre deux nœuds d'un graphe est mesurée comme la plus petite longueur de chemin entre eux. Une mesure statistique globale de la distance entre les nœuds peut alors être exprimée comme la longueur moyenne des trajets les plus courts pour tous les couples possibles de nœuds dans le réseau. Nous avons discuté dans la Section. 1.3.2 l'expérience de Stanley Milgram réalisée dans les années 1960, dans laquelle on a trouvé que le nombre d'étapes pour qu'un destinataire reçoit la lettre de l'expéditeur, via le réseau social, est égale à six en moyenne. L'expérience de Milgram est une démonstration simple, magnifique et puissante de l'effet de petit-monde.

L'effet petit-monde a des implications évidentes pour la dynamique des processus qui se



déroulent sur les réseaux. Par exemple, si l'on considère la diffusion de l'information, ou encore tout autre chose, à travers un réseau, l'effet petit-monde implique que cette propagation sera rapide sur la plupart des réseaux réels. Cela affecte le nombre de "sauts" qu'un paquet doit faire pour passer d'un ordinateur à l'autre sur Internet, le nombre d'escales d'un voyage pour un voyageur aérien ou en train, le temps qu'il faut pour qu'une maladie se propage dans une population, et ainsi de suite. L'effet petit-monde atteint également certains jeux de société bien connus, en particulier le calcul des nombres d'Erdős [66] et de Bacon.

### 1.5.2 La distribution des degrés sans-échelle

La distribution des degrés sans-échelle qu'on appelle également la loi de puissance a suscité un intérêt particulier au cours des dernières années pour ses propriétés mathématiques, ce qui entraîne parfois des conséquences physiques surprenantes et son apparence dans la diversité des réseaux naturels et artificiels, voir quelques exemples dans la Fig. 1.12.

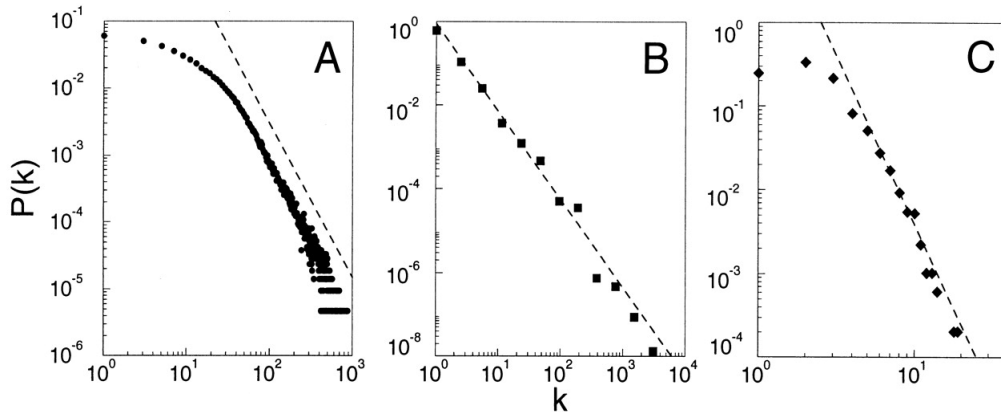


FIGURE 1.12 – Quelques exemples des réseaux sans-échelle, (A) Graphe de collaboration d'acteur avec un nombre des nœuds  $n = 212250$  et un degré moyen  $\langle k \rangle = 28.78$ , (B) WWW,  $n = 325729$ ,  $\langle k \rangle = 5.46$ , (C) Données du réseau électrique,  $n = 4941$  et  $\langle k \rangle = 2.69$ . Les lignes pointillées ont des pentes (A)  $\gamma_{actor} = 2.3$ , (B)  $\gamma_{www} = 2.1$  et (C)  $\gamma_{electrique} = 4$ .

Les distributions de la loi de puissance se produisent dans une gamme de réseaux extraordinairement diversifiée, tels que, les populations des villes [67, 68], la taille des tremblements de terre [69], des cratères de la lune [70], la fréquence d'utilisation des mots dans n'importe quelle langue humaine [71, 72], la fréquence de l'apparition de noms personnels dans la plupart

des cultures [73], le nombre de documents scientifiques écrits [74], le nombre de citations reçues par les documents [75], le nombre de visites sur les pages Web [76], les ventes de livres, les enregistrements musicaux et presque tous les autres produits de marque [77], le nombre d'espèces dans les genres biologiques [78], les revenus annuels des personnes [79], les collecteurs de réseaux quantiques complexes [80] et une foule d'autres réseaux suivent toutes des distributions de loi de puissances.

Mathématiquement, une quantité  $x$  obéit à une loi de puissance s'elle est tirée d'une distribution de probabilité

$$P(x) \propto x^{-\gamma} \quad (1.10)$$

Avec  $\gamma$  est un paramètre constant de la distribution connue sous le nom d'exposant, empiriquement cet exposant se situe généralement dans l'intervalle  $2 < \gamma < 3$ , bien qu'il existe des exceptions occasionnelles.

### 1.5.3 Structure communautaire

La société offre une grande variété d'organisations de groupes possibles : les familles, les milieux de travail et d'amitié, les villages, les villes, les nations (voir un exemple Fig. 1.13). La diffusion d'Internet a également conduit à la création de groupes virtuels directement sur le Web. Ces communautés se forment également dans de nombreux systèmes en réseau, de la biologie, de l'informatique, de l'ingénierie, de l'économie, de la politique, etc. Par exemple dans le graphe du World Wide Web les communautés peuvent correspondre à des groupes de pages portant sur les mêmes sujets ou des sujets connexes [81, 82]. Dans les réseaux d'interactions protéines-protéines, les communautés sont susceptibles de regrouper des protéines ayant la même fonction spécifique dans la cellule [83, 84].

Les communautés peuvent avoir des applications concrètes. Le Clustering des clients Web qui ont des intérêts similaires et qui sont géographiquement proches les uns des autres peut améliorer la performance des services fournis sur le WWW, en ce sens chaque groupe de clients pourrait être servi par un serveur miroir dédié [85]. Les réseaux auto-configurés formés

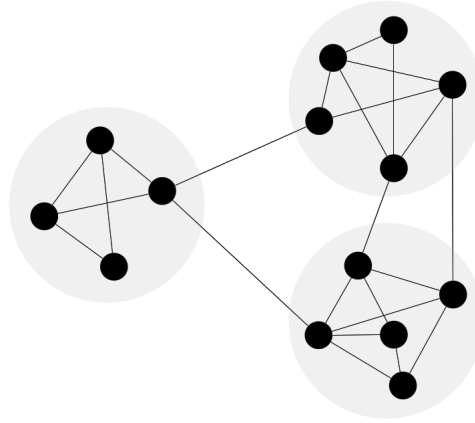


FIGURE 1.13 – Exemples d’un réseau communautaire, on voit trois groupes où chacun est plus connecté par rapport au reste du réseau.

par des nœuds de communication agissant dans la même région et changeant rapidement, ils n’ont généralement pas de tables de routage centralisées qui spécifient comment les nœuds doivent communiquer entre eux. Dans ce cas le regroupement des nœuds en clusters permet de générer des tables de routage compactes et rend le choix des chemins de communication plus efficace [86].

La détection communautaire est également importante pour d’autres raisons. L’identification des modules et de leurs limites permet une classification des nœuds, en fonction de leur position structurelle dans les modules. Donc, les nœuds avec une position centrale dans leurs grappes partagent un grand nombre des liens avec les autres partenaires du groupe, ce qui est une caractéristique importante de contrôle et de stabilité au sein du groupe, en outre, Les nœuds situés aux frontières entre les modules jouent un rôle important de médiation et mènent les relations et les échanges entre les différentes communautés [87].

## 1.6 Les modèles théoriques les plus connus

Dans cette section, nous citons les modèles théoriques les plus connus et les plus étudiés, en se limitant principalement à des réseaux non pondérés et non orientés, cependant, la plupart des réseaux présentés peuvent être généralisés facilement. Ces modèles se caractérisent chacun par la manière dont les réseaux sont créés et par plusieurs statistiques résultantes, telles que la distribution des degrés, le plus court chemin entre les nœuds et le coefficient de Clustering.

## Réseaux simples

Un réseau simple consiste des connexions régulières entre les nœuds. L'un des exemples les plus exceptionnel est le réseau à deux dimensions, comme le montre la Fig. 1.14. Ici, chaque nœud est connecté à ses voisins les plus proches. Malgré sa simplicité, de tels réseaux ont été largement utilisés, par exemple en physique, pour étudier des phénomènes tels que le ferromagnétisme avec le modèle d'Ising [88]. D'autres exemples de cette classe sont les chaînes linéaires ou les réseaux non rectangulaires tels qu'utilisés, par exemple, dans le contexte de la prédiction de la structure des protéines pour modéliser le repliement des protéines [89, 62].

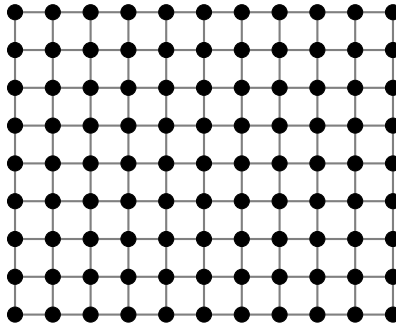


FIGURE 1.14 – Simple réseau régulier de deux dimensions

## Réseaux aléatoires

Les réseaux aléatoires ont été largement étudiés par Erdős et Rényi (ER) [90, 91, 92]. Un graphe aléatoire avec  $n$  nœuds est obtenu en connectant chaque paire de nœuds avec la probabilité  $p$ . Le nombre attendu d'arêtes pour un réseau (non orienté) construit de cette façon est  $E = p \frac{n(n-1)}{2}$ .

## Réseaux petit-monde

Mathématiquement, l'effet de petit-monde décrit les graphes dont le diamètre et la longueur moyenne de la trajectoire croissent beaucoup plus lentement que le nombre de nœuds  $n$ , typiquement en tant que  $\ln n$ , comme dans un graphe aléatoire ER. Pourtant, un

graphe aléatoire a une très faible interconnexion locale, coefficient de regroupement. Watts et Strogatz [93] ont réalisé un modèle qui rejoinde la propriété petit-monde et un coefficient de regroupement fort. Le modèle commence par placer tous les nœuds en cercle, puis connecter chaque nœud à ses premiers  $k$  voisins les plus proches et en fin avec une probabilité  $p$  en prenant chaque lien du réseau et on le reconnecte aléatoirement entre deux autres nœuds.

### Réseaux sans-échelle

Ni les réseaux aléatoires ni les réseaux petit-monde ont la propriété sans-échelle des degrés, qui est fréquemment observée dans les réseaux du monde réel. Pour expliquer cette caractéristique, Barabási et Albert ont introduit un modèle maintenant connu sous le nom de Barabási-Albert (BA) [3] ou modèle d'attachement préférentiel qui aboutit à des réseaux sans-échelle qui ont une distribution des degrés suivant une loi de puissance,  $P(k) \sim k^{-\gamma}$ . La différence majeure entre le modèle d'attachement préférentiel et les autres algorithmes décrits ci-dessus pour générer des réseaux aléatoires ou de petit-monde, est que le modèle d'attachement préférentiel ne suppose pas un nombre fixe de nœuds  $n$  et commence alors à les rebrancher avec une probabilité fixe avec d'autres nœuds.



# Chapitre 2

## Les réseaux en croissance et l'attachement préférentielle

Malgré de nombreux efforts, la théorie cohérente des réseaux en croissance manque encore un principe général prédisant la topologie d'un réseau formé. Dans le but de comprendre la formation et l'évolution des réseaux complexes, de nombreux modèles ont été introduits pour étudier les processus microscopiques impliqués dans la structure du réseau qui en résulte. Dans ce contexte, on introduit ici un modèle de réseau simple et complexe qui augmente avec un mécanisme linéaire d'attachement préférentiel et sans le mécanisme "rich get richer". L'objectif est double : d'abord vérifier si la distribution des degrés en loi de puissance reste en l'absence du mécanisme "rich get richer" et, d'autre part, voir pour éventuellement les différences microscopiques entre les réseaux sans échelle et homogènes.

### 2.1 Introduction

Au cours des dernières années, il existe un intérêt croissant à étudier l'évolution des réseaux complexes et à développer des modèles qui reflètent certaines propriétés des réseaux réels en utilisant les techniques de la mécanique statistique, la théorie des graphes et les simulations informatiques [3, 94, 95, 96]. L'une des propriétés les plus importantes étudiée en réseaux est la distribution des degrés qui est la probabilité  $P(k)$  qu'un nœud ait un degré  $k$ . On peut distinguer trois lois principales de la distribution des degrés : loi de Poisson où

$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$ , loi de puissance  $P(k) \approx k^{-\gamma}$  avec  $\gamma$  représente le degré d'exposant, et la loi exponentiel  $P(k) \approx e^{-\frac{k}{c}}$  où  $c$  est une constante.

Il semble que, dans la nature, la plupart des réseaux suivent les deux dernières lois de distribution mentionnées ci-dessus. Barabási-Albert (BA) a réinventé le réseau de distribution des degrés sans-échelle de Pric en introduisant un modèle simplifié basé à la fois sur la croissance et l'attachement préférentiel. Comme nous avons déjà signalé les réseaux sans-échelle sont largement observés dans une variété de systèmes tels que des réseaux de citation de publication, de nombreux réseaux sociaux, des réseaux de protéines et de gènes. Cependant, il existe d'autres réseaux réels qui suivent une loi exponentielle, par exemple, le Réseau mondial de transport maritime [97], le réseau nord-américain Power Gridc [98], réseaux neuronaux des *C.elegans* [99] et le réseau de messagerie à l'Université de Rovira i Virgili (ENURV) en Espagne [98].

La loi exponentielle semble être le résultat d'un réseau croissant en ajoutant de nouveaux nœuds et liens aléatoires. D'autre part, la loi de puissance semble apparaître lorsque les nœuds sont ajoutés au réseau un à la fois et en les liant à des nœuds déjà connectés. De nombreuses idées sur la formation des réseaux ont été examinées. Par exemple, Barabási, dans son modèle, a affirmé que l'attachement préférentiel et la croissance sont tous les deux nécessaires pour générer un réseau sans-échelle [100]. En réalité, il semble que la croissance n'est pas nécessaire pour un tel objectif [101]. En outre, il est intuitif, que l'attachement préférentiel sans l'effet "rich get richer" ne génère pas un réseau sans-échelle [102]. Krapivsky et al [103] ont étudié la fixation préférentielle non linéaire avec  $\Pi(k_i) \propto k^\gamma$  et ils ont montré que pour  $\gamma < 1$ , le mécanisme produit une distribution des degrés de loi exponentiel.

## 2.2 Les processus dynamiques : théorie et simulation

### 2.2.1 Introduction

Cette section est destinée à donner une brève introduction à la théorie et à la modélisation des processus équilibrés et non équilibrés sur les réseaux et à définir les approches et techniques de base de la modélisation utilisées dans la théorie des processus dynamiques. En particu-



lier, on définit le formalisme de l'équation maîtresse. Malheureusement, même si l'équation maîtresse permet une distinction et une catégorisation conceptuelles importantes, sa solution complète n'est pas assez simple même pour des processus dynamiques très simples. Pour cette raison, nous présentons des techniques qui seront va utilisées au long de cette thèse, telles que les approximations déterministes du champ moyen, qui représentent habituellement des approches pour comprendre les caractéristiques fondamentales du processus étudié. Nous discutons également des approches de modélisation basées sur des outils de Monte Carlo qui sont généralement implémentées dans des méthodes de simulation numérique à grande échelle.

Le but de ces différentes méthodes théoriques est de fournir un cadre général pour démontrer comment les interactions microscopiques entre les éléments du système conduisent à des phénomènes coopératifs et à l'émergence des propriétés des processus dynamiques. Cette stratégie, allant de l'interaction microscopique aux phénomènes collectifs émergents, a ses racines dans la méthodologie de la physique statistique et la dynamique de la population, et est actuellement considérée comme un paradigme général pour combler l'écart entre les propriétés locales et les propriétés à grande échelle des systèmes complexes. Il est important de souligner que cette présentation est une synthèse d'un énorme domaine de recherche et ne fait que rayonner la surface de la théorie statistique des processus dynamiques. Les lecteurs intéressés qui veulent plonger dans les subtilités mathématiques et formelles du sujet devraient se référer à des manuels classiques tels que ceux de Ma (1985), Chandler (1987), Huang (1987) et Balescu (1997).

### 2.2.2 Équation maîtresse

Le nom "équation maîtresse" a été inventé à l'origine par Nordsieck, Lamb et Uhlenbeck [104] dans leur étude du modèle Furry des averses de pluie cosmiques [105]. Peu de temps auparavant, Feller a appliqué une équation de la même structure à la croissance des populations [106] et Delbrück à des réactions chimiques auto-catalytiques bien mélangées [107]. Pour plus de détails vous pouvez voire ces références [108, 109, 110].

Deux schémas de modélisation principaux sont adoptés pour traiter les processus dynamiques sur les réseaux. Dans le premier, nous identifions chaque nœud du réseau avec un seul individu

ou élément du système. Dans le second cas, nous considérons les entités dynamiques telles que les personnes, les paquets d'information, l'énergie ou la matière circulant dans un réseau dont les nœuds identifient les endroits où transitent les entités dynamiques. Cependant, dans les deux cas, la description dynamique du système peut être obtenue en introduisant pour chaque nœud  $i$  la notion de variable correspondante  $\sigma_i$  caractérisant son état dynamique. Si chaque nœud représente un seul individu, la variable  $\sigma_i$  peut décrire un attribut particulier de l'individu. Sans perdre de généralité, nous pouvons énumérer tous les états possibles  $\sigma_i = 1, 2, \dots, m$  pour chaque nœud, et la connaissance de la variable d'état de tous les nœuds du réseau définit donc l'état microscopique de l'ensemble du système. Alors on peut désigner une configuration particulière du réseau à l'instant  $t$  par l'ensemble  $\sigma(t) = [\sigma_1(t), \sigma_2(t), \sigma_3(t), \dots, \sigma_N(t)]$ , où l'indice  $i = 1, \dots, N$  parcourt tous les nœuds du réseau de taille  $N$ .

L'évolution dynamique du système est simplement donnée par la dynamique de la configuration  $\sigma(t)$  dans l'espace des phases du système, définie par toutes les configurations possibles. Le processus dynamique est décrit par les transitions d'un état  $\sigma^a$  vers un autre état  $\sigma^b$ . En général, il est impossible de suivre la dynamique microscopique des systèmes à grande échelle en raison du grand nombre de variables et de la nature stochastique de la plupart des phénomènes. Pour cette raison, la description dynamique de base du système repose sur l'approche de l'équation maîtresse (EM) que nous allons brièvement introduire.

L'approche EM se concentre sur l'étude de la probabilité  $P(\sigma, t)$  de trouver le système à l'instant  $t$  dans une configuration donnée  $\sigma$ . Cette probabilité doit être normalisée,  $\sum_{\sigma} P(\sigma, t) = 1$ , et fournit une description probabiliste du système qui donne les informations les plus pertinentes

$$\frac{\partial P(\sigma, t)}{\partial t} = \sum_{\sigma'} [P(\sigma', t)W(\sigma' \rightarrow \sigma) - P(\sigma, t)W(\sigma \rightarrow \sigma')], \quad (2.1)$$

où la somme s'exécute sur toutes les configurations possibles  $\sigma$  et les termes  $W(\sigma' \rightarrow \sigma)$  représentent les taux de transition de la configuration  $\sigma'$  vers la configuration  $\sigma$  en raison de la dynamique microscopique du système.

Les EM et les simulations basées sur elles, sont maintenant utilisées dans de nombreux domaines de recherche. Elles sont appliquées dans les contextes de la dynamique de spin [111, 112, 113, 114], des réseaux régulateurs de gènes [115, 116, 117], de la propagation des

maladies [118, 119], de l'homéostasie épidermique [120], et les processus sociaux et économiques [121]. Les processus de file d'attente sont souvent modélisés en termes d'EM, mais dans ce contexte, les équations sont généralement appelées équations de Kolmogorov [122].

### 2.2.3 Modélisation et simulations numériques

La simulation numérique a commencé dans les années cinquante lorsque les ordinateurs ont été utilisés pour la première fois à des fins pacifiques, en particulier, l'ordinateur MANIAC a commencé en 1952 à Los Alamos. La simulation fournit une approche complémentaire aux méthodes théoriques. Les domaines de la physique où les approches perturbatives sont efficaces (gaz dilués, vibrations dans les solides quasi-harmoniques) ne nécessitent pas de méthodes de simulation. Inversement, la physique des états liquides, où peu de résultats exacts sont connus et où les développements théoriques ne sont pas toujours sous contrôle, a été développée en grande partie par simulation. La première simulation de liquides par la méthode Monte Carlo a été réalisée par Metropolis et al. en 1953.

Dans des modèles plus complexes, même l'approche déterministe pourrait ne pas conduire à des équations résolubles. De plus, ce cadre théorique, intrinsèquement, ne prend pas en compte l'hétérogénéité individuelle ou d'autres fluctuations possibles. L'intégration numérique sur l'ordinateur des équations obtenues ne fournit donc pas une image complète du système. Dans cette situation, des modèles informatiques microscopiques, peuvent être appliqués.

Dans ces approches, chaque nœud individuel est supposé être dans l'un des états possibles. A chaque pas de temps, la procédure de mise à jour spécifique au modèle qui dépend de la dynamique microscopique est appliquée à chaque nœud, ce qui modifie son état en fonction de l'état des nœuds voisins ou d'autres règles dynamiques. Notamment, la stochasticité du modèle peut être introduite en utilisant des simulations de Monte Carlo dans lesquelles les taux et les probabilités sont établis dans l'ordinateur avec l'utilisation de générateurs de nombres aléatoires. La perspective microscopique de cette approche est évidente dans le fait que l'on peut suivre la dynamique de chaque élément individuel. De plus, la dynamique de système se situe au niveau des interactions microscopiques entre les éléments, et les régularités statistiques et les propriétés macroscopiques du système sont étudiées en regardant les quantités globales ou moyennes. L'ordinateur est donc utilisé comme un laboratoire pour étudier des réalités

complexes non accessibles mathématiquement ou expérimentalement.

## 2.3 L'évolution dynamique et l'attachement préférentiel dans les réseaux réels

Nous commençons ce paragraphe en se demandant : Pourquoi les hubs et la loi sans-échelle sont-elles absentes dans les réseaux aléatoires ? La réponse a émergé en 1999, mettant en évidence deux hypothèses qui sont apparues dans les réseaux réels [3]. Nous discutons ensuite de ces hypothèses séparément.

### 2.3.1 L'évolution dynamique

Le modèle de réseau aléatoire suppose qu'on a toujours un nombre fixe de nœuds,  $N$ . Pourtant, dans les réseaux réels, le nombre de nœuds ne cesse de croître grâce à l'ajout de nouveaux nœuds, par exemple, En 1991, le WWW avait un seul nœud, la première page Web construite par Tim Berners-Lee, le créateur du Web. Aujourd'hui, le Web a plus d'un trillion ( $10^{12}$ ) de documents, un nombre extraordinaire qui a été atteint grâce à l'ajout continu de nouveaux documents par des millions d'individus et d'institutions (voir Fig. 2.1). Ainsi que le réseau d'acteurs qui continue de se développer [123]. Par conséquent, si nous souhaitons modéliser ces réseaux, nous ne pouvons pas recourir à un modèle statique. Notre approche de modélisation doit plutôt reconnaître que les réseaux sont le produit d'un processus de croissance constant.

### 2.3.2 Le modèle sans-échelle de Barabási-Albert

Une caractéristique commune entre le graphe aléatoire ER et les modèles de petit-monde de WS est que la distribution des degrés du réseau est homogène, avec un pic à une valeur moyenne et une décroissance exponentielle, de tels réseaux s'appellent des réseaux exponentiels. Les inconvénients de ces deux modèles précédents est qu'ils ne tiennent pas compte de deux attributs importants de la plupart des réseaux réels. Premièrement, les réseaux réels sont ouverts et ils sont dynamiquement formés par l'ajout continu de nouveaux nœuds au réseau,

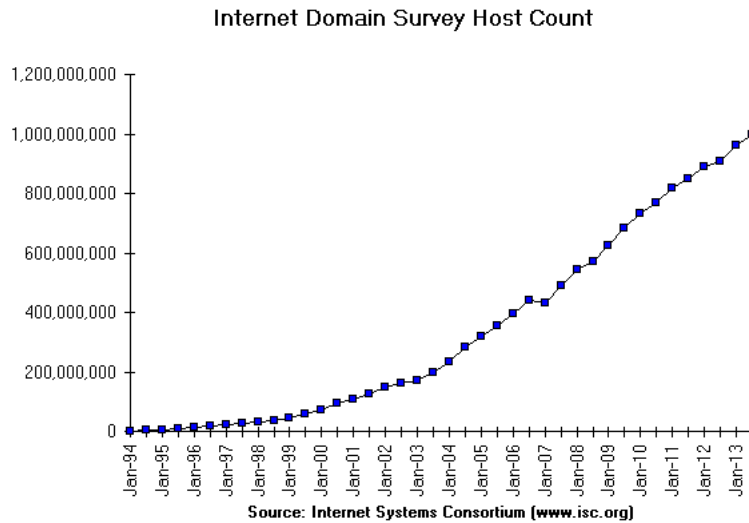


FIGURE 2.1 – L'évolution de nombre d'hôtes sur Interent de 1994 à 2013, cette figure obtenues à partir de site : <http://facesncups.com/inforev.html>

par exemple, la WWW génère continuellement de nouvelles pages Web et la littérature de recherche se développe constamment car les nouveaux articles sont en cours de publication. Deuxièmement, le graphe aléatoire ER et le modèle de petit-monde de WS prennent des probabilités uniformes lors de la création de nouvelles arêtes, mais ce n'est pas réaliste non plus.

Une découverte relativement récente et importante dans le domaine des réseaux complexes est l'observation qu'un certain nombre de réseaux complexes à grande échelle, y compris Internet, WWW, les réseaux métaboliques et plusieurs autres réseaux réels, sont sans échelle et leurs distributions de connectivité ont une forme de puissance (voir Fig. 1.12).

Nous parlons alors ici d'une classe de modèles dont l'objectif principal est de reproduire les processus de croissance qui se déroulent dans des réseaux réels. Nous nous concentrons principalement sur le modèle de croissance du réseau proposé par Barabási et Albert en 1999 [3], ce modèle repose sur deux hypothèses simples concernant l'évolution du réseau :

- **Croissance** : De nouveaux nœuds sont ajoutés au réseau, chaque nouveau nœud étant connecté à  $m$  des nœuds existants.
- **Attachement préférentielle** : Chaque nouveau nœud est connecté aux nœuds existants avec une probabilité proportionnelle à son degré.

De manière plus détaillée, considérons un réseau évoluant dans le temps,  $t$ , où à chaque fois,

un nouveau nœud est ajouté au réseau et connecté à  $m$  des nœuds existants, où la probabilité de se connecter à un nœud existant  $i$  de degré  $k_i$ , est donné par :

$$\Pi(i) = \frac{k_i}{\sum_j k_j}. \quad (2.2)$$

Le nombre de nœuds initial à l'instant  $t = 0$  est généralement supposé être connecté et doit être supérieur à  $m$ , mais les détails de sa structure n'ont qu'un faible effet sur le résultat final.

Il existe plusieurs méthodes pour analyser les résultats du modèle Barabási-Albert. La méthode la plus simple est une analyse de champ moyen décrite premièrement dans [3]. D'autres méthodes plus rigoureuse d'analyse de ce modèle utilisant des outils de la physique statistique comprennent l'approche de l'équation maîtresse [124] et l'approche de l'équation du taux [103]. La distribution des degrés trouvée est décrite par une loi de puissance avec l'exposant  $-3$ , c'est-à-dire que la probabilité de trouver un nœud avec le degré  $k$  est proportionnelle à  $k^{-3}$  (voir Fig. 2.2). Afin de trouver l'expression exacte de  $P(k)$ , nous utilisons l'approche de l'équation maîtresse.

Soit  $N(k, t)$  le nombre de nœuds de degré  $k$  à l'instant  $t$ . La distribution des degrés  $P(k, t)$  se rapporte à cette quantité via la relation  $P(k, t) = \frac{N(k, t)}{N(t)}$ . Puisque à chaque pas de temps nous ajoutons un nouveau nœud au réseau, nous avons  $N = t$ . C'est-à-dire qu'à tout moment le nombre total de nœuds est égal au nombre de pas de temps<sup>1</sup>.

L'attachement préférentiel d'un nouveau nœud avec un ancien nœud dans le réseau de degré  $k$  est selon Eq. (2.2) comme

$$\Pi(k) = \frac{k}{\sum_j k_j} = \frac{k}{2mt}, \quad (2.3)$$

où le terme  $2m$  représente le fait que dans un réseau, non orienté, chaque lien ajoute deux degrés au réseau. Notre objectif est de calculer les changements dans le nombre de nœuds de degré  $k$  après l'ajout d'un nouveau nœud au réseau. Pour cela nous inspectons les deux événements qui modifient  $N(k, t)$  et  $P(k, t)$  suite à l'arrivée d'un nouveau nœud : le premier est lorsque le nouveau nœud lie à un nœud de degré  $k$  et il le transforme en un nœud de degré  $(k + 1)$ , ce qui diminue  $N(k, t)$ , le second lorsque le nouveau nœud lie à un nœud de degré

---

1. En négligeant le nombre des nœuds au temps initial.

$(k - 1)$  et il le transforme en un nœud de degré  $k$ , augmentant ainsi  $N(k, t)$ .

Le nombre de liens attendus pour se connecter aux nœuds de degré  $k$  après l'arrivée d'un nouveau nœud est  $\frac{k}{2mt} \times NP(k, t) \times m = \frac{k}{2}P(k, t)$ , autrement dit cette formule représente le nombre des nœuds de degrés  $k$  qui acquièrent un nouveau lien et augmente leur degré à  $k + 1$ . De même logique, le nombre des nœuds de degrés  $(k - 1)$  qui acquièrent un nouveau lien et augmente leur degré à  $k$  est  $\frac{k-1}{2}P(k - 1, t)$ . En combinant ces deux formules précédentes on obtient le nombre attendu de nœuds de degré  $k$  après l'ajout d'un nouveau nœud

$$(N + 1)P(k, t + 1) = NP(k, t) + \frac{k - 1}{2}P(k - 1, t) - \frac{k}{2}P(k, t). \quad (2.4)$$

Cette équation s'applique à tous les nœuds de degré  $k > m$ . Le fait que nous manquons les nœuds de degré  $k < m$  dans le réseau ( car chaque nouveau nœud arrive avec un degré  $m$ ), nous avons alors besoin d'une équation séparée pour le cas de degré  $m$ . En suivant les mêmes arguments que nous avons suivi pour Eq. (2.4), nous obtenons

$$(N + 1)P(m, t + 1) = NP(m, t) + 1 - \frac{m}{2}P(m, t). \quad (2.5)$$

Notre but est de trouver une distribution stationnaire des degrés. Cela nous pousse à supposer que dans la limite  $N = t \rightarrow \infty, P(k, \infty) = P(k)$ . En utilisant cela nous pouvons écrire à partir de Eq. (2.4) et Eq. (2.5) que

$$P(k) = \frac{k - 1}{k + 2}P(k - 1) \quad \text{si } k > m, \quad (2.6)$$

et pour  $k = m$  on obtient

$$P(m) = \frac{2}{m + 2}. \quad (2.7)$$

Nous utilisons une approche récursive pour obtenir la distribution des degrés. De Eq. (2.6) on peut écrire

$$\begin{aligned}
 P(k) &= \frac{k-1}{k+2} P(k-1) \\
 &= \frac{(k-1)(k-2)}{(k+2)(k+1)} P(k-2) \\
 &\vdots \\
 &= \frac{(k-1)(k-2)\dots(m+1)m}{(k+2)(k+1)\dots(m+4)(m+3)} P(m) \\
 &= \frac{(m+2)(m+1)m}{(k+2)(k+1)k} P(m),
 \end{aligned} \tag{2.8}$$

de l'équation précédent et sachant que  $P(m) = \frac{2}{m+2}$  on obtient l'expression finale de la distribution des degrés dans le modèle BA,

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)}. \tag{2.9}$$

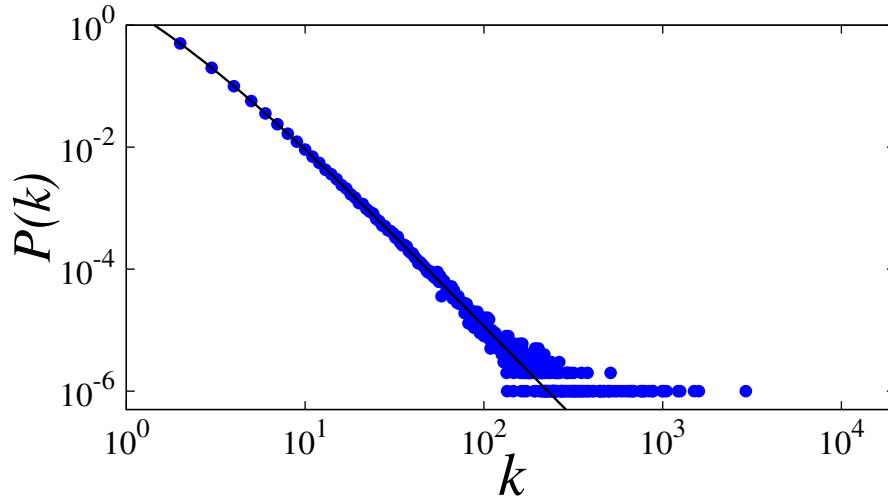


FIGURE 2.2 – La distribution des degrés du réseau BA en échelle logarithmique, où les cercles sont les simulations numériques et la ligne continue représente la formule théorique, Eq. (2.9). La taille du réseau est  $n = 10^6$  et  $m = 2$ .



L'Eq. (2.9) a d'abord été dérivée par Krapivsky et al. [103] et de manière indépendante par Dorogovtsev et al. [124]. Un traitement plus détaillé a ensuite été donné par Bollobas et al. [125], qui clarifie précisément le domaine de validité de la solution et les écarts éventuels par rapport à la valeur attendue de  $P(k)$ .

La distance moyenne dans le modèle BA est plus petite que dans le graphe aléatoire ER. Les résultats analytiques prédisent une correction de double logarithmique qui diminue la dépendance logarithmique  $L \sim \frac{\log N}{\log \log N}$  [125]. Le coefficient de regroupement décroît avec la taille du système comme  $C \sim N^{-0,75}$ . Il s'agit d'une décroissance plus lente que celle observée pour les graphes aléatoires,  $C \sim N^{-1}$ , mais le comportement reste différent par rapport aux modèles du petit monde, où  $C$  est une constante indépendante de la taille du réseau  $N$ .

### 2.3.3 L'attachement préférentielle et le mécanisme "rich get richer"

Le modèle des réseaux aléatoires suppose que nous choisissons au hasard les partenaires d'interaction d'un nœud. Pourtant, dans la plupart des réseaux réels les nouveaux nœuds préfèrent se lier aux nœuds les plus connectés, ce processus appelé attachement préférentiel. Par exemple, nous ne connaissons qu'une infime fraction du trillion ou plus de documents disponibles sur le WWW. Les nœuds que nous connaissons ne sont pas entièrement aléatoires : nous avons tous entendu parler de Google et de YouTube, mais nous rencontrons rarement les milliards de nœuds moins importants qui sont dans le Web. Comme nos connaissances sont biaisées vers les documents Web les plus populaires, nous sommes plus susceptibles de se lier à un nœud de haut niveau qu'à un nœud avec seulement quelques liens [126].

L'ancienneté, cependant, n'est pas suffisante pour expliquer la loi libre-échelle. Les Hubs nécessitent l'aide de la deuxième loi, l'attachement préférentiel. Parce que les nouveaux nœuds préfèrent se lier aux nœuds les plus connectés, les premiers nœuds avec plus de liens seront sélectionnés plus souvent et croîtront plus rapidement que les autres qui sont plus jeunes et moins connectés. Comme de plus en plus de nœuds arrivent et continuent de choisir les nœuds les plus connectés, les premiers nœuds vont inévitablement se détacher du paquet,

acquérant un très grand nombre de liens. Ils vont se transformer en hubs. Ainsi, l'attachement préférentiel induit un phénomène "rich get richer" qui aide les nœuds les plus connectés à devenir plus connectés. Alors ce phénomène "rich get richer" mène naturellement aux lois sans-échelle observées dans les réseaux réels, nous allons confirmer cette idée dans la section suivante par notre propre travail et par une façon plus rigoureuse.

## 2.4 Attachement préférentiel sans l'effet "rich get richer"

### 2.4.1 Le modèle

À l'instar du modèle BA original, notre réseau évolue selon deux mécanismes : la croissance et l'attachement préférentiel. Les nœuds entrant dans le réseau préfèrent s'attacher à des nœuds de faible degré, alors la probabilité  $\Pi(k_i)$  que l'un des liens d'un nouveau nœud se connecte au nœud  $i$  dépend de son degré  $k_i$  tel que  $\Pi(k_i) = C(1 - \frac{k_i}{\sum_j k_j})$ . où  $C$  est la constante de normalisation.

En ce qui concerne les réseaux sociaux, si nous considérons le degré de nœuds comme décrivant la richesse des gens dans une société capitaliste, nous savons que nous vivons dans un monde où les riches s'enrichissent, mais quelle sorte de société nous aurons s'il n'y a pas de faveur pour les gens riches, et il y a plutôt une subvention continue pour les pauvres ?

Pour mettre en œuvre notre idée, nous commençons par  $m_0$  nœuds, chacun avec  $m$  liens. À chaque pas de temps, nous ajoutons un nouveau nœud avec  $m$  arêtes qui lient le nouveau nœud à  $m$  différents nœuds déjà présents dans le réseau. La probabilité que le nouveau nœud soit connecté à un nœud  $i$  de degré  $k_i$  est  $\Pi(k_i) = C(1 - \frac{k_i}{\sum_j k_j})$ . La constante de normalisation  $C$  est déduite de la condition  $\sum_{i=1}^t \Pi(k_i) = 1$ , qui donne  $C = \frac{1}{t+m_0-1}$ .  $t$  est l'instant auquel le dernier nœud a été créé et représente également le nombre de nœuds ajoutés au réseau.

### 2.4.2 Distribution des degrés en utilisant l'équation maîtresse

Notons  $N(k, t)$  le nombre de nœuds de degré  $k$  à l'instant  $t$ . La distribution des degrés à un instant donné  $t$  sera écrite  $P(k, t) = \frac{N(k, t)}{N(t)}$ . Puisque à chaque pas de temps nous ajoutons

un nouveau nœud au réseau, nous avons  $N = t$ . C'est, à tout moment, le nombre total de nœuds est égal au nombre de pas de temps.

Nous écrivons l'attachement préférentiel dans notre modèle comme

$$\Pi(k) = C\left(1 - \frac{k}{\sum_j k_j}\right) = C\left(1 - \frac{k}{2mt + mm_0}\right) \quad (2.10)$$

le terme  $2m$  capture le fait que chaque lien contribue au degré de deux nœuds, et  $mm_0$  capture le fait qu'à l'instant initial nous commençons par  $m_0$  nœuds, chacun avec  $m$  liens. Notre objectif est de calculer les changements dans le nombre de nœuds de degré  $k$  après l'ajout d'un nouveau nœud au réseau. Pour cela, nous respectons les deux événements qui modifient  $P(k, t)$  suite à l'arrivée d'un nouveau nœud :

- Un nouveau nœud peut être lié à un nœud de degré  $k$ , le transformant en un nœud de degré  $(k + 1)$ , ce qui diminue  $P(k, t)$ .
- Un nouveau nœud peut être lié à un nœud de degré  $(k - 1)$ , le transformant en un nœud de degré  $k$ , augmentant ainsi  $P(k, t)$ .

Alors pour ce modèle, l'équation maîtresse peut être écrite comme suit :

$$\begin{aligned} (t + 1)P(k, t + 1) = & tP(k, t) + m\Pi(k - 1, t)tP(k - 1, t) \\ & - m\Pi(k, t)tP(k, t) + \delta_{k,m}, \end{aligned} \quad (2.11)$$

où  $\delta$  est le symbole Kronecker.

L'équation stationnaire correspondante prend la forme suivante :

$$\begin{aligned} (t + 1)P(k) = & tP(k) + m\left(1 - \frac{k - 1}{2mt + mm_0}\right)\frac{tP(k - 1)}{t - 1} \\ & - m\left(1 - \frac{k}{2mt + mm_0}\right)\frac{tP(k)}{t - 1} + \delta_{k,m}, \end{aligned} \quad (2.12)$$

pour des temps très grand on peut écrire que  $t + 1 = t$  et  $t - 1 = t$ , d'où

$$P(k)\left(1 + m\left(1 - \frac{k}{2mt + mm_0}\right)\right) = m\left(1 - \frac{k - 1}{2mt + mm_0}\right)P(k - 1) + \delta_{k,m}, \quad (2.13)$$

après quelques lignes on obtient facilement que :

$$P(k) = \begin{cases} \frac{2mt - (k-1)}{2t + 2mt - k} P(k-1), & \text{for } k > m, \\ \frac{2t}{2t + 2mt - m}, & \text{for } k = m. \end{cases} \quad (2.14)$$

La relation de récurrence ci-dessus donne la solution suivante :

$$P(k) = \begin{cases} \frac{2t}{2t + 2mt - m} \prod_{j=m+1}^k \left( \frac{2mt - j + 1}{2t + 2mt - j} \right), & \text{for } k > m, \\ \frac{2t}{2t + 2mt - m}, & \text{for } k = m. \end{cases} \quad (2.15)$$

Bien que cette équation ne soit pas fermée, l'estimation numérique de  $P(k)$  est simple comme le montre Fig. 2.3.

Nous simulons également le réseau avec des tailles allant jusqu'à  $n = 2 \times 10^6$ , le nombre initial des nœuds est  $m_0 = 3$  et  $m = 2$ . Les résultats de la simulation confirment fortement les résultats analytiques (voir Fig. 2.3).

Nous avons observé dans les simulations que  $k$  reste inférieur à 40 pour  $t = 2.10^6$ , nous prenons alors  $t \gg j$  dans Eq. (2.15) et nous obtenons

$$P(k) \approx \begin{cases} \frac{1}{1+m} \left( \frac{m}{1+m} \right)^{k-m-1}, & \text{for } k > m, \\ \frac{1}{1+m}, & \text{for } k = m. \end{cases} \quad (2.16)$$

Après la normalisation, nous obtenons la distribution des degrés exponentielle  $P(k) = Ae^{-A(k-m)}$ , avec  $A = \log\left(\frac{m+1}{m}\right)$ . L'encart de la figure Fig. 2.3 montre la forme exponentielle de  $P(k)$  et l'excellent accord entre les simulations et les résultats théoriques. Ceci confirme clairement que l'attachement préférentiel seul n'est pas suffisant pour produire des réseaux sans-échelle.

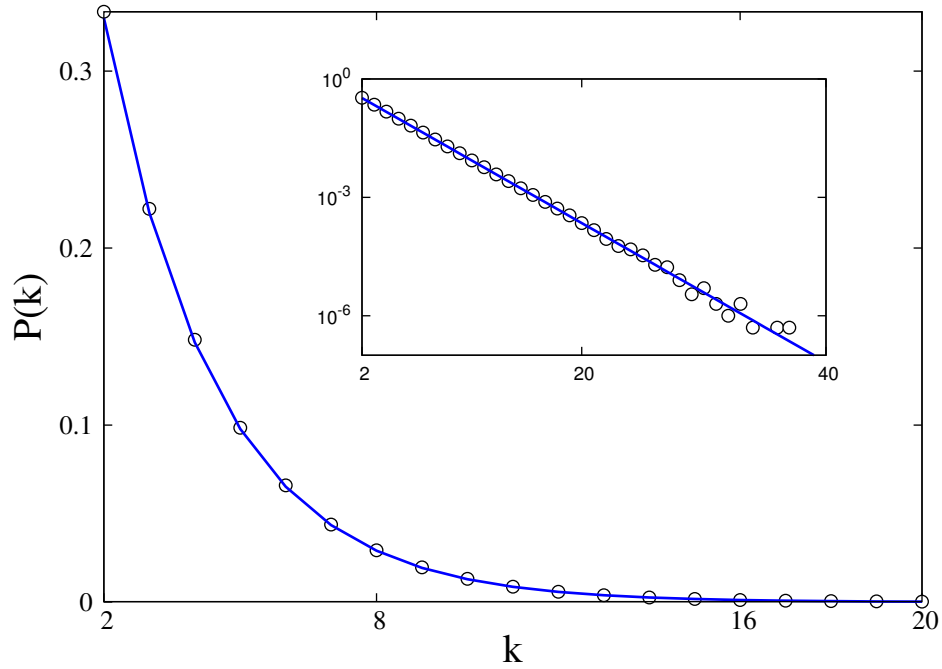


FIGURE 2.3 – les résultats de la simulation (cercles) pour  $n = 2 \cdot 10^6$ ,  $m = 2$ ,  $m_0 = 3$ , et la solution numérique (ligne continue) pour Eq. (2.15). Dans l'encart, nous tracerons les mêmes données dans l'échelle log-linéaire.

### 2.4.3 Comparaison au niveau microscopique avec le modèle de BA

Nous recherchons les différences entre réseaux hétérogènes et homogènes en comparant notre modèle au modèle BA. La distribution des degrés ne suffit pas à elle seule à caractériser les réseaux. Le calcul d'autres quantités microscopiques peut aider à mieux comprendre leur évolution et leur formation. Il s'avère que le réseau sans-échelle a des nœuds avec un degré important (hubs), tandis que le réseau aléatoire n'a pas de structure apparente. L'évaluation du degré moyen instantané du nœud cible  $\langle k_s(t) \rangle$  et de ses fluctuations peut fournir des informations quantitatives sur les nœuds du réseau. En fait,  $\langle k_s(t) \rangle$  est en quelque sorte lié au degré moyen instantané des hubs, car lors du choix aléatoire des nœuds, les hubs ont plus de chance d'être sélectionnés.

Dans un premier temps, nous analysons  $\langle k_s(t) \rangle$  et  $\langle k_s^2(t) \rangle$  dans le réseau BA

$$\langle k_s(t) \rangle = \sum_{t_i=1}^t \Pi(k_i) k_i(t) + m_0 \Pi(k_0) k_0(t), \quad (2.17)$$

où  $\Pi(k_i) = \frac{k_i(t)}{2mt+mm_0}$ ,  $t_i$  est l'instant auquel le nœud  $i$  a été créé et  $k_0(t)$  est le degré des nœuds initiaux à l'instant  $t$ .

$k_i(t)$  est facilement calculé à partir de l'équation d'évolution du degré de champ moyen :  $\frac{\partial k_i(t)}{\partial t} = m\Pi(k_i)$ , qui donne  $k_i(t) = m\left(\frac{2t+m_0}{2t_i+m_0}\right)^{\frac{1}{2}}$ .

En insérant la dernière expression dans l'Eq. (2.17), on obtient

$$\langle k_s(t) \rangle = m\left(\sum_{t_i=1}^t \frac{1}{2t_i + m_0} + 1\right) \quad (2.18)$$

$$= m\left(\ln(2t + m_0) + \gamma - a + \frac{1}{2(2t + m_0)} + O\left(\frac{1}{t^2}\right)\right), \quad (2.19)$$

où  $\gamma$  est la constante d'Euler, et  $a = \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{1+m_0}$ .

Un bon accord est obtenu comme le montre la figure Fig. 2.4 entre Eq. (2.19) et les résultats de la simulation même pour les premiers instants de l'évolution.  $\langle k_s(t) \rangle$  croît indéfiniment avec le temps et diverge pour un réseau infini ( $t \rightarrow \infty$ ) du fait que, dans les réseaux hétérogènes, les hubs sont plus susceptibles d'être sélectionnés et liés aux nouveaux nœuds.

De l'autre côté, le degré moyen du réseau reste fini [127, 128] puisque la majorité des nœuds ont un faible degré et le poids des hubs est faible.

Le deuxième moment  $\langle k_s^2(t) \rangle$  s'écrit

$$\langle k_s^2(t) \rangle = \sum_{t_i=1}^t \Pi(k_i) k_i^2(t) + m_0 \Pi(k_0) k_0^2(t) \quad (2.20)$$

$$\approx m^2(2t + m_0)^{\frac{1}{2}} \left( \sum_{t_i=1}^t \frac{1}{(2t_i + m_0)^{\frac{3}{2}}} + m_0^{-\frac{1}{2}} \right). \quad (2.21)$$

Pour des grands temps,  $\sum_{t_i=1}^t \left(\frac{1}{t_i}\right)^{\frac{3}{2}} = \zeta\left(\frac{3}{2}\right) \approx 2.612$ , on obtient

$$\langle k_s^2(t) \rangle \approx m^2 \sqrt{2t} (m_0^{-\frac{1}{2}} + 2.612 - b), \quad (2.22)$$

où  $b = 1 + \frac{1}{2^{\frac{3}{2}}} + \frac{1}{3^{\frac{3}{2}}} + \dots + \frac{1}{(1+m_0)^{\frac{3}{2}}}$ .

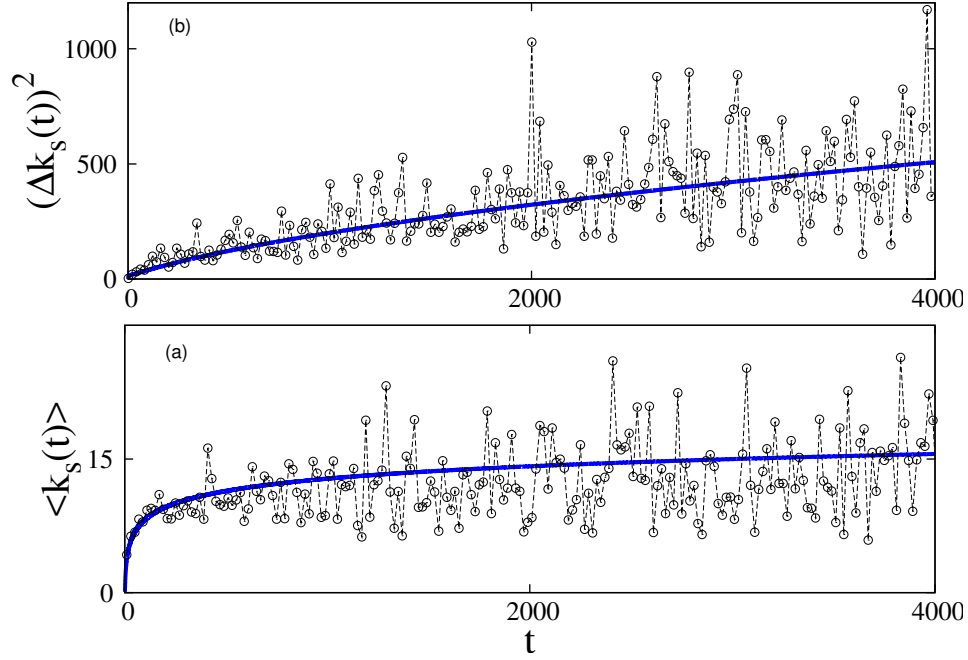


FIGURE 2.4 – (a) Évolution de  $\langle k_s(t) \rangle$  dans le modèle BA, la ligne continue représente Eq. (2.19). (b) Évolution des fluctuations de  $\langle k_s(t) \rangle$ , la ligne continue représente Eq. (2.23). Les cercles joints par des lignes pointillées dans les deux cas sont des données de simulation moyennées sur 20 réalisations pour  $m = 2$ ,  $m_0 = 3$ .

Fluctuations de  $\langle k_s(t) \rangle$  sont donnés par

$$(\Delta k_s(t))^2 \equiv \langle k_s^2(t) \rangle - \langle k_s(t) \rangle^2 \approx m^2 \left[ (m_0^{-\frac{1}{2}} + 2.612 - b) \sqrt{2t} - (\ln(2t))^2 \right], \quad (2.23)$$

qui deviennent arbitrairement large lorsque le temps augmente suffisamment.

Les données de simulation selon Eq. (2.23) ( voir Fig. 2.4(b)) montrent la tendance croissante des fluctuations de  $\langle k_s(t) \rangle$ . Cela peut s'expliquer par le fait que le degré maximum dans le réseau  $k_{max} \sim \sqrt{t}$  [128] augmente plus vite que  $\langle k_s(t) \rangle \sim \ln(t)$  (Eq. (2.19)) et la différence entre les deux quantités devient plus grande avec le temps.

Nous passons maintenant à la même analyse dans notre modèle, L'équation d'évolution du champ moyen pour  $k_i(t)$  est donnée par

$$\frac{\partial k_i(t)}{\partial t} = m \left( 1 - \frac{k_i(t)}{2mt + mm_0} \right) \frac{1}{t + m_0 - 1}, \quad (2.24)$$

La solution a la forme

$$k_i(t) = m \left( \frac{t + m_0 - 1}{2t + m_0} \right)^{\frac{1}{m_0 - 2}} \left[ \left( \frac{t_i + m_0 - 1}{2t_i + m_0} \right)^{-\frac{1}{m_0 - 2}} - A(t_i) + A(t) \right], \quad (2.25)$$

où  $A(t) = \int_1^t \frac{\left( \frac{t' + m_0 - 1}{2t' + m_0} \right)^{-\frac{1}{m_0 - 2}}}{t' + m_0 - 1} dt'.$

On sait que la valeur moyenne du nœud sélectionné  $\langle k_s(t) \rangle$  est sous la forme

$$\langle k_s(t) \rangle = \sum_{t_i=1}^t \Pi(k_i) k_i(t) + m_0 \Pi(k_0) k_0(t), \quad (2.26)$$

en remplaçant Eq. (2.25) dans cette dernière équation on obtient immédiatement pour tout temps  $t$  l'expression de  $\langle k_s(t) \rangle$ , L'équation résultante est résolue numériquement comme le montre Fig. 2.5.

Pour les temps très grands et en prenant  $t \gg m_0$ , nous trouvons  $A(t) \approx 2^{\frac{1}{m_0 - 2}} \ln(t)$ ,  $k_i(t) \approx m \left( 1 + \ln \frac{t}{t_i} \right)$ , d'où

$$\begin{aligned} \langle k_s(t) \rangle &\approx \frac{m}{t} \left( \sum_{t_i=1}^t \ln(t) - \ln(t_i) + 1 \right) \\ &\approx \frac{m}{t} \left( t \left( \ln(t) + 1 \right) - \left( \sum_{t_i=1}^t \ln(t_i) \right) \right) \\ &\approx \frac{m}{t} \left( t \left( \ln(t) + 1 \right) - \ln(t_i!) \right) \\ &\approx 2m. \end{aligned} \quad (2.27)$$

Le second moment est obtenu en substituant les expressions correspondantes de  $\Pi(k_i)$  et  $k_i(t)$  dans Eq. (2.20), nous obtenons pour les temps très grands

$$\langle k_s^2(t) \rangle \approx \frac{m^2}{t} \left( \sum_{t_i=1}^t \left( \ln \left( \frac{t}{t_i} \right) + 1 \right)^2 \right). \quad (2.28)$$

Faire les approximations  $\sum_{t_i=1}^t \ln(t_i) \approx t \ln(t) - t$ , et  $\sum_{t_i=1}^t \ln(t_i)^2 \approx t \ln(t)^2 - 2t \ln(t) + 2t - 2$ ,



on obtient  $\langle k_s^2(t) \rangle \approx 5m^2$ . Fluctuations sont

$$(\Delta k_s(t))^2 \equiv \langle k_s^2(t) \rangle - \langle k_s(t) \rangle^2 \approx m^2. \quad (2.29)$$

Ce résultat, additionné avec  $\langle k_s(t) \rangle \approx 2m$ , montrent que presque tous les nœuds ont le même degré que celui illustré sur Fig. 2.5. L'homogénéité du réseau peut s'expliquer par le fait que l'attachement préférentiel utilisé ici ne permet pas la formation des hubs, puisqu'il ne permet pas aux riches de s'enrichir, ni d'enrichir les pauvres.

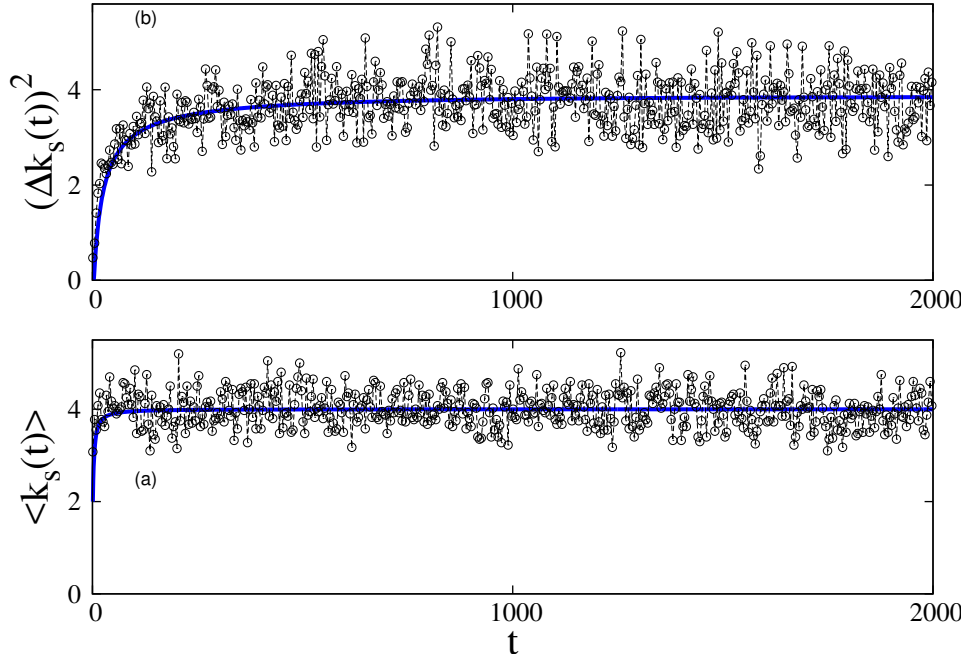


FIGURE 2.5 – (a) Évolution de  $\langle k_s(t) \rangle$  dans notre modèle, la ligne continue représente Eq. (2.26). (b) Évolution des fluctuations de  $\langle k_s(t) \rangle$ , la ligne continue représente la solution numérique de Eq. (2.26) et Eq. (2.28). Les cercles joints par des lignes pointillées dans les deux cas sont des données de simulation moyennées sur 20 réalisations pour  $m = 2$ ,  $m_0 = 3$ .

## 2.5 conclusion

Dans ce chapitre, nous avons commencé par donner quelques idées sur les processus dynamiques, l'attachement préférentiel et le mécanisme "rich get richer" dans les réseaux réels, puis nous avons introduit un simple modèle de réseau complexe avec un critère d'attachement préférentiel et sans effet "rich get richer". Le réseau obtenu est homogène, ce qui démontre le rôle crucial du "rich get richer" dans la topologie du réseau, en outre nous avons conclu que le fait de donner un traitement préférentiel aux nœuds les moins connectés équivaut à utiliser une probabilité d'attachement aléatoire.

Le Calcul du degré moyen instantané d'un nœud sélectionné et ses fluctuations fournissent plus d'informations que le degré moyen habituel du réseau, en particulier nous avons montré comment le degré moyen des hubs et ses fluctuations divergent avec le temps dans le modèle BA, et restent finis dans notre modèle.

# Chapitre 3

## Structure des réseaux sans-échelle non corrélés : couches et plus court chemin

L'objectif principal de ce chapitre vise à prédire quel sera le comportement des systèmes en réseau sur la base des propriétés structurelles mesurées et des règles locales régissant les sommets individuels. Comment, par exemple, la distribution des degrés du réseau affectera-t-elle sa structure et ses propriétés ?

### 3.1 Introduction

De nombreux réseaux du monde réel qui ont été décrits précédemment, tels que le WWW et les réseaux de collaboration, grandissent avec le temps. Par conséquent, il est raisonnable de considérer les graphes de taille croissante et d'étudier la structure de ces réseaux ayant souvent la propriété sans-échelle (voir Section.1.5.2). La communauté scientifique, en s'appuyant sur des idées issues d'une grande variété de disciplines, a fait un excellent départ sur la caractérisation et la modélisation de la structure des réseaux, mais il n'y a pas encore des progressions théoriques cruciaux dans ce domaine [56]. Ici nous allons envisager ces problèmes en donnant quelques contributions à ces études concernant la structure des réseaux sans-échelle, et en trouvant pour la première fois les expressions explicites des couches, ainsi que, des formules plus précises du plus court chemin par rapport aux anciens résultats existant dans la littérature.

## 3.2 Mesurer les lois de puissance

Identifier le comportement de la loi de puissance dans les systèmes naturels ou artificiels peut être difficile. La stratégie standard utilise un résultat qu'un courbure d'une quantité avec une distribution de loi de puissance apparaît comme une ligne droite lorsqu'elle est tracée sur des échelles logarithmiques. Il suffit toutefois de faire une simple courbure et de la tracer sur des échelles logarithmiques pour voir si elle semble droite ou non (voir Fig. 3.1).

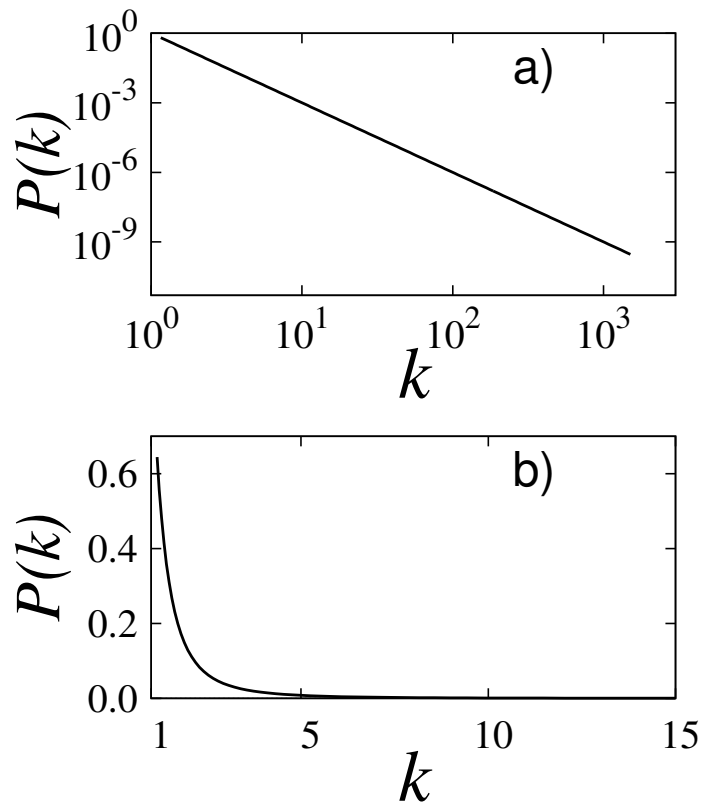


FIGURE 3.1 – Illustration de la loi de puissance pour une fonction,  $P(k) = k^{-\gamma}$ , avec l'exposant  $\gamma = 3$ . Dans a) l'échelle est logarithmique et dans b) l'échelle est linéaire.

Pour révéler une loi de puissance de la distribution, il vaut mieux, comme nous avons vu, tracer la courbure sur des échelles logarithmiques, et quand nous faisons cela pour nos données, nous voyons la forme linéaire caractéristique de la loi de puissance. Cependant, la méthode n'est pas très bonne à certains égards. En particulier, l'extrémité de la droite de la distribution est bruyante en raison d'erreurs d'échantillonnage. Le fait que la distribution diminue dans l'extrémité, cela signifie que chaque valeur mesurée ne contient que quelques

échantillons. Ainsi, les fluctuations fractionnaires dans les comptes des valeurs mesurées sont grandes et cela apparaît comme une courbe bruyante sur la figure. Une façon de traiter cela serait simplement de jeter les données dans la queue de la courbe. Mais il y a souvent des informations utiles dans ces données, car de nombreuses distributions suivent une loi de puissance seulement dans la queue.

Une solution alternative consiste à faire varier la largeur des cases dans notre courbe. Si nous allons faire cela, nous devons aussi normaliser les comptes d'échantillons par la largeur des valeurs mesurées. Autrement dit, le nombre d'échantillons dans un largeur  $\Delta x$  devrait être divisé par  $\Delta x$  pour obtenir un compte par intervalle unitaire de  $x$ . Ensuite, le nombre d'échantillons normalisés devient indépendant de la largeur de la valeur mesuré en moyenne et nous sommes libres de varier les largeurs  $\Delta x$  comme nous le souhaitons. Le choix le plus commun est de créer des largeurs tels que chacun est un multiple fixe plus large que celui qui le précède. Ceci est connu sous le nom "logarithmic binning". Cela signifie que les valeurs mesurées dans la queue de la distribution obtiennent plus d'échantillons que si les tailles des valeurs mesurées étaient fixes, ce qui réduit les erreurs statistiques dans la queue. Il a également l'effet secondaire agréable que les valeurs mesurées semblent être de largeur constante quand nous traçons la courbe sur des échelles logarithmiques.

### 3.3 Quelques notions mathématiques pour la loi de puissance

Une variable réelle continue avec une distribution de loi de puissance a une probabilité  $P(x)$  de prendre une valeur  $x$ , où

$$P(x) = Cx^{-\gamma}. \quad (3.1)$$

Comme nous avons vu dans la Section. 1.5.2, l'exposant  $\gamma > 0$ . Ainsi que la variable  $x$  est souvent une grandeur supérieure ou égale à 1.

Dans les réseaux réels, les distributions des degrés ne suivent généralement pas Eq. (3.1) sur toute leur gamme. En regardant Fig. 1.12, par exemple, nous pouvons voir que la distribution des degrés n'est pas toujours monotone pour petit  $k$ , même en tenant compte des fluctuations

statistiques dans l'histogramme. Une situation courante est que la loi de puissance est obéie dans la queue de la distribution, pour de grandes valeurs de  $k$ , mais pas dans le régime de petit  $k$ . Quand on dit qu'un réseau particulier a une distribution de degré de loi de puissance, on entend normalement seulement que la queue de la distribution a cette forme. Dans certains cas, la distribution peut également s'écarter du forme de loi de puissance pour  $k$  élevé. Par exemple, il y a souvent une coupure, de quelque type, qui limite le degré maximal de nœuds dans la queue.

### 3.3.1 Normalisation

La constante  $C$  dans Eq. (3.1) est la constante de normalisation qui obtenir par la condition que la somme des probabilités est doit être toujours égale à 1

$$P(x) = \int_{x_{min}}^{\infty} Cx^{-\gamma}dx = 1, \quad (3.2)$$

de cette formule au-dessus on trouve que l'expression de la constante  $C$  pour  $\gamma > 1$  est

$$C = (\gamma - 1)x_{min}^{\gamma-1}. \quad (3.3)$$

Alors Eq. (3.1) sera exprimée sous la forme

$$P(x) = (\gamma - 1)x_{min}^{\gamma-1}x^{-\gamma}. \quad (3.4)$$

### 3.3.2 Moments

Le  $m^{\text{ème}}$  moment de la distribution  $P(x)$  est défini comme :

$$\begin{aligned} \langle x^m \rangle &= \int_{x_{min}}^{\infty} x^m P(x) dx \\ &= (\gamma - 1)x_{min}^{\gamma-1} \int_{x_{min}}^{\infty} x^{m-\gamma} dx. \end{aligned} \quad (3.5)$$

Le premier moment  $\langle x \rangle$  est le degré moyen de la grandeur  $x$ , notez que sa valeur devient infinie si  $\gamma \leq 2$ . Le second moment  $\langle x^2 \rangle$  mesure les fluctuations de la variable mesurée,

ceci diverge si  $\gamma \leq 3$ . Sachant que dans les réseaux réels l'exposant  $\gamma$  est souvent dans l'intervalle  $2 < \gamma < 3$ , on déduit alors que les fluctuations est souvent diverge dans le réseaux réels.

### 3.3.3 Lois de puissance pour les variables discrètes

Jusqu'à présent, nous nous sommes concentrés sur les distributions de loi de puissance pour les variables réelles continues, mais beaucoup des quantités que nous traitons dans des situations pratiques sont des entiers discrets, généralement des nombres entiers positifs. Par exemple, les populations de villes ou le nombre de citations à des articles sont toutes des quantités entières. Dans la plupart des cas, la distinction n'est pas très importante. La loi de puissance est obéie seulement dans la queue de la distribution où les valeurs mesurées sont si grandes que, à toutes fins pratiques, elles peuvent être considérées comme continues. Techniquement, cependant, les lois de loi de puissance devraient être définies légèrement différemment pour les quantités entières.

Si  $k$  est une variable entière, alors la loi de puissance dans cette cas sera

$$P(k) = Ck^{-\gamma}, \quad (3.6)$$

avec  $C$  est la constante de normalisation qui vérifie la condition  $\sum_1^\infty Ck^{-\gamma} = 1$ , alors  $C = \frac{1}{\zeta(\gamma)}$ , où  $\zeta(\gamma)$  est la fonction zêta de Riemann. Alors la loi de puissance de le cas de variable discrète est

$$P(k) = \frac{1}{\zeta(\gamma)} k^{-\gamma}. \quad (3.7)$$

### 3.3.4 Degré maximale

Supposons que nous tirons  $n$  mesures d'une distribution de loi de puissance. Quelle valeur est la plus grande de ces mesures susceptibles de prendre ? La propriété définitive de la plus grande valeur dans un échantillon est qu'il n'y en a pas d'autres plus grandes. La probabilité qu'un échantillon particulier soit plus grand que  $x$  est donnée par l'équation :

$$p(x) = \int_x^\infty P(x')dx' = \frac{C}{\gamma-1}x^{1-\gamma}, \quad (3.8)$$

sachant que pour  $\gamma > 1$  on a déjà montré que  $C = (\gamma-1)x_{min}^{\gamma-1}$ , d'où

$$p(x) = \left(\frac{x}{x_{min}}\right)^{1-\gamma}, \quad (3.9)$$

La probabilité qu'un échantillon ne soit pas supérieur à  $x$  est  $1 - p(x)$ , Ainsi la probabilité qu'un échantillon particulier, échantillon  $i$ , se situera entre  $x$  et  $x + dx$  et que tous les autres ne seront pas plus grands que lui, est  $P(x)dx[1 - p(x)]^{n-1}$ . Ensuite, il y a  $n$  façons de choisir  $i$ , donnant une probabilité totale

$$\pi(x) = nP(x)[1 - p(x)]^{n-1} \quad (3.10)$$

Maintenant, nous pouvons calculer la valeur moyenne,  $x_{max}$ , du plus grand échantillon comme :

$$\begin{aligned} \langle x_{max} \rangle &= \int_{x_{min}}^\infty x\pi(x)dx \\ &= n \int_{x_{min}}^\infty xP(x)[1 - p(x)]^{n-1} \end{aligned} \quad (3.11)$$

En utilisant Eq. (3.4) et Eq. (3.4) on obtient

$$\langle x_{max} \rangle = n(\gamma-1) \int_{x_{min}}^\infty \left(\frac{x}{x_{min}}\right)^{1-\gamma} \left[1 - \left(\frac{x}{x_{min}}\right)^{1-\gamma}\right]^{n-1} dx, \quad (3.12)$$

en utilisant le changement de variable suivante,  $y = 1 - \left(\frac{x}{x_{min}}\right)^{1-\gamma}$ , on obtient que

$$\begin{aligned} \langle x_{max} \rangle &= nx_{min} \int_0^1 \frac{y^{n-1}}{(1-y)^{\frac{1}{\gamma-1}}} dy \\ &= nx_{min} B\left(n, \frac{\gamma-2}{\gamma-1}\right), \end{aligned} \quad (3.13)$$



avec  $B(a, b)$  est fonction bêta qui définie pour tous nombres complexes  $a$  et  $b$  de parties réelles strictement positives par :

$$\begin{aligned} B(a, b) &= \int_0^1 t^{a-1} (1-t)^{b-1} dt \\ &= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \end{aligned} \quad (3.14)$$

avec  $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$  est la fonction gamma.

La fonction bêta  $B(a, b)$  a une propriété intéressante que si une valeur parmi les deux est très grand par rapport à l'autre, elle suit elle-même une loi de puissance. Par exemple, pour  $a$  très grand et  $b$  fixe, on peut écrire  $B(a, b) \sim a^{-b}$ . Sachant que le nombre d'échantillons  $n$  de notre distribution de loi de puissance est toujours très grand, on déduit que

$$B(a, b) \sim n^{-\frac{\gamma-2}{\gamma-1}}, \quad (3.15)$$

d'où

$$\langle x_{max} \rangle \sim n^{\frac{1}{\gamma-1}}. \quad (3.16)$$

### 3.3.5 Fonction génératrice

La fonction génératrice est une méthode générale pour étudier la taille de la composante infinie et le réseau résiduel (les composantes finies) pour un graphe avec une distribution des degrés arbitraire, elle a été développée par Molloy et Reed [129]. Ils ont suggéré d'explorer la composante infinie comme étant explorée et utilisé des équations différentielles pour le nombre de liens non exposés et de sites non visités pour trouver la taille de la composante infinie et la distribution des degrés du graphe résiduel.

Une dérivation alternative et très puissante a été donnée par Newman, Strogatz et Watts [130]. Ils ont utilisé la méthode des fonctions génératrices pour étudier la taille de la grappe infinie ainsi que d'autres quantités (telles que le diamètre et la distribution de la taille des grappes). Ils ont également appliqué cette méthode à d'autres types de graphiques (dirigés et bipartites).

Supposons que nous avons une distribution de probabilités pour une variable entière positive, de sorte que des instances, des événements ou des tirages distincts de cette variable soient indépendants et possède la valeur  $k$  avec la probabilité  $p_k$ . Un bon exemple d'une telle distribution est la distribution des degrés de nœuds choisis au hasard dans un réseau. Si la fraction des nœuds dans un réseau de degré  $k$  est  $p_k$  alors  $p_k$  est aussi la probabilité qu'un nœud choisi au hasard du réseau a un degré  $k$ . La fonction génératrice pour la distribution de probabilité  $p_k$  est le polynôme :

$$g(z) = p_0 + p_1 z + p_2 z^2 + \dots = \sum_{k=0}^{\infty} p_k z^k \quad (3.17)$$

Si l'on connaît la fonction génératrice pour une distribution de probabilité  $p_k$  on peut alors récupérer les valeurs de  $p_k$  en différenciant :

$$p_k = \frac{1}{k!} \left. \frac{d^k g}{dz^k} \right|_0 \quad (3.18)$$

Ainsi, la fonction génératrice nous donne des informations complètes sur la distribution de probabilité et vice versa. La distribution et la fonction génératrice ne sont en réalité que deux représentations différentes de la même chose. Comme nous le verrons, dans de nombreux cas, il est plus facile de travailler avec la fonction génératrice qu'avec la distribution de probabilité, ce qui conduit à de nombreux nouveaux résultats utiles sur les réseaux.

Comme nous avons vu dans la Section. 1.5.2, un cas particulier d'intérêt particulier dans l'étude des réseaux est la distribution de la loi de puissance un certain nombre de réseaux, y compris le WWW, Internet, etc, ont des distributions des degrés suivent les lois de puissance et qui ont des conséquences intéressantes. Pour créer et résoudre des modèles de ces réseaux, il sera important pour nous d'être en mesure d'écrire des fonctions génératrices pour les distributions de loi de puissance. Selon Eq. (3.7) pour cette distribution de probabilité, la fonction génératrice est

$$g(z) = \frac{1}{\zeta(\gamma)} \sum_{k=1}^{\infty} k^{-\gamma} z^k, \quad (3.19)$$

la somme ne peut pas être exprimée sous forme fermée, mais sous forme d'une fonction spéciale

s'appelle le polylogarithme est notée  $Li_\gamma(z) = \sum_{k=1}^{\infty} k^{-\gamma} z^k$ .

Ainsi nous obtenons

$$g(z) = \frac{Li_\gamma(z)}{\zeta(\gamma)}. \quad (3.20)$$

### 3.4 Les réseaux sans-échelle non corrélés

Dans les modèles aléatoires sans-échelle, on suppose généralement qu'il n'y a pas de corrélation entre les degrés des nœuds voisins. C'est-à-dire que la probabilité d'atteindre un nœud en suivant un lien est indépendante du nœud d'où provient le lien. Cependant, dans de nombreux réseaux du monde réel, ce n'est pas le cas. Plusieurs types de corrélation existent, en fonction des propriétés internes des nœuds, les principaux types de corrélations étudiés sont les corrélations degré-degré (voir Section.1.4.4). Même si la construction de ces réseaux sans-échelle aléatoires se fait au début sans aucune corrélation, cela ne signifie pas que le réseau n'affichera pas des corrélations de degré, c'est-à-dire l'absence des corrélation lors de création des réseaux n'est pas une condition cruciale de l'absence de corrélation dans le réseau final, par exemple la Fig. 3.2 indique que les réseaux aléatoires sans-échelle génèrent des corrélations de degré, allant de l'assortativité à la disassortativité selon la valeur de l'exposant  $\gamma$ <sup>1</sup>, nous observons trois régimes d'échelle distincts :

- i) Régime assortatif :  $\gamma > 3$
- ii) Régime neutre :  $\gamma = 3$
- iii) Régime disassortatif :  $\gamma < 3$

Produire des corrélations en utilisant un modèle complètement statique est difficile car non seulement le degré d'un nœud doit être pris en compte, mais aussi sa probabilité de se connecter à chaque voisin. La méthode habituellement utilisée pour générer de tels réseaux consiste à mélanger les liens en utilisant une sorte d'algorithme de type Metropolis<sup>2</sup> [131].

Cependant, la négligence de ces corrélations n'empêche jamais de trouver des résultats

---

1.  $\gamma$  représente l'exposant de la distribution des degrés,  $P(k) \propto k^{-\gamma}$ , dans les réseaux sans-échelle

2. Inventé en 1953 par Nicholas Metropolis et ses collaborateurs du laboratoire de Los Alamos, l'algorithme de Metropolis était d'abord destiné à faire calculer par des ordinateurs les équations d'états de mélanges de molécules en interactions. Depuis lors il s'est révélé bien adapté pour résoudre de nombreux problèmes de mécanique statistique et de chimie. L'outil principal de l'algorithme est une chaîne de Markov : on tire au hasard une boule et on déplace son centre d'une distance  $d$ , le mouvement est accepté si la nouvelle configuration des boules reste sans recouvrement.

importants qui nous aideront à bien modéliser ces réseaux réels et à mieux comprendre leurs structures.

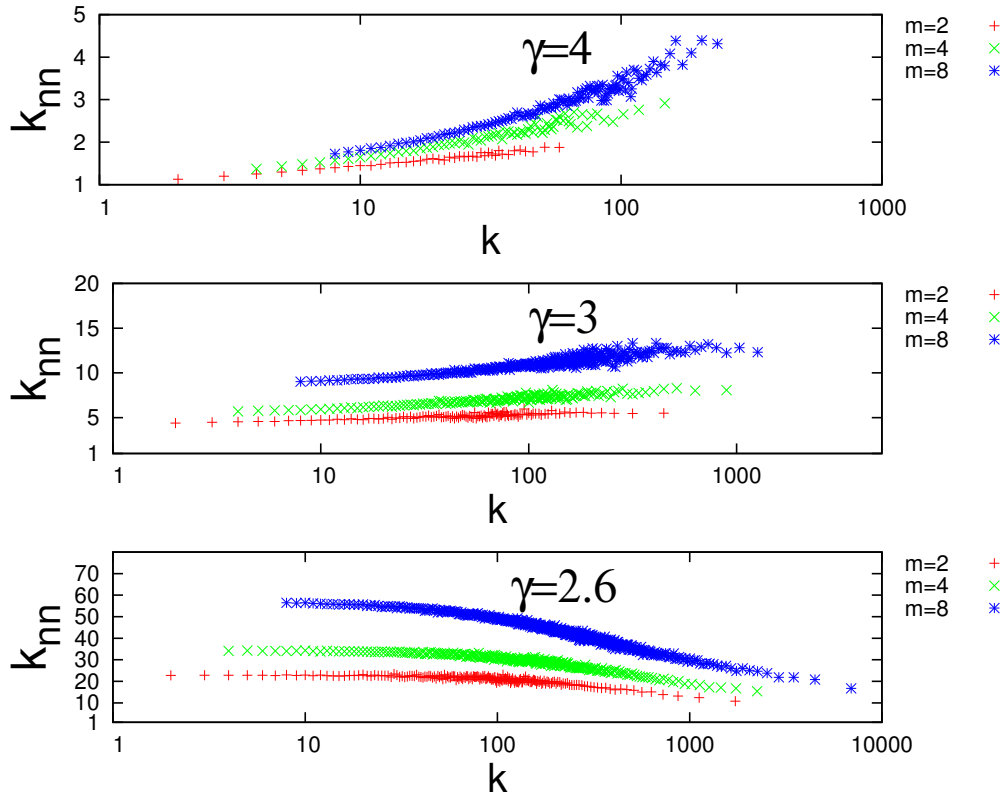


FIGURE 3.2 – Les corrélations des degrés dans le réseau aléatoire sans-échelle pour différentes valeurs de  $m$  et de  $\gamma$ . Le nombre de nœuds est  $n = 10^4$  et le nombre de réalisations pour chaque simulations est 50.

## 3.5 Les anciennes études sur les couches et le plus court chemin

### 3.5.1 Les couches

#### Contribution de Newman

Newman [132] a calculé l'expression des couches, c'est-à-dire le nombre moyen de nœuds,  $n_\ell$ , à distance  $\ell$  depuis un nœud arbitraire pour un réseau aléatoire, son calcul est long et un peu compliqué. On va essayer de résumer son calcul, soit  $p_k^{(2)}$  est la probabilité qu'un nœud a exactement  $k$  seconds voisins dans le réseau

$$p_k^{(2)} = \sum_{m=2}^{\infty} p_m P^{(2)}(k/m), \quad (3.21)$$

avec  $P^{(2)}(k/m)$  est la probabilité d'avoir  $k$  seconds voisins étant donné que nous avons  $m$  premiers voisins et  $p_m$  est la distribution des degrés ordinaire.

La fonction génératrice de cette probabilité est

$$\begin{aligned} g^{(2)}(z) &= \sum_{k=0}^{\infty} p_k^{(2)} z^k \\ &= \sum_{m=0}^{\infty} p_m \sum_{k=0}^{\infty} P^{(2)}(k/m) z^k \\ &= \sum_{m=0}^{\infty} p_m [g_1(z)]^m \\ &= g_0(g_1(z)). \end{aligned} \quad (3.22)$$

Avec  $g_0(z) = \sum_{k=0}^{\infty} p_k z^k$  et  $g_1(z) = \frac{1}{g_0'(1)} \frac{dg_0}{dz}$ .

Nous utilisons le même calcul pour calculer la distribution de probabilité du nombre de tiers voisins. Le nombre de tiers voisins est la somme des excès de chacun des seconds voisins. Ainsi, s'il y a  $m$  seconds voisins, alors la distribution de probabilité  $P^{(3)}(k/m)$  du nombre de

tiers voisins a une fonction génératrice  $[g_1(z)]^m$  et la probabilité globale d'avoir  $k$  troisième voisins est exactement analogue à l'équation

$$g^{(3)}(z) = \sum_{m=0}^{\infty} p_m^{(2)} \sum_{k=0}^{\infty} P^{(3)}(k/m) z^k \quad (3.23)$$

$$\begin{aligned} &= \sum_{m=0}^{\infty} p_m^{(2)} [g_1(z)]^m \\ &= g^{(2)}(g_1(z)) \\ &= g_0(g_1(g_1(0))). \end{aligned} \quad (3.24)$$

Par conséquence, la fonction génératrice du nombre de voisins à n'importe quelle distance  $\ell$  peut être exprimée de cette façon

$$g^{(\ell)}(z) = \sum_{m=0}^{\infty} p_m^{(\ell-1)} \sum_{k=0}^{\infty} P^{(\ell)}(k/m) z^k \quad (3.25)$$

$$\begin{aligned} &= \sum_{m=0}^{\infty} p_m^{(\ell-1)} [g_1(z)]^m \\ &= g^{(\ell-1)}(g_1(z)), \end{aligned} \quad (3.26)$$

avec  $g^{(\ell)}(z) = g_0(\dots g_1(z) \dots)$ , où dans  $g_0$  il y a  $\ell - 1$  copie de  $g_1$ .

Il est généralement assez difficile d'extraire les expressions explicites des probabilités pour les nombres de seconds voisins dans le réseau.

Si nous supposons que notre distribution de degrés est une distribution de Poisson avec une moyenne  $\langle k \rangle$ ,  $p_k = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}$ .

La moyenne d'une distribution est donnée par la dérivée première de sa fonction génératrice évaluée en  $z = 1$  et la dérivée de Eq. (3.22) est  $\frac{dg^{(2)}}{dz} = g'_0(g_1(z))g'_1(z)$ . En fixant  $z = 1$  et en rappelant que  $g_1(1) = 1$ , on trouve que le nombre moyen  $n_2$  des seconds voisins est

$$n_2 = g'_0(1)g'_1(1), \quad (3.27)$$

ainsi que  $g'_0(1) = \langle k \rangle$  et  $g'_1(1) = \frac{1}{\langle k \rangle}(\langle k^2 \rangle - \langle k \rangle)$ , d'où on obtient

$$n_2 = \langle k^2 \rangle - \langle k \rangle. \quad (3.28)$$

Nous pouvons par utiliser le même principe de trouver le nombre moyen de voisins  $n_\ell$  à n'importe quelle distance  $\ell$ . De Eq. (3.26) nous obtenons

$$\frac{dg^{(d)}}{dz} = g^{(d-1)'}(g_1(z))g'_1(z), \quad (3.29)$$

sachant que nous avons  $z = 1$  nous pouvons écrire

$$\begin{aligned} n_\ell &= g^{(d-1)'}(1)g'_1(z) \\ &= n_{\ell-1}g'_1(1) \end{aligned} \quad (3.30)$$

de Eq. (3.27) on a  $g'_1(1) = \frac{n_2}{n_1}$ , alors on obtient

$$n_{\ell-1} = n_{\ell-1} \left( \frac{n_2}{n_1} \right), \quad (3.31)$$

Nous pouvons alors écrire que la distribution des nœuds est sous la forme

$$n_\ell = \left( \frac{n_2}{n_1} \right)^{\ell-1} n_1. \quad (3.32)$$

Cela signifie que  $n_\ell$  augmente ou diminue exponentiellement avec  $\ell$  selon que  $n_2$  est supérieur ou inférieur à  $n_1$ .

### Contribution de Cohen et al

La construction de Cohen et al. [133, 134] a basé sur le modèle de Bollobás [135]. Le processus de construction tente d'exposer le réseau progressivement, en suivant une certaine méthode, ce qui permet de définir des couches dans le graphe proposé. Nous définissons le nombre de nœuds dans le réseau,  $n$ , et nous associons les degrés avec les nœuds selon la loi de puissance  $P(k) = Ck^{-\gamma}$ , comme toujours  $C = (\gamma - 1)m^{\gamma-1}$  est la constante de normalisation

avec  $m$  est le degré minimal et  $K = n^{\frac{1}{\gamma-1}}$  est le degré maximale.

A partir de ce que précède, chaque nœud du réseau dispose d'un nombre donné de liens sortants, nous appelons les liens ouverts, selon son degré choisi. Définissons  $V$  comme l'ensemble des  $n$  nœuds choisis,  $c$  comme l'ensemble des liens sortant non connectés des nœuds de  $V$ , et  $E$  comme l'ensemble des arêtes du graphe. En utilisant ces définitions, l'ensemble des liens dans  $E$  est vide à ce stade, alors que l'ensemble des liens ouverts sortants dans  $c$  contient tous les liens sortants non connectés dans le graphe. Dans la construction de Bollobás [136], les liens dans  $c$  sont appariés de manière aléatoire, de sorte qu'à la fin du processus,  $c$  est vide, et  $E$  contient tous les liens correspondants.

Au lieu de cela, nous partons du nœud de degré maximal, qui a le degré  $K$ , et le relions aléatoirement à  $K$  liaisons ouvertes disponibles, supprimant ainsi ces liens ouverts de  $c$ . Nous avons maintenant exposé la première couche de nœuds, indexée comme  $\ell = 1$ . Nous continuons maintenant à remplir la deuxième couche  $\ell = 2$  de manière similaire. Nous connectons tous les liens ouverts émergeant des nœuds de la couche 1 vers des liens ouverts choisis au hasard. Ces liens ouverts peuvent être choisis parmi les nœuds de la même couche 1 (créant ainsi une cycle) ou d'autres liens en  $c$ . Nous continuons jusqu'à ce que tous les liens ouverts émergeant de la couche 1 soient connectés, remplissant ainsi la couche  $\ell = 2$ . Généralement, pour former la couche  $\ell + 1$  à partir d'une couche arbitraire  $\ell$ , nous relions aléatoirement toutes les liaisons ouvertes émergeant de  $\ell$  à d'autres liaisons ouvertes émergeant de  $\ell$  ou choisies parmi les autres liaisons de  $c$ . Notez que lorsque nous avons formé la couche  $\ell + 1$ , la couche  $\ell$  n'a plus de liens ouverts. Le processus se poursuit jusqu'à ce que l'ensemble des liens ouverts,  $c$ , soit vide.

Le détail de ce calcul est encore plus longue mais relativement aisé, En bref Cohen et al ont considéré un réseau sans échelle spécifique et ont étudié les couches entourant le nœud le plus connecté, ils ont obtenu une relation de récurrence pour  $n_\ell$ . Ces calculs semblent bien cadrer avec les données Internet réelles dans le cas particulier de  $m = 1$ , où  $m$  est le degré minimum dans le réseau. Deux régimes sont observés pour  $n_\ell$  : le premier est caractérisé par une croissance rapide, et le second descend de façon exponentielle.



Ces deux études ne sont pas satisfaisantes, car la première expression de Newman est une fonction monotone, soit croissante ou décroissante, ce qui est conceptuellement faux [128], et dans le deuxième résultat, Cohen et al. n'ont pas réussi à trouver une expression explicite, mais plutôt une suite de récurrences sans solution. En plus leurs expressions correspondent plus aux données d'Internet et ne représentent aucune généralité.

### 3.5.2 Plus court chemin

Le plus court chemin (PCC) peut être le concept le plus intrigant dans les réseaux complexes et la théorie des graphes, principalement après la célèbre expérience de Milgram [40]. Dans cette expérience, Milgram a clairement démontré le phénomène du petit monde dans les réseaux sociaux, ce qui signifie que deux personnes dans le monde sont en moyen séparées par des petites connexions intermédiaires.

Nous pouvons citer les anciennes formules du PCC en relation avec la valeur de l'exposant  $\gamma$  comme ceci : Pour  $2 < \gamma < 3$  on dit que le réseau est ultra-petit  $\langle \ell \rangle \sim \ln \ln(n)$  [14, 15, 137, 138, 139, 140], pour  $\gamma = 3$  on dit que le réseau est petit-monde  $\langle \ell \rangle \sim \frac{\ln(n)}{\ln \ln(n)}$  [135, 138, 141, 142, 143] et pour  $\gamma > 3$  le réseau est aussi petit-monde  $\langle \ell \rangle \sim \ln(n)$  [135, 138, 141, 142, 143].

## 3.6 Structure des couches dans les réseaux sans-échelle non corrélés

### 3.6.1 L'étude théorique

La couche dans un réseau complexe est définie comme l'ensemble des nœuds à la même distance d'un nœud arbitraire choisi. Dans un réseau sans-échelle, chaque nœud est lié à d'autres nœuds  $k$  avec la probabilité  $P(k) = Ck^{-\gamma}$ ,  $k = m, m+1, \dots, K$ . Où  $C = (\gamma - 1)m^{\gamma-1}$  est la constante de normalisation,  $m$  et  $K$  sont les seuils inférieur et supérieur de la distribution.  $K = mn^{\frac{1}{\gamma-1}}$ , avec  $n$  est le nombre total de nœuds.

Nous construisons un réseau sans-échelle en choisissant aléatoirement un nœud de degré moyen  $\langle k \rangle$ , et dans chaque couche suivante, nous mettons le degré suivant le plus élevé jusqu'à ce que la couche soit pleine (Fig. 3.3).

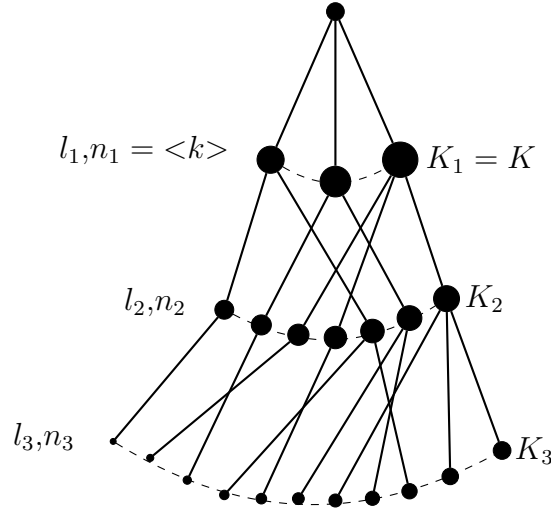


FIGURE 3.3 – Illustration du réseau construit. La taille des cercles pleins (nœuds) est le degré dépendant. Le degré maximum de couche  $\ell$  est  $K_\ell$ .

Évidemment, la première couche contiendra les premiers voisins  $\langle k \rangle$  du nœud de départ, alors  $n_1 = \langle k \rangle$ . Cela ne représente pas un réseau particulier, mais c'est juste une description idéalisée de tout réseau sans échelle.

Pour les grands réseaux non corrélés, la structure arborescente peut être supposée et les cycles dans la même couche sont négligés [14, 143]. Par la suite, la probabilité  $p_\ell$  qu'un nœud de la couche  $n_\ell$  soit lié à un autre nœud n'appartenant pas aux premiers  $\ell$  couches est  $p_\ell = \frac{\kappa_\ell - 1}{n}$ , où  $\kappa_\ell$  est le degré moyen des nœuds appartenant à la couche  $n_\ell$  et le  $-1$  est dû au lien de la couche précédente.

La probabilité que des nœuds parmi  $n - n_1 - 1$  ne soient liés à aucun nœud dans  $\ell = 1$  est  $(1 - p_1)^{n_1}$ , alors la probabilité que ces nœuds soient liés aux nœuds dans  $\ell = 1$  est  $1 - (1 - p_1)^{n_1}$ . Le nombre de nœuds dans  $\ell = 2$  est alors donné par  $n_2 = (1 - (1 - p_1)^{n_1})(n - n_1 - 1)$ . La généralisation pour  $n_\ell$  est simple, on obtient :

$$n_\ell = (1 - (1 - p_{\ell-1})^{n_{\ell-1}})(n - \sum_{j=1}^{\ell-1} n_j - 1), \quad (3.33)$$

lorsque  $n$  est grand  $p_\ell \ll 1 \implies (1 - p_{\ell-1})^{n_{\ell-1}} \simeq e^{-p_{\ell-1}n_{\ell-1}}$ . Eq. (3.33) peut être facilement

manipulé pour obtenir :

$$n_\ell = \begin{cases} \langle k \rangle & \text{si } \ell = 1 \\ (n - 1 - n_1)(e^{-\sum_{j=1}^{\ell-2} p_j n_j} - e^{-\sum_{j=1}^{\ell-1} p_j n_j}) & \text{si } \ell \geq 2. \end{cases} \quad (3.34)$$

L'équation ci-dessus peut être simplifiée en développant les sommes en exponentielles.  $p_\ell$  est  $\kappa_\ell$ -dépendante, qui dépend à son tour de  $\gamma$  et du degré maximal dans la couche  $\ell$ ,  $K_\ell$ . D'abord on donne l'expression de  $\kappa_\ell$  quand  $K_\ell \gg m$ , et ensuite on développe la somme sur  $n_j$ .

$$\kappa_\ell = \frac{\langle k_\ell^2 \rangle}{\langle k_\ell \rangle} = \begin{cases} \left(\frac{\gamma-2}{\gamma-3}\right)m & \text{si } \gamma > 3 \\ m(\log(K_\ell) - \log(m)) & \text{si } \gamma = 3 \\ \left(\frac{\gamma-2}{3-\gamma}\right)m^{\gamma-2}K_\ell^{3-\gamma} & \text{si } 2 < \gamma < 3 \\ \frac{K_\ell - m}{\log(K_\ell) - \log(m)} & \text{si } \gamma = 2. \end{cases} \quad (3.35)$$

Lorsque  $\gamma = 2$ , le degré maximum  $K_1$  est de l'ordre du nombre total de nœuds, c'est-à-dire que presque tous les nœuds sont connectés au nœud de degré maximal. Le réseau peut contenir un maximum de deux couches, avec  $n_1 = \langle k \rangle$  et  $n_2 = n - \langle k \rangle$ .

Quand  $2 < \gamma < 3$ , le réseau est encore très hétérogène.  $\kappa_\ell$  est couche-dépendent, il dépend du degré maximal  $K_\ell$  de la couche  $\ell$  (Eq. (3.35)). L'étape suivante de nos calculs consiste à relier  $\kappa_\ell$  à  $\kappa_1$ . Le nombre de nœuds entre le premier et le  $(\ell - 1)^{\text{ème}}$  couches est donné par :

$$\begin{aligned} \sum_{j=1}^{\ell-1} n_j &= n \int_{K_\ell}^{K_1} P(k) dk \\ &= nm^{\gamma-1}(K_\ell^{1-\gamma} - K_1^{1-\gamma}) \simeq nm^{\gamma-1}K_\ell^{1-\gamma}, \end{aligned} \quad (3.36)$$

où nous avons utilisé  $K_1 \gg K_\ell$  pour  $n$  assez grand. Dans ce cas où  $2 < \gamma < 3$ ,  $\kappa_\ell \gg 1$ , alors

$$\begin{aligned} \sum_{j=1}^{\ell-1} n_j &= \langle k \rangle + \langle k \rangle(\kappa_1 - 1) + \langle k \rangle(\kappa_1 - 1)(\kappa_2 - 1) + \dots + \langle k \rangle(\kappa_1 - 1)(\kappa_2 - 1) \dots (\kappa_{\ell-1} - 2) \\ &\approx \langle k \rangle \kappa_1 \kappa_2 \dots \kappa_{\ell-2}, \end{aligned} \quad (3.37)$$

on en déduit  $K_\ell = \left( \frac{\langle k \rangle \kappa_1 \kappa_2 \dots \kappa_{\ell-2}}{nm^{\gamma-1}} \right)^{\frac{1}{1-\gamma}}$ .

Le degré moyen des voisins à distance  $\ell$  peut maintenant être exprimé comme ceci :  $\kappa_\ell = \frac{\gamma-2}{3-\gamma} m \left( \frac{n}{\langle k \rangle \kappa_1 \kappa_2 \dots \kappa_{\ell-2}} \right)^{\frac{3-\gamma}{\gamma-1}}$ , ce qui conduit à la relation de récurrence  $\kappa_\ell = \kappa_{\ell-1} \left( \kappa_{\ell-2} \right)^{\frac{\gamma-3}{\gamma-1}}$ . Quand  $\gamma$  est dans  $]2, 3[$  nous avons généralement  $\kappa_{\ell-2} \gg \kappa_{\ell-1}$  mais depuis  $|\frac{\gamma-3}{\gamma-1}| < 1$ , on considère que  $\left( \kappa_{\ell-2} \right)^{\frac{\gamma-3}{\gamma-1}} \approx \left( \kappa_{\ell-1} \right)^{\frac{\gamma-3}{\gamma-1}}$ . Enfin, on obtient  $\kappa_\ell = \kappa_1^{\left( \frac{2\gamma-4}{\gamma-1} \right)^{\ell-1}}$ .

Nous sommes maintenant prêts de calculer la somme dans l'Eq. (3.34) :

$$\begin{aligned} \sum_{j=1}^l p_j n_j &= \sum_{j=1}^l \frac{\kappa_j}{n} \left( \langle k \rangle \kappa_1 \kappa_2 \dots \kappa_{j-1} \right) \\ &\approx \frac{\kappa_\ell}{n} \left( \langle k \rangle \kappa_1 \kappa_2 \dots \kappa_{\ell-1} \right) \\ &\approx \frac{\langle k \rangle}{n} \kappa_1^{1+\beta+\beta^2+\dots+\beta^{\ell-1}} \\ &\approx \frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^\ell}{1-\beta}}, \end{aligned} \quad (3.38)$$

où  $\beta = \frac{2\gamma-4}{\gamma-1}$ . En prenant  $n - n_1 - 1 \approx n$ , Eq. (3.34) peut être écrit pour  $2 < \gamma < 3$  comme :

$$n_\ell = \begin{cases} \langle k \rangle & \text{if } \ell = 1 \\ n \left( e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-2}}{1-\beta}}} - e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-1}}{1-\beta}}} \right) & \text{if } \ell \geq 2. \end{cases} \quad (3.39)$$

Quand  $\gamma > 3$ , les hubs ne sont pas grands, et les propriétés caractéristiques du réseau sont similaires à celles des réseaux aléatoires ER.  $\kappa_\ell$  est couche-indépendant comme indiqué dans Eq. (3.35). Alors  $\sum_{j=1}^{\ell-1} p_j n_j = \frac{\kappa'}{n} \sum_{j=1}^{\ell-1} n_j$ , où  $\kappa' = \kappa - 1$ .

$$\begin{aligned} \sum_{j=1}^{\ell} n_j &= \langle k \rangle + \langle k \rangle (\kappa_1 - 1) + \langle k \rangle (\kappa_1 - 1) (\kappa_2 - 1) + \dots + \langle k \rangle (\kappa_1 - 1) (\kappa_2 - 1) \dots (\kappa_{\ell-1} - 1) \\ &= \langle k \rangle + \langle k \rangle \kappa' + \langle k \rangle \kappa'^2 + \dots + \langle k \rangle \kappa'^{\ell-1} = \langle k \rangle \frac{1 - \kappa'^\ell}{1 - \kappa'}. \end{aligned} \quad (3.40)$$

Eq. (3.34) peut être écrit pour  $\gamma > 3$  comme :

$$n_\ell = \begin{cases} \langle k \rangle & \text{if } \ell = 1 \\ n(e^{-\frac{\langle k \rangle}{n} \kappa' \frac{1-\kappa' \ell-2}{1-\kappa'}} - e^{-\frac{\langle k \rangle}{n} \kappa' \frac{1-\kappa' \ell-1}{1-\kappa'}}) & \text{if } \ell \geq 2. \end{cases} \quad (3.41)$$

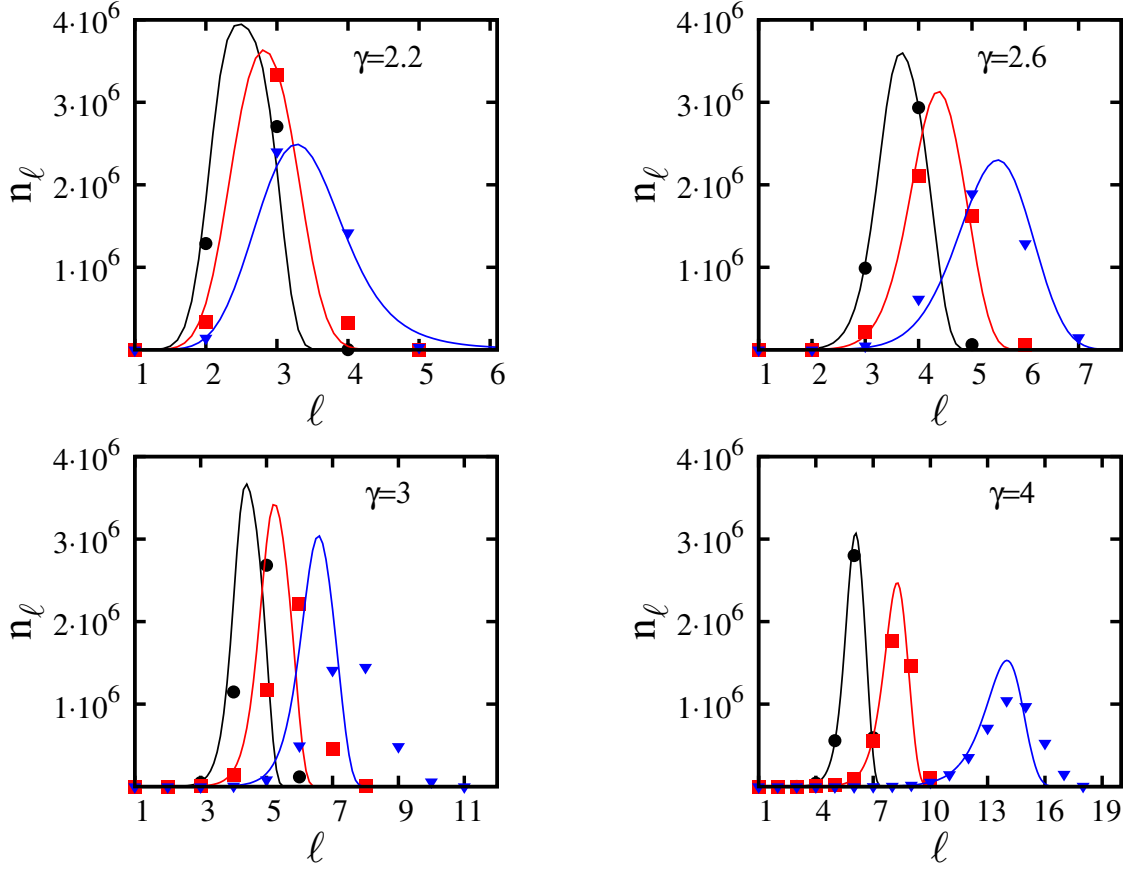


FIGURE 3.4 – Nombre de nœuds dans chaque couche pour différentes valeurs de  $\gamma$ . Les lignes pleines sont Eq. (3.39) et Eq. (3.41). Les symboles sont des simulations d'un réseau de taille  $n = 4 \times 10^6$  et une moyenne de 200 réalisation pour chaque point. Les couleurs noir, rouge et bleu se réfèrent respectivement à  $m = 8, 4$  et  $2$ .

Le cas  $\gamma = 3$  est le plus problématique. C'est le point où la structure du réseau change radicalement, nous constatons la présence de multiples grands hubs et un second moment  $\langle k^2 \rangle$  diverge pour  $\gamma < 3$ , par contre nous remarquons une absence des hubs importants et un  $\langle k^2 \rangle$  fini pour  $\gamma > 3$ . Dans notre approche, le problème se pose lors du calcul de  $\sum_{j=1}^l p_j n_j$ . Néanmoins, à ce point de transition ( $\gamma = 3$ ), certaines propriétés du réseau se comportent presque comme celles correspondant à  $\gamma > 3$ . Principalement, le PCC est mise à l'échelle avec

$n$  comme  $\frac{\log(n)}{\log(\log(n))}$  [144], et  $\langle k^2 \rangle$  reste fini. Ensuite, nous utilisons Eq. (3.41) pour calculer  $n_\ell$  dans ce cas.

Dans la Fig. 3.4, nous représentons  $n_\ell$  en fonction de  $\ell$  pour différentes valeurs de  $\gamma$  et  $m$ . En général, un excellent accord entre la théorie et les simulations est observé. Pour  $\gamma = 3$ , l'accord est moins bon, principalement due au fait que le réseau est encore hétérogène et que les hubs sont toujours présents, alors que nous avons supposé l'homogénéité du réseau pour calculer  $\sum_{j=1}^l p_j n_j$  (Eq. (3.40)). Nous observons peu de différences en  $n_\ell$  entre la théorie et les simulations lorsque  $\gamma = 4$  et  $m = 2$ . Ceci est causé par une augmentation relative de la proportion de cycles dans les mêmes couches. Ainsi, la structure arborescente parfaite, qui est l'hypothèse principale dans le calcul de  $n_\ell$ , n'est plus vraie. La figure Fig. 3.5 montre clairement l'augmentation relative des cycles quand  $m$  est abaissé, et  $\gamma$  augmenté.

Nos résultats analytiques pour  $n_\ell$  sont également comparés aux réseaux du monde réel. Nous observons à partir de la figure Fig. 3.4 que  $n_\ell$  augmente et diminue de différentes manières. Les manipulations de Eq. (3.39) et Eq. (3.41) sont nécessaires pour extraire une information détaillée sur les queues de la distribution des nœuds.

Pour  $n$  grand, Eq. (3.39) pour  $\ell > 1$  peut être approché comme :

$$\begin{aligned} n_\ell &= n \left( e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-2}}{1-\beta}}} - e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-1}}{1-\beta}}} \right) \\ &\approx -n \frac{\partial \left( e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}}} \right)}{\partial \ell} \\ &\approx -\frac{\langle k \rangle \log(\kappa_1) \log(\beta)}{1-\beta} \beta^{\ell-\frac{3}{2}} \kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}} e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}}}, \end{aligned} \quad (3.42)$$

où  $\ell - \frac{3}{2}$  est utilisé à la place de  $\ell - 1$  pour améliorer la différenciation avec la règle du point central. Quand  $\kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}} \ll \frac{n}{\langle k \rangle}$ ,  $e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}}} \approx 1$ , le terme dominant dans l'équation Eq. (3.42) est  $\kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}}$ , celle-ci peut écrire comme

$$\begin{aligned} \kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}} &= e^{\ln(\kappa_1) \frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}} \\ &= e^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta} \ln(\kappa_1)}, \end{aligned} \quad (3.43)$$

il est difficile d'extraire la loi exacte de  $n_\ell$  à partir de l'expression au-dessus, mais de façon générale on peut dire que  $n_\ell$  dans le premier régime,  $\kappa_1^{\frac{1-\beta\ell-\frac{3}{2}}{1-\beta}} \ll \frac{n}{\langle k \rangle}$ , augmente rapidement comme un double exponentiel, et après avoir atteint son maximum à  $\kappa_1^{\frac{1-\beta\ell-\frac{3}{2}}{1-\beta}} = \frac{n}{\langle k \rangle}$ ,  $n_\ell$  diminue comme un triple exponentiel  $e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta\ell-\frac{3}{2}}{1-\beta}}}$  quand  $\kappa_1^{\frac{1-\beta\ell-\frac{3}{2}}{1-\beta}} \gg \frac{n}{\langle k \rangle}$  (deuxième régime). Les résultats obtenus semblent étonnants, mais le fait que dans le cas où  $2 < \gamma < 3$  le réseau est défini comme ultra-petit monde, c'est-à-dire  $\ell \sim \ln \ln(n)$ , explique qu'il est normale que  $n_\ell$  augmente comme un double exponentiel.

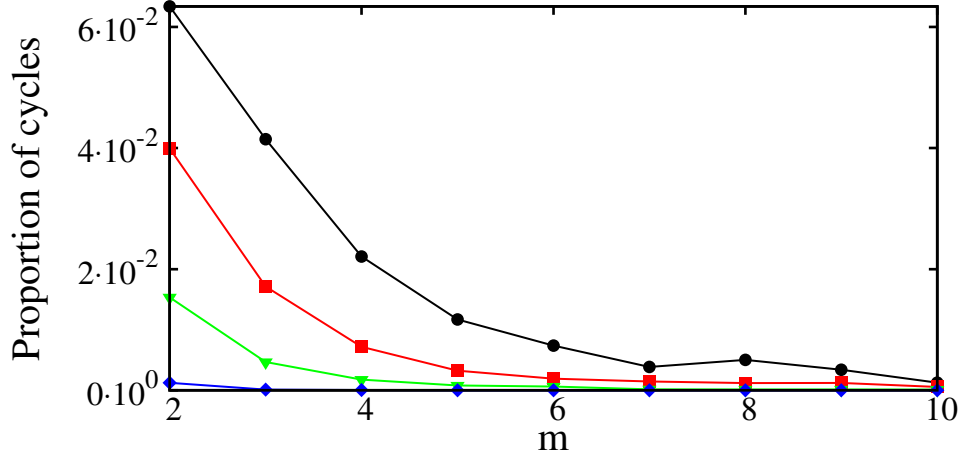


FIGURE 3.5 – Proportion de cycles dans les couches par rapport à  $m$ . De haut en bas,  $\gamma$  est respectivement 4, 3, 2.6 et 2.2. Nombre de nœuds  $n = 4.10^6$ , le nombre de points pour chaque simulation est 200.

De la même manière, Eq. (3.41) peut être écrite comme

$$\begin{aligned}
 n_\ell &= n \left( e^{-\frac{\langle k \rangle}{n} \kappa'^{\frac{1-\kappa'\ell-2}{1-\kappa'}}} - e^{-\frac{\langle k \rangle}{n} \kappa'^{\frac{1-\kappa'\ell-1}{1-\kappa'}}} \right) \\
 &\approx n \left( e^{-\frac{\langle k \rangle}{n} \kappa'^{\ell-2}} - e^{-\frac{\langle k \rangle}{n} \kappa'^{\ell-1}} \right) \\
 &\approx n \frac{\partial e^{-\frac{\langle k \rangle}{n} \kappa'^{\ell-\frac{3}{2}}}}{\partial \ell} \\
 &\approx \langle k \rangle \ln(\kappa') \kappa'^{\ell-\frac{3}{2}} e^{-\frac{\langle k \rangle}{n} \kappa'^{\ell-\frac{3}{2}}}, \tag{3.44}
 \end{aligned}$$

nous déduisons alors qu'on a une croissance de loi exponentiel pour  $\kappa'^{\ell-\frac{3}{2}} \ll \frac{n}{\langle k \rangle}$ , et une décroissance de double exponentielle pour  $\kappa'^{\ell-\frac{3}{2}} \gg \frac{n}{\langle k \rangle}$ , car  $\kappa'^{\ell-\frac{3}{2}} = e^{(\ell-\frac{3}{2})\ln(\kappa')}$ . La même remarque concernant la croissance exponentiel des couches, le fait que dans ce cas où  $\gamma \geq 3$  le réseau est défini comme petit-monde, c'est-à-dire  $\ell \sim \ln(n)$ , cela montre que la croissance exponentiel est normale. En outre, la décroissance par un triple exponentiel pour  $2 < \gamma < 3$  et double exponentiel pour  $\gamma \geq 3$  est surprenant. Néanmoins il doit savoir que ces comportements sont appaissent au loin de la couche maximale, ainsi que les comparaisons avec les simulations numériques Fig. 3.4 et les données réelles Fig. 3.6 confirment nos résultats obtenus.

### 3.6.2 Comparaison avec les données réels

La Fig. 3.6 représente une comparaison entre nos équations et le réseau Hollywoodien de 2283910 acteurs. En général nos équations sont très en accord avec ce réseau réel, ce qui donne une importance supplémentaire à notre travail concernant ce problème de couches dans les réseaux sans-échelle.

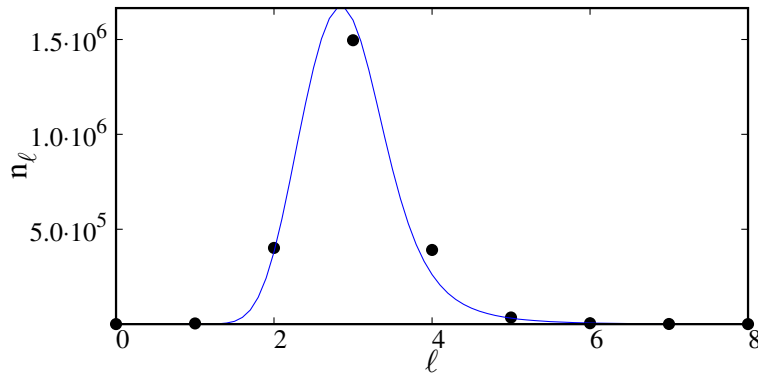


FIGURE 3.6 – Comparaison entre les données empiriques (cercle) et notre équation théorique (ligne continue) : Les cercles représentent les couches du réseau d'acteurs Hollywoodien de 2283910 acteurs et de  $\gamma = 2.13$  lorsque Kevin Bacon est au centre du réseau et la ligne continue représente l'Eq. (3.39) avec les mêmes valeurs  $\gamma$  et  $n$  du réseau Hollywoodien. Les résultats empiriques ont été extraits du site : "<http://oracleofbacon.org/center.php>".



## 3.7 Plus court chemin dans les réseaux sans-échelle non corrélés

### 3.7.1 L'étude théorique

Nous déduisons le PCC dans les réseaux sans-échelle non corrélés à partir de la distribution des nœuds,  $n_\ell$ . En fait, la forme quasi-symétrique de  $n_\ell$  dans la figure Fig. 3.4 suggère que le PCC correspond à la distance à laquelle  $n_\ell$  est maximal.

Nous commençons par le cas le plus simple  $\gamma = 2$ . Comme déjà mentionné, il y a un maximum de deux couches dans le réseau, le PCC peut être écrit comme :

$$\langle \ell \rangle = \frac{\langle k \rangle + 2(n - \langle k \rangle)}{n}, \quad (3.45)$$

qui tend vers 2 pour un grand  $n$ . Cela signifie que presque tous les nœuds sont connectés en moyen au nœud de degré maximal.

Quand  $2 < \gamma < 3$ , aucune solution pour  $\frac{\partial n_\ell}{\partial \ell} = 0$  ne peut être trouvée directement à partir de l'Eq. (3.39), à la place nous utilisons l'approximation donnée dans Eq. (3.42), où  $n_\ell$  est maximal quand  $\kappa_1^{\frac{1-\beta\langle \ell \rangle - \frac{3}{2}}{1-\beta}} = \frac{n}{\langle k \rangle}$ . Après avoir remplacé  $\kappa_1$  et  $K_1$  par leurs expressions correspondantes, on obtient :

$$\langle \ell \rangle = \frac{\log \left( 1 - \frac{\log(n) - \log(\langle k \rangle)}{\log(n) + \frac{\gamma-1}{3-\gamma} \log(\frac{\gamma-2}{3-\gamma} m)} \right)}{\log \left( \frac{2\gamma-4}{\gamma-1} \right)} + \frac{3}{2}. \quad (3.46)$$

pour  $n$  grand,  $\langle \ell \rangle \approx -\frac{\log(\log(n))}{\log(\beta)}$ . Cette forme d'échelle est derrière la nomenclature mondiale ultra-petite [14], et est largement acceptée pour cette gamme de  $\gamma$  [15, 137, 138, 139, 140].

Pour  $\gamma \geq 3$ , le PCC peut être déduit de Eq. (3.41) en résolvant  $\frac{\partial n_\ell}{\partial \ell} = 0$ . Cela donne :

$$\langle \ell \rangle = \frac{\log(n)}{\log(\kappa')} + \frac{\log(\log(\kappa')) - \log(\langle k \rangle)}{\log(\kappa')} + 1. \quad (3.47)$$

Si  $\gamma > 3$ ,  $\kappa$  est constant (Eq. (3.35)). Pour un grand  $n$ ,  $\langle \ell \rangle \approx \frac{\log(n)}{\log(\kappa')}$ , qui est la forme d'échelle rapportée dans de nombreux autres travaux [135, 138, 141, 142, 143], et connu comme le

comportement du petit monde.

Quand  $\gamma = 3$ ,  $\kappa$  dépend de  $K$ , qui dépend à son tour de  $n$ . Prenant  $\kappa = m(\log(K) - \log(m))$  et  $K = mn^{\frac{1}{\gamma-1}}$ , nous trouvons pour  $n$  grand,  $\langle \ell \rangle \approx \frac{\log(n)}{\log(\log(n))}$ . Ce résultat est en accord avec les travaux précédents [138, 14, 141, 145], et confirme le cas particulier de  $\gamma = 3$ . En effet, la présence des hubs rend les distances entre les nœuds plus petites que celles où les hubs sont absents ( $\gamma > 3$ ), en même temps les hubs ne sont pas suffisamment grands pour faire des distances ultra-petites comme  $2 < \gamma < 3$ .

Comme nous avons signalé dans la Section. 3.5.2, les travaux importants [15, 14] dans ce sujet donnent seulement l'allure du PCC pour les grandes valeurs de  $n$ . L'exception est la contribution de Fronczak et al. [141] où ils ont trouvé les expressions du PCC basées sur les paramètres du réseau, leurs équations ont prédit que le PCC tend vers une valeur constante lorsque  $n$  tend vers l'infini pour  $2 < \gamma < 3$ , mais cela n'est pas vrai selon les simulations numériques (voir Fig. 3.7) et tous les autres résultats existant dans la littérature [15, 137, 138, 139, 140, 14]. De la même figure (Fig. 3.7) on voit que nos résultats sont parfaitement en accord avec les simulations pour  $\gamma \neq 3$ , pour  $\gamma = 3$  nos résultats sont très proches et possèdent la même allure que les simulations, ils sont également très similaires aux ceux de [141].

## 3.8 conclusion

Dans ce chapitre, nous avons étudié en détail certains aspects fondamentaux des réseaux sans-échelle non corrélés. Avec des étapes et des hypothèses simples, nous avons obtenu des expressions explicites, en fonction de l'exposant de degré  $\gamma$ , pour le nombre des nœuds à une distance donnée d'un nœud arbitraire,  $n_\ell$ , ainsi qu'une description précise de la formule de distribution. En plus, nous avons montré que  $n_\ell$  augmente suivant la loi de puissance pour les quelques premières couches, et après avoir atteint un maximum il diminue exponentiellement dans les dernières couches. Profitant de la formule de distribution  $n_\ell$ , nous avons pu déduire l'expression explicite du PCC. Les expressions obtenues reproduisent les formes de mise à l'échelle connues pour différentes plages de  $\gamma$ . Autrement dit, le monde ultra-petit pour  $2 < \gamma < 3$ , et le petit monde pour  $\gamma \geq 3$ . Nos résultats théoriques concordent très bien avec

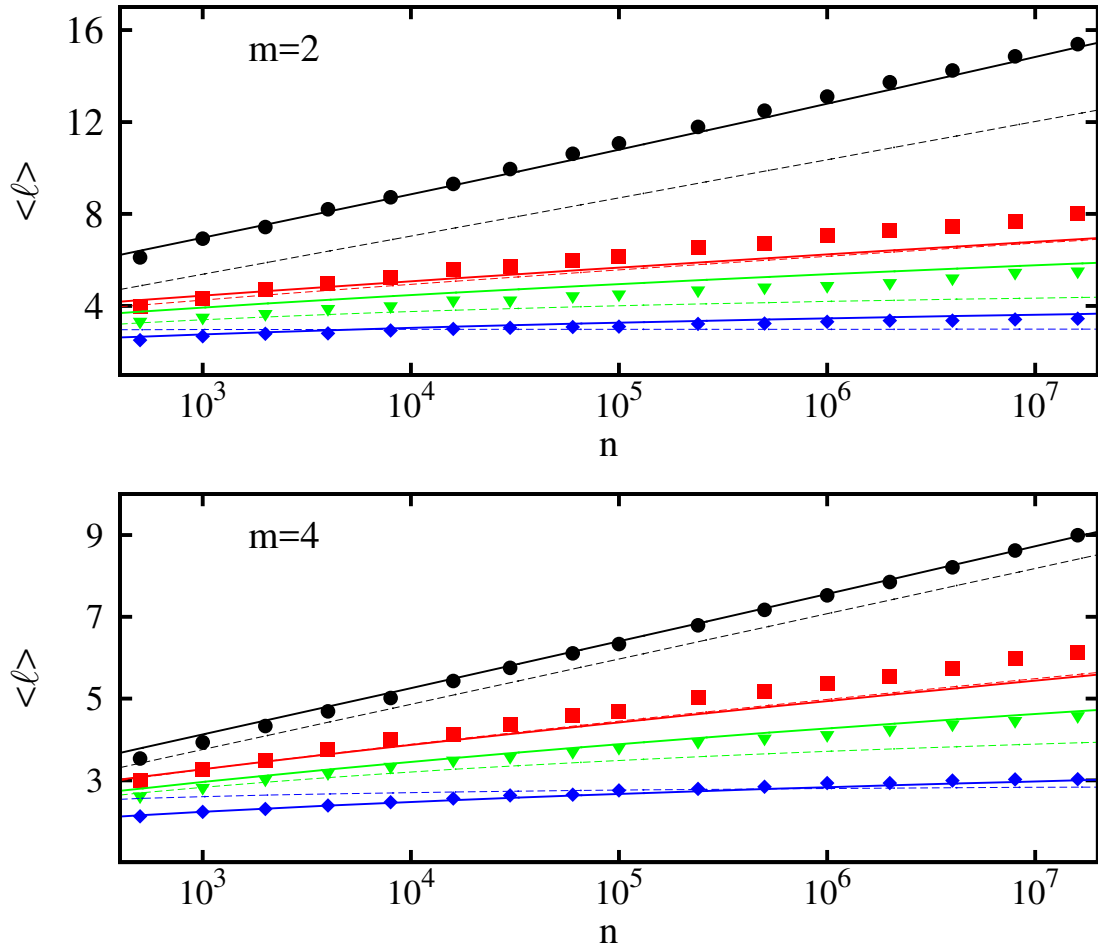


FIGURE 3.7 – La longueur moyenne du trajet en fonction du nombre de nœuds. Les valeurs de  $\gamma$  de haut en bas sont respectivement 4, 3, 2.6 et 2.2. Les lignes pleines correspondent à l'équation Eq. (3.46) et à l'équation Eq. (3.47). Chaque simulation est moyennée à plus de 200.

les simulations, sauf dans le cas de  $\gamma = 3$ , où nous avons observé la même forme, comme pour les autres valeurs de  $\gamma$ , dans les queues de  $n_\ell$ , mais avec une notable différence dans la position du maximum. Cette différence n'affecte pas la forme de mise à l'échelle du PCC pour cette valeur de  $\gamma$ .



# Chapitre 4

## Émergence de la propriété petit-monde dans le modèle petit-monde

L'émergence de la propriété petit-monde dans le modèle Newman-Watts est encore un phénomène pas bien compris, parmi ses problématiques on cite : à partir de quel point ce réseau change sa nature grand-monde vers petit-monde ? Si ce réseau se sature par les raccourcis ajoutés, comment et quand la saturation se réalise ? Est-ce-qu'il y a vraiment une transition de phase ou non ? Etc. Par des études théoriques confirmant par un grand nombre de simulations, on abordera quelques questions de ce genre.

### 4.1 Introduction

Le modèle de Watts-Storogatz (WS) retient une attention particulière depuis son apparition en 1998 [93], il a deux propriétés très intéressantes : la présence d'un coefficient de Clustering élevé et d'un petit PCC. Ces deux propriétés se trouvent dans la plupart des réseaux réels [128, 146], en plus ce modèle est très simple et combine la régularité avec l'aléatoire. En 1999 Newman et Watts [147] ont réalisé une petite modification sur le modèle WS, dont il y a  $n$  nœuds qui se distribuent dans un réseau régulier unidimensionnel sous forme d'un cercle où chaque nœud fait  $2k$  liens avec ses plus proches voisins (voir Fig. 4.1), le nombre de

nœuds reste fixe puis chaque lien se reconnecte par une probabilité  $\phi$  entre deux autres nœuds choisis aléatoirement sans supprimer aucuns liens, les liens ajoutés sont nommés raccourcis, en moyenne il y a  $x = nk\phi$  raccourcis. Ce modèle détient aussi une grande attention car il a les mêmes propriétés du modèle WS, sauf pour  $k = 1$  où il y a une certaine différence<sup>1</sup>, au même temps il est moins difficile dans l'étude.

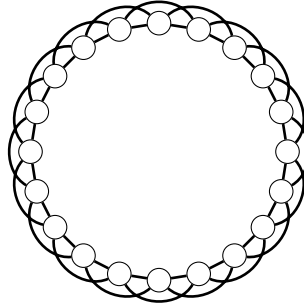


FIGURE 4.1 – Un réseau petit-monde WS pour  $n = 20$ ,  $k = 2$  sans raccourcis.

### 4.1.1 Le modèle petit-monde de Watts-Strogatz

Le réseau aléatoire ER reproduit très bien l'observation concernant la longueur moyenne relativement petite des chemins, mais le coefficient de regroupement (Clustering) dans ce réseau est très faible par rapport à ceux observés dans les systèmes réels, et la distribution des degrés est déférente de celle du réseau réel. En 1998 Watts et Strogatz [93] ont proposé un modèle qui reproduit deux propriétés parmi celles mentionnées précédemment de manière simple.

Les modèles de petit-monde peuvent être construits sur des réseaux de n'importe quelle dimension ou topologie, mais le cas le mieux étudié est l'unidimensionnel. Soit  $n$  le nombre de nœuds et  $k$  un nombre pair, le modèle Watts-Strogatz commence par la construction suivante (voir Fig.4.2) :

- Placez tous les noeuds en cercle ;
- Connectez chaque nœud à ses premiers  $k$  voisins les plus proches ;
- Avec une probabilité  $p$  en prenant chaque lien du réseau et on le reconnecte.

---

1. Le cas  $k = 1$  ne nous intériorisons jamais dans ce modèle, car dans ce cas le coefficient de Clustering est zéro, alors on prend toujours  $k > 1$ .

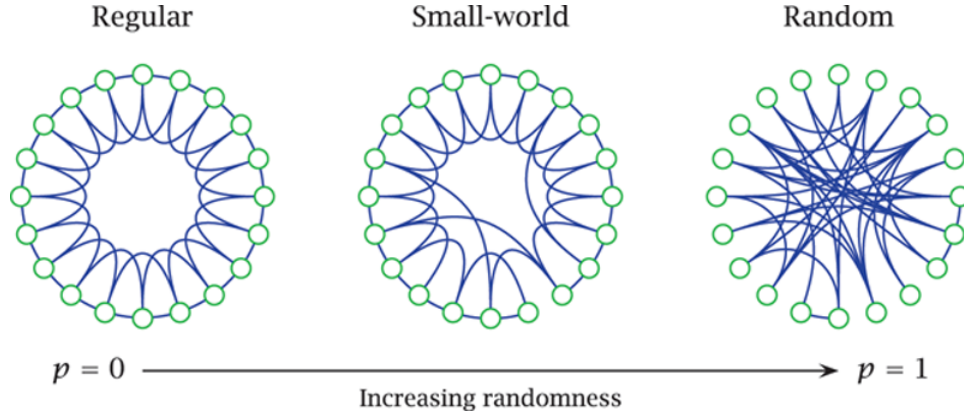


FIGURE 4.2 – Représentation schématique de l'évolution du processus de "reconnexion" dans le modèle Watts-Strogatz.

Ce modèle peut être justifié en disant que la plupart des gens sont amis avec leurs voisins immédiats, par exemple les voisins de la même rue ou des collègues du même bureau. D'autre part, beaucoup des gens ont quelques amis qui se vivent loin. Les propriétés structurales du modèle sont quantifiées par la longueur de chemin caractéristique  $L(p)$  et le coefficient de regroupement  $C(p)$  en tant que fonctions de la probabilité de reconnecter  $p$ . La longueur de chemin caractéristique  $L$  est définie comme la valeur moyenne de nombre d'arêtes dans le chemin le plus court entre deux nœuds.

Nous savons que le réseau ordinaire à  $p = 0$  est un grand monde en grappe, où  $L$  se développe linéairement avec  $n$ , tandis que le graphe aléatoire à  $p = 1$  est un petit monde faiblement regroupé où  $L$  ne se développe que logarithmiquement avec  $n$ .

Selon Fig. 4.3 on voit que l'introduction de quelques bords à longue distance entraîne une baisse immédiate de  $L(p)$ , par contre  $C(p)$  reste pratiquement inchangé pour les petites  $p$  même si  $L(p)$  chute rapidement. L'implication importante est que, au niveau local, la mise à l'échelle du petit monde de la distance géodésique moyenne entre les nœuds peut se produire pour les mêmes valeurs de  $p$  auxquelles le coefficient de regroupement local du réseau est significativement important.

Le plus court chemin moyen a été montré comme [148, 149]

$$L \sim \frac{n}{k} f(x), \quad (4.1)$$

avec  $x$  le nombre moyen des bords qui se reconnectent dans le réseau et  $f(x)$  une fonction

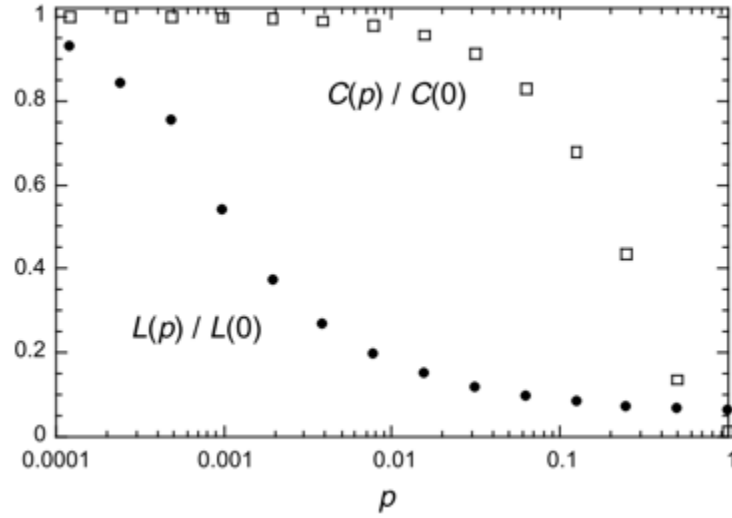


FIGURE 4.3 – Représentation de la variation de la longueur moyenne du chemin  $L(p)$  et du coefficient de regroupement  $C(p)$  avec la modification de la probabilité de reconnexion  $p$  dans le modèle Watts-Strogatz.

universelle qui a les formes asymptotiques suivantes :

$$f(x) \sim \begin{cases} \frac{1}{4} & \text{si } x \gg 1 \\ \frac{\log 2x}{4x} & \text{si } x \ll 1. \end{cases} \quad (4.2)$$

Le coefficient de regroupement  $C$  est défini par Eq. (1.10), que nous reproduisons ici :

$$C = \frac{3 \times (\text{Nombre de triangles})}{(\text{Nombre de triples connectés})}. \quad (4.3)$$

Pour évaluer  $C$ , nous devons calculer le nombre de triangles et les triplets connectés dans le réseau, alors on va trouver la valeur moyenne de  $C$  comme la suite [150] :

$$C = \frac{3(k-1)}{4(k-1) + 8kp + 4kp^2}. \quad (4.4)$$

En ce qui concerne la répartition de degré, lorsque  $p = 0$  est une fonction delta positionnée à  $k$ , alors que pour  $p = 1$  elle est similaire à celle d'un réseau ER. Pour les valeurs intermédiaires



de  $p$ , la répartition du degré est donnée par [151] :

$$P(k) = \sum_{i=0}^{\min(k-m, m)} C_i^m \frac{(pm)^{k-m-i}}{(k-m-i)!} e^{pm}. \quad (4.5)$$

## 4.2 La théorie de groupe de renormalisation et transition de phase

Dans un système physique proche d'une transition de phase, les méthodes d'approximations les plus courantes s'appuyant à négliger les corrélations entre un grand nombre de particules ou plutôt entre un grand nombre de degrés de liberté, ces approximations sont valables lorsque la longueur de corrélation est petite. En pratique, on sait traiter de façon simple les corrélations à deux particules. Le problème à trois particules est déjà beaucoup plus difficile. Ce type de méthodes est voué à l'échec quand la longueur de corrélation est grande.

D'où l'idée de réduction du nombre de degrés de liberté, on cherche à établir une correspondance entre un problème de longueur de corrélation donnée et un problème de longueur de corrélation plus petite. La théorie de groupe de renormalisation établit ainsi des correspondances entre systèmes de longueurs de corrélation différentes. Considérons par exemple un système de moments magnétique en interaction au voisinage du point critique de la transition ferromagnétique. Afin de réduire le nombre élevé de degrés de liberté, au lieu de considérer tous les moments magnétiques atomiques individuellement, on les regroupe en blocs comprenant plusieurs moments que l'on considère comme nouvelles entités de base, avec un nouveau moment (moment de bloc). On calcule alors les interactions entre ces nouvelles entités, ce qui suppose que l'on a su moyenner les fluctuations des variables internes à l'intérieur des blocs. On change ensuite d'échelle de façon que le nouveau réseau devienne le même que le précédent, enfin, on renormalise judicieusement la taille des nouveaux moments magnétiques. Une série d'opérations fait passer un système à un autre avec réduction de la longueur de corrélation dans un rapport donné. Dans cette transformation, les degrés de liberté du système sont grignotés car on a moyenné les fluctuations des variations internes des blocs. Autrement dit on a remplacé l'interaction initiale entre les anciens degrés de liberté par une nouvelle interaction effective entre les degrés de liberté réduits [152, 153].

## 4.3 Plus court chemin et transition de phase dans le modèle NW : ancienne contribution et la notre

### 4.3.1 Ancienne contribution

Malgré tous les efforts qui ont été faits, il n'existe pas encore de calcul exact des couches<sup>2</sup> et du PCC dans le modèle de NW. En 2000 Newman, Moore et Watts [149] ont trouvé pour la première fois l'expression du PCC dans ce modèle sous la forme universelle  $\ell = \frac{n}{k}f(x)$  avec  $f$  est la fonction universelle<sup>3</sup> :

$$f(x) = \frac{1}{2\sqrt{x^2 + 2x}} \tanh^{-1}\left(\sqrt{\frac{x}{x+2}}\right), \quad (4.6)$$

selon NW cette expression est bonne pour le cas de petit nombre des raccourcis et pour les réseaux larges, avec la remarque qu'au voisinage de  $\phi = 1$  on observe une déviation entre cette solution analytique et les simulations numériques (voir Fig. 4.7). D'un autre côté, le comportement de la transition entre réseau-régulier et réseau-aléatoire ou plutôt entre grand-monde et petit-monde est aussi une part très importante dans l'étude de ce modèle, Barthélémy et Amaral [148] ont trouvé que l'apparition du comportement du petit-monde n'est pas une transition de phase mais un phénomène de croisement, leur méthode tend à déterminer la taille de réseau  $n^*(\phi)$  pour lequel : si  $n < n^*$ ,  $\ell$  croît linéairement avec  $n$ , et si  $n > n^*$ ,  $\ell$  croît comme  $\ln(n)$ . Par des analyses numériques ils ont trouvé que  $n^*(\phi) \sim \phi^{-\tau}$  avec  $\tau \approx \frac{2}{3}$ . Ces résultats sont rapidement critiqués par Barrat [154] car il a démontré que  $\tau$  ne peut pas être inférieur à 1, en plus il a trouvé que  $\tau = 1$  en utilisant la même approche de Barthélémy et Amaral mais avec des tailles de système plus grandes. Newman et Watts [6, 155] confirment cette valeur de l'exposant critique unique  $\tau = 1$  en utilisant une transformation de groupe de renormalisation et ils ont dit aussi que la transition vers le petit-monde se fait d'une façon continue, c'est-à-dire par une transition de phase de deuxième ordre où le point de transition de phase se réalise lorsque la densité des raccourcis tend vers zéro ( $\phi = 0$ ). Mais M. Argollo et al. [156] ont trouvé que la transition de phase est de première ordre au point

---

2. C'est-à-dire le nombre moyen de nœuds,  $n_\ell$ , à distance  $\ell$  depuis un nœud arbitraire

3. Cette fonction universelle a été déjà démontré par Newman et Watts [147], mais sans donner aucune expression théorique, leur étude est basée sur les simulations numériques.

$\phi = 0$ , à cause d'une discontinuité d'un certain paramètre d'ordre dans ce point. On en déduit qu'il y a encore un débat à propos du type de transition de phase dans ce modèle.

### 4.3.2 Notre contribution

Notre travail ici se consacre, au début, à faire une étude sur les couches dans le modèle NW en utilisant la transformation de groupe de renormalisation en espace réel (GR), (voir Fig. 4.5). Une couche  $n_d$  est le nombre de nœuds ayant la distance  $d$  autour d'un nœud arbitraire. Sachant que ce modèle est un mélange entre la régularité et l'aléatoire, nous proposons de séparer les couches selon deux types, couches-régulières  $n_d^r$  qui représentent les nœuds restant dans leurs distances régulières initiales,  $d$ , sans aucune influence par les raccourcis ajoutés et les couches-aléatoires  $n_d^{al}$  qui représentent les nœuds changeant leurs distances régulières vers une autre plus proche du nœud arbitraire,  $d$ , à cause des raccourcis (voir Fig. 4.4). En manipulant les expressions de ces couches, on trouvera que la somme des couches aléatoires  $S_{al}$  et la somme des couches régulières  $S_r$ , pouvant également s'écrire sous forme d'une fonction universelle comme le PCC,  $S_{al} = nh(x)$  et  $S_r = n(1 - h(x))$  avec  $h(x) = 1 - \sqrt{\frac{\pi}{4x}} \text{erf}(\sqrt{x})$ . Sachant que la fonction  $h(x)$  représente la fraction des nœuds appartenant aux couches-aléatoires, on peut la considérer comme un paramètre d'ordre, car elle varie entre 0 et 1 selon le degré de l'ordre et de désordre dans le réseau. A partir de l'expression de  $h(x)$  on déduit l'absence d'une transition de phase dans ce modèle mais un phénomène de croisement qui commence depuis  $x = 0$ .

Pour calculer le PCC,  $\ell$ , nous utilisons les résultats précédents et nous considérons également, comme le cas des couches, que le PCC est la somme du PCC de réseau-régulier  $\ell_r$  et le PCC de réseau-aléatoire  $\ell_{al}$ ,  $\ell = \ell_{al} + \ell_r$ , en se basant dans le calcul de  $\ell_{al}$  sur la proposition que le PCC est la position de la couche maximale (voir Chapitre. 3) et pour calculer  $\ell_r$  on utilise une approximation qui sera décrite dans la troisième section de ce chapitre, et on va trouver une expression du PCC plus précise par rapport à l'expression de Newman et al. (voir Fig. 4.7). En plus nous allons démontrer que la formule universelle en fonction du nombre de raccourcis est valable sauf pour  $y \ll 1$ , avec  $y = 2k\phi$  un paramètre qu'on va définir plus tard. Ce résultat est très important, car on a montré que la fonction

universelle  $f(x)$  qui a été considérée conceptuellement vrai n'est pas toujours valable, en revanche on obtient une autre formule universelle  $g(y)$  qui est valable lorsque  $y$  n'est pas très inférieur à 1. Cette nouvelle formule sous la forme  $g(y) = \frac{\ln(n)}{\ell}$ , nous montre que la propriété petit-monde émerge d'une façon universelle en fonction de  $y$ .

## 4.4 Structure du réseau NW : couche et plus court chemin

### 4.4.1 Les couches aléatoires et régulières

Au début on va appliquer une transformation de groupe de renormalisation en espace réel, en transformant le réseau NW de  $k$  et  $n$  quelconque à un réseau de  $\acute{k} = 1$  et de nombre de nœuds  $\acute{n} = \frac{n}{k}$  (voir Fig. 4.5)<sup>4</sup>. On suppose que chaque ensemble de  $k$  nœuds voisins comme un seul nœud, d'où la probabilité qu'un nœud dans le nouveau réseau ( $\acute{k} = 1$  et  $\acute{n} = \frac{n}{k}$ ) soit lié aléatoirement à un autre nœud est  $q = 1 - \left(1 - \frac{2k\phi}{n}\right)^{k^2}$ , pour  $n \gg k\phi$  on peut écrire  $q = \frac{2k^3\phi}{n}$ . Puis nous calculons la probabilité  $P_r(j)$  qu'un nœud reste dans sa distance régulière  $j$  et la probabilité  $P_{al}(j)$  qu'un nœud change sa distance régulière vers une nouvelle distance  $j$  plus proche du nœud arbitraire grâce aux raccourcis ajoutés, par exemple, dans la Fig. 4.4, les nœuds noirs sont ceux qui deviennent plus proches du nœud arbitraire grâce au raccourci. Généralement les recules des nœuds vers des distances plus proche du nœud arbitraire se fassent à travers 1 raccourci, 2 ou plus, pour cette raison on distingue chaque cas différemment.

■ À travers 1 raccourci :

Soit  $\pi^1(i)$  la probabilité qu'un nœud ne change pas sa distance régulière  $j$  vers la  $i^{\text{ème}}$

---

4. On voit que le degré de liberté est réduit, car dans le nouveau réseau le paramètre  $\acute{k}$  est toujours égale à 1, c'est exactement l'idée de la théorie de groupe de renormalisation.

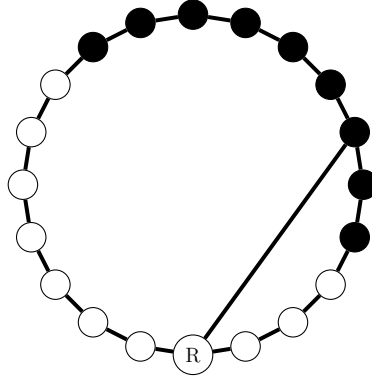


FIGURE 4.4 – Illustration des nœuds aléatoires (nœuds noirs) qui deviennent plus proches du nœud arbitraire  $R$ , et nœuds réguliers (nœuds blanches) qui restent dans leurs distance initial par rapport au même nœud arbitraire  $R$ , le réseau est de taille  $n = 20$ ,  $k = 1$  et un seul raccourci.

distance ( $i < j$ ) à travers un seul raccourci :

$$\begin{aligned}
 \pi^1(1) &= (1 - q), \\
 \pi^1(2) &= (1 - q)^4, \\
 \pi^1(3) &= (1 - q)^{4 \cdot 2}, \\
 &\vdots \\
 \pi^1(i) &= (1 - q)^{4(i-1)}.
 \end{aligned} \tag{4.7}$$

Il faut signaler que l'expression  $\pi^1(i) = (1 - q)^{4(i-1)}$  est une approximation de champs moyenne qui est vrai pour la plupart des cas et qui représente la valeur minimale de cette probabilité, car les nombres des possibilités  $4, 4 \times 2, \dots, 4 \times (i - 1)$  sont les valeurs maximales, mais à cause des interférences entres elles, le nombre de possibilités peut être moindre, la même remarque pour tous les cas suivants, on prend toujours le nombre maximal de possibilités. Par conséquent nos équations sont des approximations de champs moyenne et ne sont pas des expressions mathématiques exactes, certes comme tous les systèmes de très grand nombre de constituants, il suffit de faire la bonne approximation au niveau microscopique, c'est-à-dire local, pour trouver la description parfaite des grandeurs macroscopiques du système, ce que nous avons fait exactement. De l'Eq. (4.7) la probabilité  $P_r^1(j)$  qu'un nœud ne change pas sa distance régulière  $j$  à

travers un seul raccourci est :

$$\begin{aligned}
 P_r^1(j) &= \pi^1(1)\pi^1(2)\dots\pi^1(j-1) \\
 &= (1-q)(1-q)^4(1-q)^{4\cdot 2}\dots(1-q)^{4(j-2)} \\
 &= (1-q)^{1+4+4\cdot 2+4\cdot 3\dots 4(j-2)} \\
 &= (1-q)^{1+4\sum_{i=1}^{j-1}(i-1)}.
 \end{aligned}$$

On doit signaler qu'on peut négliger le cas  $\pi^i(i) = (1-q)^i(\acute{n}-2i)^{i-1}$ , dans le cas d'un raccourci et dans tous les cas suivants sans aucun problème, alors  $P_r^1(j)$  sera écrite :

$$P_r^1(j) = (1-q)^{4\sum_{i=1}^{j-1}(i-1)}$$

■ À travers 2 raccourcis :

La probabilité qu'un nœud ne change pas sa distance régulière  $j$  vers la  $i^{\text{ème}}$  distance ( $i < j$ ) à travers deux raccourcis, dans le cas où le nœud arbitraire lié directement par un raccourci avec un nœud intermédiaire  $z$  qui est à une distance  $i-1$  du nœud  $j$  à travers aussi un seul raccourci, est  $(1-q^2)^{4(i-2)}$ . On considère que cette probabilité est la même pour les autres possibilités d'arrangements de deux raccourcis autour du nœud intermédiaire.

Le nombre d'arrangements possibles de ces deux raccourcis autour de nœud intermédiaire  $z$  pour lequel la distance entre  $j$  et le nœud arbitraire sera toujours  $i$  est <sup>5</sup>  $(i-1)$ .

En outre le nombre maximal de positions possibles de  $z$  sur le réseau pour que  $i < j$  est  $\acute{n}-2i$ , alors la probabilité qu'un nœud ne change pas sa distance régulière  $j$  vers la  $i^{\text{ème}}$  distance ( $i < j$ ) à travers deux raccourcis est

$$\pi^2(i) = (1-q^2)^{4(i-1)(i-2)(\acute{n}-2i)},$$

---

5. Par exemple dans le cas où  $i = 4$ , on a 3 arrangements possibles pour que la distance entre le nœud  $j$  et le nombre arbitraire est 4 :

le 1<sup>ère</sup> : la distance entre le nœud arbitraire et le nœud intermédiaire  $z$  est 1 et entre  $z$  et le nœud  $j$  est 3.

le 2<sup>ème</sup> : la distance entre le nœud arbitraire et le nœud intermédiaire  $z$  est 2 et entre  $z$  et le nœud  $j$  est également 2.

le 3<sup>ème</sup> : la distance entre le nœud arbitraire et le nœud intermédiaire  $z$  est 3 et entre  $z$  et le nœud  $j$  est 1.

d'où

$$\begin{aligned} P_r^2(j) &= \pi^2(1)\pi^2(2)\dots\pi^2(j-1) \\ &= (1-q^2)^{4\sum_{i=1}^{j-1}((i-1)(i-2)(n-2i))}, \end{aligned} \quad (4.8)$$

comme nous avons déjà signalé le terme  $\pi^i(i)$  est négligeable, ici est  $\pi^2(2)$ .

■ À travers 3 raccourcis :

Dans ce cas on a deux nœuds intermédiaires  $z$  et  $\hat{z}$  sous la condition qu'entre deux nœuds intermédiaires il y a un seul raccourci. La probabilité qu'un nœud ne change pas sa distance  $j$  vers la  $i^{\text{ème}}$  distance à travers 3 raccourcis dans le cas où le nœud arbitraire est lié directement par un raccourci avec  $z$  qui est aussi lié directement par un raccourci avec  $\hat{z}$  et ce dernier à une distance de  $i-2$  au nœud  $j$  à travers aussi un seul raccourci, est  $(1-q^3)^{4(i-3)}$ . On considère que cette probabilité est la même pour les autres possibilités d'arrangements de trois raccourcis autour des deux nœuds intermédiaires.

Le nombre d'arrangements possibles de ces trois raccourcis autour des nœuds intermédiaires  $z$  et  $\hat{z}$  pour que la distance entre  $j$  et le nœud arbitraire soit toujours  $i$  est  $C_2^{i-1} = \frac{(i-1)(i-2)}{2!}$ .

En outre le nombre maximal des positions possibles de  $z$  et  $\hat{z}$  sur le réseau pour  $i < j$  est approximativement  $(n-2i)^2$ .

Alors on obtient que  $\pi^3(i) = (1-q^3)^{4\frac{(i-1)(i-2)}{2!}(i-3)(n-2i)^2}$  d'où

$$\begin{aligned} P_r^3(j) &= \pi^3(1)\pi^3(2)\dots\pi^3(j-1) \\ &= (1-q^3)^{4\sum_{i=1}^{j-1}(\frac{(i-1)(i-2)}{2!}(i-3)(n-2i)^2)}. \end{aligned} \quad (4.9)$$

■ À travers 4 raccourcis :

Dans ce cas on a trois nœuds intermédiaires  $z, z'$  et  $z''$ , toujours sous la condition qu'il y a un seul raccourci entre chaque deux nœuds intermédiaires. La probabilité qu'un

noeud ne change pas sa distance  $j$  vers la  $i^{\text{ème}}$  distance à travers 4 raccourcis dans le cas où le noeud arbitraire est lié directement par un raccourci avec  $z$  et celui-ci est aussi directement lié par un raccourci avec  $z'$  qui est de sa part lié directement par un raccourci avec le noeud  $z''$  et ce dernier à une distance de  $i - 3$  au noeud  $j$  à travers aussi un seul raccourci est  $(1 - q^4)^{4(i-4)}$ . On considère comme les cas précédent que cette probabilité est la même pour les autres possibilités d'arrangements de quatre raccourcis autour les trois noeuds intermédiaires.

Le nombre d'arrangements possibles de ces trois raccourcis autour des noeuds intermédiaires  $z, z'$  et  $z''$  pour que la distance entre  $j$  et le noeud arbitraire est  $i$  est  $C_3^{i-1} = \frac{(i-1)(i-2)(i-3)}{3!}$ . En outre le nombre maximal des positions possibles de  $z, z'$  et  $z''$  sur le réseau pour  $i < j$  est approximativement  $(n - 2i)^3$ . Alors on obtient

$$\pi^4(i) = (1 - q^4)^{4 \frac{(i-1)(i-2)(i-3)}{3!} (i-4)(n-2i)^3},$$

d'où

$$\begin{aligned} P_r^4(j) &= \pi^4(1)\pi^4(2)\dots\pi^4(j-1) \\ &= (1 - q^4)^{4 \sum_{i=1}^{j-1} \left( \frac{(i-1)(i-2)(i-3)}{3!} (i-4)(n-2i)^3 \right)}. \end{aligned} \quad (4.10)$$

■ À travers  $m$  raccourcis :

Par la même méthode, si on suppose qu'il y a  $(m - 1)$  noeuds intermédiaires, on va obtenir l'expression générale de la probabilité qu'un noeud ne change pas sa distance  $j$  vers la  $i^{\text{ème}}$  distance à travers  $m$  raccourcis sous la forme

$$\pi^m(i) = (1 - q^m)^{4 \frac{(i-1)(i-2)\dots(i-m)}{(m-1)!} (n-2i)^{(m-1)}},$$

alors la probabilité qu'un noeud ne change pas sa distance régulière  $j$  à travers  $m$  raccourcis est

$$\begin{aligned} P_r^m(j) &= \pi^m(1)\pi^m(2)\dots\pi^m(j-1) \\ &= (1 - q^m)^{4 \sum_{i=1}^{j-1} \left( \frac{(i-1)(i-2)\dots(i-m)}{(m-1)!} (n-2i)^{(m-1)} \right)}, \end{aligned} \quad (4.11)$$



sachant que  $q < 1$  on peut écrire

$$P_r^m(j) = e^{-4q^m \sum_{i=1}^{j-1} \left( \frac{(i-1)(i-2)\dots(i-m)}{(m-1)!} (\acute{n}-2i)^{m-1} \right)}. \quad (4.12)$$

Alors la probabilité  $P_r(j)$  qu'un nœud ne change pas sa distance régulière  $j$  vers la  $i^{\text{ème}}$  distance par n'importe quel nombre des raccourcis est

$$\begin{aligned} P_r(j) &= P_r^1(j) P_r^2(j) P_r^4(j) \dots P_r^m(j) \\ &= e^{-4q \sum_{i=1}^{j-1} (i-1) B(i)}, \end{aligned} \quad (4.13)$$

avec  $B(i) = (1 + [q(i-2)(\acute{n}-2i)] + [q^2 \frac{(i-2)(i-3)}{2!} (\acute{n}-2i)^2] + \dots + [q^{m-1} \frac{(i-2)\dots(i-m)}{(m-1)!} (\acute{n}-2i)^{m-1}] + \dots + [q^{i-2} (\acute{n}-2i)^{i-2}])$ , les termes de cette expression correspondent respectivement aux probabilités que le nœud  $j$  ne change pas sa distance vers le  $i^{\text{ème}}$  distance par 2, 3, 4, ...,  $i-1$  raccourcis, le cas de  $i$  raccourcis est négligé comme nous avons déjà dit. D'un autre coté on voit que  $B(i)$  peut s'écrire sous forme de Binôme

$$\begin{aligned} B(i) &= (1 + [q(i-2)(\acute{n}-2i)] + [q^2 \frac{(i-2)(i-3)}{2!} (\acute{n}-2i)^2] + \dots + [q^{i-2} (\acute{n}-2i)^{i-2}]) \\ &= \sum_{j=1}^{i-2} C_j^{i-2} [q(\acute{n}-2i)]^j 1^{i-2-j} \\ &= (q(\acute{n}-2i) + 1)^{i-2}. \end{aligned}$$

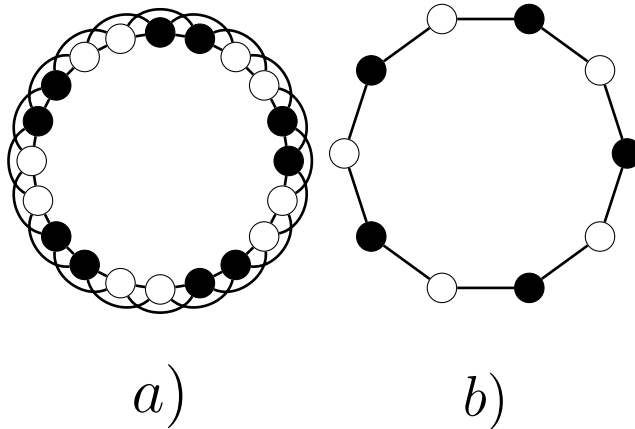


FIGURE 4.5 – La transformation GR d'un réseau (a) de  $n = 20$  et  $k = 2$  vers un réseau (b) de  $\acute{n} = 10$  et  $\acute{k} = 1$ .

Alors la probabilité qu'un nœud ne change pas sa distance  $j$  est :

$$\begin{aligned} P_r(j) &= e^{-4q \sum_{i=1}^{j-1} (i-1)(q(\acute{n}-2i)+1)^{i-2}} \\ &= e^{-4q \int_{i=1}^{j-1} (i-1)(q(\acute{n}-2i)+1)^{i-2} di}. \end{aligned} \quad (4.14)$$

En outre la probabilité  $P_{al}(j)$  qu'un nœud change sa distance régulière vers la  $j^{\text{ème}}$  est égale au produit de la probabilité que ce nœud ne change pas sa distance vers une autre inférieure à  $j$ ,  $P_r(j-1)$ , et la probabilité que le nœud change sa distance vers la  $j^{\text{ème}}$  distance,  $1 - \pi(j)$ ,

$$\begin{aligned} P_{al}(j) &= P_r(j-1)(1 - \pi(j)) \\ &= P_r(j-1) - \pi(j)P_r(j-1) \\ &= P_r(j-1) - P_r(j) \\ &= -\frac{\partial P_r(j)}{\partial j} \\ &= 4q(j-2)(q(\acute{n}-2(j-1)-1)) + 1)^{j-3} P_r(j), \end{aligned} \quad (4.15)$$

pour simplifier les calculs on peut prendre sans aucun souci<sup>6</sup> que dans l'équation précédent  $j-1 = j$ , d'où on obtient

$$P_{al}(j) = 4q(j-1)(q(\acute{n}-2j-1)) + 1)^{j-2} P_r(j). \quad (4.16)$$

Le nombre de nœuds dans la couche  $n_j^r$  est  $P_r(j)$  multiplié par 2, car initialement dans chaque couche on a deux nœuds réguliers

$$n_j^r = 2P_r(j), \quad (4.17)$$

et le nombre des nœuds dans le couche aléatoire  $n_j^{al}$  est  $P_{al}(j)$  multiplié par le nombre des nœuds qui ont une distance plus grande que  $j$

$$n_j^{al} = (\acute{n} - 2j)P_{al}(j), \quad (4.18)$$

---

6. Cette approximation signifie une augmentation de l'erreur de l'ordre  $\frac{1}{\ell}$ , cependant  $\ell$  tend vers l'infini lorsque  $n$  est très large, alors cette approximation ne pose aucun problème.

d'où la  $j^{\text{ème}}$  couche est

$$\begin{aligned}
 n_j &= n_j^r + n_j^{al} \\
 &= 2P_r(j) + (\acute{n} - 2j)P_{al}(j) \\
 &= 2P_r(j) + (\acute{n} - 2j)P_r(j-1) - (\acute{n} - 2j)P_r(j) \\
 &= (\acute{n} - 2j)P_r(j-1) - (\acute{n} - 2(j+1))P_r(j) \\
 &= v(j-1) - v(j) \quad \text{avec : } v(j) = (\acute{n} - 2(j+1))P_r(j).
 \end{aligned} \tag{4.19}$$

Maintenant on va calculer la somme des nœuds de couches-régulières  $S'_r$  et la somme des nœuds de couches-aléatoires  $S'_{al}$  afin de les utiliser dans le calcul du PCC. De Eq. (4.14) et Eq. (4.17) on obtient

$$\begin{aligned}
 S'_r &= n_1^r + \sum_2^{\frac{\acute{n}}{2}} n_i^r \\
 &= n_1^r + \int_2^{\frac{\acute{n}}{2}} n_i^r di \\
 &= n_1^r + 2 \int_2^{\frac{\acute{n}}{2}} e^{-4q \int_{j=1}^{i-1} (j-1)(q(\acute{n}-2j)+1)^{j-2} dj} di,
 \end{aligned} \tag{4.20}$$

l'intégrale  $\int_2^{\frac{\acute{n}}{2}} e^{-4q \int_{j=1}^{i-1} (j-1)(q(\acute{n}-2j)+1)^{j-2} dj} di$ , n'a aucune solution disponible, pour cette raison, on est obligé d'utiliser une approximation en tenant compte seulement des chemins à travers un seul raccourci et on néglige les autres.

L'approximation s'avère vraie dans la première région où le nombre de raccourcis est très faible, par contre dans la seconde région où le nombre de raccourcis est grand elle n'est pas vraie, néanmoins la fraction des nœuds du réseau régulier dans la seconde région est très petite, pour cette raison cette approximation reste valable pour les deux régions. Alors on considère que  $\int_2^{\frac{\acute{n}}{2}} e^{-4q \int_{j=1}^{i-1} (j-1)(q(\acute{n}-2j)+1)^{j-2} dj} di \approx \int_2^{\frac{\acute{n}}{2}} e^{-4q \int_{j=1}^{i-1} (j-1)^{j-2} dj} di$ . Pour  $\acute{n} \gg 1$  on peut noter que

$\int_2^{\frac{\acute{n}}{2}} e^{-4q \int_{j=1}^{i-1} (j-1) dj} di \approx \int_2^{\frac{\acute{n}}{2}} e^{-4q \int_{j=1}^i j dj} di$ , d'où :

$$\begin{aligned}
 S'_r &\approx n_1^r + 2 \int_2^{\frac{\acute{n}}{2}} e^{-4q \int_{j=1}^i j dj} di \\
 &\approx n_1^r + 2 \int_2^{\frac{\acute{n}}{2}} e^{-2qi^2} di \\
 &\approx 2 + 2 \left[ \frac{\sqrt{\frac{\pi}{2}} \operatorname{erf}(\sqrt{2q}i)}{2\sqrt{q}} \right]_2^{\frac{\acute{n}}{2}} \quad \text{avec } n_1^r = 2 \\
 &\approx 2 + 2\sqrt{\frac{\pi}{8q}} \left[ \operatorname{erf}(\sqrt{2q}\frac{\acute{n}}{2}) - \operatorname{erf}(2\sqrt{2q}) \right],
 \end{aligned} \tag{4.21}$$

pour  $q \ll 1$  on a  $\operatorname{erf}(2\sqrt{2q}) = \frac{2}{\sqrt{\pi}} 2\sqrt{2q}$  alors on obtient

$$\begin{aligned}
 S'_r &\approx 2 + 2\sqrt{\frac{\pi}{8q}} \left[ \operatorname{erf}(\sqrt{2q}\frac{\acute{n}}{2}) - \frac{2}{\sqrt{\pi}} 2\sqrt{2q} \right] \\
 &\approx 2 + 2\sqrt{\frac{\pi}{8q}} \operatorname{erf}(\sqrt{\frac{q\acute{n}^2}{2}}) - 4 \\
 &\approx \acute{n} \left( \sqrt{\frac{\pi}{2q\acute{n}^2}} \operatorname{erf}(\sqrt{\frac{q\acute{n}^2}{2}}) - \frac{2}{\acute{n}} \right) \\
 &\approx \acute{n} \sqrt{\frac{\pi}{2q\acute{n}^2}} \operatorname{erf}(\sqrt{\frac{q\acute{n}^2}{2}}).
 \end{aligned} \tag{4.22}$$

On sait que  $q = \frac{2k^3\phi}{n}$  et  $\acute{n} = \frac{n}{k}$  alors  $\frac{q\acute{n}^2}{2} = kn\phi$  qui est le nombre moyen des raccourcis dans le réseau, d'où la somme des nœuds réguliers s'écrit aussi comme une fonction universelle, sous la forme

$$S'_r = \frac{n}{k} (1 - h(kn\phi)), \tag{4.23}$$

avec  $h(x) = 1 - \sqrt{\frac{\pi}{4x}} \operatorname{erf}(\sqrt{x})$

En outre la somme des nœuds aléatoires est la soustraction du nombre de nœuds et la

somme des nœuds réguliers

$$\begin{aligned}
 S'_{al} &\approx \dot{n} - S'_r \\
 &\approx \dot{n} - \dot{n} \sqrt{\frac{\pi}{2q\dot{n}^2}} \operatorname{erf}\left(\sqrt{\frac{q\dot{n}^2}{2}}\right) \\
 &\approx \dot{n} \left(1 - \sqrt{\frac{\pi}{2q\dot{n}^2}} \operatorname{erf}\left(\sqrt{\frac{q\dot{n}^2}{2}}\right)\right) \\
 &\approx \frac{n}{k} h(kn\phi).
 \end{aligned} \tag{4.24}$$

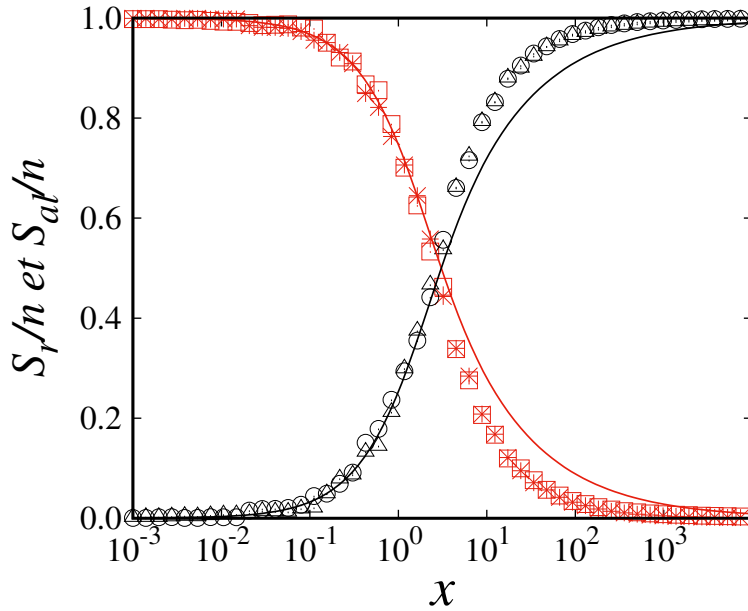


FIGURE 4.6 – Fraction des nœuds aléatoires  $S_{al}/n$  (noir) et des nœuds réguliers  $S_r/n$  (rouge) en fonction de nombre des raccourcis  $x$ . Les lignes représentent les fonction  $h(x)$  (noir) et  $1 - h(x)$  (rouge). Les symboles représentent les simulations numériques d'un réseau de taille  $n = 10^4$ , avec  $k = 1$  (étoile et cercle) et  $k = 5$  (carré et triangle). Chaque simulation est moyennée à 200 réalisations.

Si on prend en considération la transformation de GR qu'on a déjà utilisée dans nos calculs, les valeurs de  $S'_r$  et  $S'_{al}$  représentent en réalité le nombre des groupes  $k$  de nœuds, alors pour trouver le nombre réel des nœuds dans le réseau, il suffit de multiplier  $S'_r$  et  $S'_{al}$  par  $k$ , car dans chaque groupe on a  $k$  nœuds, d'où le nombre des nœuds réguliers est  $S_r = n(1 - h(kn\phi))$  et le nombre de nœuds aléatoires est  $S_{al} = nh(kn\phi)$ . Sachant que la fonction  $h(x)$  peut être

considérée comme un paramètre d'ordre et selon son expression, on peut conclure qu'il n'y a pas une transition de phase dans ce modèle mais un phénomène de croisement qui commence depuis  $x = 0$ . L'existence d'un tel paramètre d'ordre dans le modèle de NW constitue une découverte très intéressante, car il n'était pas évident de s'en apercevoir et on trouve pas de tel constat dans la littérature.

#### 4.4.2 Plus court chemin

On sait que l'expression du PCC s'écrit  $\ell = \sum_{i=1}^{\frac{n}{2}} (i \cdot n_i)$ , mais cette somme n'a pas une solution directe. Pour cette raison on va dissocier le PCC en deux types, exactement comme le cas des couches, d'où  $\ell = \ell_{al} + \ell_r$ , avec  $\ell_r$  est le PCC du réseau régulier et  $\ell_{al}$  est le PCC du réseau aléatoire.

On commence par le PCC du réseau régulier  $\ell_r$  qui s'écrit sous la forme

$$\ell_r = \frac{S'_r \int_1^{\frac{n}{2}} i \cdot n_i^r di}{\dot{n} a_r}, \quad (4.25)$$

avec  $a_r = \int_1^{\frac{n}{2}} n_i^r di = S'_r$  est la constante de normalisation, d'où  $\ell_r = \int_1^{\frac{n}{2}} i \cdot \frac{n_i^r}{\dot{n}} di$ .

De Eq. (4.14) et Eq. (4.17) on obtient

$$\ell_r = \frac{1}{\dot{n}} \int_1^{\frac{n}{2}} 2i e^{-4q \int_{j=1}^{i-1} (j-1)(q(\dot{n}-2j)+1)^{j-2} dj} di,$$

en utilisant la même approximation précédente de l'Eq. (4.21) on obtient

$$\begin{aligned} \ell_r &\approx \frac{1}{\dot{n}} \int_1^{\frac{n}{2}} 2i e^{-2qi^2} di \\ &\approx \frac{1}{\dot{n}} \frac{e^{-2q} - e^{-\frac{q\dot{n}^2}{2}}}{2q} \\ &\approx \frac{1}{\dot{n}} \frac{1 - e^{-\frac{q\dot{n}^2}{2}}}{2q} \quad \text{car : } q \ll 1 \\ &\approx \frac{\dot{n}}{4} \frac{1 - e^{-\frac{q\dot{n}^2}{2}}}{\frac{q\dot{n}^2}{2}}, \end{aligned} \quad (4.26)$$

sachant que  $\frac{q\acute{n}^2}{2} = nk\phi = x$  et  $\acute{n} = \frac{n}{k}$  alors  $\ell_r$  peut s'écrire sous la forme universelle suivante

$$\ell_r \approx \frac{n}{k} f_r(x) \quad \text{avec} \quad f_r(x) = \frac{1 - e^{-x}}{4x}. \quad (4.27)$$

Afin de calculer  $\ell_{al}$ , nous utilisons la proposition que le PCC est la position de la couche maximale (voir Section. 3.7.1). Dans notre cas ici c'est la position de la couche aléatoire maximale, de point de vue mathématique il suffit de résoudre  $\frac{\partial n'_{al}(i)}{\partial i} = 0$ . Soit  $u(i) = 4q(i-1)(q\acute{n}-2i+1)^{i-2}$ , de l'Eq. (4.14) et Eq. (4.16) on obtient que  $P_{al}(i) = u(i)e^{-\int_{j=1}^{i-1} u(j)dj}$ . D'un autre coté, il est évident que  $\ell_{al}$  prédomine dans l'expression du PCC si  $S_{al} \gg S_r$ , alors la dimension du PCC devient négligeable devant la dimension de  $\acute{n}$  car le nombre de raccourcis dans le réseau est important, on peut donc considérer que  $\acute{n} - 2i \approx \acute{n}$ , donc  $u(i) = 4q(i-1)(q\acute{n}+1)^{i-2}$ . Soit  $y = q\acute{n}$  le paramètre qui représente le degré moyen des raccourcis pour chaque nœud après la transformation de GR, car  $q$  est la probabilité pour qu'un pair de nœuds se connecte et  $\acute{n}$  est le nombre de nœuds.

De l'Eq. (4.18) et  $P_{al}(i) = u(i)e^{-\int_{j=1}^{i-1} u(j)dj}$  on obtient

$$n'_{al}(i) = \acute{n}u(i)e^{-\int_{j=1}^{i-1} u(j)dj}, \quad (4.28)$$

donc

$$\begin{aligned} \frac{\partial n'_{al}(i)}{\partial i} &= \acute{n} \frac{\partial u(i)}{\partial i} e^{-\int_{j=1}^{i-1} u(j)dj} + \acute{n}u(i) \frac{\partial e^{-\int_{j=1}^{i-1} u(j)dj}}{\partial i} \\ &= \acute{n} \frac{\partial u(i)}{\partial i} e^{-\int_{j=1}^{i-1} u(j)dj} - \acute{n}u(i)^2 e^{-\int_{j=1}^{i-1} u(j)dj}, \end{aligned} \quad (4.29)$$

alors la solution de  $\frac{\partial n'_{al}(i)}{\partial i} = 0$  est la solution de l'équation suivante

$$\frac{\partial u(i)}{\partial i} - u(i)^2 = 0. \quad (4.30)$$

On a  $\frac{\partial u(i)}{\partial i} = u(i) \left[ \frac{1}{i-1} + \ln(y+1) \right]$  et on sait que  $\ln(y+1) \gg \frac{1}{i-1}$ , car  $i$  reflète ici la valeur du PCC qui croît avec la taille du réseau  $\acute{n}$ , alors si  $\acute{n}$  croît l'expression  $\frac{1}{i-1}$  tend vers 0, par contre  $y = 2k^2\phi$  ne dépend pas de la taille du réseau, d'où on peut négliger  $\frac{1}{i-1}$  devant  $\ln(y+1)$  et on obtient  $\frac{\partial u(i)}{\partial i} = u(i)\ln(y+1)$ .

Alors l'Eq. (4.30) devient

$$u(i) = \ln(y + 1), \quad (4.31)$$

si on remplace  $u(i)$  par son expression  $u(i) = 4q(i - 1)(q\acute{n} + 1)^{i-2} = 4q(i - 1)(y + 1)^{i-2}$ , on va trouver la solution de l'Eq. (4.31) qui nous donne la position maximale de la couche-aléatoire  $i_{max}$

$$i_{max} = \frac{W\left(\frac{\ln(y+1)^2(y+1)}{4q}\right)}{\ln(y+1)} + 1, \quad (4.32)$$

avec  $W(x)$  est la fonction de Lambert.

Alors selon notre proposition, la valeur de  $\ell_{al}$  est  $i_{max}$  multipliée par la fraction des nœuds aléatoires  $\ell_{al} = i_{max}h(x)$ , d'où

$$\ell_{al} = \left(\frac{W\left(\frac{(y+1)^2(y+1)}{4q}\right)}{\ln(y+1)} + 1\right)h(x). \quad (4.33)$$

De l'Eq. (4.27) et Eq. (4.33) on obtient l'expression globale du PCC

$$\ell = \left(\frac{W\left(\frac{\ln(y+1)^2(y+1)}{4q}\right)}{\ln(y+1)} + 1\right)h(x) + \acute{n}\frac{1 - e^{-x}}{4x} \quad (4.34)$$

pour trouver une expression universelle de  $\ell$  en fonction du nombre des raccourcis  $x$ , il faut considérer que  $y \ll 1$ , d'où

$$\ell = \acute{n}f(x) \quad (4.35)$$

avec  $f(x) = \left(\frac{2W(\frac{x}{2})h(x)+1-e^{-x}}{4x}\right)$  est la nouvelle fonction universelle.

De la Fig. 4.7, on voit que l'Eq. (4.34) est en accord avec les simulations sauf dans l'intervalle  $1 < x < 20$  où il y a une certaine déviation, on explique cela par l'existence des fluctuations élevées dans cette intervalle, car il est évident que lorsque les fluctuations deviennent importantes l'approximation du champs moyen n'est plus valable, comme le cas usuel dans tous les systèmes physiques au voisinage des points critiques. Cette explication est démontrée par les simulations en utilisant le paramètre d'ordre,  $\frac{S_{al}}{n}$ . Soit  $\sigma = \frac{\sqrt{\langle \frac{S_{al}}{n} \rangle^2 - \langle \frac{S_{al}}{n} \rangle^2}}{\langle \frac{S_{al}}{n} \rangle}$  les fluctuations



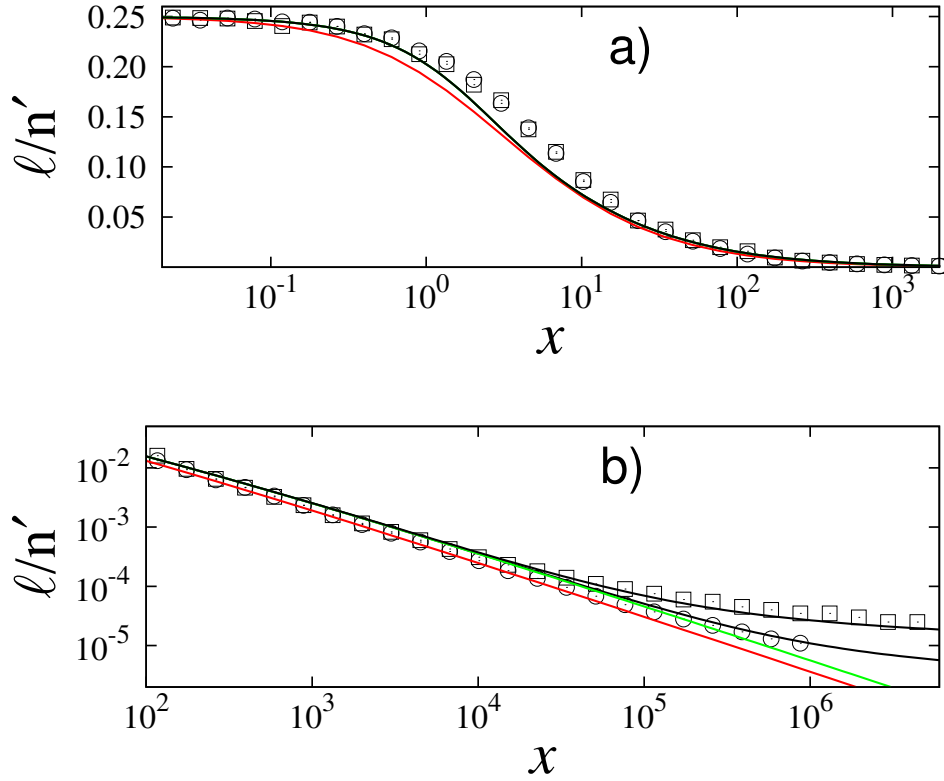


FIGURE 4.7 – L’expression universelle  $f(x) = \ell/\acute{n}$  en fonction de  $x$  pour un réseau de taille  $n = 10^6$ . La formule de Newman et watts Eq. (4.6) (ligne rouge), Eq. (4.35) (ligne vert), Eq. (4.34) (ligne noir) et les simulations numériques pour  $k = 1$  (cercle) et  $k = 5$  (carré). L’échelle est linéaire-log dans **a)** et log-log dans **b)**. Chaque simulation est moyennée à 100 réalisations.

de  $\frac{S_{al}}{n}$  par rapport à sa taille, en effet selon la Fig. 4.8 le maximum des fluctuations est dans l’intervalle  $1 < x < 20$  où l’Eq. (4.34) montre une petite déviation.

## 4.5 Émergence spectaculaire vers la propriété petit-monde

### 4.5.1 Invalidité de l’ancienne fonction universelle et la saturation des raccourcis

Selon l’Eq. (4.34), l’expression universelle  $\ell = \acute{n}f(x)$  n’est plus valable si la condition  $y \ll 1$  n’est pas vérifiée, que l’on peut aussi montrer par les simulations numériques. Soit  $\Delta$

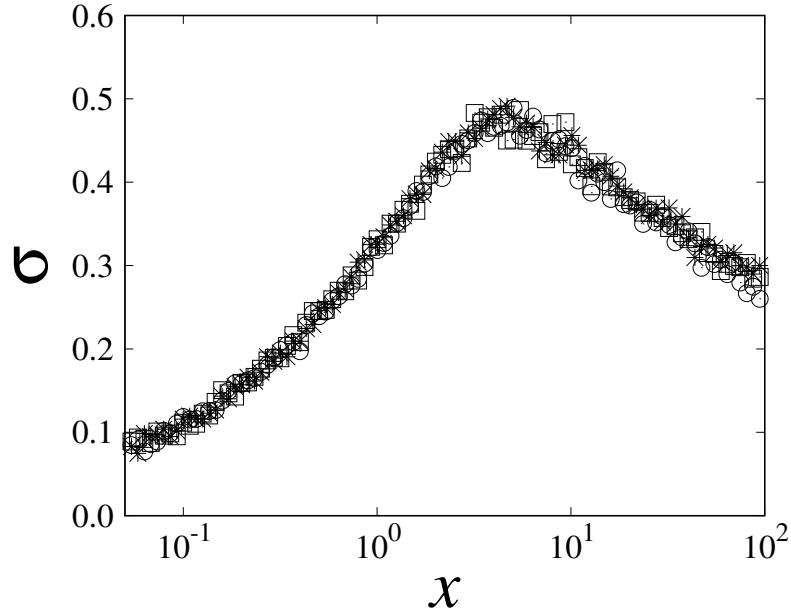


FIGURE 4.8 – Fluctuation  $\sigma$  en fonction de  $x$  d'un réseau de taille  $n = 10^4$  pour différents valeurs de  $k$ ,  $k = 1$  (étoile),  $k = 2$  (carré) et  $k = 5$  (cercle), le nombre de réalisations est 1000.

l'écart de  $\frac{\ell}{\tilde{n}}$  en fonction de  $y$  entre un réseau de  $n$ ,  $k$  et  $\phi$  donné et d'autres réseaux possédant des valeurs aléatoires  $n_{al}$ ,  $k_{al}$  et  $\phi_{al}$  sous la condition que  $nk\phi = n_{al}k_{al}\phi_{al}$ . L'expression de l'écart de  $\frac{\ell}{\tilde{n}}$  peut s'écrire comme

$$\Delta = \frac{\frac{\ell}{\tilde{n}} - \frac{\ell_{al}}{n_{al}}}{\frac{\ell}{\tilde{n}} + \frac{\ell_{al}}{n_{al}}}. \quad (4.36)$$

De la Fig. 4.9.(a) on observe une augmentation remarquable de  $\Delta$  lorsque  $y$  n'est pas très inférieur à 1, par intuition on peut expliquer ce phénomène par le début d'une saturation des raccourcis lorsque  $y$  augmente, car après la saturation le réseau change son comportement devant les raccourcis ajoutés. Afin de montrer l'existence d'un changement de comportement vis-à-vis des raccourcis ajoutés, on va simuler numériquement le nombre de raccourcis existants dans le trajet du PCC entre deux nœuds choisis au hasard par rapport au nombre moyen des raccourcis dans le réseau,  $\rho$ , en fonction de  $y$ . En effet, selon les Fig. 4.9.(a) et Fig. 4.9.(b), on remarque une saturation des raccourcis qui varie d'une façon similaire avec  $\Delta$  en fonction de  $y$ . Pour  $y \ll 1$  on voit que  $\rho \approx 1$ , alors presque tous les raccourcis existent dans les trajets du PCC, c'est-à-dire l'impact des raccourcis sur le PCC est maximale, en revanche avec l'augmentation de  $y$  le pourcentage des raccourcis existants dans les trajets diminue,

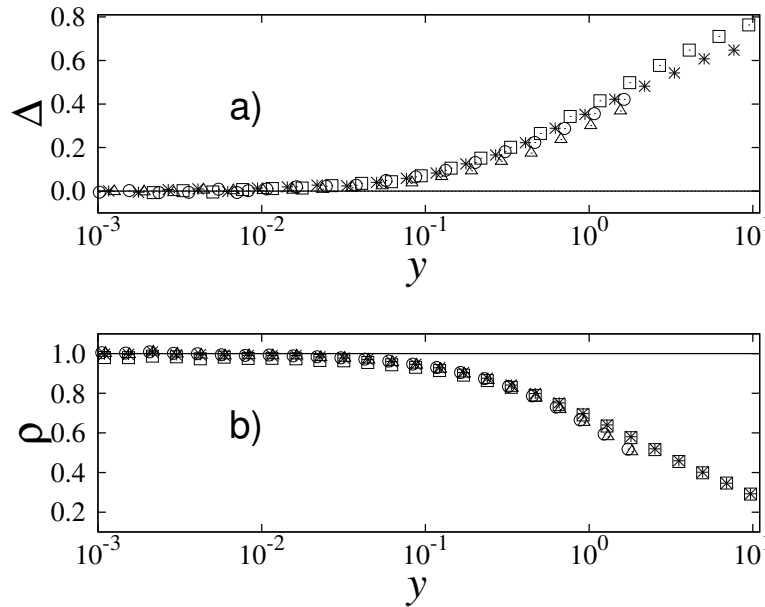


FIGURE 4.9 – La valeur de  $\Delta$  et  $\rho$  en fonction de  $y$  pour différentes valeurs de  $n$  et  $k$  :  $n = 10^{-3}$  et  $k = 1$  (triangle),  $n = 10^{-3}$  et  $k = 4$  (carré),  $n = 10^{-4}$  et  $k = 1$  (cercle),  $n = 10^{-4}$  et  $k = 4$  (étoile). Dans **a)** le nombre de réalisations pour chaque simulation est 1000 et dans **b)** le nombre de réalisations pour chaque simulation est 5000.

c'est-à-dire l'impact des raccourcis sur le PCC diminue. On conclut que la saturation des raccourcis ajoutés explique bien la non validité de la fonction universelle  $f(x)$  pour  $y \ll 1$ .

### 4.5.2 La nouvelle fonction universelle

La saturation des raccourcis peut signifier que le réseau commence à posséder la propriété petit-monde, autrement dit le PCC devient de l'ordre de  $\ln(n)$ , pour confirmer cette proposition on réalise des simulations numériques et on les compare avec notre équation. En effet, de la Fig. 4.10 on voit que le rapport  $\frac{\ln(n)}{\ell}$  émerge et devient important d'une façon universelle en fonction du même paramètre  $y$  avec lequel la saturation émerge. Alors une nouvelle équation universelle  $g$  sous la forme  $g(y) = \frac{\ln(n)}{\ell}$  apparaît.

Maintenant on essaye de formuler l'expression analytique de cette nouvelle équation universelle  $g$ , de l'Eq. (4.34) on peut trouver facilement que dans la limite de grande taille du

système on va obtenir que

$$\frac{\ln(n)}{\ell} = \ln(y + 1). \quad (4.37)$$

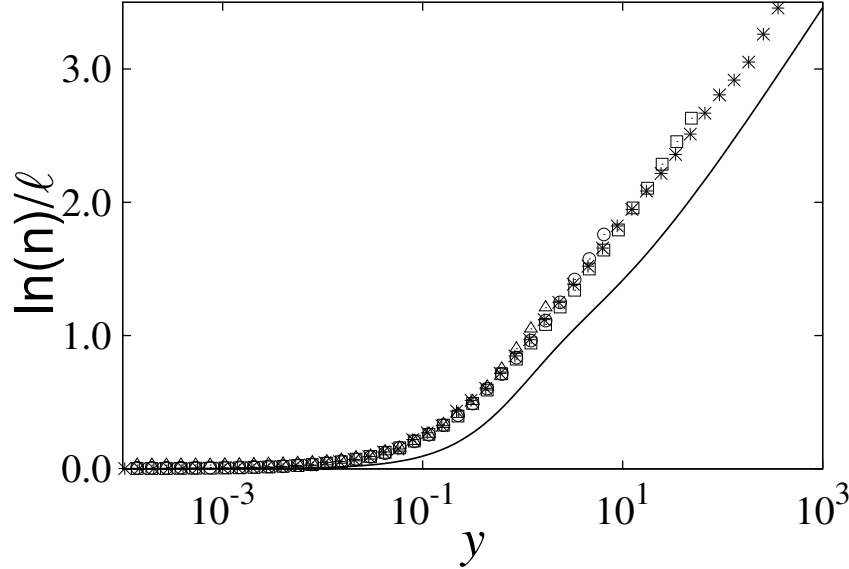


FIGURE 4.10 – L’expression universelle de PCC en fonction de  $y$ , par Eq. (4.39) (ligne continue) et par les simulations numériques pour différents valeurs de  $n$  et  $k$ ,  $n = 10^{-6}$  et  $k = 20$  (étoile),  $n = 10^{-5}$  et  $k = 5$  (carré),  $n = 10^{-4}$  et  $k = 2$  (cercle),  $n = 10^{-3}$  et  $k = 1$  (triangle), nombre de réalisations pour chaque simulations est 100. En **a)** l’ échelle est linéaire-log et en **b)** l’ échelle est log-log.

Il est évident que l’Eq. (4.37) n’est pas une expression exacte à cause de la transformation de GR que nous avons utilisée au début lorsque nous avons supposé que chaque  $k$  voisins, on le dénomme ici **amas**, est comme un seul nœud.

Sur le trajet du PCC entre deux amas quelconques dans le réseau on trouve des liens ordinaires et des raccourcis, dans le cas où il y a deux raccourcis successifs sur le trajet, c’est-à-dire on a deux raccourcis liés au même **amas**, la probabilité que les deux raccourcis soient liés au même nœud dans cette amas est  $\frac{1}{k}$  et la probabilité que les deux raccourcis ne soient pas liés au même nœud, c’est-à-dire on aura un lien en plus, est  $1 - \frac{1}{k}$ , autrement dit chaque fois qu’on a deux raccourcis successifs il y a une probabilité de  $1 - \frac{1}{k}$  que le PCC augment par 1.

D’autre part, le nombre de liens réguliers entre les **amas** est toujours  $\hat{n}$  et le nombre de

raccourcis entre les **amas** est  $\frac{q\dot{n}(\dot{n}-1)}{2} \approx \frac{q\dot{n}^2}{2}$ ,  $q$  est la probabilité qu'un pair d'amas se connecte et  $\frac{\dot{n}(\dot{n}-1)}{2}$  est le nombre des paires possibles. Alors la fraction des raccourcis par rapport au nombre total des liens est  $\frac{q\dot{n}}{2+q\dot{n}} = \frac{y}{2+y}$ , d'où la probabilité d'obtenir deux raccourcis liés au même **amas** dans le trajet entre deux nœuds quelconques est  $(\frac{y}{2+y})^2$ . Par conséquent la probabilité pour que la valeur du PCC soit doublée est  $(\frac{y}{2+y})^2(1 - \frac{1}{k})$ , où  $(\frac{y}{2+y})^2$  est la probabilité d'avoir deux raccourcis liés au même **amas**, et  $(1 - \frac{1}{k})$  est la probabilité que ces deux amas soient liés à deux nœuds différents. La probabilité que le PCC reste inchangé est  $(1 - (\frac{y}{2+y})^2) + \frac{1}{k}(\frac{y}{2+y})^2$ , le première terme représente la probabilité que deux raccourcis ne soient pas liés au même **amas** et le deuxième représente la probabilité que deux raccourcis soient liés au même **amas** et au même nœud  $\frac{1}{k}(\frac{y}{2+y})^2$ , d'où l'Eq. (4.37) devient

$$\frac{\ln(n)}{l} = \frac{\ln(y+1)}{\frac{1}{k}(\frac{y}{2+y})^2 + (1 - (\frac{y}{2+y})^2) + 2(1 - \frac{1}{k})(\frac{y}{2+y})^2} \quad (4.38)$$

pour  $k \gg 1$  on obtient que

$$\frac{\ln(n)}{\ell} = \frac{\ln(y+1)}{(1 + (\frac{y}{2+y})^2)} \quad (4.39)$$

Cette fonction montre un certain écart par rapport aux résultats numériques, car dans notre calcul nous n'avons pas pris en compte tous les cas possibles des raccourcis entre les **amas** mais nous avons fait juste une approximation de champs moyen. Cependant notre expression de  $g(y) = \frac{\ln(y+1)}{(1 + (\frac{y}{2+y})^2)}$  nous donne l'allure de la variation de la fonction universelle, qui varie d'une façon logarithmique.

Le fait que nous avons réussi à contrôler le nombre de raccourcis ajoutés pour atteindre la phase petit-monde dans le modèle de NW est un résultat intéressant et cette contribution devrait renforcer l'intérêt de ce modèle. La difficulté qu'a empêché à répondre à ce problème au cours des vingt dernières années est de trouver le paramètre avec lequel on détermine le point où le réseau est saturé, alors par un raisonnement purement théorique on a trouvé que ce paramètre est  $y$ .

## 4.6 Conclusion

En utilisant la transformation de GR sur le modèle NW, nous avons trouvé une expression analytique du PCC plus précise que l'ancienne formule, et à partir de cette nouvelle expression, nous avons conclu que le paramètre  $y$  est la variable avec laquelle la propriété petit-monde émerge, la variable avec laquelle on détermine l'erreur et l'invalidité de l'ancienne fonction universelle  $f(x) = \frac{\ell}{\tilde{n}}$ , la variable qui nous donne le degré de saturation de raccourcis, ainsi que la variable de la nouvelle fonction universelle  $g(y) = \frac{\ln(n)}{\ell}$ . Nous en déduisons que l'émergence qui se passe en fonction de ce paramètre  $y$  dans le modèle NW est une émergence spectaculaire.

# Chapitre 5

## Une nouvelle approche pour prédire les transitions de phases dans le phénomène de percolation

La détermination du type de transition de phase dans le phénomène de percolation n'est pas encore devenue une science exacte, car on manque d'une règle qui peut déterminer quelle transition sera de premier ou de second ordre. Cependant la détermination se fait au cours de la transition ou après, le but de ce chapitre est d'apporter une contribution concernant ce problème qui est un sujet très important dans le domaine de la physique fondamentale.

### 5.1 Introduction

Le but ultime de l'étude des réseaux est de mieux comprendre le comportement des réseaux des systèmes. Par exemple, nous étudions la structure d'Internet pour mieux comprendre la manière avec laquelle le trafic Internet circule, la façon du fonctionnement des protocoles de communication, ou comment organiser le réseau afin de le rendre plus performant. Nous étudions les réseaux biochimiques comme les réseaux métaboliques car nous espérons qu'ils permettent de comprendre les processus chimiques complexes qui se déroulent dans la cellule, ou des outils algorithmiques pouvant nous aider à obtenir des informations biologiques de grands volumes de données générées par les techniques de laboratoire modernes[146]. Nous

études le réseau causal, qui représente la structure à grande échelle de l'espace-temps dans notre univers accéléré, pour comprendre pourquoi ce réseau est aussi un graphe de loi libre-échelle avec un regroupement fort [157].

Notre but dans ce chapitre, à part le rappel théorique, est de mieux comprendre le comportement des transitions de phase dans les réseaux réels à partir des approches dans le domaine des réseaux complexes. En bref, nous allons établir une formule pour prévoir le type de transition de phase au voisinage du point critique.

### 5.2 Transition de phase

Comme dans la fusion de la glace, ferro-magnétisme, la supraconductivité et le repliement des protéines, la nature contient quotidiennement des transitions de phase : processus dans lesquels les systèmes changent radicalement certaines propriétés physiques lorsqu'une variation minimale des variables se produisent.

En tenant compte de la grande présence des transitions de phase dans le monde réel, ce n'est pas étonnant que les physiciens aient commencé, depuis le tout début de la mécanique statistique, à découvrir des comportements universels et critiques (scaling à proximité des points de transition), ainsi que de faire un grand effort pour la classification des transitions de phase. La première tentative à fournir une classification a été réalisée par Ehrenfest [1], mais les schémas de classification les plus modernes regroupent les transitions de phase en deux grandes catégories : de premier et de second ordre.

En thermodynamique, les transitions de phase du premier ordre sont celles qui impliquent une chaleur latente, c'est-à-dire au cours d'elles le système absorbe ou libère une quantité typiquement élevée d'énergie par volume. Normalement, elles sont caractérisées (au point de transition) par "un régime en phase mixte", dans lequel certaines parties du système ont subi la transition et d'autres non, en présence ainsi d'une sorte de coexistence des phases des deux régimes. Les signes de telle transition sont : le comportement abrupt et discontinu du paramètre d'ordre à proximité du point de transition, ainsi que l'irréversibilité intrinsèque de la transition, avec la présence (dans la majorité des cas) de boucles hystérèses. En variante, les transitions de phase de second ordre (également appelées transitions de phase continues) sont



réversibles et correspondent à des paramètres d'ordre présentant un comportement continu au voisinage du point critique. Parmi les transitions du premier ordre nous citons la fusion de la glace, l'ébullition de l'eau, et les transitions de super-refroidissement et de surchauffe. Pour le second ordre nous avons la transition ferromagnétique, la supraconductivité et la transition super-fluide.

### 5.3 Percolation

La percolation est la transition vers la connectivité à grande échelle des réseaux suite à l'ajout progressif de liens, elle se produit au cours des processus de croissance et d'évolution dans une grande variété de systèmes naturels, technologiques et sociaux [158, 159, 160, 161, 162, 163, 164]. La percolation se produit aussi en physique dans les solides atomiques et moléculaires, ainsi que dans les réseaux sociaux, biologiques et artificiels [159, 165, 166]. Autrement dit, on peut dire qu'un système de percolation consiste à une grille, dans laquelle les nœuds ou les liens sont retirés avec une certaine probabilité  $1 - p$ , ou on considère que les nœuds se lient avec une probabilité  $p$  [167, 168], le système se déconnecte en petits groupes, c'est-à-dire qu'il devient impossible de passer d'un côté de la grille à l'autre en suivant les liens conducteurs, mais au-dessus de  $p_c$ , une grappe s'étend, semblable à la composante géante dans les graphiques aléatoires (voir Fig. 5.1).

### 5.4 Percolation explosive

En percolation de liaison, on commence de  $n$  nœuds non connectés. A chaque pas de temps, un lien est ajouté entre deux nœuds sélectionnés selon une règle donnée. Le nombre des liens ( $E$ ) ajoutés au système à un certain pas de temps divisé par la taille du système est le paramètre de contrôle ( $\langle k \rangle = \frac{2E}{n}$ ) qui décrit la transition de phase. En ce qui concerne le paramètre d'ordre, on peut prendre la fraction des nœuds appartenant à la composante géante du réseau (généralement noté  $S(t)$ ). Au fur et à mesure que le temps augmente, le nombre de liens ajoutés dans le réseau augmente, dans la limite thermodynamique ( $N \rightarrow \infty$ ),  $S(t)$  présente une transition de phase de zéro à  $O(1)$  au voisinage d'un point critique ( $\langle k \rangle_c$ ). La

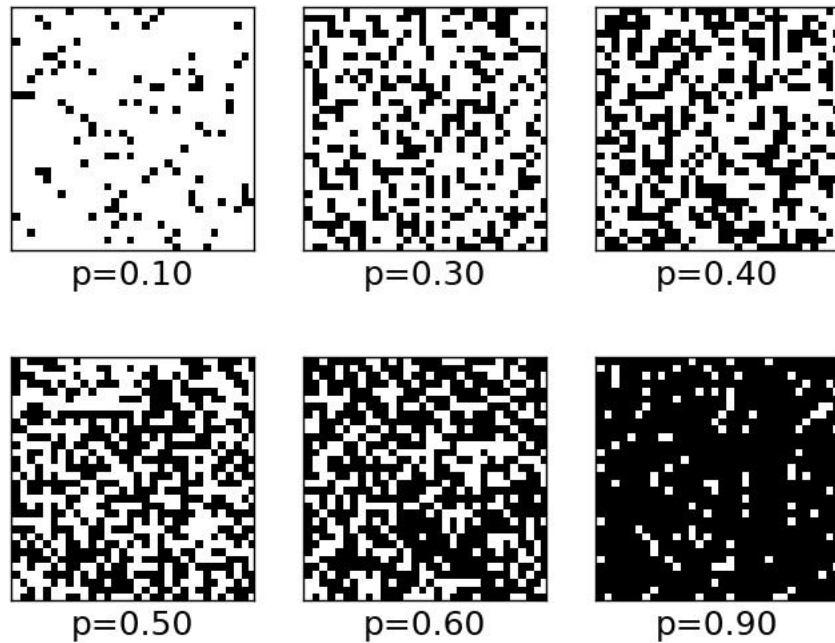


FIGURE 5.1 – Une démonstration de percolation sur une grille à deux dimensions pour différentes valeurs de  $p$ . Sous le seuil de percolation le système est composé de petites grappes, après un certain point critique  $p_c$  une grappe étendue et occupe la grille.

percolation classique est une transition de phase géométrique typique, qui a été largement étudiée depuis 1940 dans de nombreux domaines, y compris les mathématiques, la physique statistique et l'ingénierie. De nombreux modèles de percolation, tels que la percolation d'invasion, de premier passage, bootstrap et k-core. Tous ces modèles et beaucoup d'autres ont démontré que la transition de phases est étroitement liée à diverses applications, la turbulence, les modèles magnétiques, les colloïdes, la transition de Hall quantique spin, etc. Le lecteur est adressé aux revues récentes dans [169, 170, 171], ses références contiennent un ensemble des principaux résultats et concepts de la percolation classique.

Pendant des décennies, la plupart des transitions de la percolation classique (ordinaire) semblent être de second ordre (continue) comme la percolation de liaison dans les réseaux ER, pour laquelle la taille de la composante géante augmente progressivement lorsque le paramètre de contrôle,  $\langle k \rangle$ , dépasse le seuil de percolation. Il convient toutefois de noter qu'il existe également de rares exemples de modèles de percolation présentant une transition de phase de premier ordre (discontinue), tels que la percolation bootstrap [172], k-core [173] et de

brouillage [174]. Par l'observation que les règles de concurrence non locales ou globales lors des processus de fusion d'amas dans le réseau peuvent conduire à une percolation explosive PE, un certain nombre de ces processus, tels que ceux d'Achlioptas ont été étudiés [175, 176, 177].

## 5.5 Percolation dans le réseau aléatoire d'Erdős-Rényi

Un réseau aléatoire est créé en spécifiant que chaque paire de nœuds est connecté par un lien avec une probabilité uniforme  $p$ . Ce type de réseaux a été étudié d'un point de vue mathématique pure par Erdős et Rényi [90, 91, 92], il est souvent désigné par son nom mathématique  $G(n, m)$ , avec  $n$  est le nombre de nœuds et  $m$  le nombre de liens. Dans la cas où  $n$  est très grand, bon nombre de propriétés de l'ensemble des réseaux aléatoires ont été exprimées analytiquement de façon parfaite et élégante.

Les graphes aléatoires ER sont les mieux étudiés parmi les modèles de graphes, leurs propriétés structurelles varient en fonction de  $p$  montrant notamment un changement radical à une probabilité critique  $p_c = \frac{1}{n}$ , correspondant à un degré moyen critique  $\langle k \rangle_c = 1$ . Erdős et Rényi ont prouvé que :

- Si  $p < p_c$ , lorsque  $n$  tend vers l'infini, le graphe n'a quasiment pas de composante de taille supérieure à  $(\ln(n))$ .
- Si  $p = p_c$ , la composante la plus importante a quasiment la taille de  $n^{\frac{2}{3}}$ .
- Si  $p > p_c$ , le graphe a une composante de taille de l'ordre de  $n$  et aucune autre composante n'a de taille supérieure à  $(\ln(n))$ .

La transition au  $p_c$  présente les caractéristiques typiques d'une transition de phase de deuxième ordre. En particulier, si on considère la taille de la composante la plus importante comme paramètre d'ordre, la transition tombe dans la même classe d'universalité que celle des transitions de percolation de champ moyen (voir Fig. 5.2). Erdős et Rényi ont étudié la distribution des degrés minimum et maximum dans le graphe aléatoire, mais la distribution des degrés complet a été obtenue plus tard par Bollobás [178]. La probabilité qu'un nœud

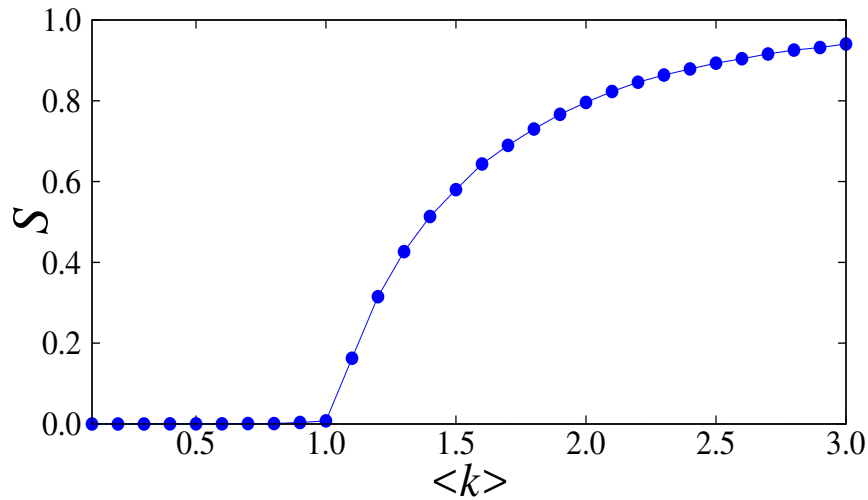


FIGURE 5.2 – Les simulations numériques du seuil de percolation dans le modèle ER en fonction du degré moyen, pour un réseau de taille  $n = 10^5$ . On observe que le point où la fraction de la composante géante,  $S$ , émerge est à  $\langle k \rangle = 1$ .

possédant le degré  $k$  est la distribution binomiale :

$$P(k) = C_{n-1}^k p^k (1-p)^{n-1-k} \quad (5.1)$$

Pour  $n$  très grand, la répartition du degré est bien approchée par une distribution de Poisson :

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \quad (5.2)$$

Le coefficient de regroupement,  $C$ , est une quantité très simple à calculer pour le graphe aléatoire de Poisson, rappelons qu'il est défini comme la probabilité que deux voisins d'un nœud du réseau soient également voisins les uns des autres. Dans un graphe aléatoire, la probabilité que les deux nœuds soient voisins est égale à  $p = \frac{\langle k \rangle}{(n-1)}$ . Par conséquent :

$$C = \frac{\langle k \rangle}{(n-1)} \quad (5.3)$$

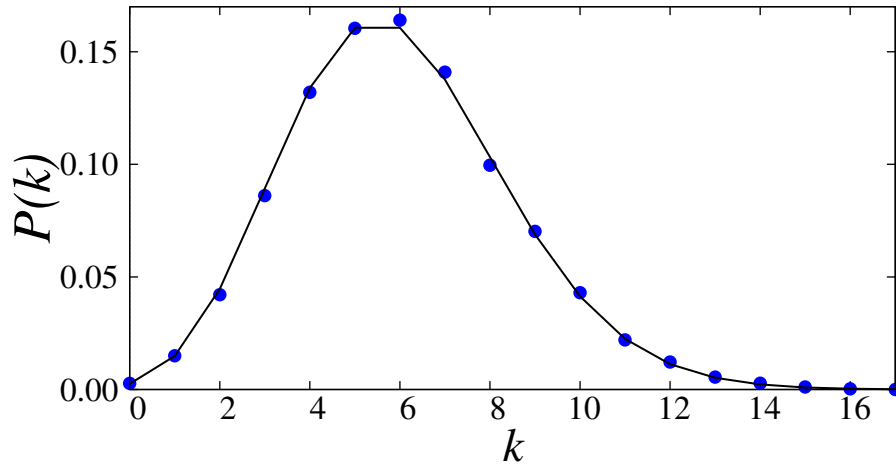


FIGURE 5.3 – Illustration de la distribution des degrés d’un réseau aléatoire ER de degré moyen égale 6, les cercles représentent les simulations numériques et la ligne noire représente Eq. (5.2).

Cette valeur de Clustering qui tend vers 0 quand  $n$  tend vers l’infini est l’un des nombreux aspects dans lesquels le graphe aléatoire diffère fortement de la plupart des réseaux réels, dont beaucoup ont des coefficients de regroupement assez élevés.

### 5.5.1 Percolation explosive dans le processus d’Achlioptas

L’idée principale du processus Achlioptas (PA) [179] est de modifier la règle pour générer des graphes ER. Au lieu d’ajouter des liens aléatoires, dans le processus Achlioptas on choisit deux liens au hasard puis on utilise une règle pour sélectionner l’un ou l’autre. Le lien sélectionné sera ajouté au graphe alors que l’autre lien sera rejeté. Considérons la Fig. 5.4 (adapté de [179]) : **(A)** l’évolution du réseau selon le modèle ER, à chaque étape deux nœuds sont choisis au hasard et reliés par un lien (représenté par la ligne pointillée). Dans cet exemple, deux composantes de taille 7 et 2 sont fusionnées. **(B)** Deux liens aléatoires  $\{e_1, e_2\}$  sont sélectionnés à chaque étape mais un seul sera ajouté au réseau suivant une règle de sélection, tandis que l’autre sera abandonné. Soumettant à la règle de produit (RP), le lien sélectionné est celui qui minimise le produit des tailles des composantes qui sont fusionnées. Pour cet

exemple,  $e_1$  (avec le produit  $2 \times 7 = 14$ ) sera choisi et  $e_2$  rejeté (parce que  $4 \times 4 = 16$ ). En revanche, la règle avec laquelle on sélectionne le lien minimisant la somme des tailles des composantes choisi  $e_2$  au lieu de  $e_1$ . (C) Évolution de fraction de la composante géante  $\frac{C}{n}$  pour : ER, BF (une règle de taille bornée, voir [180]), et RP. Le réseau étudié est de taille  $n = 512000$ .

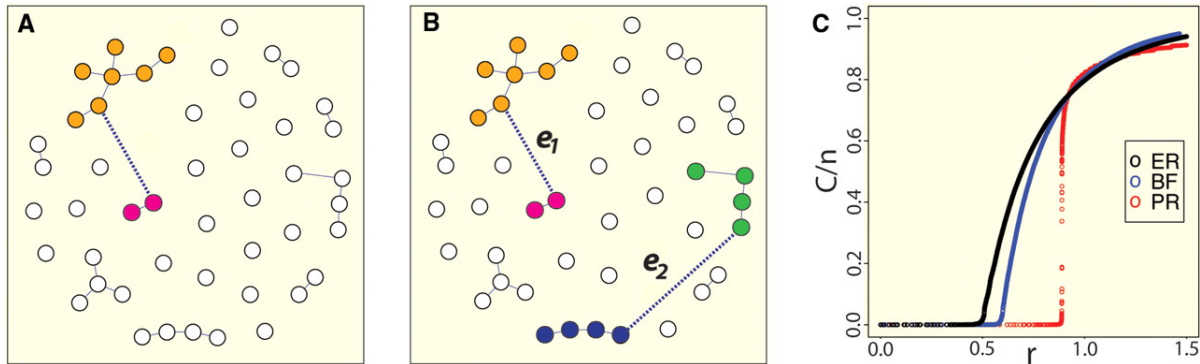


FIGURE 5.4 – Comparaison entre la percolation aléatoire et PA. (A) Percolation ER classique, où les liens sont ajoutés au hasard dans le réseau. (B) Percolation PA, où à chaque étape deux liens sont en compétition pour être établis. (C) Paramètre d'ordre : la taille relative de la composante géante par rapport au nombre des liens ajoutés normalisés par la taille du système.

Achlioptas s'est demandé s'il était possible de déplacer le point critique de cette transition de phase suivant une règle de sélection appropriée. Une règle qui peut naturellement être imaginée c'est la règle du produit : Parmi les liens donnés, choisir celui qui minimise le produit des tailles des composantes contenant les quatre extrémités de  $\{e_1, e_2\}$ . Cette règle a été suggérée dans [181] comme la plus susceptible de retarder le point critique. Une autre règle est la règle de somme, où la taille de la nouvelle composante formée est minimisée. La Fig. 5.4.(c) est suffisante pour motiver l'idée pour que la transition de percolation RP soit discontinue. Achlioptas et al. ont utilisé une méthode intéressante pour vérifier que la transition est bien discontinue. Ils mesurent numériquement le pas de temps  $t_0$  auquel la taille de la composante géante ( $C$ ) devient plus grande que  $n^{\frac{1}{2}}$  et  $t_1$  où  $C$  devient plus grand que  $\frac{1}{2} \times n$ . Un pas de temps correspond à une valeur de  $r$  où  $r = \frac{E}{n} = \frac{t}{n}$  (puisque le nombre de liens à l'instant  $t$  est  $E$ ). Pour le modèle ER et d'autres transitions de percolation continues, on trouve que  $\Delta \equiv t_1 - t_0$  est extensive, c'est-à-dire il évolue avec  $n$ . En outre, pour le modèle RP, ils ont trouvé que  $\Delta \sim n^{\frac{2}{3}}$ . Puisque  $\Delta$  est sous-linéaire en  $n$ , c'est un argument numérique qui

montre que dans la limite thermodynamique  $n \rightarrow \infty$  la transition de percolation est en effet discontinue.

Cependant, cette conjecture a été rapidement et rigoureusement désapprouvée par Riordan et Warnke [182, 183], en montrant qu’il ne s’agit pas d’une transition discontinue mais continue. En effet, leur argument montre des transitions de phase continues pour une classe encore plus grande de processus PA [183]. Leurs résultats indiquent que la continuité de la transition de phase est une caractéristique assez robuste et donc tous les processus Achlioptas ont une transition de phase continue. En revanche, Nagler et ses collègues [184] ont montré qu’il est également vrai que certains processus de percolation basés sur le choix d’un nombre fixe de sommets aléatoires sont discontinus, et ils ont résolu ce paradoxe apparent. En identifiant et analysant un processus continu dans le sens défini par Riordan et Warnke [183] mais qui présente encore un nombre infini de sauts discontinus dans un voisinage arbitraire du point de transition : un escalier du Diable. Ils ont démontré analytiquement que la continuité à la première transition de connectivité et la discontinuité du processus de percolation sont compatibles pour certains systèmes de percolations compétitifs.

La percolation explosive est un problème scientifique très intéressant dans des champs de recherche plus larges que la percolation. Il détient un potentiel considérable pour faire progresser notre connaissance des transitions de phase et des phénomènes critiques. Le fait qu’une légère modification des règles de sélection classiques puisse modifier de manière significative la classe d’universalité de la transition de phase sous-jacente est impressionnant. De plus, il semble que dans des cas extrêmes, l’ordre de transition lui-même puisse également changer.

### 5.5.2 Percolation explosive dans d’autres modèles

Au-delà de la percolation ordinaire, de nombreux autres modèles ont été développés en se basant sur diverses contraintes concernant l’occupation des liens afin de produire PE. Dans ce qui suit, nous examinons brièvement certains de ces modèles selon la référence [169] :

### Les modèles de probabilité

Ce modèle consiste à donner pour tout amas  $i$  un poids  $s_i$ , pour chaque pas de temps un lien entre une paire d'amas  $ij$  est occupé par une probabilité proportionnelle à  $(s_i s_j)^\alpha$ . La somme de tous les poids est déterminée comme une constante de normalisation :  $w = \sum_i \sum_j (s_i s_j)^\alpha$ , alors la probabilité pour que l'amas  $i$  soit lié à un autre amas  $j$  est  $\frac{(s_i s_j)^\alpha}{w}$  [185].

Un autre exemple de ces modèles est le modèle hamiltonien de la référence [186], on commence par un réseau de  $n$  nœuds sans aucun lien, de sorte que chaque nœud appartient initialement à un amas différent. Tout d'abord, un hamiltonien  $H$  simple est défini en fonction des tailles des grappes et du nombre de liens ajoutés à ces grappes. Ensuite, un nouveau lien  $e$  entre toute paire de nœuds non encore connectés est ajouté, avec une probabilité proportionnelle à  $e^{(-\beta \Delta H_e)}$ , où  $\Delta H_e$  est le changement d'énergie après l'ajout d'un lien.

### Les modèles Hybrides

Dans les modèles hybrides la composition des arêtes suit à la fois des règles de percolation ordinaire et des règles concurrentielles : à chaque pas de temps, un lien peut être choisi aléatoirement avec probabilité  $p$ , alors qu'il peut aussi être choisi suivant une règle concurrentielle donnée par la probabilité  $1 - p$ . Par exemple, dans [187] RP est utilisée, dans la référence [188] un mélange de  $k = 2$  et  $k = 3$ -core sur le réseau aléatoire, un mélange de  $k = 3$ -core et des modèles de percolation de brouillage dans le réseau 2D sont étudiés. Les modèles hybrides ont été analysés dans des réseaux 2D [188, 187] et des réseaux ER [188, 187, 189].

### Les modèles bootstrap

Dans le cadre du PE, certains modèles bootstrap ont été proposés [190]. Le processus standard de percolation bootstrap sur un treillis en supposant que les sites sont actifs ou inactifs et que l'état d'un site dépend de ses voisins. Initialement, chaque site est actif avec une probabilité  $p$ , et inactif avec probabilité  $1 - p$ . Chaque site activé reste dans son état, alors que chaque site inactif peut devenir actif (et rester actif pour toujours) si ses  $k$  voisins les plus proches sont actifs (avec  $k = 2, 3, \dots$ ). La procédure est poursuivie jusqu'à ce que le système atteigne la configuration stable. Dans l'état final du processus, on s'intéresse à savoir



s'il existe une composante géante occupée par des sites actifs. La percolation bootstrap peut montrer des transitions continues ou discontinues, selon le type du réseau et sa dimension.

## 5.6 Percolation au voisinage du point critique : approche théorique et simulation

### 5.6.1 Introduction

La détermination du type de la transition de phase dans le phénomène de percolation n'est pas encore devenue une science exacte, c'est-à-dire, il manque une règle avec laquelle on peut déterminer qu'une telle transition est de premier ou de second ordre indépendamment des mécanismes locaux d'évolution du système ou modèle étudié, même si la percolation est l'un des modèles les plus simples et les plus anciens dans la physique statistique. Cependant pour quelques modèles spécifiques, par exemple dans [191], une approche de l'équation du taux a été utilisée pour déterminer le type de transition de phase de la percolation, comme exemple dans l'évolution des réseaux ER, l'équation différentielle est écrite dans la limite thermodynamique par :

$$\frac{\partial n_s(t)}{\partial t} = \sum_{i+j=s} \frac{k_i n_i}{c(t)} \frac{k_j n_j}{c(t)} - 2 \frac{k_s n_s}{c(t)}, \quad (5.4)$$

où  $c(t) = \sum_s k_s n_s$  est la constante de normalisation,  $\frac{k_i n_i}{c(t)} \frac{k_j n_j}{c(t)}$  est la probabilité qu'un amas de taille  $i$  et de poids  $k_i$  sera connecté à un amas de taille  $j$  et de poids  $k_j$  et  $2 \frac{k_s n_s}{c(t)}$  est la probabilité qu'un amas de taille  $s$  et de poids  $k_s$  sera connecté à n'importe quel amas dans le réseau.

Dans [185, 192] les auteurs ont généralisé l'Eq. (5.4) en prenant  $k_i = i^\omega$ , le cas  $\omega = 1$  est exactement le cas ER. En utilisant la technique de fonction génératrice, ils ont trouvé l'expression de distribution d'amas au point critique,  $n_s(t_c) \sim s^{-\gamma}$ , où  $\gamma$  est déterminé comme :

$$\gamma = \begin{cases} 1 + 2\omega & \text{si } 0 < \omega < \frac{1}{2} \\ \frac{1}{2} + \omega & \text{si } \frac{1}{2} < \omega < 1. \end{cases} \quad (5.5)$$

En se basant sur des simulations numériques, on a trouvé que pour  $\frac{1}{2} < \omega < 1$  la transition de phase est continue et pour  $0 < \omega < \frac{1}{2}$  la transition est discontinue. On peut déterminer par simulation le type de transition de percolation dans ce modèle en mesurant l'exposant  $\omega$  en termes de processus d'agrégation de cluster. En 1983, un modèle unidimensionnel de percolation a été introduit [193] dans lequel les sites  $i$  et  $j$  sont reliés par la probabilité  $p_{ij} = \frac{p}{|i-j|^s}$ , où  $p$  est un paramètre défini dans l'intervalle  $0 \leq p \leq 1$ , et  $s$  est aussi un paramètre. Dans la référence [194] les auteurs ont prouvé que la transition est discontinue pour  $1 < s \leq 2$  et ils ont fait d'autres progrès notables associés à la percolation à longue distance, pour plus de détails voir [195].

Dans ces modèles la caractérisation des transitions de phase dépend de certains paramètres intrinsèques et n'est pas une caractérisation générale qu'on peut appliquer à tous les différents modèles de percolation. La question qui se pose : Est-il possible d'unifier la loi qu'on peut utiliser pour déterminer le type de la transition ? Pour répondre à cette question, il faut d'abord trouver les paramètres ou les grandeurs avec lesquelles on va établir cette unification. Pour cette raison, on propose une méthode qui peut prédire le type de la transition de phase au voisinage du point critique en admettant que la distribution d'amas dans le réseau au voisinage du point critique suit toujours une loi de puissance, on définit deux paramètres : le premier est l'exposant  $\gamma$  de la distribution d'amas au voisinage du point critique,  $P(s) = cs^{-\gamma}$ , avec  $c$  est la constante de normalisation. Le deuxième paramètre est  $\gamma'$  qui dépend de la façon de choisir les amas, plus précisément il dépend des corrélations existantes dans le réseau, car dans les réseaux corrélés la probabilité de choisir une amas de taille  $s$  est différente de sa distribution réelle. Par exemple, dans la Fig. 5.5 on observe que dans le cas d'un processus Achlioptas, l'exposant de la distribution d'amas avec laquelle les amas sont choisies,  $\gamma'$ , est différente à la distribution d'amas réelle  $\gamma = 2.5$ , par contre dans le cas aléatoire les distributions restent exactement les mêmes,  $\gamma' = \gamma = 2.5$ . Autrement dit, dans les réseaux non-corrélés la probabilité de choisir un amas est exactement liée à sa distribution, ce qui n'est pas le cas avec l'existence des corrélations. D'où on définit la distribution dont les amas sont choisies par  $P(s)' = c's^{-\gamma'}$ , avec  $c'$  est la constante de normalisation.

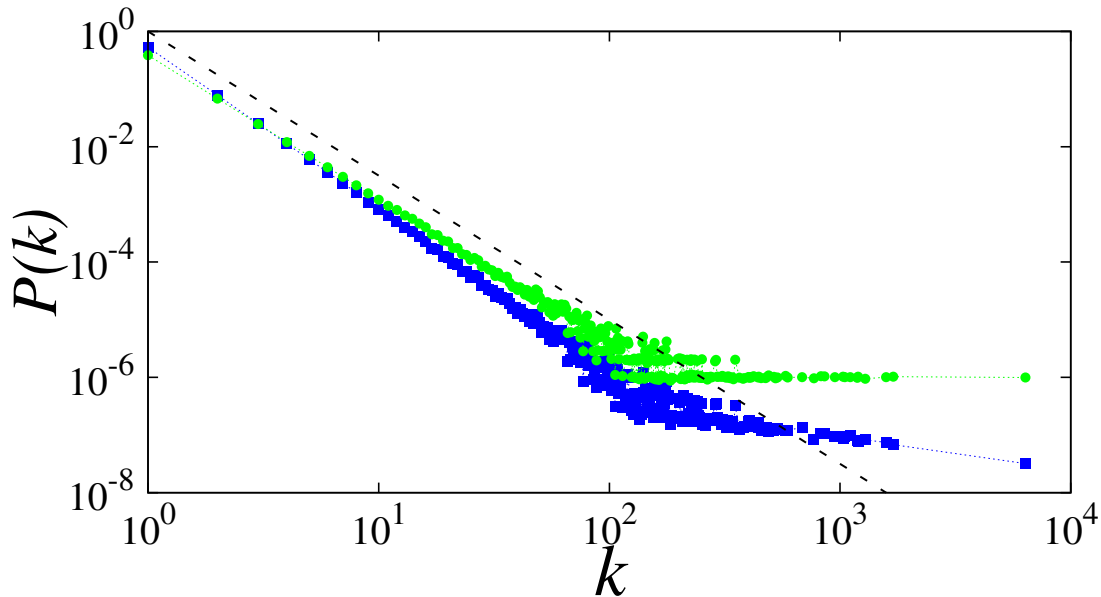


FIGURE 5.5 – La distribution d’amas selon deux méthodes différentes pour choisir les liens dans le même réseau, où l’exposant de la distribution d’amas dans le réseau est  $\gamma = 2.5$ , la première méthode est aléatoire comme le processus ER (cercles verts), la deuxième est par le processus Achlioptas [179] (carrés bleus). La ligne pointillée a pour but de montrer l’exposant  $\gamma = 2.5$ . La taille du réseau est  $n = 10^6$ , chaque simulation a une moyenne de  $n$  réalisations.

### 5.6.2 Approche théorique

#### Composante géante :

Dans cette partie nous allons élaborer une expression de la composante géante, beaucoup plus générale que celle de ER, au voisinage du point critique en fonction de deux paramètres  $\gamma$  et  $\gamma'$ . Nous supposons que le point critique est notre temps initial, puis nous ajouterons des liens au réseau. Pour qu’un amas de taille  $i$  n’appartienne pas à la composante géante, il ne doit pas être connecté à la composante géante via un autre amas de taille  $j$ . Cela signifie que : (a) l’amas de taille  $i$  n’est pas connecté à l’amas de taille  $j$  par un lien, ou (b) l’amas de taille  $i$  est connecté à l’amas de taille  $j$ , mais l’amas de taille  $j$  n’est pas elle-même membre de la composante géante.

Soit  $p = \frac{\langle k \rangle}{n-1}$  la probabilité que deux nœuds se connectent, où  $\langle k \rangle = \frac{2 \times (\text{nombre de liens ajoutés})}{n}$

est la densité des liens ajoutés par rapport à la taille du réseau. La probabilité du résultat (a) est  $(1 - p)^{ij}$ , et la probabilité du résultat (b) est  $(1 - (1 - p)^{ij})u$ , où  $u$  est la probabilité qu'un nœud n'appartienne pas à la composante géante. D'où la probabilité qu'un amas de taille  $i$  n'appartienne pas à la composante géante à travers n'importe quel amas de taille  $j$  dans le réseaux est

$$\begin{aligned} q_{ij} &= [(1 - p)^{ij} + (1 - (1 - p)^{ij})u]^{n_a P(j)'} \\ &\approx [1 - ijp(1 - u)]^{n_a P(j)'} \\ &\approx e^{-ijp(1-u)n_a P(j)'}, \end{aligned} \quad (5.6)$$

avec  $P(j)' = c'j^{-\gamma'}$  est la probabilité de choisir un amas de taille  $j$  et  $n_a$  est le nombre des amas. D'où la probabilité qu'un amas de taille  $i$  n'appartienne pas à la composante géante à travers n'importe quel amas est

$$\begin{aligned} \pi_i &= \prod_{j=1}^K q_{ij} \\ &= e^{-ipn_a(1-u)\sum_{j=1}^K jP(j)'}, \end{aligned} \quad (5.7)$$

où  $K = n_a^{\frac{1}{\gamma'-1}}$  est l'amas maximal au voisinage du point critique ( $\langle k \rangle_c$ ).

D'où  $\sum_{j=1}^K jP(j)' = \langle k_a \rangle'$  est le degré moyen avec lequel les amas sont choisis au réseau, ainsi que  $\langle k_a \rangle$  est la valeur moyenne des amas, alors  $n = n_a \langle k_a \rangle$ , et on sait que  $p = \frac{\langle k \rangle}{n-1}$ , on peut écrire

$$\pi_i = e^{\frac{-i(1-u)\langle k \rangle \langle k_a \rangle'}{\langle k_a \rangle}}. \quad (5.8)$$

La fraction des amas de taille  $i$  qui n'appartiennent pas à la composante géante par rapport à la taille de réseau  $n$  est

$$u_i = \frac{n_a i P(i) \pi_i}{n}, \quad (5.9)$$

d'où la fraction des nœuds qui n'appartiennent pas à la composante géante est

$$u = \sum_{i=1}^K u_i. \quad (5.10)$$

Alors la fraction des nœuds qui appartiennent à la composante géante est

$$S = 1 - u = 1 - \sum_{i=1}^K u_i. \quad (5.11)$$

De Eq. (5.8), Eq. (5.9) et sachant que  $n = n_a \langle k_a \rangle$  et  $S = 1 - u$  on obtient

$$\begin{aligned} S &= 1 - \frac{1}{\langle k_a \rangle} \sum_{i=1}^K i P(i) e^{\frac{-iS \langle k \rangle \langle k_a \rangle'}{\langle k_a \rangle}} \\ &= 1 - \frac{1}{\langle k_a \rangle} \int_{i=1}^K i P(i) e^{\frac{-iS \langle k \rangle \langle k_a \rangle'}{\langle k_a \rangle}} di. \end{aligned} \quad (5.12)$$

Cette formule ci-dessus est une expression beaucoup plus générale que celle de ER [90]. En effet, dans le cas où on ajoute aléatoirement des liens sans corrélations,  $\gamma = \gamma'$ , et lorsqu'au temps initial tout amas a la taille 1, on va trouver exactement l'expression de ER,  $S = 1 - e^{-S \langle k \rangle}$ .

### Détermination du point critique :

A partir de Eq. (5.12) et en se basant sur une solution graphique citée dans [196] on peut obtenir facilement le point critique  $\langle k \rangle_c$  où la composante géante apparaît, cette méthode s'appuie sur la condition que le point critique  $\langle k \rangle$  justifie

$$\frac{\partial \left( 1 - \frac{1}{\langle k_a \rangle} \sum_{i=1}^K i P(i) e^{\frac{-iS \langle k \rangle \langle k_a \rangle'}{\langle k_a \rangle}} \right)}{\partial S} \Big|_{S=0} = 1, \quad (5.13)$$

la solution de cette équation ci-dessus nous donne l'expression suivante

$$\langle k \rangle_c = \frac{\langle k_a \rangle^2}{\langle k_a^2 \rangle \langle k_a \rangle'}, \quad (5.14)$$

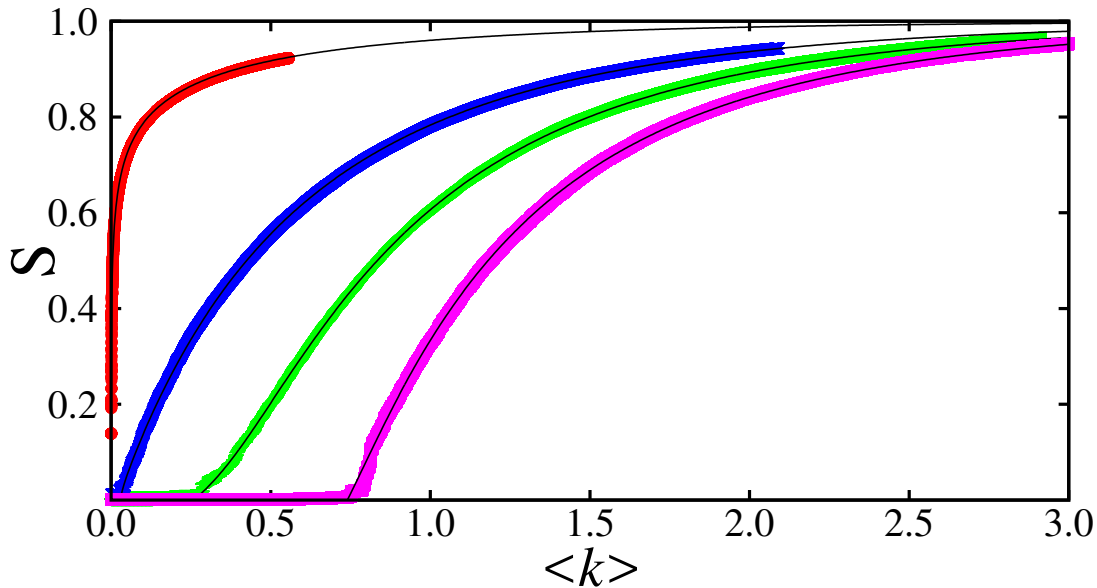


FIGURE 5.6 – Fraction de la composante géante  $S$  en fonction de la densité des liens  $\langle k \rangle$  pour différentes valeurs de  $\gamma$ , les couleurs de gauche à droite représentent respectivement les simulations numériques pour  $\gamma = 2, 2.5, 3, 4$ , les lignes noire représentent la solution numérique de l'Eq. (5.12). Le nombre des nœuds est  $n = 10^5$ .

dans le cas où il n'y a pas de corrélations  $\gamma = \gamma'$ , c'est-à-dire  $\langle k_a \rangle = \langle k_a \rangle'$ , on obtient

$$\langle k \rangle_c = \frac{\langle k_a \rangle}{\langle k_a^2 \rangle} = \frac{1}{\kappa_a}, \quad (5.15)$$

cette expression est beaucoup plus exacte de celle donnée par [185], où les auteurs ont trouvé que  $\langle k \rangle_c = \frac{1}{\langle k_a^2 \rangle}$ , mais en se basant sur les simulations numériques (voir Fig. 5.7) il est clair que notre formule est plus exacte.

Il convient de mentionner que la détermination du point critique par les simulations n'est pas une opération facile, il est assez compliqué et délicat de la trouver avec une grande précision. Dans nos simulations présentées dans la Fig. 5.7, nous avons estimé le point critique par la détermination d'un point où la valeur moyenne des petits amas dans le réseau

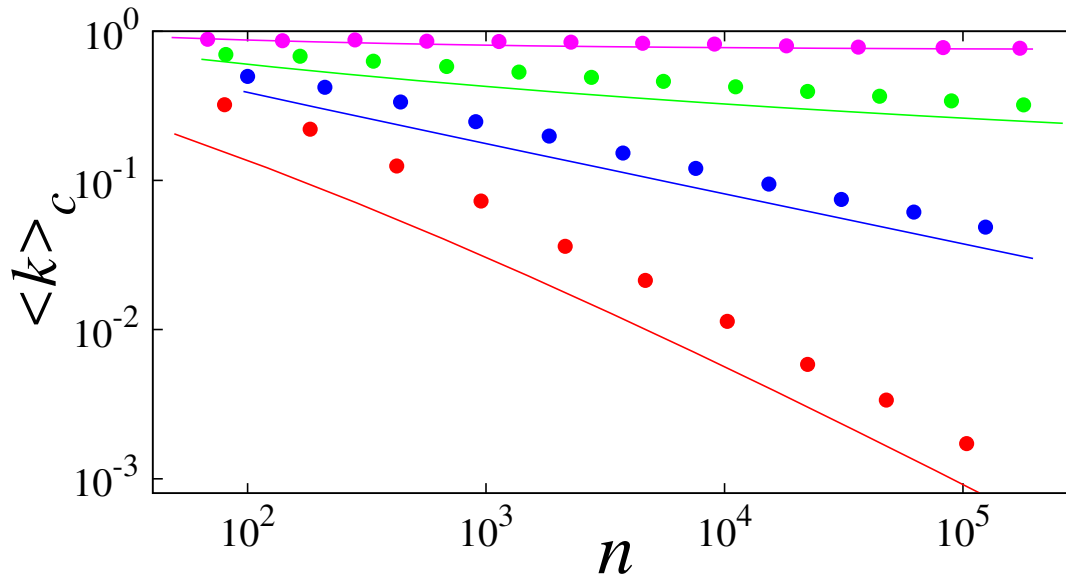


FIGURE 5.7 – Estimation de point critique par les simulations numériques (cercles) et par Eq. (5.15) (ligne). Les valeurs de  $\gamma$  de haut en bas sont respectivement 4, 3, 2.5, 2. Chaque simulation est moyennée à 500 réalisations.

est maximale, cependant cette estimation est devenue plus précise avec l'augmentation du nombre des nœuds du réseau, car les effets d'échelle diminuent avec la taille. Alors que l'expression théorique semble exacte, le problème reste la détermination du point critique par les simulations numériques.

### Détermination du type de transition de phase :

Pour déterminer le type de la transition de phase on va essayer de calculer la valeur de  $\langle k \rangle$  pour que  $S$  atteigne une valeur à l'échelle macroscopique,  $O(1)$ , (par exemple  $S = \frac{1}{4}$ ), si  $\langle k \rangle$  tend vers zéro on dit que la transition est de premier ordre sinon on dit qu'elle est de second ordre.

A partir de Eq. (5.12) et sachant que  $P(i) = ci^\gamma$ , on obtient

$$S = 1 - \frac{c}{\langle k_a \rangle} \int_{i=1}^K i^{1-\gamma} e^{\frac{-iS\langle k \rangle \langle k_a \rangle'}{\langle k_a \rangle}} di, \quad (5.16)$$

l'intégrale dans cette expression n'a pas de solution explicite connu, alors il est impossible d'extraire la valeur de  $\langle k \rangle$  en fonction de  $S$ , pour cette raison on propose une autre approche, on va construire la composante géante par la manière la plus rapide possible, puis on déduit le type de la transition de phase selon  $\gamma$  et  $\gamma'$ .

Pour chaque nouveau lien, on lie les deux amas les plus grands dans le réseau<sup>1</sup>, mathématiquement on traduit cela par l'expression suivante

$$S = \sum_{i=1}^t k(t) = \int_1^t k(i) di, \quad (5.17)$$

où  $t$  représente le temps et également le nombre des liens ajoutés, et  $k(t)$  est  $t^{\text{ème}}$  degré maximal, avec  $k(1) = K$ .

$k(t)$  peut être calculé par l'expression suivante

$$\sum_{i=k(t)}^K n_a P(i) = \int_{i=k(t)}^K n_a P(i) di = t, \quad (5.18)$$

d'où on obtient

$$\begin{aligned} k(t) &= \left( K^{1-\gamma} + \frac{(\gamma-1)t}{cn_a} \right)^{\frac{1}{1-\gamma}} \\ &= \left( \frac{t+1}{n_a} \right)^{\frac{1}{1-\gamma}} \quad \text{car : } c = \gamma - 1 \quad \text{et} \quad K = n_a^{\frac{1}{\gamma-1}} \\ &= \left( \frac{t}{n_a} \right)^{\frac{1}{1-\gamma}} \quad \text{car : } t+1 \approx t, \end{aligned} \quad (5.19)$$

de Eq. (5.17) et Eq. (5.19) on trouve que

$$S = \frac{\gamma-1}{(\gamma-2)\langle k_a \rangle} \left( \left( \frac{t}{n_a} \right)^{\frac{2-\gamma}{1-\gamma}} - \left( \frac{1}{n_a} \right)^{\frac{2-\gamma}{1-\gamma}} \right), \quad (5.20)$$

---

1. Cette approche est valable sauf au voisinage du point critique lorsque la distribution d'amas suit une loi de puissance, car lorsqu'on commence à ajouter les liens par cette méthode loin du point critique, cela peut ralentir la transition au point critique au lieu de l'accélérer.



sachant que  $\langle k_a \rangle = \frac{\gamma-1}{\gamma-2}$  et  $t = \frac{n\langle k \rangle}{2}$ , on obtient que Eq. (5.20) devient

$$S = \left( \frac{\langle k \rangle (\gamma - 1)}{2(\gamma - 2)} \right)^{\frac{2-\gamma}{1-\gamma}} - \left( \frac{1}{n_a} \right)^{\frac{2-\gamma}{1-\gamma}}, \quad (5.21)$$

d'où on obtient

$$\langle k \rangle = \frac{2(\gamma - 2)}{\gamma - 1} \left( S + \left( \frac{1}{n_a} \right)^{\frac{2-\gamma}{1-\gamma}} \right)^{\frac{1-\gamma}{2-\gamma}}. \quad (5.22)$$

Pour  $\gamma > 2$  et une valeur  $S$  donnée qui est différente de 0, on peut écrire dans la limite thermodynamique que  $\langle k \rangle = \frac{2(\gamma-2)}{\gamma-1} \left( S \right)^{\frac{1-\gamma}{2-\gamma}}$ , sous les conditions précédentes cette dernière expression ne peut jamais tendre vers zéro, ce qui signifie absolument que la transition de phase ne peut pas être discontinue pour  $\gamma > 2$ , certes elle est de second ordre, ou plutôt elle est continue.

On a démontré ici qu'il est impossible de trouver une transition de phase de premier ordre, lorsque l'exposant  $\gamma > 2$  au voisinage du point critique indépendamment de l'autre exposant  $\gamma'$ , c'est-à-dire indépendamment des corrélations collectives dans le réseau.

Maintenant on va étudier le cas  $\gamma = 2$  pour déterminer quel type de transition de phase apparaîtrait ici.

Dans ce cas, l'intervalle de Eq. (5.16) se calcule en s'appuyant sur l'approximation suivante  $e^{\frac{-iS\langle k \rangle \langle k_a \rangle'}{\langle k_a \rangle}} \approx \frac{1}{1 + \frac{iS\langle k \rangle \langle k_a \rangle'}{\langle k_a \rangle}}$ , d'où

$$\begin{aligned} \int_{i=1}^K \frac{i^{-1}}{1 + \frac{iS\langle k \rangle \langle k_a \rangle'}{\langle k_a \rangle}} di &= \left[ \ln(i) - \ln(i\langle k_a^2 \rangle \langle k \rangle S + \langle k_a \rangle) \right]_{i=1}^{i=K} \\ &= \ln(K) - \ln(K\langle k_a^2 \rangle \langle k \rangle S + \langle k_a \rangle) \quad \text{car : } K \gg 1 \end{aligned} \quad (5.23)$$

si on remplace Eq. (5.23) dans Eq. (5.16) on obtient

$$S = 1 - \frac{c}{\langle k_a \rangle} (\ln(K) - \ln(K\langle k_a^2 \rangle \langle k \rangle S + \langle k_a \rangle)), \quad (5.24)$$

d'où l'expression de  $\langle k \rangle$  s'écrit

$$\langle k \rangle = \frac{e^{\frac{c \ln(K) + \langle k_a \rangle (S-1)}{c}} - \langle k_a \rangle}{K \langle k_a \rangle' S}, \quad (5.25)$$

sachant que pour  $\gamma = 2$  on a  $K = n_a$  et  $\langle k_a \rangle = c \ln(n_a)$ , on obtient

$$\langle k \rangle = \frac{n_a^S - c \ln(n_a)}{n_a \langle k_a \rangle' S}, \quad (5.26)$$

pour  $n_a \gg 1$  on a  $n_a \gg c \ln(n_a)$  et lorsque  $S > 0$  on peut écrire

$$\langle k \rangle = \frac{n_a^{S-1}}{\langle k_a \rangle' S}, \quad (5.27)$$

il est clair que cette expression ci-dessus tend toujours vers 0 pour toutes les valeurs de  $S < 1$  et pour n'importe quelle valeur de  $\gamma'$ . Le fait que  $\langle k \rangle$  tends vers 0 cela signifie que la transition est discontinue dans la limite thermodynamique, c'est-à-dire, pour  $\gamma = 2$  il s'agit d'une transition de phase de premier ordre.

Alors nous avons montré comment le type de la transition de phase dépend seulement de la distribution d'amas au voisinage du point critique (l'exposant  $\gamma$ ), et ne dépend pas des corrélations collectives entre les amas (l'exposant  $\gamma'$ ).

En outre, on peut déduire que les corrélations n'ont aucune importance au voisinage du point critique dans la détermination du type de la transition de phase, plutôt la transition sera de premier ordre si les corrélations réussissent à mener le réseau à atteindre un exposant inférieur ou égal à 2, sinon la transition sera de second ordre.

## 5.7 Conclusion

Dans ce chapitre nous avons introduit une méthode pour prévoir le type de la transition de phase dans le phénomène de percolation au voisinage du point critique. Cette méthode est basée sur deux paramètres, le premier est l'exposant de la distribution d'amas au voisinage du point critique,  $\gamma$ , et le deuxième est l'exposant de la distribution d'amas avec laquelle les amas sont choisis,  $\gamma'$ , ce dernier reflète les corrélations dans le réseau. Nous avons trouvé qu'une transition de phase est de premier ordre si  $\gamma \leq 2$ , sinon la transition est de second ordre, indépendamment de la valeur de  $\gamma'$ . On en déduit que les corrélations collectives entre les amas au voisinage du point critique n'ont aucune influence sur le type de la transition de phase, mais elles jouent un rôle important pour que le réseau acquiert un certain exposant de

distribution d'amas.



# Conclusion

Dans cette thèse, quatre contributions ont été réalisées pour le domaine des réseaux complexes, permettant de résoudre des problèmes de calculs théoriques, et de proposer des nouvelles méthodes. Dans la première contribution, nous avons introduit un simple modèle de réseau complexe avec un critère d'attachement préférentiel sans l'effet "rich get richer", le réseau obtenu est homogène, ce qui démontre le rôle crucial de l'effet "rich get richer" dans la topologie du réseau. Nous en avons déduit qu'un traitement préférentiel aux nœuds les moins connectés équivaut à utiliser une probabilité d'attachement aléatoire. En calculant le degré moyen instantané d'un nœud sélectionné et ses fluctuations, nous avons montré comment le degré moyen de hubs et ses fluctuations divergent avec le temps dans le modèle BA, et restent finis dans notre modèle.

Dans la deuxième contribution, nous avons étudié en détail certains aspects fondamentaux des réseaux libre-échelle aléatoires non corrélés. Avec des étapes et des hypothèses simples, nous avons obtenu les expressions explicites du nombre des nœuds à une distance donnée d'un nœud arbitraire,  $n_\ell$ . Nous avons obtenu également la description précise de la forme de la distribution. Profitant de la forme de la distribution  $n_\ell$ , nous avons pu déduire l'expression explicite du PCC. Les expressions obtenues reproduisent les formes de mise à l'échelle connues pour différentes plages de  $\gamma$ . Autrement dit, le monde ultra-petit pour  $2 < \gamma < 3$ , et le petit-monde pour  $\gamma \geq 3$ . Nos résultats théoriques concordent très bien avec les simulations, sauf dans le cas de  $\gamma = 3$ , où nous avons observé la même forme, dans les queues de  $n_\ell$ , mais avec une petite déviation dans la position du maximum. Cette différence n'affecte pas la forme de mise à l'échelle du PCC pour cette valeur de  $\gamma$ , ainsi nos expressions restent les plus précises dans la littérature.

Dans la troisième contribution, en utilisant la transformation de groupe de renormalisation GR sur le modèle NW, nous avons élaboré une expression analytique du PCC plus perfectionnée par rapport à l'existant. En effet, à partir de cette nouvelle expression, nous avons montré que suivant le paramètre  $y = 2k^2\phi$  la propriété petit-monde émerge, et selon sa valeur l'erreur de la fonction universelle  $\ell = \dot{n}f(x)$  se révèle plus ou moins tandis que la nouvelle fonction se comporte inversement. Alors c'est une émergence spectaculaire qui se passe en fonction du paramètre  $y$ .

Dans la quatrième contribution, nous avons introduit une méthode pour anticiper le type de transition de phase dans le phénomène de percolation au voisinage de point critique, nous avons trouvé qu'une transition de phase du premier ordre est impossible si l'exposant de distribution d'amas  $\gamma > 2$ , quelle que soit les corrélations dans le système donné, en revanche elle est du premier ordre si  $\gamma \leq 2$ . De façon générale, il s'avère qu'il suffit de connaître  $\gamma$  au voisinage de point critique pour en déduire le type de transition de phase.

# Bibliographie

- [1] Jeong H., B.Tombor, R.Albert, Z.N.Oltvai, and A.L.Barabasi. The large-scale organization of a metabolic networks. *Nature*, 407 :651–654, (2000).
- [2] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *Computer Communication Review*, 29 :251, (1999).
- [3] A.-L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286 :509–512, (1999).
- [4] Cohen, J. E., Briand F., and Newman, C. M. Community food webs : Data and theory. *Biomathematics*, 20, (1990).
- [5] Pimm, S. L. *Food Webs*. University of Chicago Press, Chicago, 2nd ed. (2002).
- [6] M. E. J. Newman and D. J. Watts. Renormalization group analysis of the small-world network model. *Phys. Lett.A*, 263 :341, (1999).
- [7] F. R. Pitts. A graph theoretic approach to historical geography. *The Professional Geographer*, 17 :15–20, (1965).
- [8] L. A. N. Amaral, A. Scala, M. Barthélemy, and H. E. Stanley. Classes of small-world networks. *Proc. Natl. Acad. Sci, USA* 97 :11149–11152, (2000).
- [9] I. Dobson, B. A. Carreras, V. E. Lynch, and D. Newman. Complex systems analysis of series of blackouts : Cascading failure, critical points, and self-organization. *Chaos*, 17 :026103, (2007).
- [10] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee, and S. S. Manna. Small-world properties of the indian railway network. *Phys.Rev. E*, 67 :036106, (2003).

- [11] A. Broido and K. C. Claffy. Internet topology : Connectivity of ip graphs, in s. fahmy and k. park, eds., scalability and traffic control in ip networks, no. 4526 in proc. spie. *International Society for Optical Engineering*, page 172–187, (2001).
- [12] G. B. West, J. H. Brown, and B. J. Enquist. A general model for the origin of allometric scaling laws in biology. *Science*, 276 :122–126, (1997).
- [13] S. N. Dorogovtsev and J. F. F. Mendes. *Evolution of Networks : From Biological Nets to the Internet and WWW*. Oxford Univ Press, Oxford, (2003).
- [14] R. Cohen and S. Havlin. Scale-free networks are ultrasmall. *Phys. Rev. Lett*, 90 :058701, (2003).
- [15] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Metric structure of random networks. *Nucl. Phys. B*, 653 :307, (2003).
- [16] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Math*, 1 :91–113, (2003).
- [17] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs : simple building blocks of complex networks. *Science*, 298 :824–827, (2002).
- [18] J. Scott. *Social Network Analysis : A Handbook*. Sage Publications, London, 2nd ed. (2000).
- [19] W. W. Zachary. An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33 :452–473, (1977).
- [20] J. L. Moreno. *Who Shall Survive ?* Beacon House, Beacon, (1934).
- [21] J. Galaskiewicz and P. V. Marsdeny. Interorganizational resource networks : Formal patterns of overlap. *Social Science Research*, 7 :89–107, (1978).
- [22] Krioukov, D., Kitsak, M., Sinkovits, R. S., Rideout, D., Meyer, D., and Boguna, M. Network cosmology. *Scientific Reports*, 2 :793, (2012).
- [23] E. Bullmore and O. Sporns. Complex brain networks : Graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, 10 :168–198, (2009).
- [24] D. Lazer and al. Computational social science. *Science*, 323 :721–723, (2009).



- [25] Y.Y. Liu, J.J. Slotine, and A.L. Barabasi. Controllability of complex networks. *Nature*, 473 :167–173, (2011).
- [26] L. Smolin. *The life of the cosmos*. Oxford University Press, Oxford, (1997).
- [27] C. A. Trugenberger. Quantum gravity as an information network : Self-organization of a 4d universe. *Phys.Rev. D*, 92 :084014, (2015).
- [28] Bianconi, G., Rahmede, C., and Wu, Z. Complex quantum network geometries : Evolution and phase transitions. *Phys. Rev. E*, 92 :022815, (2015).
- [29] B. Bollobas and O. Riordan. *Percolation*. Cambridge University Press, Cambridge, (2006).
- [30] A. Acin, J. I. Cirac, and M. Lewenstein. Entanglement percolation in quantum networks. *Nature Physics*, 3 :256, (2007).
- [31] S. Sachdev. *Quantum Phase Transitions*. Cambridge University Press, Cambridge, (1999).
- [32] J. Novotny, G. Alber, and I. Jex. Universality in random quantum networks. *Phys. Rev. A*, 92 :062335, December (2015).
- [33] D. Bruß and G. Leuchs (eds.). *Lectures on Quantum Information*. Weinheim, Wiley-VCH, (2007).
- [34] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, (2000).
- [35] M. Mitchell. Complex systems : Network thinking. *Artificial Intelligence*, 170 :1194–1212, (2006).
- [36] I.A. Basheer M. Hajmeer. Artificial neural networks : fundamentals, computing, design and application. *Journal of Microbiological Methods*, 43 :3–31, (2000).
- [37] A. Passerini, G. Tack, and T. Guns. Introduction to the special issue on combining constraint solving with mining and learning. *Artificial Intelligence*, 244 :1–5, (2017).
- [38] F. Harary. *Graph Theory*. Perseus, Cambridge, MA (1995).
- [39] D. B. West. *Introduction to Graph Theory*. Prentice Hall, Upper Saddle River, NJ (1996).

- [40] S. Milgram. The small world problem. *Psychology Today*, page 56–67, May (1967).
- [41] J. Travers and S. Milgram. *Sociometry*, 32 :425, (1969).
- [42] P. Dodds, R. Muhamad, and D. Watts. An experimental study of search in global social networks. *Science*, 301(5634) :827–829, (2003).
- [43] R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet : A Statistical Physics Approach*. Cambridge University Press, Cambridge, (2004).
- [44] S. Wasserman and K. Faust. *Social Network Analysis : Methods and Applications*. Cambridge University Press, Cambridge, (1994).
- [45] Boccaletti, Stefano, and al. Complex networks : Structure and dynamics. *Physics reports*, 424.4 :175–308, (2006).
- [46] Q. Ye, B. Wu, and B. Wang. Distance distribution and average shortest path length estimation in real-world networks. *In Proceedings of the 6th international conference on Advanced data mining and applications : Part I*, volume 6440 of Lecture Notes in Computer Science :322–333, (2010).
- [47] V. Latora and M. Marchiori. Efficient behavior of small-world networks. *Phys. Rev. Lett*, 87(19) :198701, Oct (2001).
- [48] S. Melnik, A. Hackett, M. A. Porter, P. J. Mucha, and J. P. Gleeson. The unreasonable effectiveness of tree-based theory for networks with clustering. *Phys. Rev. E*, 83 :036112, (2011).
- [49] M. Zhao, T. Zhou, B.-H. Wang, G. Yan, H.-J. Yang, and W.-J. Bai. Relations between average distance, heterogeneity and network synchronizability. *Physica A*, 371 :377, (2006).
- [50] R. Albert, H. Jeong, and A. L. Barabási. Error and attack tolerance of complex networks. *Nature (London)*, 406 :378, (2000).
- [51] L. A. Adamic. The small world web. in proceedings of the third european conference on research and advanced technology for digital libraries. *Sociometry*, 1696 of Lecture Notes in Computer Science :443–452, (1999).
- [52] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee, and S. S. Manna. *Small-world properties of the indian railway network*. (2002).

- [53] A. Wagner and D. Fell. *The small world inside large metabolic networks*. (August 21 2000).
- [54] S. Bilke and C. Peterson. Topological properties of citation and metabolic networks. *Rev. E*, 64 :036106, (2001).
- [55] D. J. Watts and S. H. Strogatz. Collective dynamics of “small-world” networks. *Nature*, 393 :440–442, (1998).
- [56] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev*45, 2 :167–256, (2003).
- [57] G. Bianconi and A. Capocci. Number of loops of size  $h$  in growing scale-free networks. *Phys. Rev. Lett*, 90 :art. no. 078701, (2003).
- [58] A. Fronczak, J. A. Holyst, M. Jedynak, and J. Sienkiewicz. Higher order clustering coefficients in barabasi-albert networks. *Phys. A*, 316 :688–694, (2002).
- [59] P. M. Gleiss, P. F. Stadler, A. Wagner, and D. A. Fell. Relevant cycles in chemical reaction networks. *Adv. in Complex Systems*, 4 :207–226, (2001).
- [60] M. Boguñá and R. Pastor-Satorras. *Phys. Rev. E*, 66 :047104, (2002).
- [61] S. Maslov and K. Sneppen. Specificity and stability in topology of protein networks. *Science*, 296 :910–913, (2002).
- [62] M. Dehmer. *Structural Analysis of Complex Networks*. Springer, (2011).
- [63] P. Hage and F. Harary. Eccentricity and centrality in networks. *Soc Networks*, 17 :57–63, (1995).
- [64] S. Wasserman and K. Faust. *Social network analysis : methods and applications. Structural analysis in the social sciences*. Cambridge University Press, Cambridge, (1994).
- [65] L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40 :35–41, (1977).
- [66] Famous trails to Paul Erdos. Relevant cycles in chemical reaction networks. *Math. Intelligencer*, 51 :51–63, (1999).

- [67] M. E. J. Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46 :323–351, (2005).
- [68] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51 n.4 :661–703, November (2009).
- [69] B. Gutenberg and R. F. Richter. Energy ratio of reflected and refracted seismic waves. *Bull. Seismol. Soc. Am*, 34 :85–102, (1944).
- [70] G. Neukum and B.A. Ivanov. in *Hazards Due to Comets and Asteroids*. edited by T. Gehrels, (University of Arizona Press, Tucson, AZ, 1994),pp. 359 – 416.
- [71] G.K. Zipf. *Human Behaviour and the Principle of Least Effort*. Addison-Wesley, Reading, MA, 1949.
- [72] J.B. Estoup. *Gammes Stenographiques*. Institut Stenographique de France, Paris, (1916).
- [73] D.H. Zanette and S.C. Manrubia. Vertical transmission of culture and the distribution of family names. *Physica A*, 295 :1, (2001).
- [74] A.J. Lotka. The frequency distribution of scientific productivity. *Acad. Sci*, 16 :317–323, (1926).
- [75] D.J. de S. Price. Networks of scientific papers. *Science*, 149 :510–515, (1965).
- [76] L. A. Adamic and B. A. Huberman. Q. j. electron. *Commerce*, 1 :512, (2000).
- [77] R.A.K. Cox, J.M. Felton, and K.C. Chung. The concentration of commercial success in popular music : an analysis of the distribution of gold records. *J. Cult. Econ*, 19 :333–340, (1995).
- [78] J.C. Willis and G.U. Yule. Some statistics of evolution and geographical distribution in plants and animals and their significance. *Nature*, 109 :177, (1922).
- [79] V. Pareto. *Cours d'Economie Politique*. Droz, Geneva, (1896).
- [80] G. Bianconi and C. Rahmede. Complex quantum network manifolds in dimension  $d > 2$  are scale-free. *Scientific Reports*, 5 :13979, (2015).
- [81] Dourisboure, Y., F. Geraci,, and M. Pellegrini. in WWW'07 : Proceedings of the 16th international conference on the World Wide Web (ACM, New York, NY, USA). pages 461–470, (2007).

- [82] Flake, G. W., S. Lawrence, C. Lee Giles, and F. M. Coetzee. Self-organization and identification of web communities. *IEEE Computer*, 35 :66, (2002).
- [83] J. Chen and B. Yuan. Detecting functional modules in the yeast protein–protein interaction network. *Bioinformatics*, 22(18) :2283, (2006).
- [84] A. W. Rives and T. Galitski. Modular organization of cellular networks. *Proc. Natl. Acad. Sci. USA*, 100(3) :1128, (2003).
- [85] B. Krishnamurthy and J. Wang. On network-aware clustering of web clients. *Proceedings of the conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, 30(4) :97, (2000).
- [86] M. Steenstrup. Cluster-based networks. *Addison Wesley, Reading, USA*, chapter 4 :75–138, (2001).
- [87] P. Csermely. Cluster-based networks. *Trends Biochem. Sci*, 33(12) :569, (2008).
- [88] D. Chowdhury and D. Stauffer. *Principles of equilibrium statistical mechanics*. Wiley-VCH, Weinheim, (2000).
- [89] H-P. Hsu, V. Mehra, and P. Grassberger. Structure optimization in an off-lattice protein model. *Phys Rev E*, 68 :037703, (2003).
- [90] P. Erdos and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6 :290–297, (1959).
- [91] P. Erdos and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5 :17–61, (1960).
- [92] P. Erdos and A. Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Scientia Hungary*, 12 :261–267, (1961).
- [93] D.J. Watts and S.H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393 :440–442, (1998).
- [94] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys*, 74 :47, (2002).
- [95] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Adv. Phys*, 51 :1079, (2002).

- [96] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev*, 45 :167, (2003).
- [97] W. B. Deng, L. Guo, W. Li, and X. Cai. World wide marine transportation network : Efficiency and container throughput. *Chin. Phys. Lett*, 26 :118901, (2009).
- [98] R. Albert, I. Albert, and G. L. Nakarado. Structural vulnerability of the north american power grid. *Phys. Rev. E*, 69 :025103, (2004).
- [99] T. B. Achacoso and W. S. Yamamoto. *AY's Neuroanatomy of C. elegans for Computation*. CRC Press, Boca Raton, FL, (1992).
- [100] A.-L. Barabási, H. Jeong, and R. Albert. Statistical mechanics and its applications. *Physica A*, 272 :173, (1999).
- [101] Y.-B. Xie, T. Zhou, and B.-H. Wang. Scale-free networks without growth. *Physica A*, 387 :1683, (2008).
- [102] V. K. Samalam. *arXiv :1202.1498*.
- [103] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Physical Review Letters*, 85 :4629–4632, (2000).
- [104] A. Nordsieck, W. E. Lamb Jr., and G. E. Uhlenbeck. On the theory of cosmic-ray showers i : The furry model and the fluctuation problem. *Physica*, 7(4) :344–360, (1940).
- [105] W. H. Furry. On fluctuation phenomena in the passage of high energy electrons through lead. *Phys. Rev*, 52(6) :569–581, (1937).
- [106] W. Feller. Die grundlagen der volterraschen theorie des kampfes ums dasein in wahr-scheinlichkeitstheoretischer behandlung. *Acta Biotheor*, 52(6) :11–40, (1939).
- [107] M. Delbrück. Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys*, 8(1) :120–124, (1940).
- [108] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library. Elsevier, Amsterdam, 3rd edition, 2007.
- [109] C. Gardiner. *Stochastic Methods - A Handbook for the Natural and Social Sciences*. volume 13 of Springer Series in Synergetics, Amsterdam, Springer, Berlin, Heidelberg, 4th edition, 2009.

- [110] Weber M F and Frey E. Master equations and the theory of stochastic path integrals. *Rep. Prog. Phys*, 80 :046601, (2017).
- [111] R. J. Glauber. Time-dependent statistics of the ising model. *J. Math. Phys*, 4(2) :294–307, (1963).
- [112] K. Kawasaki. Diffusion constants near the critical point for time-dependent ising models. i. *Phys. Rev*, 145(1) :224–230, (1966).
- [113] K. Kawasaki. Diffusion constants near the critical point for time-dependent ising models. ii. *Phys. Rev*, 148(1) :375–381, (1966).
- [114] K. Kawasaki. Diffusion constants near the critical point for time-dependent ising models. iii. *Phys. Rev*, 150 :285–290, (1966).
- [115] A. M. Walczak, A. Mugler, and C. H. Wiggins. A stochastic spectral analysis of transcriptional regulatory cascades. *Proc. Natl. Acad. Sci. USA*, 106 :6529–6534, (2009).
- [116] C. V. Rao, D. M. Wolf, , and A. P. Arkin. Control, exploitation and tolerance of intracellular noise. *Nature*, 420 :231–237, (2002).
- [117] L. S. Tsimring. Noise in biology. *Rep. Prog. Phys*, 77 :026601, (2014).
- [118] N. T. J. Bailey. A simple stochastic epidemic. *Biometrika*, 37 :193–202, (1950).
- [119] K. Rock, S. Brand, J. Moir, and M. J. Keeling. Dynamics of infectious diseases. *Rep. Prog. Phys*, 77 :026602, (2014).
- [120] E. Clayton, D. P. Doupé, A. M. Klein, D. J. Winton, B. D. Simons, and P. H. Jones. A single type of progenitor cell maintains normal epidermis. *Nature*, 446 :185–189, (2007).
- [121] W. Weidlich and M. Braun. The master equation approach to nonlinear economics. *J. Evol. Econ*, 2 :233–265, (1992).
- [122] D. Gross, J. F. Shortle, J. M. Thompson, and C. M. Harris. Fundamentals of queueing theory. *John Wiley and Sons, Hoboken, NJ, 4th edition*, (2008).
- [123] A. L. Barabási. *Network Science*. Cambridge University Press, (2015).
- [124] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Structure of growing networks with preferential linking. *Phys. Rev. Lett*, 85 :4633–4636, (2000).

- [125] B. Bollobás and O. Riordan. A polynomial of graphs on surfaces. *Math. Ann*, 323 :1–35, (2002).
- [126] A. L. Barabási. *Linked : The New Science of Networks*. Perseus, Cambridge, 2002.
- [127] A. Barrat et al. *Dynamical Processes on Complex Networks*. (Cambridge University Press, 2008).
- [128] R. Cohen and S. Havlin. *Complex Networks : Structure, Robustness and Function*. (Cambridge University Press, 2010).
- [129] M. Molloy and B. Reed. Combinatorics. *Probability and Computing*, 7 :295, (1998).
- [130] M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64 :026118, (2001).
- [131] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21 :1087–1092, (1953).
- [132] M. E. J. Newman. *Networks : An Introduction*. University of Michigan and Santa Fe Institute Press, (2010)p.456.
- [133] R. Cohen and S. Havlin. *Complex Networks : Structure, Robustness and Function*. (Cambridge University Press, 2010).pp :72-77.
- [134] T. Kalisky, R. Cohen, O. Mokryn, D. Dolev, Y. Shavitt, and S. Havlin. Equation of state calculations by fast computing machines. *Phys. Rev. E*, 74 :066108, (2006).
- [135] B. Bollobás. *Random Graphs*. Academic Press, London, (1985).
- [136] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1 :311–316, (1980).
- [137] R. Cohen, S. Havlin, and D. ben Avraham. *Structural properties of scale free networks, in : S. Bornholdt and H. G. Schuster, editors, Handbook of graphs and networks, chapter 4*. Wiley-VCH, Berlin, (2002).
- [138] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *PNAS*, 99 :15879–15882, (2002).



- [139] R. J. Fox and D. R. Bellwood. Herbivores in a small world : network theory highlights vulnerability in the function of herbivory on coral reefs. *Functional Ecology*, 28 :642, (2014).
- [140] R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electronic Journal of Probability*, 12 :703766, (2007).
- [141] A. Fronczak, P. Fronczak, and J. A. Holyst. Average path length in random networks. *Phys. Rev. E*, 70 :056110, (2004).
- [142] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. Random graphs with arbitrary i.i.d. degrees. *Random Structures and Algorithms*, 26 :76123, (2006).
- [143] R. Cohen and S. Havlin. *Scaling properties of complex networks and spanning trees, Chapter 3 in "Handbook of Large-Scale Random Networks", Eds. Bela Bollobas, Robert Kozma and Dezsó Miklos.* (Bolyai Society of Mathematical Studies and Springer, 2009).
- [144] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24 :5–34, (2004).
- [145] B. Bollobás and O. Riordan. *Handbook of graphs and networks, eds. S. Bornholdt and H. G. Schuster.* Wiley-VCH, Berlin, (2002).
- [146] M. E. J. Newman. *Networks : An Introduction.* University of Michigan and Santa Fe Institute Press, (2010).
- [147] M. Newman and D. Watts. Scaling and percolation in the small-world network model. *Phys. Lett. A*, 341, (1999).
- [148] M. Barthélemy and L. A. N. Amaral. Small-world networks : Evidence for a crossover. *Phys. Rev. Lett*, 82 :3180, (1999).
- [149] M. E. J. Newman, C. Moore, and D.J. Watts. Mean-field solution of the small-world network model. *Phys. Rev. Lett*, 84 :3201, (2000).
- [150] M.E.J. Newman. *Networks : An Introduction.* University of Michigan and Santa Fe Institute Press, (2010)p.558- 560.
- [151] A. Barrat and M. Weigt. On the properties of small-world network models. *Eur. Phys. J. B*, 13 :547, (2000).

- [152] A. Pelissetto and E. Vicari. Critical phenomena and renormalization-group theory. *Phys. Rept*, 368 :549, (2002).
- [153] K. G. Wilson. The renormalization group : Critical phenomena and the kondo problem. *Rev. Mod. Phys*, 47 :773–840, (1975).
- [154] A. Barrat. Comment on ‘Small-world networks : crossover picture’.
- [155] M. E. J. Newman and D. J. Watts. Scaling and percolation in the small-world network model. *Phys. Rev.E*, 60 :7332, (1999).
- [156] M. Argollo de Menezes, C. Moukarzel, and T.J.P. Penna. *Europhys. Lett*, 50 :574, (2000).
- [157] S. Hong and A. Dey. Network analysis of cosmic structures : network central- ity and topological environment. *Monthly Notices of the Royal Astronomical Society*, 450 :1999–2015, June (2015).
- [158] S. H. Strogatz. Exploring complex networks. *Nature*, 410 :268–276, (2001).
- [159] M. E. J. Newman, D. J. Watts, and S. H. Strogatz. Random graph models of social networks. *Proc. Natl Acad. Sci*, 99 :2566–2572, (2002).
- [160] C. Song, S. Havlin, and H. A. Makse. Origins of fractality in the growth of complex networks. *Nature Phys*, 2 :275–281, (2006).
- [161] R. Parshani, S. V. Buldyrev, and S. Havlin. Interdependent networks : Reducing the coupling strength leads to a change from a first to second order percolation transition. *Phys. Rev. Lett*, 105 :048701, (2010).
- [162] R. Parshani, S. V. Buldyrev, H. E. Stanley, and S. Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464 :1025–1028, (2010).
- [163] D. Ben-Avraham and S. Havlin. *Diffusion and Reactions in Fractals and Disordered Systems*. Cambridge Univ. Press, (2001).
- [164] A. A. Saberi. Recent advances in percolation theory and its applications. *Phys.Rep*, 578, (2015).
- [165] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Critical phenomena in complex networks. *Rev. Mod. Phys*, 80 :1275–1335, (2008).

- [166] H. D. Rozenfeld, L. K. Gallos, and H. A. Makse. Explosive percolation in the human protein homology networks. *Eur. Phys. J.*, E 75 :305–310, (2010).
- [167] A. Bunde and S. Havlin. *Fractals and Disordered Systems*. Springer, Berlin, (1996).
- [168] D. Stauffer and A. Aharony. *Introduction to Percolation Theory*. Taylor and Francis, London, (1994).
- [169] S. Boccaletti, J. Almendral, S. Guan, I. Leyva, Z. Liu, I. Sendina-Nadal, Z. Wang, and Y. Zou. Explosive transitions in complex networks’ structure and dynamics : percolation and synchronization. *Physics Reports*, 660 :1, (2016).
- [170] R. M. D’Souza and J. Nagler. Anomalous critical and supercritical phenomena in explosive percolation. *Nat. Phys*, 11 :531–538, (2015).
- [171] N. Araújo, P. Grassberger, B. Kahng, K. Schrenk, and R. Ziff. Recent advances and open challenges in percolation. *Eur. Phys. J. Spec. Top*, 223 (11) :2307–2321, (2014).
- [172] J. Adler. Bootstrap percolation. *Physica A*, 453 :171, (1991).
- [173] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. k-core organization of complex networks. *Phys. Rev. Lett*, 96 :040601, (2006).
- [174] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno. Dynamics of jamming transitions in complex networks. *EPL (Europhysics Letters)*, 71 (2) :325–331, (2005).
- [175] R. A. da Costa, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Explosive percolation transition is actually continuous. *Phys. Rev. Lett*, 105 :255701, (2010).
- [176] R. A. da Costa, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Inverting the achlioptas rule for explosive percolation. *Phys. Rev. E*, 91, (2015).
- [177] Y. S. Cho and B. Kahng. Suppression effect on explosive percolation. *Phys. Rev. Lett*, 107 :275703, (2011).
- [178] B. Bollobás. Modern graph theory. *Springer*, (1998,New York).
- [179] D. Achlioptas, R. M. D’Souza, and J. Spencer. Explosive percolation in random networks. *Science*, 323 (5920) :1453–1455, (2009).
- [180] T. Bohman and A.M. Frieze. Avoiding a giant component. *Random Structures and Algorithms*, 19 :75–85, (2001).

- [181] B. Bollobas. The evolution of random graphs. *Trans. Amer. Math. Soc*, 286 :257—274, (1984).
- [182] O. Riordan and L. Warnke. Explosive percolation is continuous. *Science*, 333 :322, (2011).
- [183] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *The Ann. Appl.Prob*, 22 :1450–1464, (2012).
- [184] J. Nagler, T. Tiessen, and H. W. Gutch. Continuous percolation with discontinuities. *Phys. Rev. X*, 2 :031009, (2012).
- [185] Y. S. Cho, B. Kahng, and D. Kim. Cluster aggregation model for discontinuous percolation transitions. *Phys. Rev. E*, 81 :030103(R), (2010).
- [186] A. A. Moreira, E. A. Oliveira, S. D. S. Reis, H. J. Herrmann, and J. S. Andrade. Hamiltonian approach for explosive percolation. *Phys. Rev. E*, 81 :040101, (2010).
- [187] N. Bastas, P. Giazitzidis, M. Maragakis, and K. Kosmidis. Explosive percolation : Unusual transitions of a simple model. *Physica A : Statistical Mechanics and its Applications*, 407 :54–65, (2014).
- [188] L. Cao and J. M. Schwarz. Correlated percolation and tricriticality. *Phys. Rev. E*, 86 :061131, (2012).
- [189] J. Fan, M. Liu, L. Li, and X. Chen. Continuous percolation phase transitions of random networks under a generalized achlioptas process. *Phys. Rev. E*, 85 :061110, (2012).
- [190] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Bootstrap percolation on complex networks. *Phys. Rev. E*, 82 (1), (2010).
- [191] R. M. Ziff, E. M. Hendriks, and Ernst.M.H. Kinetics of gelation and universality. *J. Phys. A : Math. Gen*, 16 :2293, (1983).
- [192] Y. S. Cho and B. Kahng. Discontinuous percolation transitions in real physical systems. *Phys. Rev. E*, 84 :050102(R), (2011).
- [193] L. S. Schulman. Long range percolation in one dimension. *Journal of Physics A : Mathematical and General*, 16 :L639, (1983).
- [194] M. Aizenman and C. M. Newman. Discontinuity of the percolation density in one dimensional  $\frac{1}{|x-y|^2}$  percolation models. *Commun. Math. Phys*, 107 :611, (1986).

## BIBLIOGRAPHIE

---

- [195] D. Lee, Y.S. Cho, and B. Kahng. Diverse types of percolation transitions. *J. Stat. Mech. Theor. Exp*, 107 :124002, (2016).
- [196] M.E.J. Newman. *Networks : An Introduction*. University of Michigan and Santa Fe Institute Press, (2010)p.403-408.