



UNIVERSITE ABDELMALEK ESSAÂDI
FACULTÉ DES SCIENCES
TÉTOUAN

Centre d'Etudes Doctorales
« Sciences et Technologies »

Formation Doctorale : « Mathématiques, Physiques et Nouvelles Technologies »

THÈSE

Présentée à la Faculté des Sciences de TÉTOUAN pour obtenir le grade de :

Docteur en Physique
Spécialité : Physique statistique

Titre de la Thèse : CONTRIBUTION À LA MODELISATION DES RÉSEAUX
COMPLEXES

Par :
Ahmed LACHGAR

Soutenue le devant le Jury

Pr. Président

Pr. Rapporteur

Pr. Rapporteur

Pr. Rapporteur

Pr. Examineur

Pr. Examineur

Pr. Abdelfattah ACHAHBAR Université Abdelmalek Essaâdi, Faculté des sciences de Tétouan) Directeur

Structure de recherche accréditée d'accueil : UAE/U11FS
Équipe de Physique de la matière condensée/ l'université Abdelmalek Essaâdi, Faculté des Sciences-Tétouan

This doctoral the

My mother and my father two inexhaustible source
sacrifice. Your prayers and ble
my life.

May Almighty Allah pre
happine

It is al

Remerciement

Français—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.

Mots clés : Réseaux complexes, Théorie des graphes, Réseau en évolution, Distribution des degrés sans échelle, Topologie du réseau, Réseaux petit monde, Transition de phase, Monte Carlo, Champ moyen.

Équipe de Physique de la Matière Condensée,
Département de Physique, Faculté des Sciences,
Université Abdelmalek ESSAADI
B.P 2121, Tétouan, Maroc

English— 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.

Keywords : Complex networks, Graph theory, Rich get richer effect, Evolving network, Scale-free degree distribution, Network topology, Small-world networks, Phase transition, Monte Carlo, Mean field.

Condensed Matter Physics team
Department of Physics, Faculty of Sciences
Abdelmalek ESSAADI University
B.P 2121, Tétouan, Morocco

Table des matières

Résumé	iv
Table des matières	ix
Liste des tableaux	x
Table des figures	xvi
Introduction	1
1 Notions et préliminaire	4
1.1 Les réseaux complexes	5
1.1.1 Introduction	5
1.1.2 Définition	5
1.1.3 Différents types des Réseaux complexes	5
1.2 Les réseaux complexes, la physique moderne et l'unification	9
1.3 La théorie des graphes	11
1.3.1 Bref historique	12
1.3.2 L'expérience de Milgram	12
1.3.3 Représentation d'un graphe	13
1.4 Caractéristiques des réseaux complexes	15

1.4.1	Plus court chemin	15
1.4.2	Coefficient de regroupement (Clustering)	16
1.4.3	Distribution des degrés	17
1.4.4	Degré de corrélation	18
1.4.5	Mesures de la centralité	19
1.5	Propriétés des réseaux réels	20
1.5.1	La propriété petit-monde	20
1.5.2	La distribution des degrés sans échelle	21
1.5.3	Structure communautaire	22
1.6	Les modèles théoriques les plus connus	23
1.7	Recours à la physique statistique	25
2	Modèle de l'attachement préférentiel en faveur des "pauvres"	27
2.1	Introduction	27
2.2	Les processus dynamiques : théorie et simulation	29
2.2.1	Introduction	29
2.2.2	Équation maîtresse	30
2.2.3	Modélisation et simulations numériques	31
2.3	L'évolution dynamique et l'attachement préférentiel dans les réseaux réels	32
2.3.1	L'évolution dynamique	32
2.3.2	Le modèle sans échelle BA	33
2.3.3	L'attachement préférentielle et le mécanisme "rich get richer"	36
2.4	Attachement préférentiel sans l'effet "rich get richer"	37
2.4.1	Le modèle	37
2.4.2	Distribution des degrés en utilisant l'équation maîtresse	38
2.4.3	Comparaison au niveau microscopique avec le modèle de BA	39
2.5	conclusion	44

3	Structure des réseaux sans échelle non corrélés : couches et plus court chemin	45
3.1	Introduction	45
3.2	Caractérisation des réseaux sans échelle	46
3.3	Quelques notions mathématiques de la loi de puissance	47
3.4	Les réseaux sans échelle non corrélés	49
3.5	Structure des couches dans les réseaux sans échelle non corrélés	50
3.5.1	Quelques contributions importantes	50
3.5.2	Nouvelle approche	53
3.6	Plus court chemin	59
3.7	conclusion	62
4	Émergence de la propriété petit-monde dans le modèle de Newman-Watts	63
4.1	Introduction	63
4.1.1	Modèle de Watts-Strogatz	64
4.1.2	Modèle de Newman-Watts	66
4.2	Structure du réseau de NW : couches de voisins et plus court chemin	68
4.2.1	Couches de voisins	68
4.2.2	Plus court chemin	78
4.3	Émergence de la propriété petit-monde dans le modèle de NW	81
4.3.1	Validité de la fonction universelle de NW	81
4.4	Une nouvelle fonction universelle	82
4.5	Conclusion	84
5	Percolation dans les réseaux complexes	86
5.1	Introduction	86
5.2	Percolation	87
5.2.1	Percolation dans le réseau aléatoire d'Erdős-Rényi	89
5.2.2	Percolation explosive	91

5.2.3	Percolation explosive dans d'autres modèles	93
5.3	Nouvelle approche pour prédire le type de la transition dans les systèmes percolatifs	94
5.3.1	État de l'art	94
5.3.2	Nouvelle approche	95
5.4	Conclusion	103
	Conclusion	105
	Bibliographie	107

Liste des tableaux

3.1	Principaux résultats sur le plus court chemin en fonction de γ	60
-----	---	----

Table des figures

1.1	Représentation de quelques types de réseaux composés par des nœuds et des liens.	6
1.2	Affichage hyperbolique de blogs en utilisant à la fois les ensembles de données WWE et ICWSM 2007 (http://datamining.typepad.com/).	6
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).	7
1.4	Le graphe social derrière Facebook, graphe des relations d'amis de 500 million de personnes, image par Paul Butley, 2010.	8
1.5	La structure de l'Internet au niveau des systèmes autonomes. Les nœuds de cette représentation sont des systèmes autonomes et les arêtes 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	12
1.7	Exemple d'un graphe non-orienté et sa matrice d'adjacence M_G avec 5 nœuds et 6 liens.	13
1.8	Exemple d'un graphe orienté et sa matrice d'adjacence M_G avec 5 nœuds et 6 liens.	14
1.9	Exemple d'un graphe non-orienté et sa liste d'adjacence avec 5 nœuds et 6 liens.	14
1.10	Exemple de coefficient de regroupement.	16
1.11	Exemple des réseaux assortatif et disassortatif.	19

1.12	Quelques exemples des réseaux sans échelle, (A) Graphe de collaboration d'acteur avec un nombre des noeuds $n = 212250$ 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) $\gamma_{electrique} = 4$.	21
1.13	Exemples d'un réseau communautaire, on voit trois groupes où chacun est plus connecté par rapport au reste du réseau.	23
1.14	Simple réseau régulier de deux dimensions	24
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	32
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 $m = 2$.	36
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 l'Eq. (2.15). Dans l'encart, nous traçons les mêmes données dans l'échelle log-linéaire.	40
2.4	(a) Évolution de $\langle k_s(t) \rangle$ dans le modèle BA, la ligne continue représente l'Eq. (2.19). (b) Évolution des fluctuations de $\langle k_s(t) \rangle$, la ligne continue représente l'Eq. (2.23). Les cercles joints par des lignes pointillées dans les deux cas sont des données de simulation moyennées sur 20 réalisations pour $m = 2$, $m_0 = 3$.	41
2.5	(a) Évolution de $\langle k_s(t) \rangle$ dans notre modèle, la ligne continue représente l'Eq. (2.26). (b) Évolution des fluctuations de $\langle k_s(t) \rangle$. La ligne continue représente la solution numérique de l'Eq. (2.26) et de l'Eq. (2.28). Les cercles joints par des lignes pointillées dans les deux cas sont des données de simulation moyennées sur 20 réalisations pour $m = 2$, $m_0 = 3$.	43
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.	46

3.2	Les corrélations des degrés dans le réseau aléatoire sans échelle pour différents valeurs de m et de γ . Le nombre de nœuds est $n = 10^4$ et le nombre de réalisations pour chaque simulation est 50.	50
3.3	Illustration du réseau construit. La taille des cercles (nœuds) schématise l'importance des degrés. Le degré maximum de la couche ℓ est K_ℓ	54
3.4	Dépendance de κ_ℓ en ℓ . Les cercles sont les résultats de la simulation d'un réseau sans échelle. Les étoiles représentent la simulation du réseau de la Fig. 3.3. Le nombre de nœuds dans les deux cas est 4×10^6 , chaque point est la moyenne sur 200 réalisations.	55
3.5	Nombre de nœuds dans chaque couche pour différentes valeurs de γ . Les lignes continues correspondent aux Eq. (3.26) et Eq. (3.28). Les symboles sont les simulations numériques d'un réseau de taille $n = 4 \times 10^6$. Chaque point est la moyenne sur 200 réalisations. Les couleurs noir, rouge et bleu représentent respectivement les cas $m = 8, 4$ et 2	57
3.6	Proportion de cycles dans les couches par rapport à m . De haut en bas, γ est respectivement 4, 3, 2.6 et 2.2. Nombre de nœuds $n = 4.10^6$, le nombre de réalisations pour chaque point est 200.	58
3.7	Comparaison entre les données empiriques (cercle) et notre théorie (ligne continue) : Les cercles représentent les couches du réseau d'acteurs Hollywoodien de 392732 acteurs avec $\gamma = 2.17$. L'acteur Kevin Bacon est au centre du réseau. La ligne continue représente l'Eq. (3.26) avec les mêmes valeurs pour γ et n . Les résultats empiriques ont été extraits du site : " http://oracleofbacon.org/center.php ".	59
3.8	Le plus court chemin en fonction du nombre de nœuds. Les valeurs de γ de haut en bas sont respectivement 4, 3, 2.6 et 2.2. Les lignes continues correspondent à l'Eq. (3.32) et à l'Eq. (3.33). Les lignes discontinues correspondent aux équations dans [1]. Chaque simulation est moyennée sur plus de 200 réalisations.	61
4.1	Représentation schématique de l'évolution du processus de "reconnexion" dans le modèle Watts-Strogatz. Le nombre de nœuds est $n = 20$ et $k = 2$	65
4.2	Représentation de la variation de la longueur moyenne du chemin et du coefficient de regroupement avec la probabilité de reconnexion ϕ dans le modèle Watts-Strogatz.	65

4.3	Représentation schématique de l'évolution du processus de "reconnexion" dans le modèle NW. Le nombre de nœuds est $n = 20$ et $k = 2$	67
4.4	Illustration des nœuds appartenant aux couches aléatoires (nœuds sombres) dont la distance au nœud racine R a été changée par les raccourcis. Les nœuds appartenant aux couches réguliers (nœuds clairs) restent à la même distance de R après introduction des raccourcis. Le réseau est de taille $n = 20$, $k = 1$, et avec un seul raccourci.	69
4.5	Représentation du cas $\{1, i - 1\}$. Les nœuds sombres et les nœuds j et R sont les positions qui ne peuvent pas être occupés par le nœud intermédiaire. Étant donné que $\hat{n} = 20$ et $i = 3$, le nombre de positions possibles pour le nœud intermédiaire est $\hat{n} - 2i = 14$	71
4.6	La transformation du GR d'un réseau (a) de $n = 20$ et $k = 2$ vers un réseau (b) de $\hat{n} = 10$ et $\hat{k} = 1$	73
4.7	Fraction des nœuds aléatoires $\frac{S_{al}}{n}$ (noir) et des nœuds réguliers $\frac{S_r}{n}$ (rouge) en fonction du nombre des raccourcis x . Les lignes représentent les deux fonctions $h(x)$ (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 sur 200 réalisations. L'échelle est linéaire-log	77
4.8	Fluctuation σ en fonction de x d'un réseau de $k = 1$ pour différents valeurs de n , $n = 10^3$ (étoile), $n = 10^4$ (carré) et $n = 10^5$ (cercle), le nombre de réalisations est 1000. L'échelle est semi-logarithmique	78
4.9	Comportement de $\frac{\langle \hat{\ell} \rangle}{\hat{n}}$ en fonction de x pour un réseau de taille $n = 10^6$. La formule de Newman et watts Eq. (4.5) (ligne rouge), Eq. (4.41) (ligne verte), Eq. (4.40) (ligne noire) et les simulations numériques pour $k = 1$ (cercle) et $k = 5$ (carré). L'échelle est semi-logarithmique dans a) et log-log dans b) . Chaque simulation est moyennée sur 100 réalisations.	81
4.10	Variations de Δ en fonction de y (a) et x (b) pour différentes valeurs de n et k . $k = 1$ et $n = 10^5$ (triangle), $k = 2$ et $n = 10^4$ (plus), $k = 2$ et $n = 10^6$ (cercle), $k = 2$ et $n = 10^5$ (étoile), $k = 5$ et $n = 10^6$ (carré). Le nombre de réalisations pour chaque simulation est 1000. L'échelle est semi-logarithmique	82

4.11	$\frac{\ln(n)}{\langle \ell \rangle}$ en fonction de y . La ligne continue est l'Eq. (4.46), les symboles représentent 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. L'échelle est semi-logarithmique.	83
4.12	Distance entre deux amas ayant un amas commun (intermédiaire) dans le réseau de NW, avec deux raccourcis (ligne continue) entres eux et $k = 3$. Les lignes en pointillés représentent les liens régulières. Dans a) la distance entre le noeud i et j est 3 car les deux raccourcis ne sont pas liés au même noeud de l' amas intermédiaire, dans b) la distance entre le noeud i et j est 2 car les deux raccourcis sont liés au même noued de l' amas intermédiaire.	85
5.1	Une démonstration de percolation par site sur une grille à deux dimensions pour différentes valeurs de p . Sous le seuil de percolation le système est composé de petits amas, après un certain point critique p_c un amas "infini" occupe la grille.	88
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 $\langle k \rangle = 1$	90
5.3	Distribution des degrés d'un réseau aléatoire ER de degré moyen $\langle k \rangle = 6$, les cercles représentent les simulations numériques et la ligne noire représente l'Eq. (5.2).	90
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.	92
5.5	La distribution des amas selon deux méthodes différentes pour choisir les noeuds dans le même réseau, la première méthode est aléatoire comme le processus ER (cercles noirs), la deuxièmes est le processus Achlioptas [2] (cercles bleus). La ligne en pointillée montre l'exposant $\beta = 2.5$ qu'on trouve lorsque le choix des noeuds est aléatoire. La taille du réseau est $n = 10^6$, chaque simulation est une moyenne sur n réalisations.	97

5.6	Fraction de la composante géante S en fonction de la densité des liens $\langle k \rangle$ pour différentes valeurs de β , les couleurs de gauche à droite représentent les simulations numériques respectivement pour $\beta = 2, 2.5, 3, 4$, les lignes noire représentent la solution numérique de l'Eq. (5.12). Le nombre des nœuds est $n = 10^5$	98
5.7	Estimation de point critique par les simulations numériques (cercles) et par l'Eq. (5.15) (ligne). Les valeurs de β de haut en bas sont respectivement 4, 3, 2.5, 2. Chaque simulation est moyennée sur 500 réalisations.	100
5.8	Illustration de la variation de Δk en fonction de n lorsque la transition est continue. Δk est la valeur de $\langle k \rangle$ quand $S = 0.5$	101
5.9	Illustration de la variation de Δk en fonction de n lorsque la transition est discontinue. Δk est la valeur de $\langle k \rangle$ quand $S = 0.5$	101

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 [3], Internet relie un grand nombre de serveurs et d'ordinateurs dans le monde entier échangeant constamment des quantités gigantesques de paquets d'informations [4], le World Wide Web est un réseau virtuel de sites Web liés avec des hyperliens [5], et les réseaux alimentaires relient également, via des relations trophiques, un grand nombre d'espèces interdépendantes [6, 7]. 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 sans échelle**¹, 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. sans-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. 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 compréhension de quelques aspects fondamentaux des réseaux complexes, en abordant les problèmes concernant la dynamique de la croissance dans le modèle de Barabási-Albert (BA) [5], la structure des réseaux sans échelle, l'émergence de la propriété petit-monde dans le modèle de Newman-Watts [8] (NW), et la caractérisation des transitions de phases de percolation lors de l'apparition de la composante géante dans les réseaux complexes aléatoires.

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, et propose un modèle de croissance 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, et qui sont comparées à ceux du modèle 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 sans 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 claire sur l'émergence de la propriété petit-monde dans ce modèle, on trouve à partir d'une prédiction purement théorique qu'en fonction d'un certain paramètre il y a émergence de la phase petit-monde. Nous étudions également la validité de la fonction universelle usuelle de NW,

et nous introduisons une nouvelle fonction universelle liée à ce paramètre. Le dernier chapitre est consacré à la transition de phase de percolation, qui reste un des problèmes ouverts de la physique statistique. En admettant qu'au point critique les grandeurs pertinentes du système suivent toujours une loi 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 suggérant 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 langage de programmation Fortran 90, 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.

Notions et préliminaire

Sommaire

1.1 Les réseaux complexes	5
1.1.1 Introduction	5
1.1.2 Définition	5
1.1.3 Différents types des Réseaux complexes	5
1.2 Les réseaux complexes, la physique moderne et l'unification	9
1.3 La théorie des graphes	11
1.3.1 Bref historique	12
1.3.2 L'expérience de Milgram	12
1.3.3 Représentation d'un graphe	13
1.4 Caractéristiques des réseaux complexes	15
1.4.1 Plus court chemin	15
1.4.2 Coefficient de regroupement (Clustering)	16
1.4.3 Distribution des degrés	17
1.4.4 Degré de corrélation	18
1.4.5 Mesures de la centralité	19
1.5 Propriétés des réseaux réels	20
1.5.1 La propriété petit-monde	20
1.5.2 La distribution des degrés sans échelle	21
1.5.3 Structure communautaire	22
1.6 Les modèles théoriques les plus connus	23
1.7 Recours à la physique statistique	25

1.1 Les réseaux complexes

1.1.1 Introduction

La recherche sur les réseaux complexes peut être conceptualisée comme se situant à l'intersection entre la théorie des graphes et la mécanique statistique, ce qui lui confère une nature multidisciplinaire. Alors que son origine remonte aux travaux pionniers sur la percolation et les graphes aléatoires de Flory [9] et Erdős-Rényi [10, 11, 12], la recherche dans les réseaux complexes n'est devenue une priorité que récemment. La principale raison est la découverte que les réseaux réels ont des caractéristiques qui ne sont pas expliquées par une connectivité aléatoire uniformément distribuée. En effet, les réseaux réelles possèdent des distributions de degré en loi de puissance et des hubs, parmi d'autres caractéristiques structurelles. Trois développements particuliers ont fortement contribué aux développements en cours : l'étude de Watts et Strogatz sur les réseaux petit monde [13], la caractérisation par Barabási-Albert (BA) des modèles sans échelle [5], et l'identification par Girvan et Newman des structures communautaires présentes dans de nombreux réseaux [14].

1.1.2 Définition

Un **réseau** est une collection de points réunis sous forme de paires, les points sont appelés nœuds ou sommets et les lignes sont appelées liens ou arêtes. Le mot **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, informatiques, biologiques et sociales peuvent être considérés comme des réseaux complexes. Un réseau en général 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 peuvent être accompagnés d'informations supplémentaires pour capturer plus de détails du système.

1.1.3 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 Technologiques, Réseaux Biologiques, Réseaux Sociaux

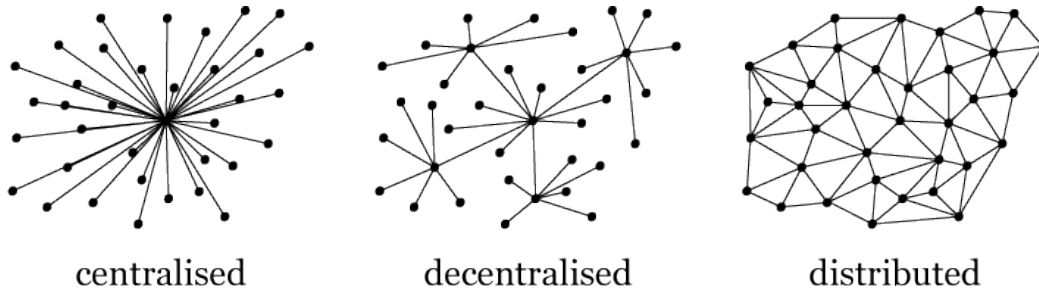


FIGURE 1.1 – Représentation de quelques types de réseaux composés par des nœuds et des liens.

et Réseaux d'Informations.

1.1.3.1 Réseaux Technologiques

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 [15, 16, 17, 18]. Le réseau 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 [4, 19]. 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.

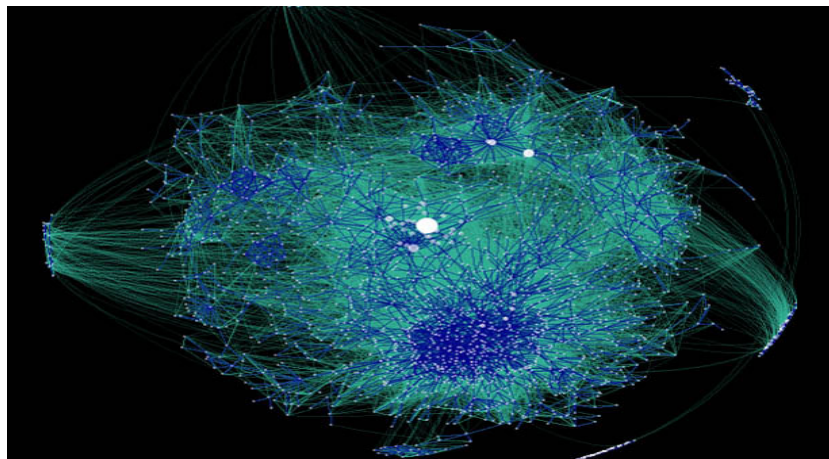


FIGURE 1.2 – Affichage hyperbolique de blogs en utilisant à la fois les ensembles de données WWE et ICWSM 2007 ([http ://datamining.typepad.com/](http://datamining.typepad.com/)).

Il existe un certain nombre d'importantes 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 il accomplit efficacement cette fonction et si nous connaissons la structure du réseau, nous pouvons aborder de nombreuses questions et problèmes de pertinence pratique.

1.1.3.2 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 [20]. Une autre classe de réseau, ce sont les réseaux d'interactions gène-gène, protéines-gène et protéines-protéines [21], 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

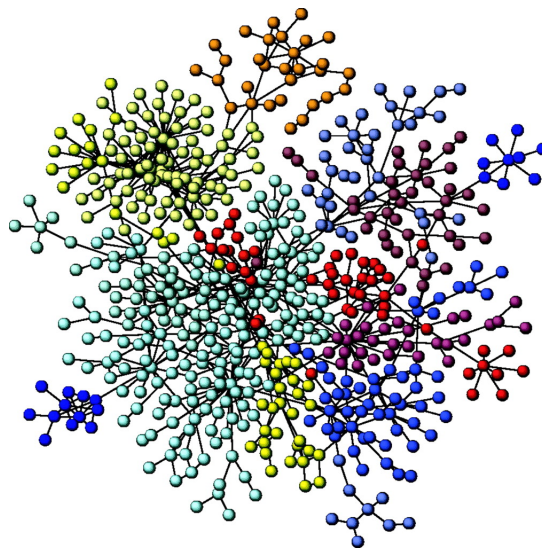


FIGURE 1.3 – Un réseau modulaire est illustré au moyen du protéome humain (données obtenues à partir de la base de données DIP : <http://dip-doe-mbi.ucla.edu>). Les nœuds sont des protéines et des liens indiquent leur interaction physique (protéine-protéine).

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 [3]. 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 le fait que certain classe de réseaux sans échelle est Ultra-small [22, 23, 24]. Nos résultats du troisième chapitre renforcent cette hypothèse et montrent que la dépendance du diamètre avec la taille du réseau peut devenir extrêmement 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*¹ est un des réseaux génétiques qui est bien étudiés dans la littérature [25].

1.1.3.3 Réseaux Sociaux

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 [26], 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 [27, 28], les relations d'affaires entre les entreprises [29]. 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 Facebook qui relie des millions de personnes à travers le monde (voir Fig. 1.4).



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

Bien avant que l'Internet commence à jouer un rôle important dans la vie de beaucoup de

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.

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.

1.1.3.4 Réseaux d'informations

Les réseaux de citation et le World Wide Web (WWW), comme dans la 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 arêtes 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.

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

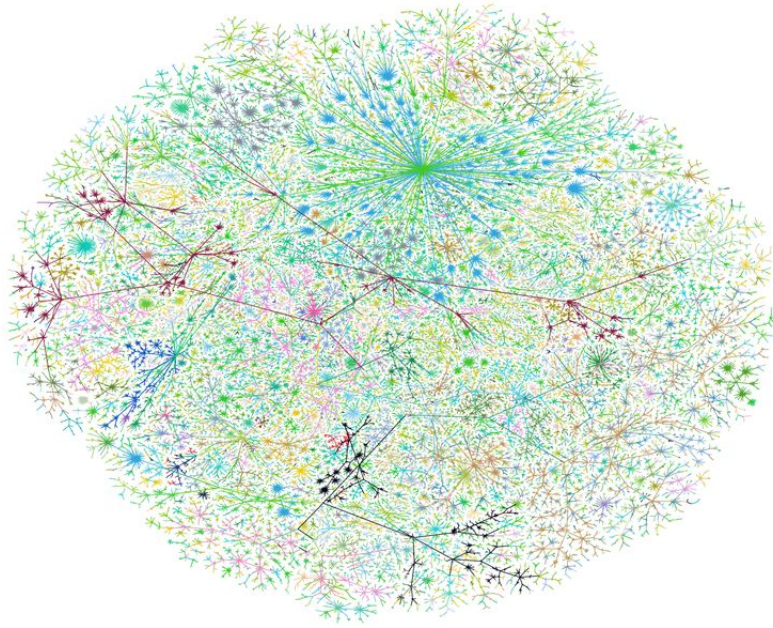


FIGURE 1.5 – La structure de l’Internet au niveau des systèmes autonomes. Les nœuds de cette représentation sont des systèmes autonomes et les arêtes montrent les itinéraires empruntés par les données qui circulent entre eux. La photo a été créée par Hal Burch et Bill Cheswick en 2009.

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 [30, 31, 32, 33]. 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 essaient de découvrir. Une telle **unification** va être, sans doute, un grand pas dans notre compréhension de la nature. En outre, l’existence de cette idée au cœur de l’unification montre le pouvoir mystérieux que les êtres humains peuvent découvrir derrière les apparences de la nature [34]. 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 [30]. 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 [35, 36]. 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 [37], il a été montré que ces réseaux quantiques peuvent présenter une transition de phase de percolation

d'enchevêtrement [38, 39]. 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 [40, 41, 42].

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 est importante pour l'orientation de la direction de l'intelligent, l'équilibrage de charge, la recherche de toutes sortes et le déploiement d'agents² 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 [43, 44, 45].

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 importante et nous n'allons pas fournir une présentation formelle et complète de celle-ci (voir [46, 47] par exemple). 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.

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

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).

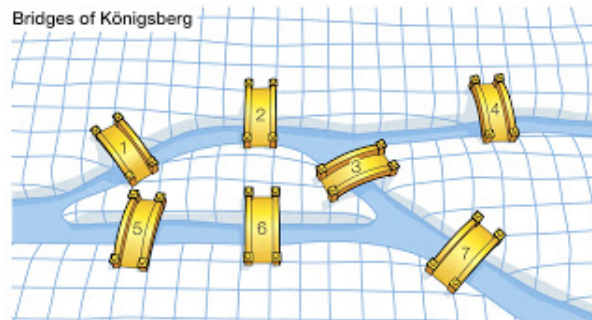


FIGURE 1.6 – Ponts de Königsberg, 1736

Afin de trouver une solution à ce problème 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, ce qui permet 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 [48, 49]). 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 moyenne. 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 [50].

1.3.3 Représentation d'un graphe

Un graphe est un ensemble de points nommés nœuds ou sommets reliés par des traits nommés arêtes (ou liens ou arcs). L'ensemble des arêtes entre nœuds forme une figure similaire à un réseau. Un graphe $G(V, E)$ est alors défini par un couple (V, E) , où V (vertices en anglais) désigne l'ensemble des nœuds et E (edges en anglais) l'ensemble des liens.

1.3.3.1 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-orienté l'élément $M_{i,j} = 1$ si un lien existe entre le i^{eme} et le j^{eme} nœud sinon $M_{i,j} = 0$ (voir Fig. 1.7), ainsi la matrice sera en général symétrique et les éléments diagonaux sont nécessairement 0.

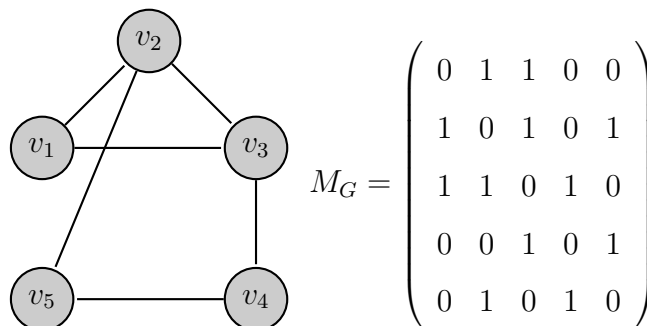


FIGURE 1.7 – Exemple d'un graphe non-orienté et sa matrice d'adjacence M_G avec 5 nœuds et 6 liens.

Un graphe orienté est un graphe où les arêtes sont dirigées, 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).

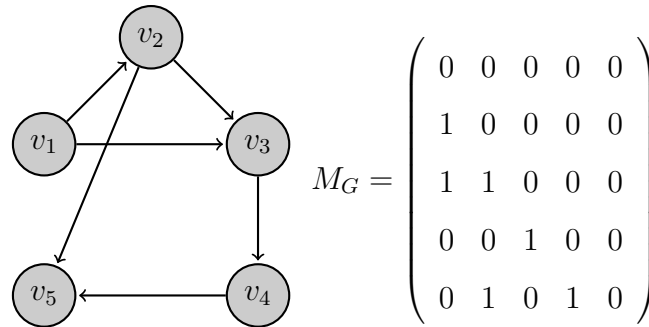


FIGURE 1.8 – Exemple d’un graphe orienté et sa matrice d’adjacence M_G avec 5 nœuds et 6 liens.

1.3.3.2 La liste d’adjacence

Une représentation par liste d’adjacence d’un graphe associe, à chaque sommet du graphe, la collection de ses voisins, comme sommets ou comme arêtes.

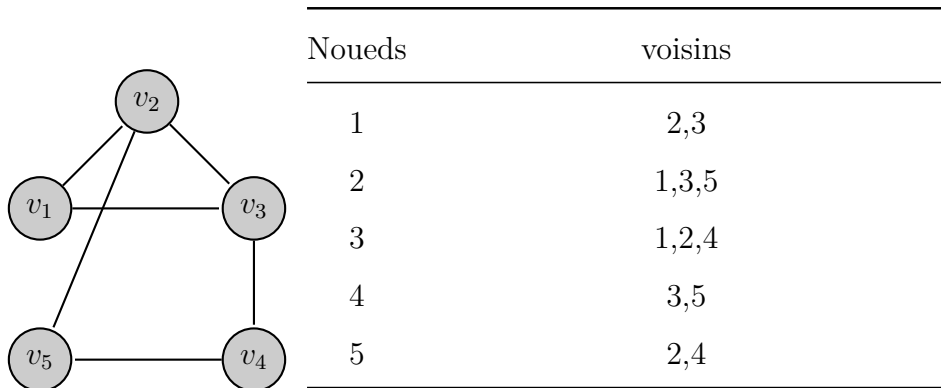


FIGURE 1.9 – Exemple d’un graphe non-orienté et sa liste d’adjacence avec 5 nœuds et 6 liens.

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.

1.4 Caractéristiques des réseaux complexes

Chaque réseau complexe présente des caractéristiques topologiques spécifiques qui caractérisent sa connectivité et influencent fortement la dynamique des processus exécutés sur le réseau. L'analyse et la synthèse des réseaux complexes reposent donc sur l'utilisation de mesures capables d'exprimer les caractéristiques topologiques les plus pertinentes. La description quantitative des propriétés des réseaux fournit des outils fondamentaux pour la classification des réseaux théoriques et réels en grandes catégories. Dans ce paragraphe nous introduisons quelques grandeurs “microscopiques” pour quantifier et analyser les réseaux complexes.

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

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

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 popularisée dans le contexte sociologique où elle est parfois appelée *six degrés de séparation* [48].

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 [51]. 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 [52, 26, 53].

Les ensembles de données massifs du réseau s'accumulent à un rythme énorme dans des champs variés [54]. 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 [55]. 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 [56], en examinant l'apparition de la synchronisation [57] et en évaluant la résilience des réseaux de communication aux attaques et aux échecs [58].

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 Internet [59], les réseaux de transport [60] et les réseaux métaboliques [61, 62]. 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 [52].

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 $0 \leq C \leq 1$.

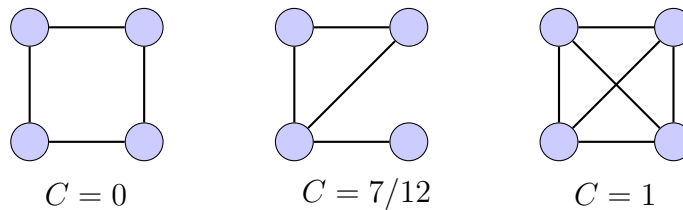


FIGURE 1.10 – Exemple de coefficient de regroupement.

Une autre définition du coefficient de regroupement, également largement utilisé, a été

donnée par Watts et Strogatz [13], 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 [63, 64, 65, 66], 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 arêtes, 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 dans le 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' [67]. 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

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

S'il n'y a pas de corrélation de degré, l'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 [53].

Selon l'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 [68, 63].

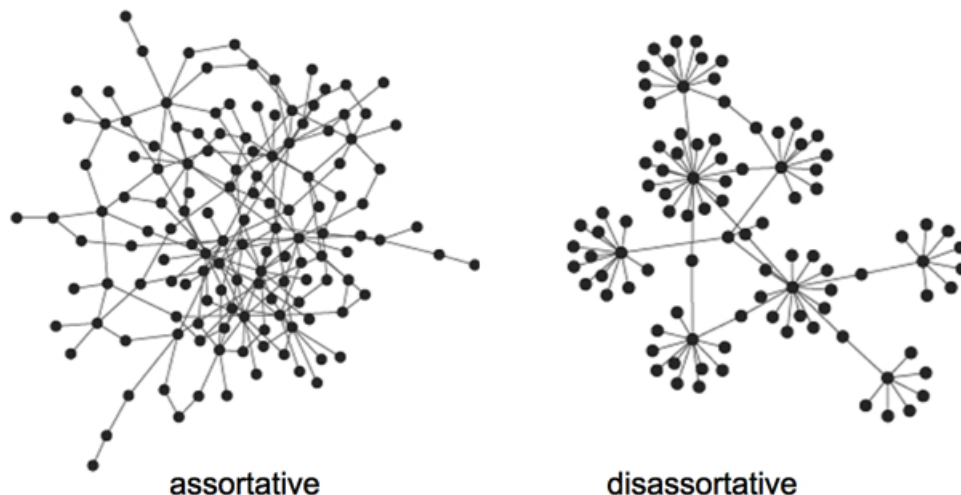


FIGURE 1.11 – Exemple des réseaux assortatif et disassortatif.

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 [69]. 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 [70, 71] 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 [72]. 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'intermédiation 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 σ_{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 mentionné 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 reçoive 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 [73] 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.

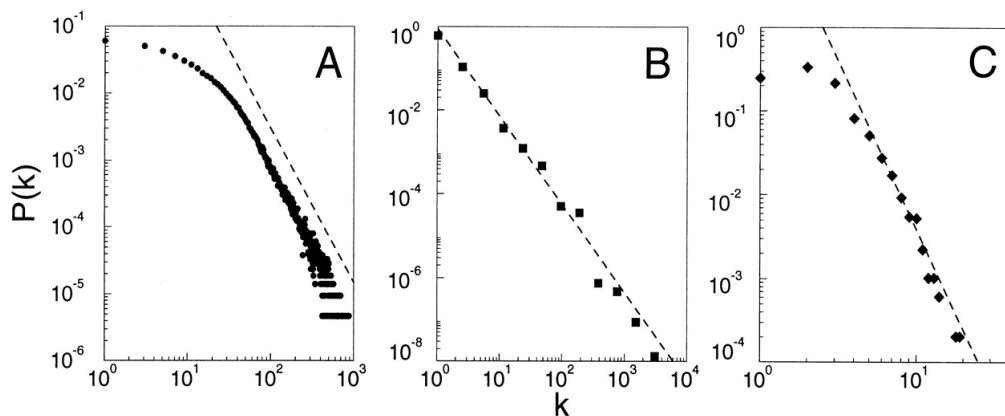


FIGURE 1.12 – Quelques exemples des réseaux sans échelle, (A) Graphe de collaboration d'acteur avec un nombre des nœuds $n = 212250$ et un degré moyen $\langle k \rangle = 28.78$, (B) WWW, $n = 325729$, $\langle k \rangle = 5.46$, (C) Données du réseau électrique, $n = 4941$ et $\langle k \rangle = 2.69$. 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 [74, 75], la taille des tremblements de terre [76], des cratères de la lune [77], la fréquence d'utilisation des mots dans n'importe quelle langue humaine [78, 79], la fréquence de l'apparition de noms personnels dans la plupart des cultures [80], le nombre de documents scientifiques écrits [81], le nombre de citations reçues par les documents [82], le nombre de visites sur les pages Web [83], les ventes de livres, les enregistrements musicaux et presque tous les autres produits de marque [84], le nombre d'espèces dans les genres biologiques [85], les revenus annuels des personnes [86], les collecteurs de réseaux quantiques complexes [87] 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 γ 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 dans la 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 [88, 89]. Dans les réseaux d'interactions protéines-protéines, les communautés sont susceptibles de regrouper des protéines ayant la même fonction spécifique dans la cellule [90, 91].

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

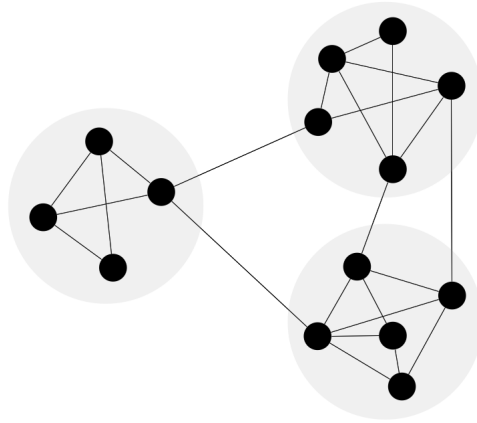


FIGURE 1.13 – Exemples d’un réseau communautaire, on voit trois groupes où chacun est plus connecté par rapport au reste du réseau.

améliorer la performance des services fournis sur le WWW, en ce sens chaque groupe de clients pourrait être servi par un serveur miroir dédié [92]. Les réseaux auto-configurés formés 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 [93].

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 [94].

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.

1.6.0.1 Réseaux simples

Un réseau simple consiste en des connexions régulières entre les nœuds. L'un des exemples les plus exceptionnels 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 [95]. 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 de celles-ci [96, 69].

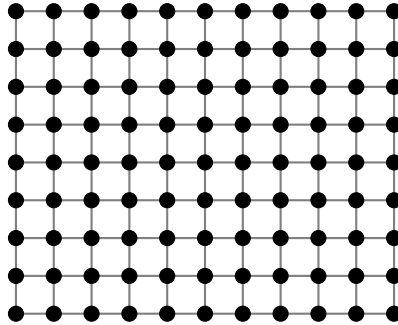


FIGURE 1.14 – Simple réseau régulier de deux dimensions

1.6.0.2 Réseaux aléatoires

Les réseaux aléatoires ont été largement étudiés par Erdős et Rényi (ER) [10, 11, 12]. 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 $E = p \frac{n(n-1)}{2}$.

1.6.0.3 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, c'est-à-dire le coefficient de regroupement. Watts et Strogatz [97] ont réalisé un modèle qui rejoint la propriété petit-monde et ayant 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 avec une probabilité p en prend chaque lien du réseau et on le reconnecte aléatoirement entre deux autres nœuds.

1.6.0.4 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) [5] ou modèle d'attachement préférentiel qui aboutit à des réseaux sans échelle qui sont représentés par une distribution des degrés en 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 . Dans ce cas, le réseau croît en ajoutant à chaque instant un nœud qui se connecte aux nœuds déjà existants avec la probabilité d'attachement préférentielle.

1.7 Recours à la physique statistique

Tous les systèmes mentionnés ci-dessus partagent certaines caractéristiques clés qui les rendent semblables et appartiennent à la classe des réseaux complexes. On s'intéresse à comment les relations entre les nombreux composants de ces systèmes donnent lieu à des phénomènes ou des comportements exprimés par le système dans son ensemble, et non par ses éléments uniques. Ceci indique clairement que la mécanique statistique est un candidat idéal pour l'investigation et la modélisation de tels systèmes. Le but de cette branche de la physique est de décrire l'émergence de certaines propriétés macroscopiques à partir de la description des interactions des nombreux constituants microscopiques. Bien que généralement, dans les relations des système complexes entre les agents n'ont pas besoin d'être microscopique, nous pouvons toujours penser que ces interactions sont sur une échelle différente (plus petite) par rapport à l'échelle des phénomènes qui émergent du système dans son ensemble, et donc la physique statistique peut être utilisée comme un outil puissant pour étudier et dégager les propriétés émergentes pertinentes.

L'attribut complexe ne signifie pas nécessairement que ces types de systèmes sont compliqués. Cette distinction est cruciale car les caractéristiques et le comportement des systèmes complexes peuvent différer sensiblement de ceux des systèmes compliqués. Différentes branches de la science

définissent la complexité de différentes manières selon ce qui correspond le mieux aux objectifs de leur recherche [98, 99, 100, 101]. Pour cette raison, il n’y a pas d’accord unique sur cette terminologie et pour ce qui concerne les systèmes complexes, nous citons ici la définition de Neil Johnson sur la science complexe comme “l’étude des phénomènes qui émergent d’une collection d’objets en interaction“ [102].

Modèle de l'attachement préférentiel en faveur des "pauvres"

Sommaire

2.1	Introduction	27
2.2	Les processus dynamiques : théorie et simulation	29
2.2.1	Introduction	29
2.2.2	Équation maîtresse	30
2.2.3	Modélisation et simulations numériques	31
2.3	L'évolution dynamique et l'attachement préférentiel dans les réseaux réels	32
2.3.1	L'évolution dynamique	32
2.3.2	Le modèle sans échelle BA	33
2.3.3	L'attachement préférentielle et le mécanisme "rich get richer"	36
2.4	Attachement préférentiel sans l'effet "rich get richer"	37
2.4.1	Le modèle	37
2.4.2	Distribution des degrés en utilisant l'équation maîtresse	38
2.4.3	Comparaison au niveau microscopique avec le modèle de BA	39
2.5	conclusion	44

2.1 Introduction

L'étude de l'évolution des réseaux complexes a connu un intérêt croissant durant les dernières années, grâce au développement de modèles qui reflètent certaines propriétés des réseaux réels, par utilisation des techniques de la mécanique statistique, de la théorie des graphes et des simulations numériques [5, 103, 104, 105]. L'une des propriétés les plus importantes étudiée

dans ces réseaux est la distribution des degrés, qui est la probabilité $P(k)$ qu'un nœud ait un degré k . On distingue 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 $P(k) \approx k^{-\gamma}$ où γ est l'exposant caractéristique, et la loi exponentielle $P(k) \approx e^{-\frac{k}{c}}$ 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 Price [?]. Ce dernier a remarqué en 1965 que le nombre de liens vers des articles scientifiques, c'est-à-dire le nombre de citations qu'ils reçoivent, suit une loi de puissance, et donc le réseau de citations est sans échelle. Price n'a cependant pas utilisé le terme réseau sans échelle, qui n'a été inventé que quelques décennies plus tard par BA.

BA a introduit un modèle simplifié basé à la fois sur la croissance et l'attachement préférentiel. L'évolution du réseau consiste à commencer avec m_0 nœuds connectés, et à chaque instant un nouveau nœud est introduit dans le réseau et se connecte à m nœuds déjà bien connectés dans le réseau. Cet attachement préférentiel se réalise avec la probabilité $\Pi(k_i) = \frac{k_i}{\sum_j k_j}$, où i est l'un des m nœuds choisis pour se connecter au nouveau nœud et k_i est son degré. Cette probabilité se traduit par le fait que les nœuds les plus connectés auraient toujours plus de chances de se connecter aux nouveaux nœuds, ou selon le principe de Pareto [106], les riches deviennent encore plus riche “rich get richer”.

Comme nous l'avons déjà signalé (voir la section 1.5.2), les réseaux sans échelle sont largement observés dans une variété de systèmes. Cependant, il existe d'autres réseaux réels qui suivent une loi exponentielle, par exemple, le Réseau mondial de transport maritime [107], le réseau électrique des USA [108], réseaux de neurones des *C.elegans* [109] et le réseau de messagerie à l'Université de Rovira i Virgili (ENURV) en Espagne [108].

On croit que la distribution des degrés d'un réseau complexe en croissance, suit une loi exponentielle, lorsque les nouveaux nœuds ajoutés au réseau sont liés aléatoirement avec une probabilité uniforme aux nœuds déjà existants. D'autre part, la loi de puissance se manifeste lorsque les nœuds ajoutés au réseau sont liés à des nœuds déjà bien connectés avec la probabilité d'attachement préférentielle.

De nombreuses idées sur la formation des réseaux ont été examinées. Par exemple, Barabási a affirmé que l'attachement préférentiel et la croissance sont tous les deux nécessaires pour générer un réseau sans échelle [110]. Actuellement, on sait que la croissance n'est pas nécessaire pour un tel objectif, et que un réseau statique avec un nombre de nœuds fixe peut avoir une distribution de degré en loi de puissance [111]. En outre, il est intuitif, que l'attachement préférentiel sans

l'effet "rich get richer" ne génère pas un réseau sans échelle [112]. Krapivsky et al [113] ont étudié l'attachement préférentiel non linéaire avec $\Pi(k_i) \propto k^\gamma$ et ils ont montré que pour $\gamma < 1$, la distribution des degrés suit une loi exponentielle.

Malgré de nombreux efforts, une théorie cohérente et rigoureuse des réseaux en croissance n'est pas encore établie, et il n'y a pas encore de principe général pour prédire la topologie finale d'un réseau en évolution. Dans le but de comprendre la formation et l'évolution des réseaux complexes, plusieurs modèles ont été introduits pour étudier les processus microscopiques impliqués dans les réseaux en formation.

Dans ce contexte, on introduit dans ce chapitre un modèle de réseau complexe qui croît avec un mécanisme d'attachement préférentiel favorisant les noeuds les moins connectés. L'objectif est de vérifier d'une part si la distribution des degrés en loi de puissance persiste en l'absence du mécanisme "rich get richer" et d'étudier, d'autre part, les différences microscopiques entre les réseaux sans échelle (hétérogènes) et les réseaux homogènes.

Avant d'aborder notre modèle, nous allons traiter en détail le modèle de BA, ce qui permettra d'introduire les différentes méthodes et techniques essentielles dans nos calculs.

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 dynamiques et stationnaires des réseaux, et à définir les approches et techniques de base. En particulier, on va introduire le formalisme de l'équation maîtresse (EM).

En général, la solution de EM n'est pas facile même pour des processus dynamiques très simples. Pour cette raison, nous présentons des techniques qu'on va utiliser au long de cette thèse, telles que les approximations déterministes du champ moyen, qui représentent des approches pour comprendre les caractéristiques qualitatives du processus étudié, et la modélisation avec la méthode de Monte Carlo qui est généralement implémentée dans des simulations numériques 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 elle est actuellement considérée comme un paradigme général pour combler l'écart entre les propriétés locales et les propriétés à grande échelle des systèmes complexes.

2.2.2 Équation maîtresse

Le nom "équation maîtresse" a été inventé à l'origine par Nordsieck, Lamb et Uhlenbeck [114] dans leur étude du modèle Furry des averses de pluie cosmiques [115]. Peu de temps auparavant, Feller a appliqué une équation de la même structure à la croissance des populations [116], et Delbrück à des réactions chimiques auto-catalytiques bien mélangées [117]. Pour plus de détails voir les références [118, 119, 120].

La description dynamique du système est obtenue en introduisant pour chaque constituant i la variable correspondante σ_i , qui décrit un attribut particulier de i , et caractérise son état dynamique. Nous pouvons énumérer tous les états possibles $\sigma_i = 1, 2, \dots, m$ pour chaque constituant, on désigne alors une configuration particulière du système 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 constituants du système 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 σ^a vers un autre état σ^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 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 σ . 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 σ et les termes $W(\sigma' \rightarrow \sigma)$ représentent les taux de transition de la configuration σ' vers la configuration σ 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 [121, 122, 123, 124],

des réseaux régulateurs de gènes [125, 126, 127], de la propagation des maladies [128, 129], de l'homéostasie épidermique [130], et les processus sociaux et économiques [131]. 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 [132].

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 élément individuel est supposé être dans l'un des états possibles. À chaque pas de temps, la procédure de mise à jour spécifique au modèle qui dépend de la dynamique microscopique est appliquée à chaque élément, ce qui modifie son état en fonction de l'état de ses voisins ou d'autres règles dynamiques. Notamment, la stochasticité du modèle peut être introduite en utilisant des simulations 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 du 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

2.3.1 L'évolution dynamique

Le modèle de réseau aléatoire ER, 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 à se développer [133]. Par conséquent, si nous souhaitons modéliser

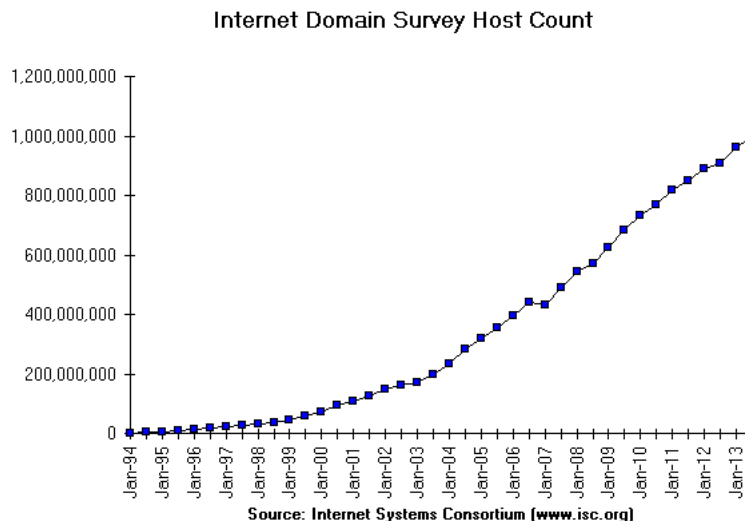


FIGURE 2.1 – L'évolution de nombre d'hôtes sur Interent de 1994 à 2013, cette figure obtenues à partir de site : <http://facesncups.com/inforev.html>

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 BA

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, 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, 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 nous intéressons à 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 BA, 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 préférentiel** : 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 BA. La méthode la plus simple est une analyse de champ moyen décrite premièrement dans [5]. D'autres méthodes plus

rigoureuse d'analyse de ce modèle utilisant des outils de la physique statistique comprennent l'approche de l'équation maîtresse [134] et l'approche de l'équation du taux [113]. 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 temps¹.

L'attachement préférentiel d'un nouveau nœud avec un ancien nœud dans le réseau de degré k est selon l'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 se 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 se lie à un nœud de degré $(k - 1)$ 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 $\frac{k}{2mt} \times m \times nP(k, t) = \frac{k}{2}P(k, t)$, 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 présidentes 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

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

arguments que nous avons suivi pour l'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 l'Eq. (2.4) et l'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 l'Eq. (2.6) on peut écrire

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

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

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

L'Eq. (2.9) a d'abord été dérivée par Krapivsky et al. [113] et de manière indépendante par Dorogovtsev et al. [134]. Un traitement plus détaillé a ensuite été donné par Bollobas et al. [135], 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 $L \sim \frac{\log n}{\log \log n}$ [135]. 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 .

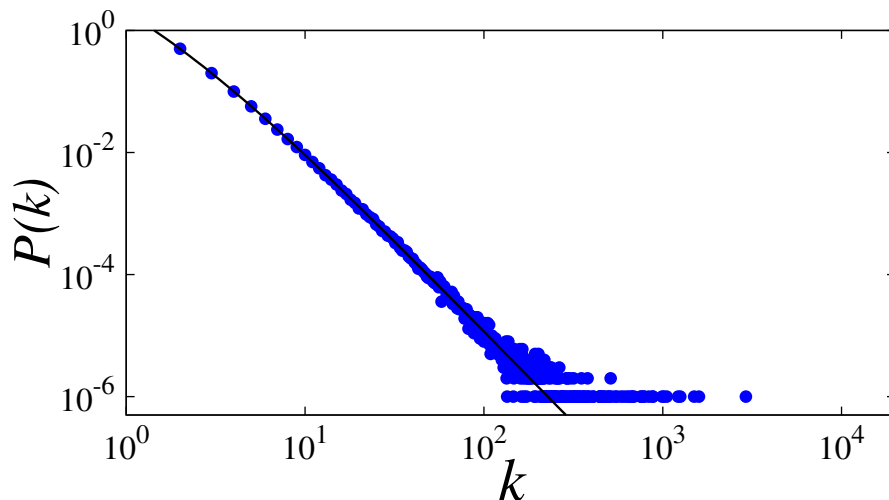


FIGURE 2.2 – La distribution des degrés du réseau BA en échelle logarithmique, où les cercles sont les simulations numériques et la ligne continue représente la formule théorique, Eq. (2.9). La taille du réseau est $n = 10^6$ et $m = 2$.

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 est 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 nous lier à un nœud de haut niveau qu'à un nœud avec seulement quelques liens [136].

L'ancienneté, cependant, n'est pas suffisante pour expliquer la distribution sans 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 distributions sans échelle observées dans les réseaux réels.

Dans la section suivante on va introduire un modèle de réseau qui croît avec l'attachement préférentiel, mais en faveur des nœuds les moins connectés, et on va le comparer quantitativement et "microscopiquement" avec le modèle de BA.

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

2.4.1 Le modèle

Malgré de nombreux efforts, une théorie cohérente et rigoureuse des réseaux en croissance n'est pas encore établie, et il n'y a pas encore de principe général pour prédire la topologie finale d'un réseau en évolution. Dans le but de comprendre la formation et l'évolution des réseaux complexes, plusieurs modèles ont été introduits pour étudier les processus microscopiques impliqués dans les réseaux en formation

Dans ce contexte, on introduit un modèle de réseau complexe qui croît avec un mécanisme d'attachement préférentiel favorisant les nœuds les moins connectés. L'objectif est de vérifier d'une part si la distribution des degrés en loi de puissance persiste en l'absence du mécanisme "rich get richer" et d'étudier, d'autre part, les différences microscopiques entre les réseaux sans échelle (hétérogènes) et les réseaux homogènes. À l'instar du modèle BA original, notre réseau évolue selon deux mécanismes : la croissance et l'attachement préférentiel. Les nœuds entrants 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 des 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 aurons-nous 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 noeud 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 noeud a été créé et représente également le nombre de noeuds ajoutés au réseau.

2.4.2 Distribution des degrés en utilisant l'équation maîtresse

Notons $n(k, t)$ le nombre de noeuds 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 noeud au réseau, nous avons $n = t$. C'est, à tout moment, le nombre total de noeuds 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 noeuds, et mm_0 capture le fait qu'à l'instant initial nous commençons par m_0 noeuds, chacun avec m liens. Notre objectif est de calculer les changements dans le nombre de noeuds de degré k après l'ajout d'un nouveau noeud au réseau. Pour cela, nous respectons les deux événements qui modifient $P(k, t)$ suite à l'arrivée d'un nouveau noeud :

- Un nouveau noeud peut être lié à un noeud de degré k , le transformant en un noeud de degré $(k + 1)$, ce qui diminue $P(k, t)$.
- Un nouveau noeud peut être lié à un noeud de degré $(k - 1)$, le transformant en un noeud 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 δ 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 la 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 $t = 2 \cdot 10^6$, nous prenons alors $t \gg j$ dans l'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 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, et il doit y avoir un "traitement de faveur" pour les riches pour produire la distribution en loi de puissance.

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 à celui de BA. La distribution des degrés n'est pas suffisante pour caractériser

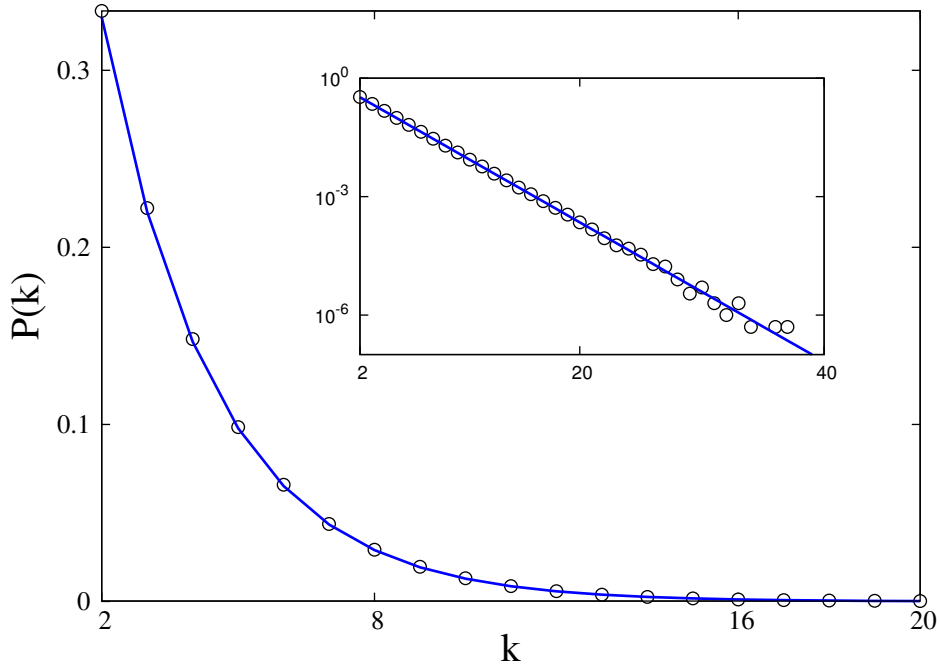


FIGURE 2.3 – les résultats de la simulation (cercles) pour $N = 2.10^6$, $m = 2$, $m_0 = 3$, et la solution numérique (ligne continue) pour l'Eq. (2.15). Dans l'encart, nous traçons les mêmes données dans l'échelle log-linéaire.

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 avec le champ moyen : $\frac{\partial k_i(t)}{\partial t} = m \Pi(k_i)$, qui donne $k_i(t) = m \left(\frac{2t+m_0}{2t_i+m_0} \right)^{\frac{1}{2}}$.

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

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 γ 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 Fig. 2.4, entre l'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 [137, 138] puisque la majorité des nœuds ont un faible degré et le poids des hubs est faible.

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

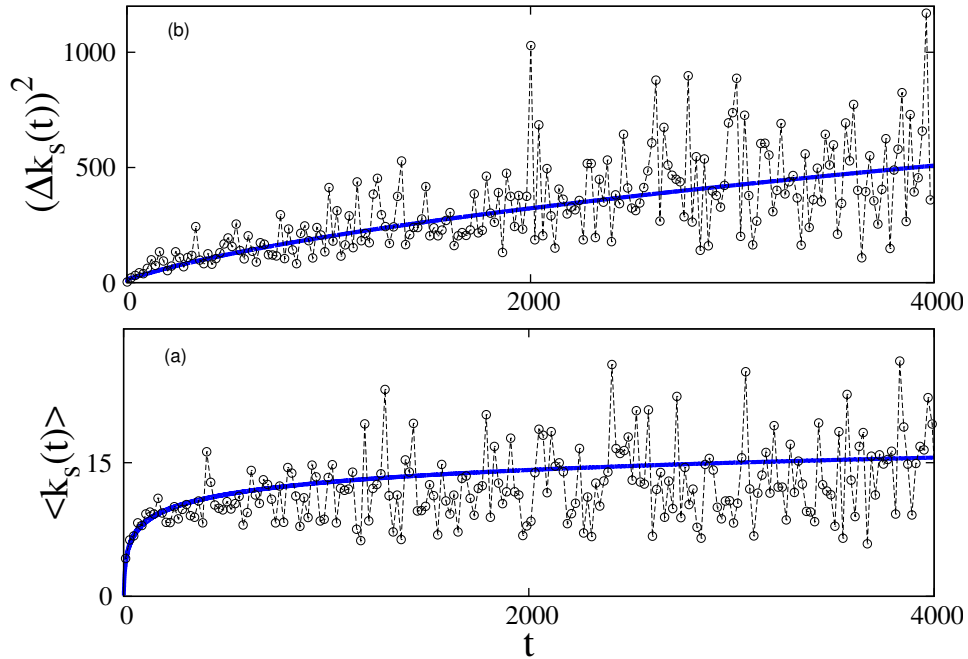


FIGURE 2.4 – (a) Évolution de $\langle k_s(t) \rangle$ dans le modèle BA, la ligne continue représente l'Eq. (2.19). (b) Évolution des fluctuations de $\langle k_s(t) \rangle$, la ligne continue représente l'Eq. (2.23). Les cercles joints par des lignes pointillées dans les deux cas sont des données de simulation moyennées sur 20 réalisations pour $m = 2$, $m_0 = 3$.

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

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

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

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

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

Les fluctuations de $< k_s(t) >$ sont données par

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

qui deviennent arbitrairement large lorsque le temps augmente suffisamment.

Les données de simulation selon l'Eq. (2.23) (voir Fig. 2.4(b)) montrent la tendance croissante des fluctuations de $< k_s(t) >$. Cela peut s'expliquer par le fait que le degré maximum dans le réseau $k_{max} \sim \sqrt{t}$ [138] augmente plus vite que $< k_s(t) > \sim \ln(t)$ (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 $A(t) = \int_1^t \frac{\left(\frac{t' + m_0 - 1}{2t' + m_0} \right)^{-\frac{1}{m_0-2}}}{t' + m_0 - 1} dt'$.

Or la valeur moyenne du nœud sélectionné $< k_s(t) >$ est :

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

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

Pour les temps 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

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

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

Le second moment est obtenu en substituant les expressions correspondantes de $\Pi(k_i)$ et $k_i(t)$ dans l'Eq. (2.20), nous obtenons pour les temps grands

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

En utilisant 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 $\langle k_s^2(t) \rangle \approx 5m^2$. L'expression finale des fluctuations est alors

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

Ce résultat, additionné avec $\langle k_s(t) \rangle \approx 2m$, montrent que presque tous les nœuds ont le même degré comme le montre la 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 encore plus ni aux pauvres de devenir riches.

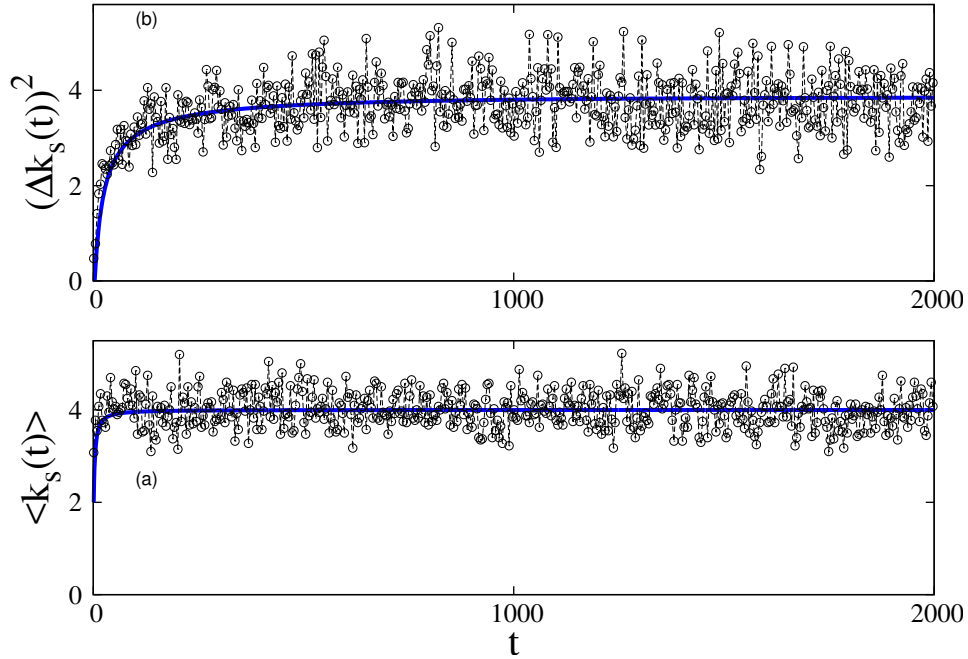


FIGURE 2.5 – (a) Évolution de $\langle k_s(t) \rangle$ dans notre modèle, la ligne continue représente l'Eq. (2.26). (b) Évolution des fluctuations de $\langle k_s(t) \rangle$. La ligne continue représente la solution numérique de l'Eq. (2.26) et de l'Eq. (2.28). Les cercles joints par des lignes pointillées 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 le critère d'attachement préférentiel et 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. 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.

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

Sommaire

3.1	Introduction	45
3.2	Caractérisation des réseaux sans échelle	46
3.3	Quelques notions mathématiques de la loi de puissance	47
3.4	Les réseaux sans échelle non corrélés	49
3.5	Structure des couches dans les réseaux sans échelle non corrélés	50
3.5.1	Quelques contributions importantes	50
3.5.2	Nouvelle approche	53
3.6	Plus court chemin	59
3.7	conclusion	62

3.1 Introduction

L'objectif principal de ce chapitre vise à prédire la structure des couches des réseaux sans échelles en nous basant sur les propriétés locales régissant les sommets individuels.

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 pas encore de formalisme théorique bien établi [63]. Dans ce chapitre nous allons expliquer notre contribution à la compréhension de

la structure intrinsèque des réseaux sans échelle. En particulier, nous allons calculer le nombre moyen de voisins d'un nœud donné se trouvant à une distance précise, et nous allons en déduire le plus court chemin dans le réseau.

3.2 Caractérisation des réseaux sans échelle

Un réseau sans échelle (ou réseau invariant d'échelle, ou encore scale-free network en anglais) est un réseau dont la distribution des degrés suit une loi de puissance, c'est-à-dire, $P(k) = Ck^{-\gamma}$, où C est une constante. Donc, dans un tel réseau, la proportion de nœuds de degré k est proportionnelle à $k^{-\gamma}$. Cette loi traduit l'absence d'échelle caractéristique. En effet, si on multiplie k par une constance α , $P(\alpha k) = C\alpha^{-\gamma}k^{-\gamma} = C'k^{-\gamma}$, donc les propriétés sont les mêmes si on change l'échelle.

Identifier la loi de puissance dans les systèmes naturels ou artificiels peut être difficile. La stratégie standard utilise le fait que la courbe d'une distribution sans échelle apparaît comme une ligne droite lorsqu'elle est tracée sur des échelles logarithmiques (voir Fig. 3.1). Cependant,

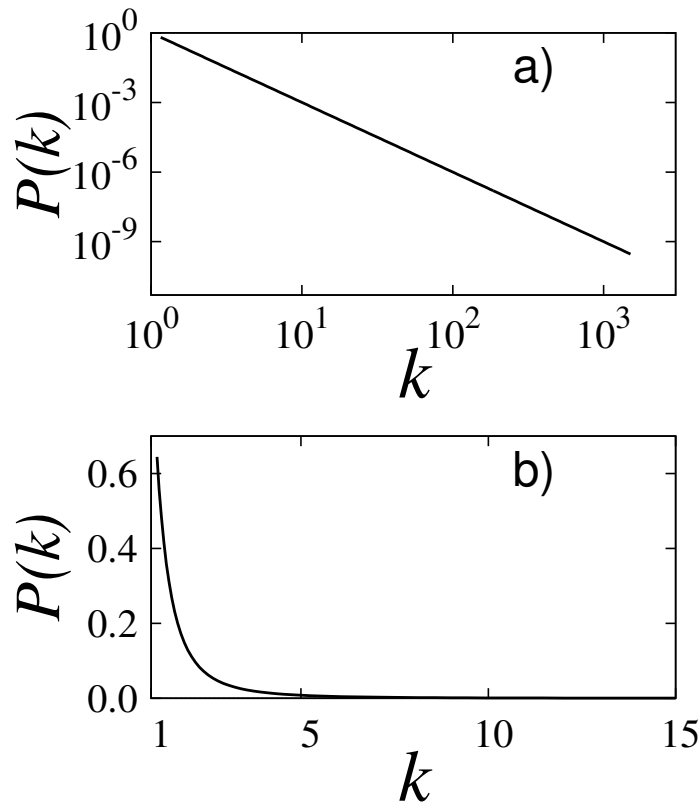


FIGURE 3.1 – Illustration de la loi de puissance pour une fonction, $P(k) = k^{-\gamma}$, avec l'exposant $\gamma = 3$. Dans a) l'échelle est logarithmique et dans b) l'échelle est linéaire.

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 (voir Fig. 2.2). Du fait que la distribution diminue dans l'extrémité, chaque valeur mesurée ne contient que quelques échantillons, et les fluctuations dans les valeurs mesurées sont grandes, et cela apparaît comme une courbe bruyante sur la figure. Une façon de traiter cela est de simplement 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. Le nombre d'échantillons dans une largeur Δx devrait être divisé par Δx pour obtenir une moyenne par intervalle unitaire de x . Ensuite, le nombre d'échantillons normalisé devient indépendant de la largeur de la valeur mesurée en moyenne, et nous sommes libres de varier les largeurs Δx comme nous le souhaitons. Le choix le plus commun est de créer des largeurs telles que chacune est un multiple fixe plus large que celui qui le précède, ce qui est connu sous le nom "logarithmic binning". Ceci réduit les erreurs statistiques dans la queue. En outre, les valeurs mesurées semblent être de largeur constante quand nous traçons la courbe sur des échelles logarithmiques.

3.3 Quelques notions mathématiques de la loi de puissance

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

$$P(k) = Ck^{-\gamma}. \quad (3.1)$$

Comme nous avons vu dans la Section 1.5.2, l'exposant γ et la variable k sont positifs.

Dans les réseaux réels, les distributions de degrés ne suivent généralement pas une loi de puissance pure sur toute leur gamme. En regardant la Fig. 1.12, par exemple, nous observons que la distribution de degrés n'est pas toujours monotone pour les petits k , même en tenant compte des fluctuations statistiques dans l'histogramme. Quand on dit qu'un réseau a une distribution de degré sans échelle, on se réfère implicitement à la queue de la distribution, qui doit avoir cette forme. Cependant, et dans certains cas, la distribution peut s'écarter de la loi d'échelle pour k élevé. En fait, il y a parfois une coupure qui limite le degré maximum de nœuds dans la queue.

La constante C dans l'Eq. (3.1) est la constante de normalisation qui s'obtient en intégrant :

$$P(k) = \int_m^\infty C k^{-\gamma} dk = 1, \quad (3.2)$$

où m est le degré minimum dans le réseau, on obtient

$$C = (\gamma - 1)m^{\gamma-1}, \quad (3.3)$$

alors l'Eq. (3.1) devient :

$$P(k) = (\gamma - 1)m^{\gamma-1}k^{-\gamma}. \quad (3.4)$$

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

$$\begin{aligned} \langle k^j \rangle &= \int_m^\infty k^j P(k) dk \\ &= (\gamma - 1)m^{\gamma-1} \int_m^\infty k^{j-\gamma} dk. \end{aligned} \quad (3.5)$$

Le premier moment $\langle k \rangle$ est le degré moyen de k , notons que $\langle k \rangle$ devient infini si $\gamma \leq 2$. Le second moment $\langle k^2 \rangle$ mesure les fluctuations de k , il diverge si $\gamma \leq 3$. Sachant que dans les réseaux réels l'exposant γ est souvent dans l'intervalle $2 < \gamma < 3$, ceci implique que les fluctuations divergent dans ces réseaux.

En réalité, dans un réseau, les degrés sont des entiers naturels, et dans nos calculs les intégrales doivent être remplacées par des sommes. Mais, puisque la loi d'échelle n'est vérifiée que pour les grands degrés, ceux-ci peuvent être considérés comme représentant une variable continue.

Techniquement, si k est une variable discrète, et $P(k) = Ck^{-\gamma}$, alors la constante de normalisation C doit vérifier $\sum_{k=1}^\infty Ck^{-\gamma} = 1$, donc $C = \frac{1}{\zeta(\gamma)}$, et $\zeta(\gamma)$ est la fonction zêta de Riemann. Donc la loi de puissance, si k est discret, s'écrit :

$$P(k) = \frac{1}{\zeta(\gamma)} k^{-\gamma}. \quad (3.6)$$

Dans les calculs on a souvent besoin de connaître la valeur du degré maximum K , mais dans un réseau dans échelle en croissance, ce n'est pas possible de connaître sa valeur exacte. Pour estimer la valeur de K , on suppose que dans un réseau fini avec n nœud, on peut avoir au plus un nœud dans l'intervalle $[K, \infty[$, c'est-à-dire, la probabilité d'observer un nœud avec un degré supérieur à K est $\frac{1}{n}$, donc :

$$\int_K^\infty P(k) dk = \frac{1}{n}, \quad (3.7)$$

d'où on obtient $K = mn^{\frac{1}{\gamma-1}}$.

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 corrélations existent en fonction des propriétés des nœuds. Les principaux types de corrélations étudiées sont les corrélations degré-degré (voir Section.1.4.4). Même si la construction de ces réseaux sans échelle aléatoires se fait au départ sans aucune corrélation, cela ne signifie pas que le réseau n'affichera pas des corrélations de degré, c'est-à-dire l'absence de corrélations lors de la création des réseaux n'entraîne pas l'absence de corrélations dans le réseau formé. La Fig. 3.2 indique que les réseaux aléatoires sans échelle présentent des corrélations de degré, allant de l'assortativité à la dissassortativité selon la valeur de l'exposant γ . Nous observons trois situations :

- i) $\gamma > 3$, le réseau est dit assortatif. Dans ce cas, la corrélation entre les degrés des sommets et ceux de leurs voisins est élevée. Autrement dit, les sommets ayant un degré élevé sont connectés entre eux ; les sommets ayant un degré faible sont connectés entre eux et il y a peu de liens entre les sommets de degré élevé et ceux de degré faible
- ii) $\gamma = 3$, le réseau est neutre.
- iii) $\gamma < 3$, le réseau est dissassortatif. C'est le cas contraire à l'assortativité. Les sommets voisins sont peu corrélés, et les sommets avec un degré élevé sont connectés à ceux avec un degré faible.

Générer 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 ces réseaux corrélés consiste à mélanger les liens en utilisant une sorte d'algorithme de type Metropolis¹ [139].

Cependant, la négligence de ces corrélations n'empêche jamais de trouver des résultats importants, qui nous aideront à bien modéliser ces réseaux réels et à mieux comprendre leurs structures.

1. 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. 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.

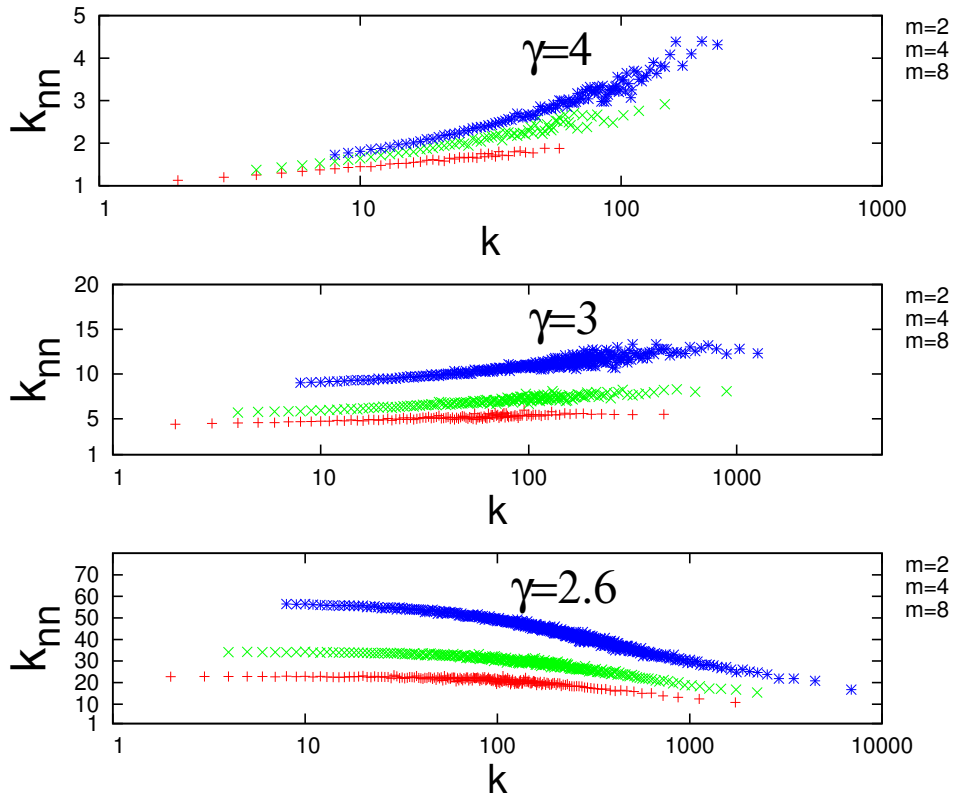


FIGURE 3.2 – Les corrélations des degrés dans le réseau aléatoire sans échelle pour différents valeurs de m et de γ . Le nombre de nœuds est $n = 10^4$ et le nombre de réalisations pour chaque simulation est 50.

3.5 Structure des couches dans les réseaux sans échelle non corrélés

3.5.1 Quelques contributions importantes

Newman est l'un des premiers à chercher une expression explicite pour les couches dans les réseaux complexes [140], où le mot couche désigne le nombre moyen de nœuds à une distance donnée d'un nœud arbitraire choisi au hasard, et qu'on appelle nœud racine.

Le calcul de Newman se base sur la fonction génératrice qui suscite un intérêt particulier dans les réseaux complexes. En effet, celle-ci donne toute l'information sur la distribution des degrés $P(k)$. La fonction génératrice est le polynôme :

$$g(z) = P(0) + P(1)z + P(2)z^2 + \dots = \sum_{k=0}^{\infty} P(k)z^k. \quad (3.8)$$

Si on connaît la fonction génératrice, on peut récupérer les valeurs de $P(k)$ en différenciant :

$$P(k) = \frac{1}{k!} \left. \frac{d^k g(z)}{dz^k} \right|_{z=0}. \quad (3.9)$$

Ainsi, la distribution de probabilité et la fonction génératrice ne sont en réalité que deux représentations différentes de la même chose. Dans de nombreux cas, il est plus facile de travailler avec la fonction génératrice qu'avec la distribution de probabilité.

Soit $p_k^{(2)}$ la probabilité qu'un nœud a exactement k seconds voisins dans le réseau

$$p_k^{(2)} = \sum_{s=0}^{\infty} P(s) P^{(2)}(k/s), \quad (3.10)$$

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

La fonction génératrice de cette probabilité s'écrit :

$$\begin{aligned} g^{(2)}(z) &= \sum_{k=0}^{\infty} p_k^{(2)} z^k \\ &= g_0(g_1(z)), \end{aligned} \quad (3.11)$$

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 utiliserons la notation g' , car elle s'avère beaucoup moins lourde que la notation plus courante $\frac{dg}{dz}$).

En généralisant la procédure, Newman a trouvé la fonction génératrice du nombre de voisins à n'importe quelle distance ℓ , qui s'écrit :

$$g^{(\ell)}(z) = \sum_{s=0}^{\infty} p_s^{(\ell-1)} \sum_{k=0}^{\infty} P^{(\ell)}(k/s) z^k \quad (3.12)$$

$$= g^{(\ell-1)}(g_1(z)), \quad (3.13)$$

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 difficile d'extraire les expressions explicites des probabilités à partir de l'équation précédente, et ce même pour le cas des seconds voisins. Cependant, pour une distribution de degrés suivant la loi de Poisson avec une moyenne $\langle k \rangle$, $P(k) = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}$, et si on suppose que $z = 1$, le calcul se simplifie pour le nombre de seconds voisins, et on trouve :

$$n_2 = g'_0(1) g'_1(1), \quad (3.14)$$

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

$$n_2 = \langle k^2 \rangle - \langle k \rangle. \quad (3.15)$$

Le nombre moyen de voisins n_ℓ à n'importe quelle distance ℓ , pour une distribution de poisson, est obtenu à partir de l'Eq. (3.13) :

$$\frac{dg^{(\ell)}}{dz} = g^{(\ell-1)'}(g_1(z))g_1'(z), \quad (3.16)$$

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

$$\begin{aligned} n_\ell &= g^{(\ell-1)'}(1)g_1'(1) \\ &= n_{\ell-1}g_1'(1), \end{aligned} \quad (3.17)$$

de l'Eq. (3.14), et sachant que $g_0'(1) = \langle k \rangle = n_1$ et que $g_1'(1) = \frac{n_2}{n_1}$, on obtient :

$$n_\ell = n_{\ell-1} \left(\frac{n_2}{n_1} \right), \quad (3.18)$$

nous pouvons écrire finalement que la distribution des nœuds voisins à une distance ℓ d'un nœud choisi au hasard est :

$$n_\ell = \left(\frac{n_2}{n_1} \right)^{\ell-1} n_1. \quad (3.19)$$

Cela signifie que n_ℓ augmente ou diminue exponentiellement avec ℓ selon que n_2 est supérieur ou inférieur à n_1 .

La solution de Newman pour la $\ell^{\text{ème}}$ couche est une fonction monotone, c'est-à-dire soit croissante soit décroissante, ce qui est loin du comportement des réseaux réels [138].

Une autre contribution importante dans le sujet est celle de Cohen et al. [141, 142]. Le calcul dans ce cas est basé sur le travail de Bollobás [143]. On commence par attribuer à chaque nœud un degré, et donc des liens disponibles. Le réseau est ensuite construit en commençant par connecter le nœud de degré maximum K aléatoirement avec K liens du réseau, ce qui constitue les nœuds de la première couche. Dans la deuxième étape, les liens disponibles de la première couche se connectent avec d'autres liens disponibles d'une façon aléatoire, remplissant ainsi la deuxième couche. Cette procédure continue jusqu'à ce qu'il n'y ait plus de liens disponibles. Le traitement analytique de Cohen et al. n'aboutit pas à une expression explicite sous forme d'équation fermée. Leur résultat est donné sous forme d'une équation de récurrence qui se compare bien avec le réseau Internet dans le cas particulier où le degré minimum dans le réseau est 1.

Autres auteurs ont utilisés l'approche de la cavité pour déduire des équations de récurrence pour le nombre moyen de nœuds se trouvant à une distance donnée d'un nœud arbitraire [144].

Dans ce cas, le réseau étudié est le modèle de configuration ².

Plus récemment, des résultats analytiques pour le réseau d'Erdős-Rényi ont été développés dans le régime sous-critique [145, 146].

3.5.2 Nouvelle approche

Notre calcul est basé sur le fait que pour les grands réseaux non corrélés, la structure arborescente est vérifiée, et les cycles dans une même couche sont absents [22, 147]. Par la suite, la probabilité p_ℓ qu'un nœud de la couche n_ℓ soit lié à un autre nœud n'appartenant pas aux premières ℓ couches est $p_\ell = \frac{\kappa_\ell - 1}{n}$, où κ_ℓ est le degré moyen des nœuds appartenant à la couche ℓ , et le -1 est dû au lien de la couche précédente.

La probabilité qu'un nœud parmi les $n - n_1 - 1$ ne soit lié à aucun nœud dans la couche $\ell = 1$ est $(1 - p_1)^{n_1}$, alors la probabilité que ce nœud soit lié à la couche $\ell = 1$ 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 la couche ℓ est simple, on obtient :

$$n_\ell = (1 - (1 - p_{\ell-1})^{n_{\ell-1}})(n - \sum_{j=1}^{\ell-1} n_j - 1). \quad (3.20)$$

Lorsque n est grand $p_\ell \ll 1 \implies (1 - p_{\ell-1})^{n_{\ell-1}} \simeq e^{-p_{\ell-1}n_{\ell-1}}$. L'Eq. (3.20) peut être facilement manipulée pour obtenir :

$$n_\ell = \begin{cases} \langle k \rangle & \text{si } \ell = 1 \\ (n - 1 - n_1)(e^{-\sum_{j=1}^{\ell-2} p_j n_j} - e^{-\sum_{j=1}^{\ell-1} p_j n_j}) & \text{si } \ell \geq 2. \end{cases} \quad (3.21)$$

L'équation ci-dessus peut être simplifiée en développant les sommes dans les exponentielles. p_ℓ dépend de κ_ℓ , qui dépend à son tour de γ et du degré maximum dans la couche ℓ , K_ℓ .

κ_ℓ est facilement calculée, quand $K_\ell \gg m$ on trouve :

$$\kappa_\ell = \frac{\langle k_\ell^2 \rangle}{\langle k_\ell \rangle} = \begin{cases} \left(\frac{\gamma-2}{\gamma-3}\right)m & \text{si } \gamma > 3 \\ m(\ln(K_\ell) - \ln(m)) & \text{si } \gamma = 3 \\ \left(\frac{\gamma-2}{3-\gamma}\right)m^{\gamma-2}K_\ell^{3-\gamma} & \text{si } 2 < \gamma < 3 \\ \frac{K_\ell - m}{\ln(K_\ell) - \ln(m)} & \text{si } \gamma = 2. \end{cases} \quad (3.22)$$

Lorsque $\gamma = 2$, le degré maximum dans le réseau K est de l'ordre du nombre total de nœuds, c'est-à-dire presque tous les nœuds sont connectés au nœud de degré maximum. Le réseau peut

2. Le modèle de configuration est une méthode pour générer des réseaux aléatoires à partir d'une séquence des degrés de nœuds qui sont préalablement fixés.

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 très hétérogène. κ_ℓ dépend de la couche ℓ à travers le degré maximum K_ℓ (Eq. (3.22)). L'étape suivante de nos calculs consiste à relier κ_ℓ à κ_1 via K_ℓ . Pour réaliser cette tâche non triviale, nous remplaçons le réseau sans échelle ordinaire par un réseau hypothétique sans échelle et ordonné (Fig. 3.3). Nous commençons par choisir aléatoirement un nœud, appelé racine, ayant un degré moyen $\langle k \rangle$. Ses voisins premiers, seconds, tiers, ..., sont placés respectivement dans les couches $\ell = 1, 2, 3, \dots$, allant progressivement et dans l'ordre décroissant, du nœud ayant le plus grand degré dans la première couche au nœud avec le plus petit degré dans la dernière couche. Le nombre de nœuds dans chaque couche est la somme des degrés des nœuds de la couche précédente. Par construction, la première couche contiendra $\langle k \rangle$ nœuds, donc $n_1 = \langle k \rangle$.

L'idée sous-jacente à cette construction est que pour $\gamma \in]2, 3[$, les réseaux sans échelle sont dissassortatifs, donc les nœuds sont plus probablement liés aux hubs.

Nous avons vérifié la validité de cette construction en comparant avec la simulation numérique

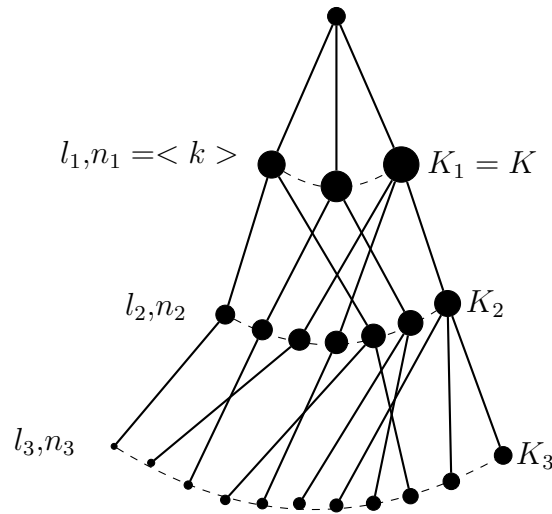


FIGURE 3.3 – Illustration du réseau construit. La taille des cercles (nœuds) schématise l'importance des degrés. Le degré maximum de la couche ℓ est K_ℓ .

d'un réseau sans échelle généré par le modèle de configuration (cercles sur la Fig. 3.4). Dans le réseau construit (étoiles sur la Fig. 3.4), à chaque nœud est attribué un degré choisi au hasard selon une loi de puissance. Le nœud racine est alors connecté comme expliqué ci-dessus et illustré sur la Fig. 3.3.

La Fig. 3.4(a) montre que le réseau construit (Fig. 3.3) représente une bonne approximation du réseau sans échelle, et peut être utilisé pour surmonter d'autres problèmes analytiques dans les réseaux hétérogènes. Pour $\ell = 1$, une différence significative a été observée entre les deux

réseaux à cause des mauvaises statistiques, en effet, la première couche contient un nombre relativement petit de nœuds. Une autre observation importante à partir de la Fig. 3.4 est que pour $2 < \gamma < 3$ (réseau hétérogène) κ_ℓ dépend fortement des couches, ce qui est en accord avec l'Eq. (3.22), et pour $\gamma = 4$ (réseau homogène) κ_ℓ est indépendant des couches.

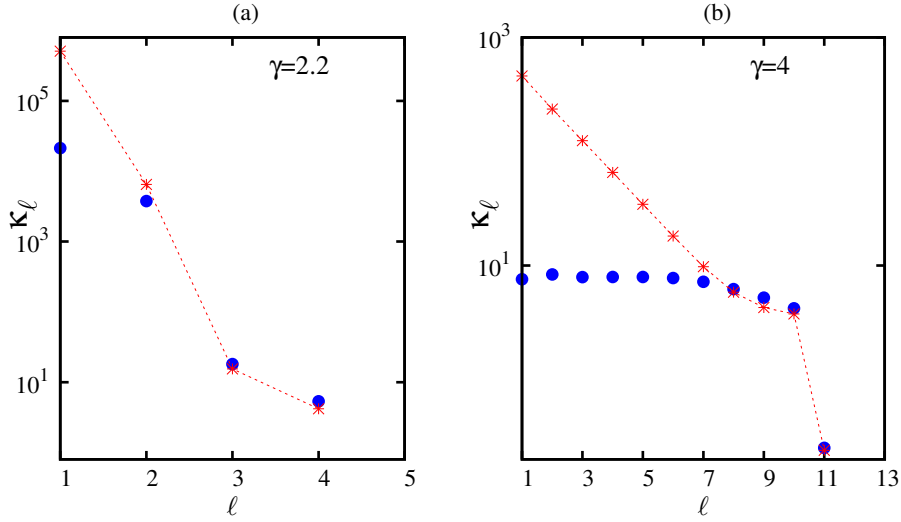


FIGURE 3.4 – Dépendance de κ_ℓ en ℓ . Les cercles sont les résultats de la simulation d'un réseau sans échelle. Les étoiles représentent la simulation du réseau de la Fig. 3.3. Le nombre de nœuds dans les deux cas est 4×10^6 , chaque point est la moyenne sur 200 réalisations.

Selon la construction de la Fig. 3.3, le nombre de nœuds entre la première et la $(\ell - 1)$ ème couche est donné par :

$$\begin{aligned} \sum_{j=1}^{\ell-1} n_j &= n \int_{K_\ell}^{K_1} P(k) dk \\ &= nm^{\gamma-1} (K_\ell^{1-\gamma} - K_1^{1-\gamma}) \simeq nm^{\gamma-1} K_\ell^{1-\gamma}, \end{aligned} \quad (3.23)$$

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

$$\begin{aligned} \sum_{j=1}^{\ell-1} n_j &= \langle k \rangle + \langle k \rangle (\kappa_1 - 1) + \langle k \rangle (\kappa_1 - 1)(\kappa_2 - 1) + \dots + \langle k \rangle (\kappa_1 - 1)(\kappa_2 - 1) \dots (\kappa_{\ell-1} - 2) \\ &\approx \langle k \rangle \kappa_1 \kappa_2 \dots \kappa_{\ell-2}, \end{aligned} \quad (3.24)$$

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 de voisins à distance ℓ peut maintenant être exprimé comme : $\kappa_\ell = \frac{\gamma-2}{3-\gamma} m \left(\frac{n}{\langle k \rangle \kappa_1 \kappa_2 \dots \kappa_{\ell-2}} \right)^{\frac{3-\gamma}{\gamma-1}}$,

ce qui conduit à la relation de récurrence $\kappa_\ell = \kappa_{\ell-1} \left(\kappa_{\ell-2} \right)^{\frac{\gamma-3}{\gamma-1}}$. Quand γ est dans $]2, 3[$ nous avons généralement $\kappa_{\ell-2} \gg \kappa_{\ell-1}$, mais puisque $\left| \frac{\gamma-3}{\gamma-1} \right| < 1$, on considère alors 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 pouvons maintenant calculer les sommes dans l'Eq. (3.21) :

$$\begin{aligned} \sum_{j=1}^l p_j n_j &= \sum_{j=1}^l \frac{\kappa_j}{n} \left(\langle k \rangle \kappa_1 \kappa_2 \dots \kappa_{j-1} \right) \\ &\approx \frac{\kappa_\ell}{n} \left(\langle k \rangle \kappa_1 \kappa_2 \dots \kappa_{\ell-1} \right) \\ &\approx \frac{\langle k \rangle}{n} \kappa_1^{1+\beta+\beta^2+\dots+\beta^{\ell-1}} \\ &\approx \frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^\ell}{1-\beta}}, \end{aligned} \quad (3.25)$$

où $\beta = \frac{2\gamma-4}{\gamma-1}$. En prenant $n - n_1 - 1 \approx n$, l'Eq. (3.21) peut être écrite pour $2 < \gamma < 3$ comme :

$$n_\ell = \begin{cases} \langle k \rangle & \text{if } \ell = 1 \\ n \left(e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-2}}{1-\beta}}} - e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-1}}{1-\beta}}} \right) & \text{if } \ell \geq 2. \end{cases} \quad (3.26)$$

Quand $\gamma > 3$, les hubs ne sont plus présents, et les propriétés caractéristiques du réseau sont similaires à celles du réseau aléatoire d'ER. κ_ℓ est indépendant de ℓ comme indiqué dans l'Eq. (3.22) et illustré dans la Fig. 3.4, alors $\sum_{j=1}^{\ell-1} p_j n_j = \frac{\kappa'}{n} \sum_{j=1}^{\ell-1} n_j$, où $\kappa' = \kappa_1 - 1$.

$$\begin{aligned} \sum_{j=1}^{\ell} n_j &= \langle k \rangle + \langle k \rangle (\kappa_1 - 1) + \langle k \rangle (\kappa_1 - 1)(\kappa_2 - 1) + \dots + \langle k \rangle (\kappa_1 - 1)(\kappa_2 - 1) \dots (\kappa_{\ell-1} - 1) \\ &= \langle k \rangle + \langle k \rangle \kappa' + \langle k \rangle \kappa'^2 + \dots + \langle k \rangle \kappa'^{\ell-1} = \langle k \rangle \frac{1 - \kappa'^\ell}{1 - \kappa'}. \end{aligned} \quad (3.27)$$

L'Eq. (3.21) s'écrit pour $\gamma > 3$ comme :

$$n_\ell = \begin{cases} \langle k \rangle & \text{if } \ell = 1 \\ n \left(e^{-\frac{\langle k \rangle}{n} \kappa'^{\frac{1-\kappa'^{\ell-2}}{1-\kappa'}}} - e^{-\frac{\langle k \rangle}{n} \kappa'^{\frac{1-\kappa'^{\ell-1}}{1-\kappa'}}} \right) & \text{if } \ell \geq 2. \end{cases} \quad (3.28)$$

Le cas $\gamma = 3$ est le plus problématique. C'est le point où la structure du réseau change radicalement, on passe de la présence de multiples grands hubs et un second moment $\langle k^2 \rangle$ qui diverge pour $\gamma < 3$, à l'absence de 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, en ce point de la transition ($\gamma = 3$), certaines propriétés du réseau se comportent presque comme celles pour $\gamma > 3$. Principalement, le plus court chemin est de l'ordre de $\frac{\ln(n)}{\ln(\ln(n))}$ [148], et $\langle k^2 \rangle$ reste fini.

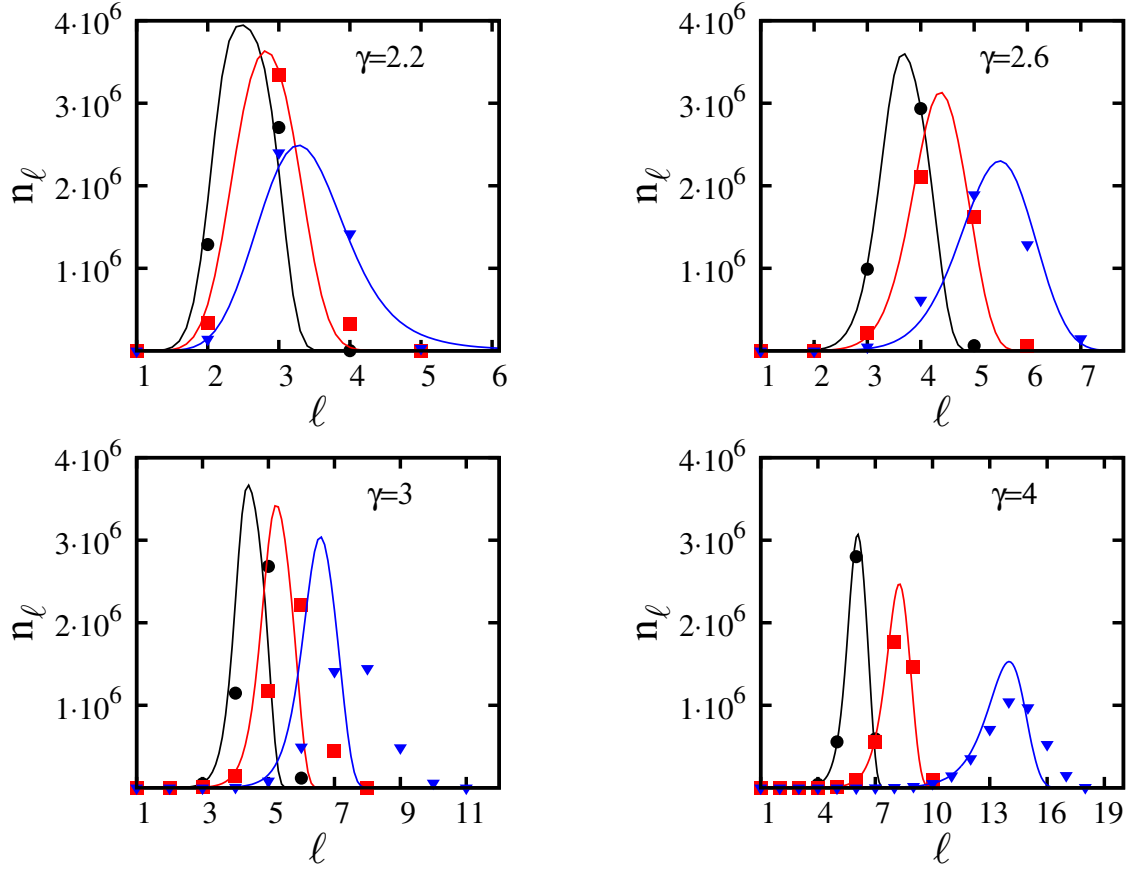


FIGURE 3.5 – Nombre de nœuds dans chaque couche pour différentes valeurs de γ . Les lignes continues correspondent aux Eq. (3.26) et Eq. (3.28). Les symboles sont les simulations numériques d'un réseau de taille $n = 4 \times 10^6$. Chaque point est la moyenne sur 200 réalisations. Les couleurs noir, rouge et bleu représentent respectivement les cas $m = 8, 4$ et 2 .

Nous utilisons, donc, l'Eq. (3.28) pour calculer $n_{\hat{\ell}}$ dans ce cas.

Dans la Fig. 3.5, nous représentons $n_{\hat{\ell}}$ en fonction de ℓ pour différentes valeurs de γ 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 $\sum_{j=1}^l p_j n_j$ dans l'Eq. (3.27). Nous observons aussi une légère différence entre la théorie et les simulations lorsque $\gamma = 4$ et $m = 2$. Ceci est à cause de l'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 $n_{\hat{\ell}}$, n'est pas très vérifiée. La Fig. 3.6 montre clairement l'importance relative des cycles pour les petites valeurs de m . Cet effet est accentué au fur et à mesure qu'on augmente γ .

Nous observons (Fig. 3.5) que $n_{\hat{\ell}}$ possède deux allures différentes lors de la croissance et de la décroissance. Pour extraire une information quantitative, concernant les queues de $n_{\hat{\ell}}$, à partir

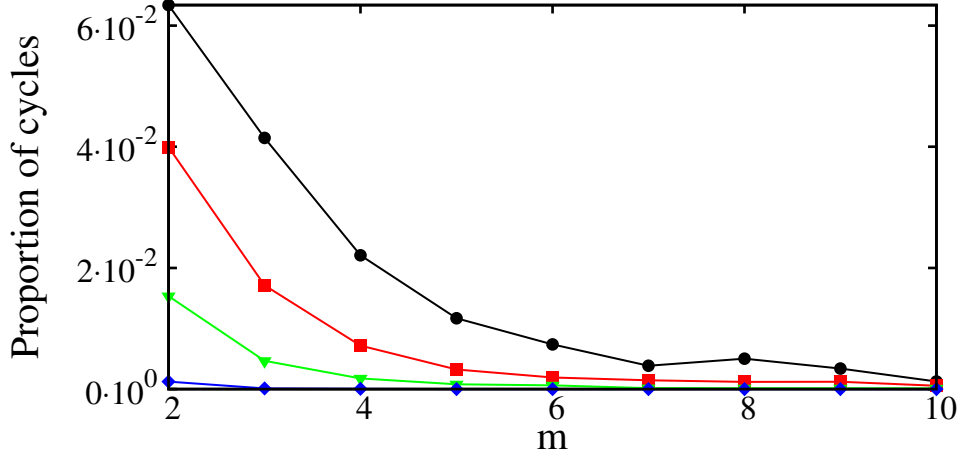


FIGURE 3.6 – Proportion de cycles dans les couches par rapport à m . De haut en bas, γ est respectivement 4, 3, 2.6 et 2.2. Nombre de nœuds $n = 4.10^6$, le nombre de réalisations pour chaque point est 200.

des Eq. (3.26) et Eq. (3.28), on utilise la règle du point milieu.

Pour n grand, lorsque $2 < \gamma < 3$, l'Eq. (3.26) pour $\ell > 1$ peut être approximée comme :

$$\begin{aligned}
 n_\ell &= n \left(e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-2}}{1-\beta}}} - e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-1}}{1-\beta}}} \right) \\
 &\approx -n \frac{\partial \left(e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}}} \right)}{\partial \ell} \\
 &\approx -\frac{\langle k \rangle \ln(\kappa_1) \ln(\beta)}{1-\beta} \beta^{\ell-\frac{3}{2}} \kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}} e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}}}, \tag{3.29}
 \end{aligned}$$

où $\ell - \frac{3}{2}$ est utilisé à la place de $\ell - 1$ pour améliorer la différenciation avec la règle du point milieu.

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'Eq. (3.29), $\kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}}$, peut s'écrire comme $\kappa_1^{\frac{1}{1-\beta}} e^{\frac{-\ln(\kappa_1)(\ell-\frac{3}{2})\ln(\beta)}{1-\beta}}$. Ce dernier terme montre que n_ℓ croît comme une double exponentielle. Après avoir atteint son maximum en $\kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}} = \frac{n}{\langle k \rangle}$, le terme dominant dans l'Eq. 3.29 devient $e^{-\frac{\langle k \rangle}{n} \kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}}}$. Pour $\kappa_1^{\frac{1-\beta^{\ell-\frac{3}{2}}}{1-\beta}} \gg \frac{n}{\langle k \rangle}$, n_ℓ décroît comme une triple exponentielle.

3.6. Plus court chemin

De la même manière, dans le cas où $\gamma > 3$, l'Eq. (3.28) peut s'écrire comme

$$\begin{aligned}
 n_\ell &= n \left(e^{-\frac{\langle k \rangle}{n} \kappa' \frac{1-\kappa'^{\ell-2}}{1-\kappa'}} - e^{-\frac{\langle k \rangle}{n} \kappa' \frac{1-\kappa'^{\ell-1}}{1-\kappa'}} \right) \\
 &\approx n \left(e^{-\frac{\langle k \rangle}{n} \kappa'^{\ell-2}} - e^{-\frac{\langle k \rangle}{n} \kappa'^{\ell-1}} \right) \\
 &\approx n \frac{\partial e^{-\frac{\langle k \rangle}{n} \kappa'^{\ell-\frac{3}{2}}}}{\partial \ell} \\
 &\approx \langle k \rangle \ln(\kappa') \kappa'^{\ell-\frac{3}{2}} e^{-\frac{\langle k \rangle}{n} \kappa'^{\ell-\frac{3}{2}}}, \tag{3.30}
 \end{aligned}$$

En suivant les mêmes démarches que précédemment, on obtient une croissance exponentielle pour n_ℓ lorsque $\kappa'^{\ell-\frac{3}{2}} \ll \frac{n}{\langle k \rangle}$, et une décroissance en double exponentielle lorsque $\kappa'^{\ell-\frac{3}{2}} \gg \frac{n}{\langle k \rangle}$. Nos résultats montrent que n_ℓ décroît toujours plus rapidement qu'elle croît, ceci a été signalé dans le cas du réseau Internet dans [142].

Nos résultats analytiques sont aussi comparés aux réseaux du monde réel (voir Fig. 3.7). La Fig. 3.7 confirme le bon accord entre nos équations et le réseau Hollywoodien de 392732 acteurs.

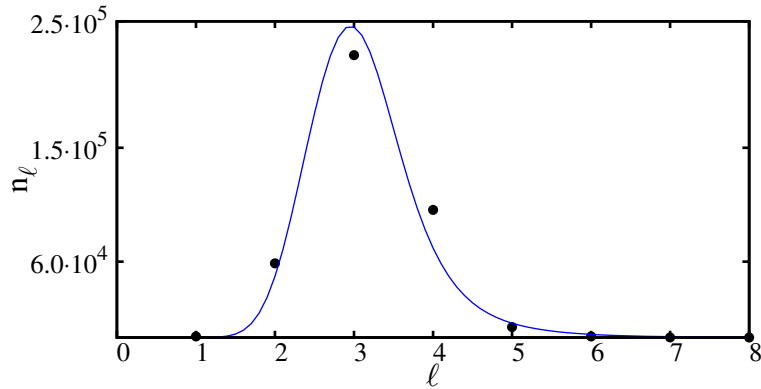


FIGURE 3.7 – Comparaison entre les données empiriques (cercle) et notre théorie (ligne continue) : Les cercles représentent les couches du réseau d'acteurs Hollywoodien de 392732 acteurs avec $\gamma = 2.17$. L'acteur Kevin Bacon est au centre du réseau. La ligne continue représente l'Eq. (3.26) avec les mêmes valeurs pour γ et n . Les résultats empiriques ont été extraits du site : "<http://oracleofbacon.org/center.php>".

3.6 Plus court chemin

Le plus court chemin, qu'on désignera par $\langle \ell \rangle$, est peut être le concept le plus intrigant des réseaux complexes, et ce après la célèbre expérience de Milgram [48]. Dans cette expérience,

Milgram a clairement montré l'effet de petit monde dans les réseaux sociaux, ce qui signifie que deux personnes dans le monde, a priori sans aucun lien, sont séparées en moyenne par un nombre petit de connexions intermédiaires.

De nombreux efforts théoriques ont été déployés pour calculer $\langle \ell \rangle$ [149, 23, 22, 148, 1, ?, ?], la plupart d'entre eux se sont focalisés sur la manière dont $\langle \ell \rangle$ varie avec le nombre de nœuds n . Dans le Tableau. 3.1 on résume les principaux résultats dans la littérature.

Une expression analytique de la distance moyenne intégrant les paramètres pertinents du réseau est cruciale pour calculer les centralités d'intermédierité et de proximité, et peut également présenter un grand intérêt pour des situations réelles. Dans le réseau urbain, par exemple, il a été observé que le trafic des véhicules empruntant les voies de communication les plus courtes est celui de la communicabilité [150].

Calculer la distance moyenne dans des réseaux complexes est généralement une tâche ardue. Ici, on déduit $\langle \ell \rangle$ dans un réseau sans échelle non corrélé à partir de la distribution des nœuds, n_{ℓ} . En fait, la forme quasi symétrique de n_{ℓ} sur la Fig. 3.5 suggère que $\langle \ell \rangle$ correspond à la distance où n_{ℓ} est maximum, ou également, à la distance où $\frac{\partial n_{\ell}}{\partial \ell} = 0$.

Nous commençons par le cas le plus simple $\gamma = 2$. Comme déjà mentionné, il y a un maximum

Valeur de γ	Ordre de grandeur	Désignation
$2 < \gamma < 3$	$\sim \ln \ln(n)$	ultra-petit monde [22, 23, 151, 149, 152, 153]
$\gamma = 3$	$\sim \frac{\ln(n)}{\ln \ln(n)}$	petit monde [143, 149, 1, 154, 147]
$\gamma > 3$	$\sim \ln(n)$	petit monde [143, 149, 1, 154, 147]

TABLE 3.1 – Principaux résultats sur le plus court chemin en fonction de γ .

de deux couches dans le réseau, $\langle \ell \rangle$ s'écrit comme :

$$\langle \ell \rangle = \frac{\langle k \rangle + 2(n - \langle k \rangle)}{n}, \quad (3.31)$$

qui tend vers 2 pour un grand n . Cela signifie que presque tous les nœuds sont connectés entre eux à travers le nœud de degré maximum.

Quand $2 < \gamma < 3$, aucune solution pour $\frac{\partial n_{\ell}}{\partial \ell} = 0$ ne peut être trouvée directement à partir de l'Eq. (3.26), alors nous utilisons l'approximation donnée dans l'Eq. (3.29), où n_{ℓ} est maximum quand $\kappa_1^{\frac{1-\beta\langle \ell \rangle - \frac{3}{2}}{1-\beta}} = \frac{n}{\langle k \rangle}$. Après avoir remplacé κ_1 et K_1 par leurs expressions correspondantes, on obtient :

$$\langle \ell \rangle = \frac{\ln \left(1 - \frac{\ln(n) - \ln(\langle k \rangle)}{\ln(n) + \frac{\gamma-1}{3-\gamma} \ln(\frac{\gamma-2}{3-\gamma} m)} \right)}{\ln \left(\frac{2\gamma-4}{\gamma-1} \right)} + \frac{3}{2}. \quad (3.32)$$

3.6. Plus court chemin

pour n grand, $\langle \ell \rangle \approx -\frac{\ln(\ln(n))}{\ln(\beta)}$. Cette forme d'échelle est derrière la désignation ultra-petit monde [22], et elle est largement acceptée pour cet intervalle de γ [23, 151, 149, 152, 153].

Pour $\gamma \geq 3$, $\langle \ell \rangle$ peut être déduit de l'Eq. (3.28) en résolvant $\frac{\partial n_\ell}{\partial \ell} = 0$. Cela donne :

$$\langle \ell \rangle = \frac{\ln(n)}{\ln(\kappa')} + \frac{\ln(\ln(\kappa')) - \ln(\langle k \rangle)}{\ln(\kappa')} + 1. \quad (3.33)$$

Si $\gamma > 3$, κ est constant (Eq. (3.22)). Pour un grand n , $\langle \ell \rangle \approx \frac{\ln(n)}{\ln(\kappa')}$, qui est la forme d'échelle rapportée dans de nombreux autres travaux [143, 149, 1, 154, 147], et connu comme l'effet de petit monde.

Quand $\gamma = 3$, κ dépend de K , qui dépend à son tour de n . Prenant $\kappa_1 = m(\ln(K) - \ln(m))$ et

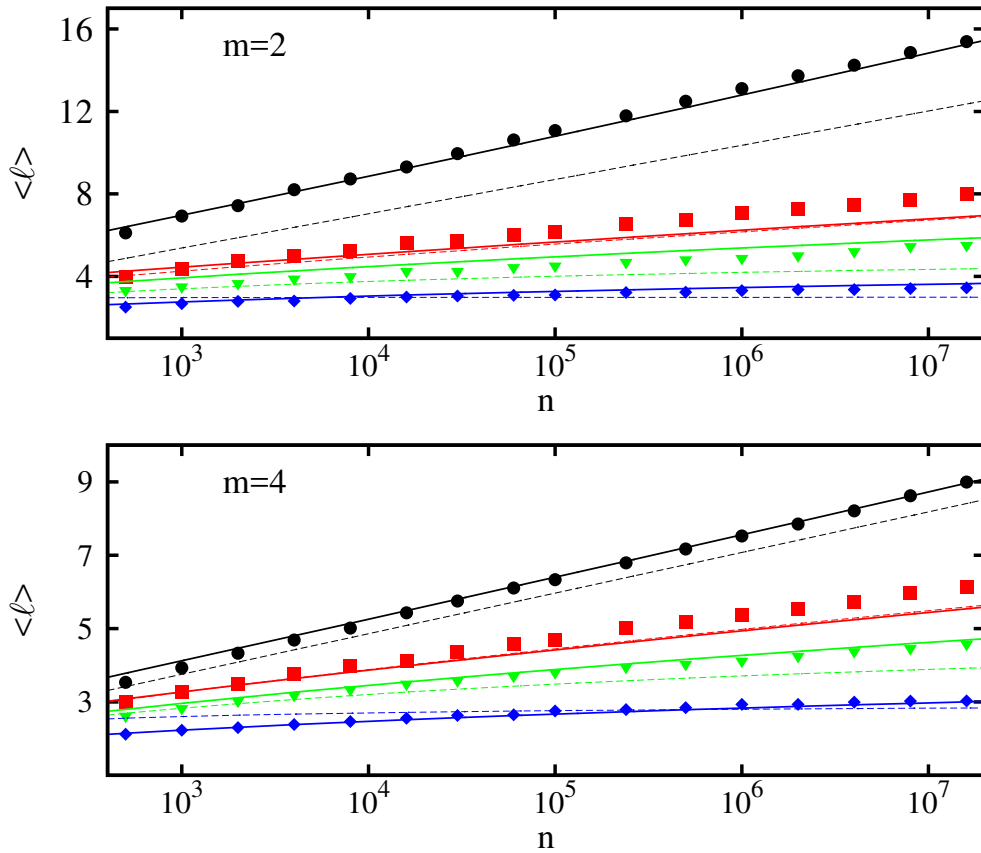


FIGURE 3.8 – Le plus court chemin en fonction du nombre de nœuds. Les valeurs de γ de haut en bas sont respectivement 4, 3, 2.6 et 2.2. Les lignes continues correspondent à l'Eq. (3.32) et à l'Eq. (3.33). Les lignes discontinues correspondent aux équations dans [1]. Chaque simulation est moyennée sur plus de 200 réalisations.

$K = mn^{\frac{1}{\gamma-1}}$, nous trouvons pour n grand, $\langle \ell \rangle \approx \frac{\ln(n)}{\ln(\ln(n))}$. Ce résultat est en accord avec les travaux précédents [149, 22, 1, 155], 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 lorsque $2 < \gamma < 3$.

Comme nous avons signalé dans la Section. ??, les travaux importants [23, 22] dans ce sujet donnent seulement l'allure de $\langle \ell \rangle$ pour les grandes valeurs de n . L'exception est la contribution de Fronczak et al. [1] où ils ont trouvé des équations fermées en fonction des paramètres du réseau, ces équations prédisent que $\langle \ell \rangle$ tend vers une valeur constante lorsque n tend vers l'infini pour $2 < \gamma < 3$, ce qui est en contradiction avec les principaux résultats dans la littérature [22, 23, 151, 149, 152, 153], et avec la simulation numérique (voir Fig. 3.8).

Nos résultats (Fig. 3.8) montrent un très bon accord avec les simulations pour $\gamma \neq 3$. Pour $\gamma = 3$, notre approche donnent des valeurs de $\langle \ell \rangle$ relativement proches des simulations numériques, et ils sont également très similaires aux résultats de [1].

3.7 conclusion

Dans ce chapitre, nous avons étudié en détail certains aspects fondamentaux des réseaux sans échelle non corrélés. À partir d'arguments probabilistes, nous avons obtenu des expressions explicites, en fonction de l'exposant de degré γ , pour le nombre de nœuds, n_ℓ , à une distance donnée d'un nœud arbitraire. Nous avons montré que le réseau hétérogène ($2 < \gamma < 3$) peut être approximé par un réseau sans échelle ordonné où les nœuds voisins du nœud racine (choisi aléatoirement) sont placés du plus grand au plus petit degré. Ceci peut simplifier les calculs, qui sont en général assez difficiles, dans plusieurs situations des réseaux hétérogènes.

Nous avons aussi décrit avec précision les formes des queues de la distribution n_ℓ .

Profitant de la forme de celle-ci, nous avons pu déduire les expressions explicites du plus court chemin pour différentes valeurs de γ . Les expressions obtenues reproduisent les comportements asymptotiques connues pour n grand, à savoir, l'ultra-petit monde pour $2 < \gamma < 3$, et le petit monde pour $\gamma \geq 3$. Nos résultats théoriques concordent très bien avec les simulations numériques et le nombre de Bacon, sauf dans le cas $\gamma = 3$, où nous avons observé la même forme, comme pour les autres valeurs de γ , dans les queues de n_ℓ , mais avec une notable différence (entre la théorie et la simulation) dans la position du maximum. Cette différence n'affecte pas l'ordre de grandeur du plus court chemin pour cette valeur de γ .

Émergence de la propriété petit-monde dans le modèle de Newman-Watts

Sommaire

4.1	Introduction	63
4.1.1	Modèle de Watts-Strogatz	64
4.1.2	Modèle de Newman-Watts	66
4.2	Structure du réseau de NW : couches de voisins et plus court chemin	68
4.2.1	Couches de voisins	68
4.2.2	Plus court chemin	78
4.3	Émergence de la propriété petit-monde dans le modèle de NW	81
4.3.1	Validité de la fonction universelle de NW	81
4.4	Une nouvelle fonction universelle	82
4.5	Conclusion	84

4.1 Introduction

Le modèle de Watts-Strogatz (WS) retient une attention particulière depuis son apparition en 1998 [97], il a deux propriétés très intéressantes : la présence d'un coefficient de Clustering élevé et d'un petit plus court chemin (propriété ou effet petit monde). Ces deux propriétés se trouvent dans la plupart des réseaux réels [138, 156], en plus ce modèle est très simple et combine la régularité avec l'aléatoire. En 1999 Newman et Watts [157] ont réalisé une petite modification sur ce modèle sans presque perdre aucunes caractéristiques, mais avec une meilleure facilité sur l'analyse théorique.

L'émergence de la propriété petit-monde dans ces modèles est encore un phénomène pas bien

compris, parmi ses problématiques on cite : à partir de quel point le réseau change sa nature grand-monde vers petit-monde ?, si le réseau se sature par les raccourcis ajoutés, comment et quand la saturation se réalise ?, existe-t-elle une transition de phase ?.

Dans ce chapitre, nous aborderons quelques-unes de ces questions. Dans nos calculs, on utilisera les techniques du groupe de renormalisation, et de la théorie de champ moyen. Les résultats analytiques sont comparés aux simulations.

4.1.1 Modèle de Watts-Strogatz

Le modèle d'ER (voir section) reproduit l'effet petit monde (distance moyenne entre les nœuds est petite) observé dans les réseaux complexes, mais le coefficient de Clustering dans ce modèle est très faible par rapport à celui observé dans les systèmes réels, et la distribution des degrés est différente de celles des réseaux réels. En 1998 Watts et Strogatz [97] ont proposé un modèle qui reproduit deux propriétés parmi celles mentionnées précédemment de manière simple. Ce modèle, nommé souvent "petit-monde", peut être construit 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 entier, le modèle Watts-Strogatz commence par la construction suivante (voir Fig.4.1) :

- Placez tous les nœuds en cercle ;
- Connectez chaque nœud à ses premiers $2k$ voisins les plus proches (k voisins à droite, et k voisins à gauche) ;
- Avec une probabilité ϕ , chaque lien du réseau est enlevé, et ensuite placé entre deux nœuds choisis au hasard, ces nouveaux liens sont appelés raccourcis.

Ce modèle peut être justifié par le fait 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 de gens ont quelques amis qui habitent loin de leur quartier. Les propriétés structurelles du modèle sont quantifiées par le plus court chemin et le coefficient de clustering en fonction de la probabilité de reconnection ϕ .

Le réseau régulier avec $\phi = 0$ est un grand-monde, c'est-à-dire que la distance moyenne entre les nœuds se fait grande lorsque le nombre de nœud n devient grand, en effet, $\langle \ell \rangle = \frac{2n}{k}$. Si $\phi = 1$, le modèle de WS est équivalent au réseau aléatoire d'ER. Dans ce cas, le réseau est un petit-monde car $\langle \ell \rangle \sim \ln(n)$.

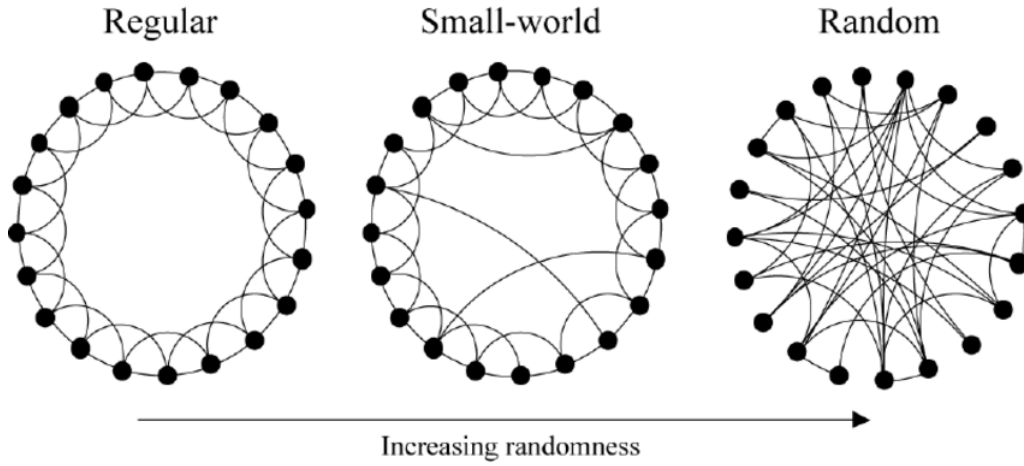


FIGURE 4.1 – Représentation schématique de l'évolution du processus de "reconnexion" dans le modèle Watts-Strogatz. Le nombre de nœuds est $n = 20$ et $k = 2$.

La Fig. 4.2 [97], montre que la présence de quelques raccourcis entraîne la diminution de $\langle \ell \rangle$.

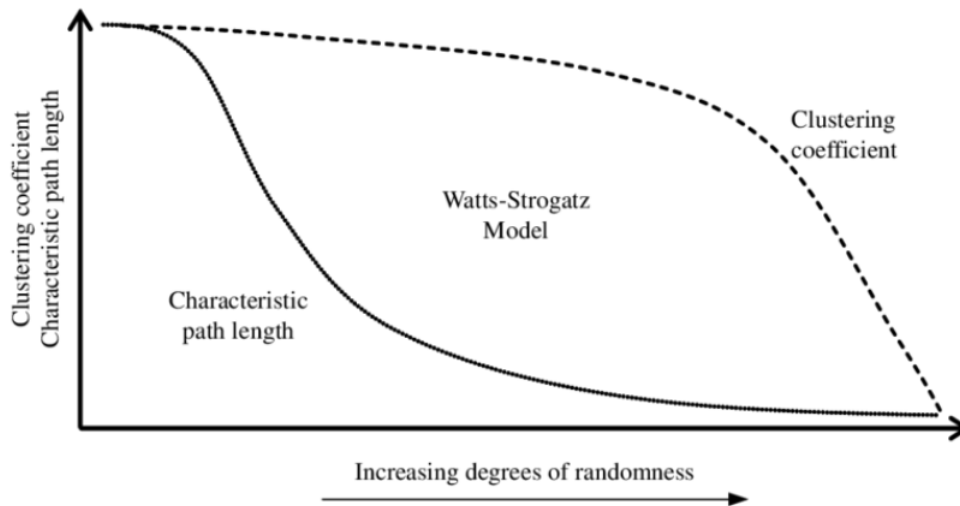


FIGURE 4.2 – Représentation de la variation de la longueur moyenne du chemin et du coefficient de regroupement avec la probabilité de reconnexion ϕ dans le modèle Watts-Strogatz.

Le coefficient de clustering reste pratiquement inchangé pour les petites valeurs de ϕ , même si $\langle \ell \rangle$ chute rapidement. L'implication importante est que la distance moyenne entre les nœuds chute rapidement et devient petite pour les mêmes valeurs de ϕ auxquelles le Clustering local du réseau est significativement important. Ce modèle permet donc de construire des réseaux exhibant des propriétés observées dans les réseaux complexes naturels, court plus court chemin et haut coefficient de clustering, dépassant ainsi le modèle d'ER.

Newman et Watts, en utilisant la transformation de groupe de renormalisation ont montré que

le plus court chemin suit la loi universelle [8] :

$$\langle \ell \rangle = \frac{n}{k} f(x), \quad (4.1)$$

avec $x = nk\phi$ est le nombre moyen de raccourcis dans le réseau et $f(x)$ une fonction universelle. Le coefficient de clustering, C , est défini par l'Eq. (1.10). Pour évaluer C , nous devons calculer le nombre de triangles et de triplets connectés dans le réseau, Barrat a trouvé une très bonne approximation pour C qui s'écrit sous la forme [158] :

$$C = \frac{3(k-1)}{2(2k-1)}(1-\phi)^3. \quad (4.2)$$

Lorsque $\phi = 0$, la distribution de degrés est une fonction delta centrée sur la valeur $2k$, alors que pour $\phi = 1$ elle est similaire à celle d'un réseau ER. Pour les valeurs intermédiaires de ϕ , la distribution de degrés est donnée par [158] :

$$P(j) = \sum_{i=0}^{\min(j-2k, m)} C_i^{2k} \frac{(2k\phi)^{j-2k-i}}{(j-2k-i)} e^{2k\phi}. \quad (4.3)$$

4.1.2 Modèle de Newman-Watts

Newman et Watts (NW) [157] ont introduit une petite modification sur le modèle WS. Sans supprimer des liens, un raccourci connectant 2 nœuds choisis au hasard est ajouté avec une probabilité ϕ . En moyenne il y a $x = nk\phi$ raccourcis. Ce modèle a les mêmes propriétés que celui de WS, sauf pour $k = 1$ où il y a une certaine différence (Expliquer, Quelles propriétés?)¹. En outre, le modèle de NW est plus abordable de point de vue traitement analytique. En effet, par exemple, Newman a donné l'expression exacte du coefficient de clustering [156] :

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

En ce qui concerne le plus court chemin, il n'existe pas encore un calcul exacte dans le modèle de NW, et ça malgré sa simplicité apparente. Le traitement le plus important date de l'année 2000 où Newman et al. [159], et avec un calcul de champ moyen très élaboré, ont trouvé une approximation de la fonction universelle Eq. (4.1) sous la forme :

$$f(x) = \frac{1}{2\sqrt{x^2 + 2x}} \tanh^{-1}\left(\sqrt{\frac{x}{x+2}}\right), \quad (4.5)$$

1. Le cas $k = 1$ est sans intérêt dans ce modèle car le coefficient de clustering est zéro

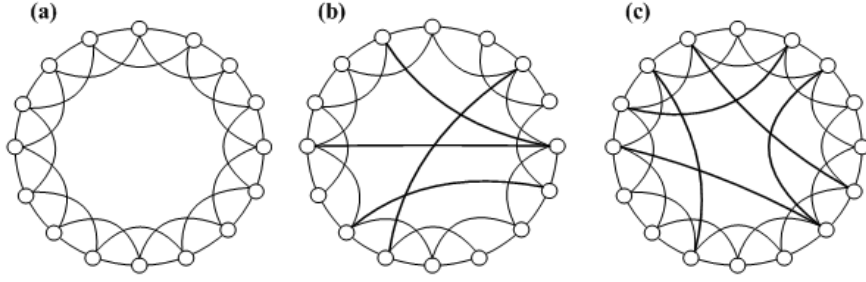


FIGURE 4.3 – Représentation schématique de l'évolution du processus de "reconnexion" dans le modèle NW. Le nombre de nœuds est $n = 20$ et $k = 2$.

les formes asymptotiques de cette fonction sont :

$$f(x) \sim \begin{cases} \frac{1}{4} & \text{si } x \ll 1 \\ \frac{\log 2x}{4x} & \text{si } x \gg 1. \end{cases} \quad (4.6)$$

La fonction $f(x)$ s'avère exacte dans la limite de n grand pour une densité de raccourcis donnée, ou dans la limite de la densité de raccourci faible pour une taille n donnée. Dans le cas où la densité des raccourcis est faible avec un nombre total de raccourcis grand (car le réseau lui-même est également grand), $f(x)$ est également une solution exacte. Au voisinage de $\phi = 1$ une déviation entre cette solution et les simulations numériques est observée (voir Fig. 4.9). Dans la section suivante, nous proposons une explication de l'origine de cette déviation non traitée auparavant.

Une des questions très étudiée dans le modèle de WS est l'existence de la transition de phase lors du passage du grand-monde pour $\phi = 0$ au petit-monde pour $\phi = 1$. En 1999 Barthélémy et Amaral [160] propose que la distance moyenne $\langle \ell \rangle$ entre deux sommets quelconques du réseau est une fonction de $\frac{n}{n^*}$, où la taille de croisement(crossover) n^* au-dessus de laquelle le réseau vérifie l'effet petit-monde ($\langle \ell \rangle \sim \ln(n)$) se comporte pour $p \ll 1$ comme $n^* \sim p^{-\tau}$ avec $\tau \approx \frac{2}{3}$. Ces auteurs ont affirmés qu'il s'agit plutôt d'un phénomène de crossover, et non d'une transition de phase. Barrat [161] argumente que τ ne peut pas être inférieur à 1, il donne la valeur $\tau = 1$ en utilisant la même approche que Barthélémy et Amaral mais avec des tailles du système plus grandes. Newman et Watts [8, 162] confirment cette valeur de l'exposant critique ($\tau = 1$) en utilisant une transformation de groupe de renormalisation. Ces derniers suggèrent que la transition vers le petit-monde se fait d'une façon continue, c'est-à-dire une transition de phase de deuxième ordre où le point critique correspond à une densité de raccourcis tendant vers zéro ($\phi \rightarrow 0$). Autres auteurs [163] trouvent un résultat différent des précédents, ceux-ci

obtiennent une transition de phase de première ordre au point $\phi = 0$, à cause d'une discontinuité d'un certain paramètre d'ordre en ce point.

On en déduit qu'il existe encore un débat à propos du type de la transition de phase dans ce modèle. Dans la section suivante on va traiter cette question en détail en nous basant sur le fait que le réseau de NW est un mélange de régularité et de l'aléatoire.

4.2 Structure du réseau de NW : couches de voisins et plus court chemin

4.2.1 Couches de voisins

On se propose d'étudier en détail le nombre de voisins $\hat{n}_{\hat{\ell}}$ se trouvant à une distance $\hat{\ell}$ d'un nœud arbitraire en appliquant la transformation du groupe de renormalisation en espace réel (GR) sur le réseau (voir Fig. 4.6). Étant donné que le modèle de NW est un mélange de régularité et de l'aléatoire, les couches vont être séparées en deux catégories :

- Couches régulières $\hat{n}_{\hat{\ell}}^r$ où la distance $\hat{\ell}$ entre le nœud racine et ses voisins n'a pas été changée par l'introduction des raccourcis.
- Couches aléatoires $\hat{n}_{\hat{\ell}}^{al}$ où la distance $\hat{\ell}$ entre le nœud racine et ses voisins a été changée par l'introduction des raccourcis (voir Fig. 4.4).

La transformation du GR remplace le réseau de n nœuds avec chacun un degré k par un autre réseau avec un nombre de nœuds $\hat{n} = \frac{n}{k}$ de degré $\hat{k} = 1$ (voir Fig. 4.6)². Chaque ensemble de k nœuds voisins (**amas**) est considéré comme un seul nœud. Après l'ajout des raccourcis avec la probabilité ϕ , la probabilité qu'un nœud (**amas**) dans le nouveau réseau ($\hat{k} = 1$ et $\hat{n} = \frac{n}{k}$) soit lié aléatoirement à un autre nœud (**amas**) est $\hat{\phi} = 1 - \left(1 - \frac{2k\phi}{n}\right)^{k^2}$, pour $n \gg k\phi$ on peut écrire $\hat{\phi} \approx \frac{2k^3\phi}{n}$. On note $P_r(\hat{\ell})$ la probabilité que la distance $\hat{\ell}$ entre un nœud quelconque et le nœud racine ne soit pas modifiée après l'ajout des raccourcis, et soit $P_{al}(\hat{\ell})$ la probabilité que la distance entre un nœud quelconque et le nœud racine devienne égale à $\hat{\ell}$ après l'ajout des raccourcis. Dans l'exemple de la Fig. 4.4, les nœuds sombres sont ceux qui deviennent plus proches du nœud racine R grâce au raccourci.

2. Le degré de liberté est réduit, car dans le nouveau réseau le paramètre \hat{k} est toujours égale à 1, c'est exactement l'idée principale de la théorie du groupe de renormalisation.

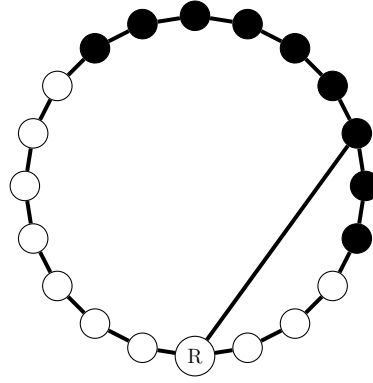


FIGURE 4.4 – Illustration des nœuds appartenant aux couches aléatoires (nœuds sombres) dont la distance au nœud racine R a été changée par les raccourcis. Les nœuds appartenant aux couches réguliers (nœuds clairs) restent à la même distance de R après introduction des raccourcis. Le réseau est de taille $n = 20$, $k = 1$, et avec un seul raccourci.

Généralement la réduction de la distance d'un nœud par rapport au nœud racine peut se faire à l'aide d'un seul, de deux ou plus de raccourcis. On note $\pi^m(i)$ la probabilité que la distance initiale d'un nœud ne soit pas changée vers une distance plus petite i à travers m raccourcis. On distingue les cas suivants :

■ Réduction de la distance par un raccourci :

Soit $\pi^1(i)$ la probabilité qu'un nœud j ne change pas sa distance initiale $\hat{\ell}$ vers la distance i ($i < \hat{\ell}$) à travers un seul raccourci :

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

dans $\pi^1(1) = (1 - \hat{\phi})$ l'exposant est 1 car il a une seule possibilité pour que le nœud j se lie directement au nœud racine R , pour $\pi^1(2) = (1 - \hat{\phi})^4$ l'exposant est 4 car pour que la distance entre j et R soit 2, il faut que le nœud j se lie directement avec l'un des deux premiers voisins du nœud R ou l'inverse, c'est-à-dire que R se lie directement avec l'un des premiers voisins du j , donc en totale 4 possibilités. La déduction des autres probabilités se fait par le même raisonnement. Le nombre de possibilités (sauts) pour construire un chemin tel que $\hat{\ell} = i$ est $4(i - 1)$. Cette valeur est en général supérieure à la valeur réelle,

car parmi ces possibilités il y en a qui correspondent à des distances $\hat{\ell} < i$ et qui sont déjà comptabilisées pour les distances inférieures. Lorsque le nombre de raccourcis est faible, le nombre de sauts $(4(i-1))$ est exacte.

De l'Eq. (4.7) on déduit la probabilité $P_r^1(\hat{\ell})$ qu'un nœud ne change pas sa distance $\hat{\ell}$ à travers un seul raccourci :

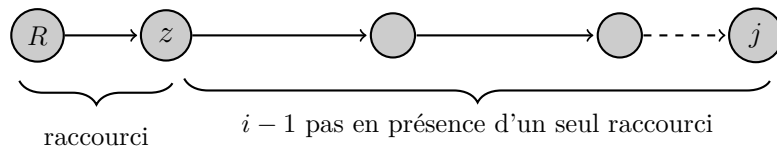
$$\begin{aligned}
 P_r^1(\hat{\ell}) &= \pi^1(1)\pi^1(2)\dots\pi^1(\hat{\ell}-1) \\
 &= (1-\hat{\phi})(1-\hat{\phi})^4(1-\hat{\phi})^{4\cdot 2}\dots(1-\hat{\phi})^{4(\hat{\ell}-2)} \\
 &= (1-\hat{\phi})^{1+4+4\cdot 2+4\cdot 3\dots 4(\hat{\ell}-2)} \\
 &= (1-\hat{\phi})^{1+4\sum_{i=1}^{\hat{\ell}-1}(i-1)}.
 \end{aligned} \tag{4.8}$$

Le terme $\pi^i(i) = (1-\hat{\phi})^i(\hat{n}-2i)^{i-1}$, et par commodité de calcul, peut être négligé. Ceci n'a pas d'influence sur le résultat car c'est une seule possibilité parmi beaucoup d'autres. Alors $P_r^1(\hat{\ell})$ s'écrit :

$$P_r^1(\hat{\ell}) = (1-\hat{\phi})^{4\sum_{i=1}^{\hat{\ell}-1}(i-1)}. \tag{4.9}$$

■ À travers 2 raccourcis :

Soit le nœud racine R et un nœud j quelconque, pour compter le nombre de possibilités pour aller de R à j sachant qu'il y a deux raccourcis entre les deux nœuds, on introduit le nœud arbitraire z qui se trouve entre les deux raccourcis. Dans ce cas, si la distance entre R et j en passant par z est i , alors le nombre de possibilités est $i-1$, c'est-à-dire les cas $\{\{1, i-1\}, \{2, i-2\}, \dots, \{i-1, 1\}\}$. Pour $i=4$, par exemple, on a les trois cas suivants : $\{\{1, 3\}, \{2, 2\}, \{3, 1\}\}$ ³. On peut schématiser le cas particulier $\{1, i-1\}$ comme : On constate que pour ce cas $(\{1, i-1\})$:



— Le nombre de chemins possibles entre R et z est 1 car ils sont directement liés à travers un raccourci.

3. Les combinaisons possibles pour que la distance entre le nœud j et le nœud racine R soit 4 :

- la distance entre le nœud R et le nœud intermédiaire z est 1 et entre z et le nœud j est 3.
- la distance entre le nœud R et le nœud intermédiaire z est 2 et entre z et le nœud j est également 2.
- la distance entre le nœud R et le nœud intermédiaire z est 3 et entre z et le nœud j est 1.

- Le nombre de chemins possibles entre z est j est $4(i-2)$, car la distance entre eux est $i-1$ (voir le cas d'un seul raccourci).

Dans ce cas, on déduit la probabilité pour que la distance entre les noeuds R et j ne soit pas égale à i est $(1 - \hat{\phi}^2)^{4(i-2)}$, où on a utilisé le fait que la probabilité pour que deux raccourcis se trouvent entre deux noeuds quelconques est $\hat{\phi}^2$.

Pour simplifier les calculs on suppose que tous les autres cas $\{\{2, i-2\}, \dots, \{i-1, 1\}\}$ possèdent la même probabilité que le cas $\{1, i-1\}$, ce qui représente une approximation de type champ moyen. Étant donné que le nombre de ces cas est $i-1$ alors la probabilité qu'un noeud ne change pas sa distance initiale vers la distance i via un certain noeud intermédiaire est $(1 - \hat{\phi}^2)^{4(i-2) \times (i-1)}$. Le nombre de positions possibles du noeud intermédiaire z dans le réseau est $\hat{n} - 2i$ (Fig.4.5), alors la probabilité qu'un noeud ne change pas sa distance initiale vers la distance i à travers deux raccourcis est

$$\pi^2(i) = (1 - \hat{\phi}^2)^{4(i-1)(i-2)(\hat{n}-2i)}, \quad (4.10)$$

d'où

$$\begin{aligned} P_r^2(\hat{\ell}) &= \pi^2(1)\pi^2(2)\dots\pi^2(\hat{\ell}-1) \\ &= (1 - \hat{\phi}^2)^{4 \sum_{i=1}^{\hat{\ell}-1} ((i-1)(i-2)(\hat{n}-2i))}, \end{aligned} \quad (4.11)$$

le terme $\pi^2(2)$ est exclu de la somme comme expliqué dans le cas d'un seul raccourci.

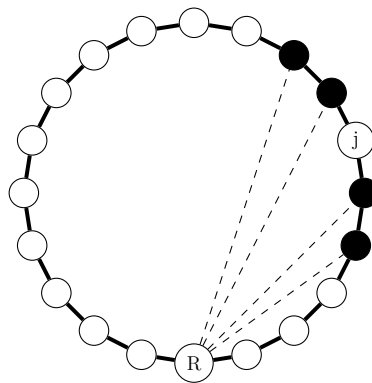
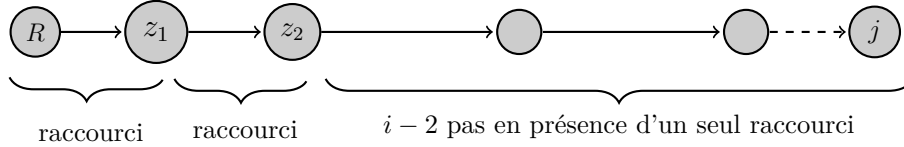


FIGURE 4.5 – Représentation du cas $\{1, i-1\}$. Les nœuds sombres et les nœuds j et R sont les positions qui ne peuvent pas être occupés par le nœud intermédiaire. Étant donné que $\hat{n} = 20$ et $i = 3$, le nombre de positions possibles pour le nœud intermédiaire est $\hat{n} - 2i = 14$.

■ À travers 3 raccourcis :

Soit le nœud racine R et un nœud j quelconque, pour compter le nombre de possibilités

pour aller de R à j sachant qu'il y a trois raccourcis entre les deux noeuds, on introduit les noeuds arbitraires z_1 et z_2 qui séparent les trois raccourcis (voir figure) . Dans ce cas, si la distance entre R et j en passant par z_1 et z_2 est i , alors le nombre de possibilités est $C_2^{i-1} = \frac{(i-1)(i-2)}{2!}$, c'est-à-dire les cas $\{\{1, 1, i-2\}, \{1, 2, i-3\}, \dots, \{i-2, 1, 1\}\}$. Pour $i = 4$, par exemple, on a les trois cas suivants : $\{\{1, 1, 2\}, \{1, 2, 1\}, \{2, 1, 1\}\}$. On peut schématiser le cas particulier $\{1, 1, i-2\}$ par



Pour le cas de la figure $(\{1, 1, i-2\})$, on a :

- Le nombre de chemins possibles entre R et z_1 est 1, car la distance entre eux est 1.
- Le nombre de chemins possibles entre z_1 et z_2 est 1, car la distance entre eux est 1.
- Le nombre de chemins possibles entre z_2 et j est $4(i-3)$, car la distance entre eux est $i-2$ (voir le cas d'un seul raccourci).

On déduit que la probabilité du cas $\{1, 1, i-2\}$ est $(1 - \hat{\phi}^3)^{4(i-3)}$. Comme dans le cas de deux raccourcis, pour simplifier les calculs on suppose que tous autres cas $\{\{1, 2, i-3\}, \dots, \{i-2, 1, 1\}\}$ possèdent la même probabilité que le cas $\{1, 1, i-2\}$. Étant donné que le nombre de ces cas est $C_2^{i-1} = \frac{(i-1)(i-2)}{2!}$ alors la probabilité qu'un noeud ne change pas sa distance initiale vers la distance i via deux noeud intermédiaires quelconques est $(1 - \hat{\phi}^3)^{4(i-2) \times \frac{(i-1)(i-2)}{2!}}$. Le nombre de positions possibles des noeuds intermédiaires z_1 et z_2 dans le réseau est borné supérieurement par $(\hat{n} - 2i)^2$, alors la probabilité qu'un noeud ne change pas sa distance initiale vers la distance i à travers trois raccourcis est

$$\pi^3(i) = (1 - \hat{\phi}^3)^{4 \frac{(i-1)(i-2)}{2!} (i-3)(\hat{n}-2i)^2} \quad (4.12)$$

d'où

$$\begin{aligned} P_r^3(\hat{\ell}) &= \pi^3(1)\pi^3(2)\dots\pi^3(\hat{\ell}-1) \\ &= (1 - \hat{\phi}^3)^{4 \sum_{i=1}^{\hat{\ell}-1} \left(\frac{(i-1)(i-2)}{2!} (i-3)(\hat{n}-2i)^2 \right)}. \end{aligned} \quad (4.13)$$

■ À travers m raccourcis :

En suivant la même démarche que les cas précédents, on généralise facilement pour le cas de m raccourcis. Dans ce cas on introduit $m-1$ noeuds intermédiaires, et on obtient :

$$\pi^m(i) = (1 - \hat{\phi}^m)^{4 \frac{(i-1)(i-2)\dots(i-m)}{(m-1)!} (\hat{n}-2i)^{(m-1)}}, \quad (4.14)$$

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

$$\begin{aligned} P_r^m(\hat{\ell}) &= \pi^m(1)\pi^m(2)\dots\pi^m(\hat{\ell}-1) \\ &= (1 - \hat{\phi}^m)^4 \sum_{i=1}^{\hat{\ell}-1} \left(\frac{(i-1)(i-2)\dots(i-m)}{(m-1)!} (\hat{n}-2i)^{m-1} \right), \end{aligned} \quad (4.15)$$

puisque $\hat{\phi} < 1$ on peut écrire

$$P_r^m(\hat{\ell}) = e^{-4q^m \sum_{i=1}^{\hat{\ell}-1} \left(\frac{(i-1)(i-2)\dots(i-m)}{(m-1)!} (\hat{n}-2i)^{m-1} \right)}. \quad (4.16)$$

Alors la probabilité $P_r(\hat{\ell})$ qu'un nœud ne change pas sa distance régulière $\hat{\ell}$ après l'ajout d'un nombre quelconque de raccourcis est

$$\begin{aligned} P_r(\hat{\ell}) &= P_r^1(\hat{\ell})P_r^2(\hat{\ell})P_r^4(\hat{\ell})\dots P_r^{\hat{\ell}-1}(\hat{\ell}) \\ &= e^{-4q \sum_{i=1}^{\hat{\ell}-1} (i-1)B(i)}, \end{aligned} \quad (4.17)$$

avec

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

qui n'est autre que la formule du binôme.

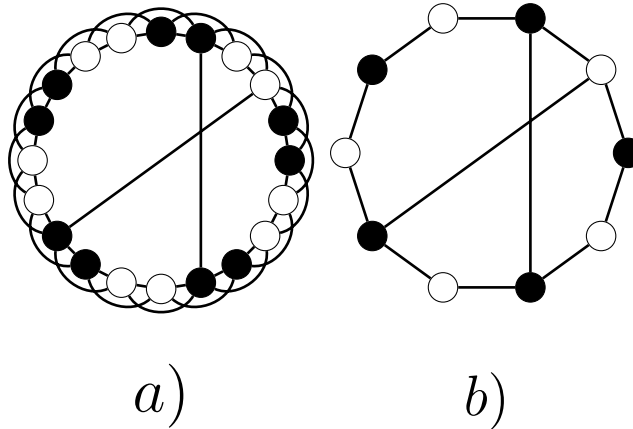


FIGURE 4.6 – La transformation du GR d'un réseau (a) de $n = 20$ et $k = 2$ vers un réseau (b) de $\hat{n} = 10$ et $\hat{k} = 1$.

Alors la probabilité qu'un nœud ne change pas sa distance $\hat{\ell}$ est :

$$\begin{aligned} P_r(\hat{\ell}) &= e^{-4q \sum_{i=1}^{\hat{\ell}-1} (i-1)(\hat{\phi}(\hat{n}-2i)+1)^{i-2}} \\ &\approx e^{-4\hat{\phi} \int_{i=1}^{\hat{\ell}-1} (i-1)(\hat{\phi}(\hat{n}-2i)+1)^{i-2} di}. \end{aligned} \quad (4.18)$$

La probabilité $P_{al}(\hat{\ell})$ qu'un nœud soit à une distance $\hat{\ell}$ du nœud racine après l'ajout des raccourcis est égale au produit de la probabilité que ce nœud ne change pas sa distance vers une autre inférieure à $\hat{\ell}$, $P_r(\hat{\ell}-1)$, par la probabilité que le nœud change sa distance vers la distance $\hat{\ell}$, $1 - \pi(\hat{\ell})$, avec $\pi(\hat{\ell}) = \prod_{i=1}^{\hat{\ell}-1} (\pi^i(\hat{\ell}))$.

$$\begin{aligned} P_{al}(\hat{\ell}) &= P_r(\hat{\ell}-1)(1 - \pi(\hat{\ell})) \\ &= P_r(\hat{\ell}-1) - \pi(\hat{\ell})P_r(\hat{\ell}-1) \\ &= P_r(\hat{\ell}-1) - P_r(\hat{\ell}) \\ &= -\frac{\partial P_r(\hat{\ell})}{\partial \hat{\ell}} \\ &= 4q(\hat{\ell}-2)(\hat{\phi}(\hat{n}-2(\hat{\ell}-1)-1)+1)^{\hat{\ell}-3}P_r(\hat{\ell}), \end{aligned} \quad (4.19)$$

pour simplifier les calculs on peut prendre $\hat{\ell}-1 \approx \hat{\ell}$ ⁴, on obtient

$$P_{al}(\hat{\ell}) = 4q(\hat{\ell}-1)(\hat{\phi}(\hat{n}-2\hat{\ell}-1)+1)^{\hat{\ell}-2}P_r(\hat{\ell}). \quad (4.20)$$

Le nombre moyen de nœuds dans la couche régulière $n_{\hat{\ell}}^r$ est $P_r(\hat{\ell})$ multiplié par 2, car initialement dans chaque couche on a deux nœuds réguliers à une distance $\hat{\ell}$

$$n_{\hat{\ell}}^r = 2P_r(\hat{\ell}), \quad (4.21)$$

et le nombre de nœuds dans la couche aléatoire $n_{\hat{\ell}}^{al}$ est $P_{al}(\hat{\ell})$ multiplié par le nombre de nœuds qui sont à une distance plus grande que $\hat{\ell}$

$$n_{\hat{\ell}}^{al} = (\hat{n} - 2\hat{\ell})P_{al}(\hat{\ell}), \quad (4.22)$$

d'où la $\hat{\ell}^{\text{ème}}$ couche qui est la somme des deux contributions s'écrit :

4. Cette approximation représente une erreur de l'ordre de $\frac{1}{\hat{\ell}}$, or $\hat{\ell}$ tend vers l'infini lorsque n est très large

$$\begin{aligned}
 n_{\hat{\ell}} &= n_{\hat{\ell}}^r + n_{\hat{\ell}}^{al} \\
 &= 2P_r(\hat{\ell}) + (\hat{n} - 2\hat{\ell})P_{al}(\hat{\ell}) \\
 &= 2P_r(\hat{\ell}) + (\hat{n} - 2\hat{\ell})P_r(\hat{\ell} - 1) - (\hat{n} - 2\hat{\ell})P_r(\hat{\ell}) \\
 &= (\hat{n} - 2\hat{\ell})P_r(\hat{\ell} - 1) - (\hat{n} - 2(\hat{\ell} + 1))P_r(\hat{\ell}) \\
 &= v(\hat{\ell} - 1) - v(\hat{\ell}) \quad \text{avec : } v(\hat{\ell}) = (\hat{n} - 2(\hat{\ell} + 1))P_r(\hat{\ell}).
 \end{aligned} \tag{4.23}$$

Pour quantifier la contribution des nœuds appartenant aux différentes couches (régulières et aléatoires) on calcule leurs sommes \hat{S}_r et \hat{S}_{al} , ce qui permettra aussi de comparer les deux contributions :

$$\begin{aligned}
 \hat{S}_r &= \hat{n}_1^r + \sum_{\hat{\ell}=2}^{\frac{\hat{n}}{2}} \hat{n}_{\hat{\ell}}^r \\
 &= \hat{n}_1^r + \int_2^{\frac{\hat{n}}{2}} \hat{n}_{\hat{\ell}}^r d\hat{\ell} \\
 &= \hat{n}_1^r + 2 \int_2^{\frac{\hat{n}}{2}} e^{-4q \int_{j=1}^{\hat{\ell}-1} (j-1)(\hat{\phi}(\hat{n}-2j)+1)^{j-2} dj} d\hat{\ell},
 \end{aligned} \tag{4.24}$$

l'intégrale $\int_2^{\frac{\hat{n}}{2}} e^{-4q \int_{j=1}^{\hat{\ell}-1} (j-1)(\hat{\phi}(\hat{n}-2j)+1)^{j-2} dj} d\hat{\ell}$ n'a aucune solution explicite. Or lorsque le nombre de raccourcis $\hat{\phi}$ est grand $\hat{S}_r \rightarrow 0$ car presque tous les nœuds sont affectés par les raccourcis, on peut donc approximer l'intégrale précédente par le cas où $\hat{\phi}$ est faible, on a $\int_2^{\frac{\hat{n}}{2}} e^{-4q \int_{j=1}^{\hat{\ell}-1} (j-1)(\hat{\phi}(\hat{n}-2j)+1)^{j-2} dj} d\hat{\ell} \approx \int_2^{\frac{\hat{n}}{2}} e^{-4\hat{\phi} \int_{j=1}^{\hat{\ell}-1} (j-1) dj} d\hat{\ell}$. Pour $\hat{n} \gg 1$ $\int_2^{\frac{\hat{n}}{2}} e^{-4\hat{\phi} \int_{j=1}^{\hat{\ell}-1} (j-1) dj} d\hat{\ell} \approx \int_2^{\frac{\hat{n}}{2}} e^{-4\hat{\phi} \int_{j=1}^{\hat{\ell}} j dj} d\hat{\ell}$, d'où :

$$\begin{aligned}
 \hat{S}_r &\approx \hat{n}_1^r + 2 \int_2^{\frac{\hat{n}}{2}} e^{-4\hat{\phi} \int_{j=1}^{\hat{\ell}} j dj} d\hat{\ell} \\
 &\approx \hat{n}_1^r + 2 \int_2^{\frac{\hat{n}}{2}} e^{-2\hat{\phi} \hat{\ell}^2} d\hat{\ell} \\
 &\approx 2 + 2 \left[\frac{\sqrt{\frac{\pi}{2}} \operatorname{erf}(\sqrt{2\hat{\phi}} \hat{\ell})}{2\sqrt{\hat{\phi}}} \right]_2^{\frac{\hat{n}}{2}} \quad \text{avec } \hat{n}_1^r = 2 \\
 &\approx 2 + 2 \sqrt{\frac{\pi}{8\hat{\phi}}} \left[\operatorname{erf}(\sqrt{2\hat{\phi}} \frac{\hat{n}}{2}) - \operatorname{erf}(2\sqrt{2\hat{\phi}}) \right],
 \end{aligned} \tag{4.25}$$

étant donné que $\hat{\phi} \ll 1$ on a $\operatorname{erf}(2\sqrt{2\hat{\phi}}) = \frac{2}{\sqrt{\pi}} 2\sqrt{2\hat{\phi}}$, alors on obtient

$$\begin{aligned}
 \hat{S}_r &\approx 2 + 2\sqrt{\frac{\pi}{8\hat{\phi}}}\left[\operatorname{erf}\left(\sqrt{2\hat{\phi}}\frac{\hat{n}}{2}\right) - \frac{2}{\sqrt{\pi}}2\sqrt{2\hat{\phi}}\right] \\
 &\approx 2 + 2\sqrt{\frac{\pi}{8\hat{\phi}}}\operatorname{erf}\left(\sqrt{\frac{\hat{\phi}\hat{n}^2}{2}}\right) - 4 \\
 &\approx \hat{n}\left(\sqrt{\frac{\pi}{2\hat{\phi}\hat{n}^2}}\operatorname{erf}\left(\sqrt{\frac{\hat{\phi}\hat{n}^2}{2}}\right) - \frac{2}{\hat{n}}\right) \\
 &\approx \hat{n}\sqrt{\frac{\pi}{2\hat{\phi}\hat{n}^2}}\operatorname{erf}\left(\sqrt{\frac{\hat{\phi}\hat{n}^2}{2}}\right).
 \end{aligned} \tag{4.26}$$

Or $\hat{\phi} = \frac{2k^3\phi}{n}$ et $\hat{n} = \frac{n}{k}$ alors $\frac{\hat{\phi}\hat{n}^2}{2} = kn\phi$, qui n'est autre que le nombre moyen de raccourcis dans le réseau. La somme des nœuds réguliers s'écrit sous la forme :

$$\hat{S}_r = \frac{n}{k}(1 - h(kn\phi)), \tag{4.27}$$

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

La somme des nœuds aléatoires est le nombre de nœuds total moins le nombre des nœuds réguliers, donc :

$$\begin{aligned}
 \hat{S}_{al} &\approx \hat{n} - \hat{S}_r \\
 &\approx \hat{n} - \hat{n}\sqrt{\frac{\pi}{2\hat{\phi}\hat{n}^2}}\operatorname{erf}\left(\sqrt{\frac{\hat{\phi}\hat{n}^2}{2}}\right) \\
 &\approx \hat{n}\left(1 - \sqrt{\frac{\pi}{2\hat{\phi}\hat{n}^2}}\operatorname{erf}\left(\sqrt{\frac{\hat{\phi}\hat{n}^2}{2}}\right)\right) \\
 &\approx \frac{n}{k}h(kn\phi).
 \end{aligned} \tag{4.28}$$

Les valeurs de \hat{S}_r et \hat{S}_{al} représentent en réalité le nombre des "amas" constitués de k nœuds dans la transformation du GR, alors pour trouver le nombre réel des nœuds dans le réseau il suffit de multiplier \hat{S}_r et \hat{S}_{al} par k , d'où le nombre des nœuds réguliers est $S_r = n(1 - h(kn\phi))$ et le nombre de nœuds aléatoires est $S_{al} = nh(kn\phi)$, où $h(x)$ est une fonction universelle qui représente la fraction des nœuds aléatoires. La Fig. 4.7 représente la variation des fractions des nœuds aléatoires et réguliers en fonction du nombre de raccourcis. Pour différentes valeurs du paramètre k , on a le même comportement de $\frac{S_r}{n}$ et $\frac{S_{al}}{n}$, ce qui confirme l'universalité de ces deux grandeurs. Une observation importante est que pour $x \leq 1$ nos équations sont parfaitement en accord avec la simulation. Lorsque $x > 1$ cet accord devient moins bon, ceci n'a rien de surprenant puisque les expressions de S_r et S_{al} ont été obtenus à partir de l'Eq. 4.24 en supposant que $\hat{\phi}$ (donc x) est petit.

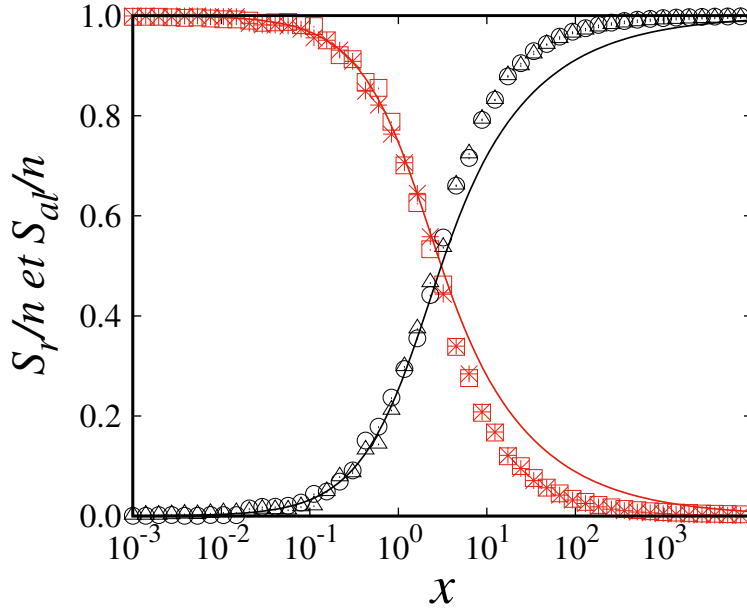


FIGURE 4.7 – Fraction des nœuds aléatoires $\frac{S_{al}}{n}$ (noir) et des nœuds réguliers $\frac{S_r}{n}$ (rouge) en fonction du nombre des raccourcis x . Les lignes représentent les deux fonctions $h(x)$ (noir) et $1 - h(x)$ (rouge). Les symboles représentent les simulations numériques d'un réseau de taille $n = 10^4$, avec $k = 1$ (étoile et cercle) et $k = 5$ (carré et triangle). Chaque simulation est moyennée sur 200 réalisations. L'échelle est linéaire-log

Comme a été déjà commenté, l'une des questions les plus intéressante dans le modèle de NW est l'existence ou non d'une transition de phase lors du passage du petit vers le grand monde. La quantité $\frac{S_r}{n}$ peut être considérée comme un éventuel paramètre d'ordre. L'ordre dans ce cas est la régularité du réseau, c'est-à-dire l'absence des raccourcis avec $\frac{S_r}{n} = 1$. À partir d'un certain nombre de raccourcis $\frac{S_r}{n} \rightarrow 0$, le réseau est desordonné. Il est pratiquement impossible de conclure sur l'existence de la transition de phase en étudiant le comportement de $\frac{S_r}{n}$ (Fig. 4.7). En général ce sont les fluctuations des grandeurs pertinentes qui peuvent donner l'information sur l'existence et la nature des transitions de phase. La Fig. 4.8 représente les fluctuations de $\frac{S_{al}}{n}$, soit $\sigma = \frac{\sqrt{\langle (\frac{S_{al}}{n})^2 \rangle - \langle \frac{S_{al}}{n} \rangle^2}}{\langle \frac{S_{al}}{n} \rangle}$, en fonction du nombre de raccourcis pour différentes valeurs de n . On remarque l'absence d'une divergence typique d'une transition de phase dans σ , qui présente un maximum au voisinage de $x = 4.6$. Cette allure de σ rappelle la forme de la chaleur spécifique en fonction de la température pour les systèmes paramagnétiques, et qui présente aussi un maximum appelé anomalie de Schottky.

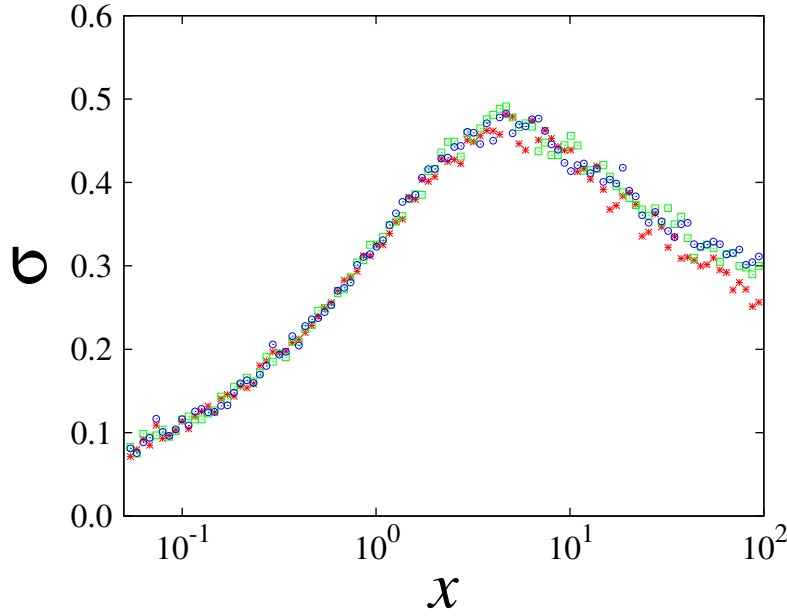


FIGURE 4.8 – Fluctuation σ en fonction de x d'un réseau de $k = 1$ pour différents valeurs de n , $n = 10^3$ (étoile), $n = 10^4$ (carré) et $n = 10^5$ (cercle), le nombre de réalisations est 1000. L'échelle est semi-logarithmique

4.2.2 Plus court chemin

Le plus court chemin dans le réseau circulaire est donné par $\langle \hat{\ell} \rangle = \sum_{\hat{\ell}=1}^{\frac{\hat{n}}{2}} (\hat{\ell} \cdot \hat{n}_{\hat{\ell}}) \approx \int_1^{\frac{\hat{n}}{2}} \hat{\ell} \cdot \hat{n}_{\hat{\ell}} d\hat{\ell}$, où $\hat{n}_{\hat{\ell}}$ est le nombre de noeud a une distance $\hat{\ell}$. Pour faciliter les calculs on introduit $\langle \hat{\ell}_r \rangle$ et $\langle \hat{\ell}_{al} \rangle$ qui sont respectivement le plus court chemin des réseaux réguliers et aléatoires. Le réseau régulier est formé des noeuds dont la distance n'a pas été modifiée par l'ajout des raccourcis, et le réseau aléatoire est formé des noeuds dont la distance initiale a été changée après l'ajout des raccourcis. Le plus court chemin du réseau global est $\langle \hat{\ell} \rangle = \langle \hat{\ell}_r \rangle + \langle \hat{\ell}_{al} \rangle$. Le plus court chemin du réseau régulier $\hat{\ell}_r$ s'écrit sous la forme

$$\langle \hat{\ell}_r \rangle = \frac{\hat{S}_r}{\hat{n}} \frac{\int_1^{\frac{\hat{n}}{2}} \hat{\ell} \cdot \hat{n}_{\hat{\ell}}^r d\hat{\ell}}{a_r}, \quad (4.29)$$

avec $a_r = \int_1^{\frac{\hat{n}}{2}} \hat{n}_{\hat{\ell}}^r d\hat{\ell} = \hat{S}_r$ est la constante de normalisation, d'où $\langle \hat{\ell}_r \rangle = \int_1^{\frac{\hat{n}}{2}} \hat{\ell} \cdot \frac{\hat{n}_{\hat{\ell}}^r}{\hat{n}} d\hat{\ell}$.

En utilisant les Eq. (4.18) et Eq. (4.21) on obtient

$$\langle \hat{\ell}_r \rangle = \frac{1}{\hat{n}} \int_1^{\frac{\hat{n}}{2}} 2\hat{\ell} e^{-4\hat{\phi} \int_{j=1}^{\hat{\ell}-1} (j-1)(\hat{\phi}(\hat{n}-2j)+1)^{j-2} dj} d\hat{\ell},$$

comme dans le cas de l'Eq. (4.24) on prend $\hat{\phi}$ petit, ce qui donne

$$\begin{aligned}
 \langle \hat{\ell}_r \rangle &\approx \frac{1}{\hat{n}} \int_1^{\frac{\hat{n}}{2}} 2\hat{\ell} e^{-2\hat{\phi}\hat{\ell}^2} d\hat{\ell} \\
 &\approx \frac{1}{\hat{n}} \frac{e^{-2\hat{\phi}} - e^{-\frac{\hat{\phi}\hat{n}^2}{2}}}{2\hat{\phi}} \\
 &\approx \frac{1}{\hat{n}} \frac{1 - e^{-\frac{\hat{\phi}\hat{n}^2}{2}}}{2\hat{\phi}} \quad \text{car : } \hat{\phi} \ll 1 \\
 &\approx \frac{\hat{n}}{4} \frac{1 - e^{-\frac{\hat{\phi}\hat{n}^2}{2}}}{\frac{\hat{\phi}\hat{n}^2}{2}}.
 \end{aligned} \tag{4.30}$$

Or $\frac{\hat{\phi}\hat{n}^2}{2} = nk\phi = x$ et $\hat{n} = \frac{n}{k}$, alors $\langle \hat{\ell}_r \rangle$ peut s'écrire sous la forme universelle suivante

$$\langle \hat{\ell}_r \rangle \approx \frac{n}{k} \hat{f}_r(x) \quad \text{avec} \quad \hat{f}_r(x) = \frac{1 - e^{-x}}{4x}. \tag{4.31}$$

Pour calculer $\langle \hat{\ell}_{al} \rangle$, nous cherchons la position de la couche contenant le plus de noeud (voir Section. ??), dans notre cas c'est le maximum de \hat{n}_{al} . Explicitelement, on résout l'équation $\frac{\partial \hat{n}_{al}(\hat{\ell})}{\partial \hat{\ell}} = 0$. Des Eq. (4.18) et Eq. (4.20) on obtient que

$$P_{al}(\hat{\ell}) = u(\hat{\ell}) e^{-\int_{j=1}^{\hat{\ell}-1} u(j) dj}, \tag{4.32}$$

où

$$u(\hat{\ell}) = 4\hat{\phi}(\hat{\ell} - 1)(\hat{\phi}(\hat{n} - 2\hat{\ell}) + 1)^{\hat{\ell}-2}. \tag{4.33}$$

Il est évident que $\langle \hat{\ell}_{al} \rangle$ prédomine dans $\langle \hat{\ell} \rangle$ si $S_{al} \gg S_r$, or la valeur du plus court chemin dans les réseaux aléatoires est de l'ordre de $\log(n)$ qui est très inférieure devant \hat{n} . On peut donc considérer que $\hat{n} - 2\hat{\ell} \approx \hat{n}$, donc $u(\hat{\ell}) = 4\hat{\phi}(\hat{\ell} - 1)(\hat{\phi}\hat{n} + 1)^{\hat{\ell}-2}$. Soit $y = \hat{\phi}\hat{n} = 2k^2\phi$, le paramètre qui représente le nombre moyen de raccourcis pour chaque **amas** après la transformation de GR, car $\hat{\phi}$ est la probabilité pour qu'un pair d'**amas** se connecte et \hat{n} est le nombre d'**amas**.

Des Eq. (4.22) et Eq. (4.32) on obtient

$$\hat{n}_{al}(\hat{\ell}) = \hat{n} u(\hat{\ell}) e^{-\int_{j=1}^{\hat{\ell}-1} u(j) dj}, \tag{4.34}$$

donc

$$\begin{aligned}
 \frac{\partial \hat{n}_{al}(\hat{\ell})}{\partial \hat{\ell}} &= \hat{n} \frac{\partial u(\hat{\ell})}{\partial \hat{\ell}} e^{-\int_{j=1}^{\hat{\ell}-1} u(j) dj} + \hat{n} u(\hat{\ell}) \frac{\partial e^{-\int_{j=1}^{\hat{\ell}-1} u(j) dj}}{\partial \hat{\ell}} \\
 &= \hat{n} \frac{\partial u(\hat{\ell})}{\partial \hat{\ell}} e^{-\int_{j=1}^{\hat{\ell}-1} u(j) dj} - \hat{n} u(\hat{\ell})^2 e^{-\int_{j=1}^{\hat{\ell}-1} u(j) dj},
 \end{aligned} \tag{4.35}$$

alors le maximum de \hat{n}_{al} est donnée par la solution de l'équation suivante

$$\frac{\partial u(\hat{\ell})}{\partial \hat{\ell}} - u(\hat{\ell})^2 = 0. \quad (4.36)$$

D'après l'Eq. (4.33) on a $\frac{\partial u(\hat{\ell})}{\partial \hat{\ell}} = u(\hat{\ell}) \left[\frac{1}{\hat{\ell}-1} + \ln(y+1) \right]$. Puisque $y = 2k^2\phi$ ne dépend pas de la taille du réseau \hat{n} alors que $\hat{\ell}$ croît avec \hat{n} on peut négliger $\frac{1}{\hat{\ell}-1}$ devant $\ln(y+1)$, on obtient $\frac{\partial u(\hat{\ell})}{\partial \hat{\ell}} = u(\hat{\ell}) \ln(y+1)$. L'Eq. (4.36) devient

$$u(\hat{\ell}) = \ln(y+1), \quad (4.37)$$

En remplaçant $u(\hat{\ell})$ par son expression (Eq. (4.33)) on trouve la solution de l'Eq. (4.37) qui donne la position $\hat{\ell}_{max}$ de \hat{n}_{al} maximale :

$$\hat{\ell}_{max} = \frac{W\left(\frac{\ln(y+1)^2(y+1)}{4\phi}\right)}{\ln(y+1)} + 1, \quad (4.38)$$

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

Pour obtenir l'expression de $\langle \hat{\ell}_{al} \rangle$ on multiplie $\hat{\ell}_{max}$ par la fraction des nœuds aléatoires : $\langle \hat{\ell}_{al} \rangle = \hat{\ell}_{max} h(x)$, d'où

$$\langle \hat{\ell}_{al} \rangle = \left(\frac{W\left(\frac{\ln(y+1)^2(y+1)}{4\phi}\right)}{\ln(y+1)} + 1 \right) h(x). \quad (4.39)$$

On peut écrire finalement l'expression du plus court chemin dans le modèle de NW en sommant les Eq. (4.31) et Eq. (4.39) :

$$\langle \hat{\ell} \rangle = \left(\frac{W\left(\frac{\ln(y+1)^2(y+1)}{4\phi}\right)}{\ln(y+1)} + 1 \right) h(x) + \hat{n} \frac{1 - e^{-x}}{4x}. \quad (4.40)$$

Pour trouver une expression universelle de $\langle \hat{\ell} \rangle$ en fonction du nombre de raccourcis x , il faut prendre $y \ll 1$ dans l'équation Eq. (4.40), soit :⁵.

$$\langle \hat{\ell} \rangle = \hat{n} \hat{f}(x), \quad (4.41)$$

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

De la Fig. 4.9, on voit que l'Eq. (4.40) est en très bon accord avec les simulations en dehors de l'intervalle $]1, 10]$. La Fig. 4.9(b) montre que l'Eq. (4.40) est plus précise que l'Eq. (4.5) de Newman et al. notamment pour les grandes valeurs de x . L'intervalle $]1, 10]$ où notre théorie est en désaccord avec la simulation correspond dans la Fig. 4.8 au maximum de fluctuations du paramètre $\frac{S_{al}}{n}$, donc c'est l'intervalle où notre approximation qui est de champ moyen est moins précise.

5. Cette condition n'est pas toujours vraie, ceci sera discuté par la suite.

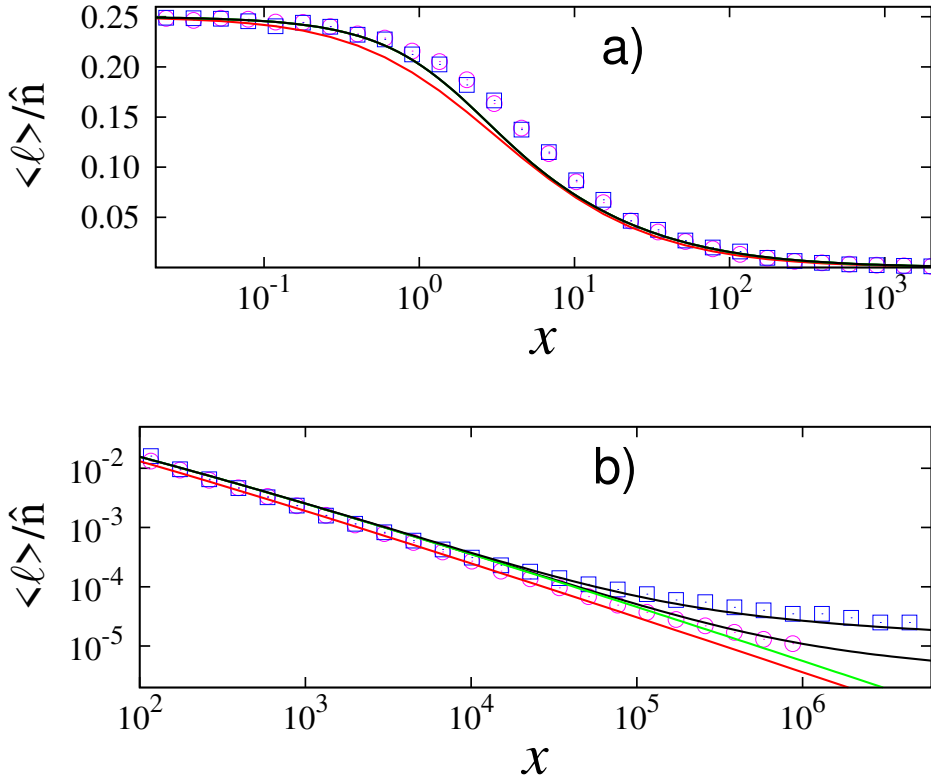


FIGURE 4.9 – Comportement de $\frac{\langle \ell \rangle}{\hat{n}}$ en fonction de x pour un réseau de taille $n = 10^6$. La formule de Newman et watts Eq. (4.5) (ligne rouge), Eq. (4.41) (ligne verte), Eq. (4.40) (ligne noire) et les simulations numériques pour $k = 1$ (cercle) et $k = 5$ (carré). L'échelle est semi-logarithmique dans a) et log-log dans b). Chaque simulation est moyennée sur 100 réalisations.

4.3 Émergence de la propriété petit-monde dans le modèle de NW

4.3.1 Validité de la fonction universelle de NW

L'expression universelle $\frac{\langle \ell \rangle}{\hat{n}} = f(x)$ est obtenue pour $y \ll 1$, ceci va être confirmé dans ce qui suit par les simulations. Pour cela on définit Δ tel que :

$$\Delta = \frac{\frac{\langle \ell \rangle}{\hat{n}}}{f(x)}, \quad (4.42)$$

de sorte que si la fonction $f(x) = \frac{\langle \ell \rangle}{\hat{n}}$ est universelle, alors Δ doit être égale 1.

Dans la Fig. 4.10 on représente Δ en fonction de y et x , $\langle \ell \rangle$ est obtenu par simulation du modèle de NW, et $f(x)$ est l'Eq. (4.5). On observe que Δ est effectivement égale 1 pour $y \ll 1$. Pour les valeurs supérieures de y , Δ est nettement supérieur à 1, par conséquent $f(x)$

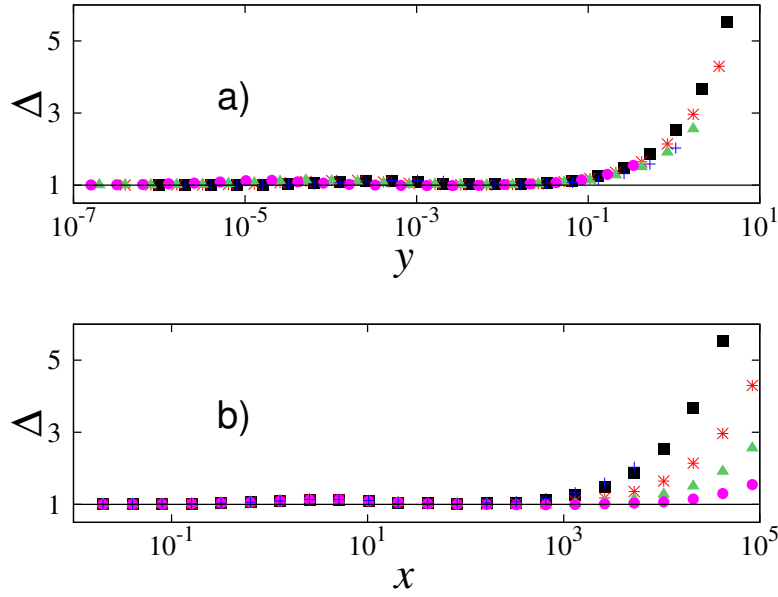


FIGURE 4.10 – Variations de Δ en fonction de y (a) et x (b) pour différentes valeurs de n et k . $k = 1$ et $n = 10^5$ (triangle), $k = 2$ et $n = 10^4$ (plus), $k = 2$ et $n = 10^6$ (cercle), $k = 2$ et $n = 10^5$ (étoile), $k = 5$ et $n = 10^6$ (carré). Le nombre de réalisations pour chaque simulation est 1000. L'échelle est semi-logarithmique

n'est plus valable.⁶ La Fig. 4.10(b) montre aussi que la fonction $f(x)$ n'est pas universelle à partir d'une certaine valeur de x qui dépend des paramètres du réseau. Les Fig. 4.10(b) et Fig. 4.9 suggèrent que le rapport $\frac{\langle \ell \rangle}{n}$ ne peut être écrit sous forme d'une fonction universelle du paramètre $x = nk\phi$. En d'autres termes, et en suivant l'argumentation de Newman [159], il faut chercher à exprimer le plus court chemin en fonction d'autres paramètres.

4.4 Une nouvelle fonction universelle

L'universalité de Δ en fonction du paramètre y (Fig. 4.10(a)) et le changement de son comportement à partir d'une certaine valeur de y nous a poussé à représenter $\frac{\ln(n)}{\langle \ell \rangle}$ en fonction de ce paramètre (Fig. 4.11). L'idée est de voir si ce changement est lié au passage du réseau grand monde lorsque $\langle \ell \rangle \propto n$ au petit monde lorsque $\langle \ell \rangle \propto \ln(n)$. On observe dans la Fig. 4.11 que lorsque $y \ll 1$ on a $\frac{\ln(n)}{\langle \ell \rangle} \rightarrow 0$, ce qui correspond au comportement grand monde du réseau car dans ce cas $\langle \ell \rangle \propto n$, par contre lorsque y n'est plus très inférieur à 1, $\langle \ell \rangle \propto \ln(n)$, ce qui correspond au comportement petit monde. Un résultat très important ici est que lorsque le

6. Newman a déjà soulevé le problème avec $f(x)$ lorsque $\phi = 1$ [159].

4.4. Une nouvelle fonction universelle

réseau est petit monde c'est $\frac{\langle \ell \rangle}{\log(n)}$ qui est universelle et non pas $\frac{\langle \ell \rangle}{n}$ comme dans [8]. En outre $\frac{\langle \ell \rangle}{\log(n)}$ doit être exprimée en fonction de y et non pas en fonction de x . Pour trouver l'expression analytique de la nouvelle fonction universelle $g(y)$ telle que $g(y) = \frac{\langle \ell \rangle}{\log(n)}$ on prend n grand dans l'Eq. (4.40). Dans ce cas $W\left(\frac{\ln(y+1)^2(y+1)}{4\hat{\phi}}\right) \approx \ln\left(\frac{\ln(y+1)^2(y+1)}{4\hat{\phi}}\right)$, $h(x) = 1 - \sqrt{\frac{\pi}{4x}} \text{erf}(\sqrt{x}) \rightarrow 1$, et $\frac{y+1}{4\hat{\phi}} = n^{\frac{2k^2\phi+1}{8k^3\phi}}$. D'autre part, le réseau est petit monde si y n'est pas très inférieur à 1 (Fig. 4.11). L'eq. (4.40) devient :

$$\begin{aligned} \langle \hat{\ell} \rangle &= \frac{\ln \ln(y+1)^2 + \ln \frac{y+1}{4\hat{\phi}}}{\ln(y+1)} + 1 + \frac{1 - e^{-x}}{2y} \\ &\approx \frac{\ln n}{\ln(y+1)} \end{aligned} \quad (4.43)$$

$\langle \hat{\ell} \rangle$ est le plus court chemin entre les **amas** du réseau issus de la transformation GR, pour

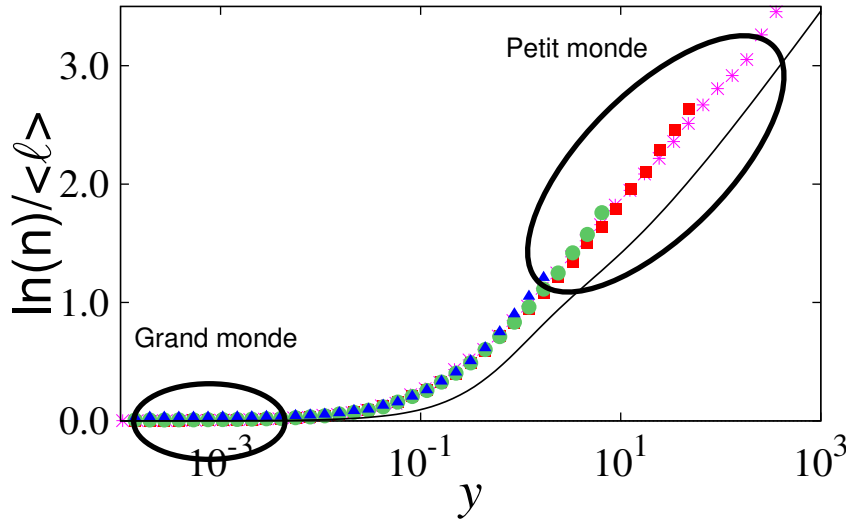


FIGURE 4.11 – $\frac{\ln(n)}{\langle \ell \rangle}$ en fonction de y . La ligne continue est l'Eq. (4.46), les symboles représentent les simulations numériques pour différents valeurs de n et k : $n = 10^{-6}$ et $k = 20$ (étoile), $n = 10^{-5}$ et $k = 5$ (carré), $n = 10^{-4}$ et $k = 2$ (cercle), $n = 10^{-3}$ et $k = 1$ (triangle), nombre de réalisations pour chaque simulations est 100. L'échelle est semi-logarithmique.

obtenir le plus court chemin $\langle \ell \rangle$ entre deux noeuds dans le réseau de NW, on cherche la relation entre les deux quantités en utilisant l'approximation suivante : Entre deux **amas** quelconques dans le réseau il peut y avoir des liens ordinaires et/ou des raccourcis. Dans le cas où il y a deux raccourcis liés au même **amas** intermédiaire (Fig. 4.12), la probabilité qu'ils soient liés au même noeud dans cet **amas** est $\frac{1}{k}$ (Fig. 4.12(b)), et la probabilité que les deux raccourcis ne soient pas

liés au même nœud (Fig. 4.12(a)), donc faisant intervenir un autre lien, est $1 - \frac{1}{k}$, autrement dit chaque fois qu'on a deux raccourcis pointant vers le même **amas** il y a une probabilité de $1 - \frac{1}{k}$ que le plus court chemin augmente de 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 $\frac{\hat{\phi}\hat{n}(\hat{n}-1)}{2} \approx \frac{\hat{\phi}\hat{n}^2}{2}$, $\hat{\phi}$ est la probabilité qu'un pair d'**amas** se connecte par un raccourci et $\frac{\hat{n}(\hat{n}-1)}{2}$ est le nombre de paires possibles. Alors la fraction de raccourcis par rapport au nombre total de liens est $\frac{\hat{\phi}\hat{n}}{2+\hat{\phi}\hat{n}} = \frac{y}{2+y}$, d'où la probabilité d'obtenir deux raccourcis liés au même **amas** entre deux nœuds quelconques du réseau est $(\frac{y}{2+y})^2$. Par conséquent la probabilité pour que la valeur du plus court chemin entre deux **amas** soit doublée est $(\frac{y}{2+y})^2(1 - \frac{1}{k})$. La probabilité que le plus court chemin reste inchangé est $(1 - (\frac{y}{2+y})^2) + \frac{1}{k}(\frac{y}{2+y})^2$, le premier 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 également au même nœud, ceci se traduit par :

$$\langle \ell \rangle = \langle \hat{\ell} \rangle \left[\frac{1}{k} \left(\frac{y}{2+y} \right)^2 + \left(1 - \left(\frac{y}{2+y} \right)^2 \right) + 2 \left(1 - \frac{1}{k} \right) \left(\frac{y}{2+y} \right)^2 \right], \quad (4.44)$$

en remplaçant dans l'Eq. (4.43), on déduit :

$$\frac{\langle \ell \rangle}{\ln(n)} = \frac{\frac{1}{k} \left(\frac{y}{2+y} \right)^2 + \left(1 - \left(\frac{y}{2+y} \right)^2 \right) + 2 \left(1 - \frac{1}{k} \right) \left(\frac{y}{2+y} \right)^2}{\ln(y+1)}, \quad (4.45)$$

en prenant $\frac{1}{k} \ll 1$ on obtient :

$$\frac{\langle \ell \rangle}{\ln(n)} = \frac{\left(1 + \left(\frac{y}{2+y} \right)^2 \right)}{\ln(y+1)} = g(y) \quad (4.46)$$

Cette fonction montre un certain écart par rapport aux résultats numériques, car dans la déduction de $\langle \ell \rangle$ en fonction de $\langle \hat{\ell} \rangle$ nous avons seulement pris en compte le cas de deux raccourcis liés au même **amas**, il s'agit donc d'une approximation de type champs moyen. Cependant l'expression de $g(y)$ se compare qualitativement bien avec les simulations (Fig. 4.11). Il convient de signaler que la fonction universelle h est fonction de x et pas de y , car nous avons ignoré dans le calcul de $h(x)$ le cas où $\hat{\phi}$ est grand (donc x est grand), alors cette expression n'est valable que pour un petit nombre de raccourcis.

4.5 Conclusion

Notre étude du modèle de NW est extensive, et nous avons introduit une nouvelle façon d'aborder ce système. En effet, et à fin d'obtenir plus d'informations sur les différents aspects

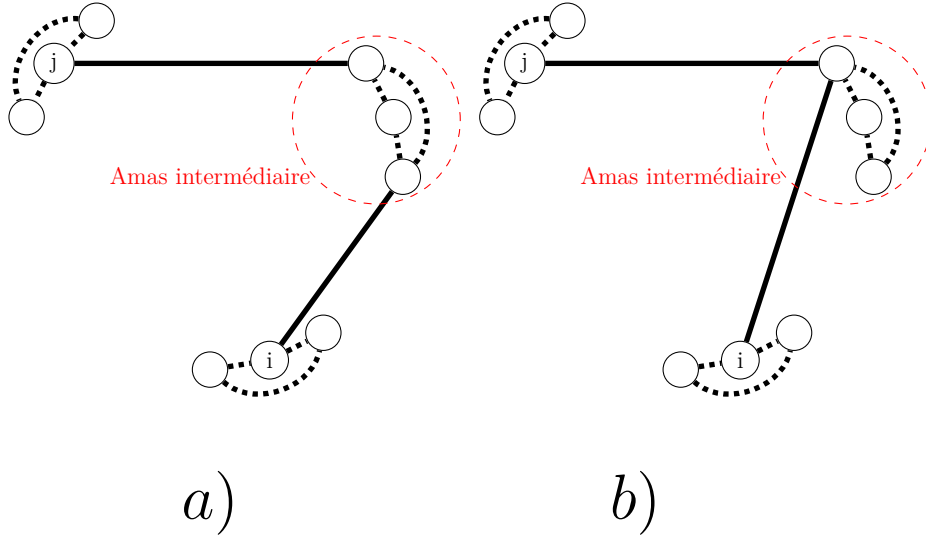


FIGURE 4.12 – Distance entre deux amas ayant un amas commun (intermédiaire) dans le réseau de NW, avec deux raccourcis (ligne continue) entres eux et $k = 3$. Les lignes en pointillés représentent les liens régulières. Dans a) la distance entre le noeud i et j est 3 car les deux raccourcis ne sont pas liés au même noeud de l’amas intermédiaire, dans b) la distance entre le noeud i et j est 2 car les deux raccourcis sont liés au même noued de l’amas intermédiaire.

du réseau, nous avons séparé la contribution des noeuds “réguliers” et “aléatoires”. En utilisant la transformation de GR, nous avons pu calculer la fraction des noeuds “réguliers” (S_r) et “aléatoires” (S_{al}). En calculant les fluctuations de $\frac{S_r}{n}$ qu’on a considéré comme paramètre d’ordre, on a conclu qu’il n’y pas de transitions de phase de grand monde vers petit monde. À partir de S_r et S_{al} nous avons déduit le plus court chemin $\langle \ell \rangle$. Notre expression pour $\langle \ell \rangle$ est plus précise que celle donnée par Newman et al. [159]. En outre, cette expression nous amène à conclure que la fonction universelle de Newman et al. [8] n’est valable que lorsque le réseau est grand monde, ce qui se traduit dans nos équations par $y = 2k^2\phi \ll 1$. Lorsque le réseau est petit monde, nous avons introduit une nouvelle fonction universelle quit s’écrit en fonction de y . En général, nous avons déduit que dans le modèle de NW, lorsque le réseau est grand le monde, la fonction universelle ($f(x)$) doit s’écrire en fonction de x et le plus court chemin en fonction de n et x , alors que lorsque le réseau est petit monde, la fonction universelle ($g(y)$) doit être fonction de y et le plus court chemin en fonction de y et $\ln(n)$.

Percolation dans les réseaux complexes

Sommaire

5.1	Introduction	86
5.2	Percolation	87
5.2.1	Percolation dans le réseau aléatoire d'Erdős-Rényi	89
5.2.2	Percolation explosive	91
5.2.3	Percolation explosive dans d'autres modèles	93
5.3	Nouvelle approche pour prédire le type de la transition dans les systèmes percolatifs	94
5.3.1	État de l'art	94
5.3.2	Nouvelle approche	95
5.4	Conclusion	103

5.1 Introduction

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

La grande présence des transitions de phase dans le monde réel a poussé les physiciens, depuis le tout début de la mécanique statistique, à découvrir des comportements universels et critiques ainsi que de faire un grand effort pour la classification de ces transitions de phases. La première tentative à fournir une classification a été réalisée par Ehrenfest [?]. Actuellement, 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 hystérésis. Par contre, 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 de transition. 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.

5.2 Percolation

La percolation est l'émergence d'une connectivité à grande échelle. L'ampleur de la connectivité qui apparaît soudainement à un point de transition a un impact profond sur les comportements macroscopiques d'un système.

Le terme de percolation a été introduit pour la première fois en 1956 par Broadbent et Hammersley [?], lors de leur étude de l'écoulement d'un fluide dans un milieu poreux. La théorie développée depuis lors, dite de la percolation, s'appuie sur des considérations statistiques pour caractériser le comportement d'un ensemble d'objets incomplètement connectés. Dans un tel système, la communication à longue distance est soit possible soit impossible selon le nombre d'objets et de contacts entre eux. La transition entre ces deux régimes est repérée par le point où le paramètre décrivant la connectivité prend une valeur critique, qu'on appelle le seuil de percolation. Ainsi, on passe d'un état de connectivité à courte distance à un état de connectivité à longue distance lorsqu'on franchit cette valeur seuil, en faisant croître continûment la concentration des objets et/ou de leurs liens. Les propriétés du système changent alors brusquement en ce point, et la transition de percolation se comporte comme une transition de phase du deuxième ordre (ou continue). Cette théorie, qui est désormais entrée dans le cadre connu des transitions de phases, s'est fortement développée au cours des dernières décennies, compte tenu du nombre croissant de ses domaines d'applications. On peut par exemple citer

les transitions para-ferromagnétique dans les alliages magnétiques dilués, états localisés-états étendus dans les semi-conducteurs amorphes, sol-gel dans les polymères, métal-supraconducteur ou isolant-conducteur dans les composites [?]. Actuellement, la percolation est également au centre des investigations sur les réseaux sociaux, biologiques, et artificiels [164, 165, 166].

Un processus de percolation standard peut avoir lieu, en général, par deux types : par sites ou par liens. La percolation par sites sur un graphe donné signifie que les sommets sont vides avec une probabilité donnée p (ou occupée avec une probabilité $1 - p$), tandis que la percolation par liens fait référence à l'existence ou non d'un bord entre deux nœuds choisis arbitrairement. Une fois la suppression (ou le placement) aléatoire des nœuds ou des arêtes effectuée, plusieurs quantités permettent la caractérisation des propriétés du réseau. On examine généralement l'existence et la taille de l'amas infini, appelée aussi composante géante, en fonction p , ainsi que la taille moyenne et les fluctuations de la taille des amas finis (ou petites composantes). De cette manière, on définit une probabilité critique p_c en dessous de laquelle le réseau percole, et un ensemble d'exposants critiques caractérisant la transition de phase (Fig. 5.1).

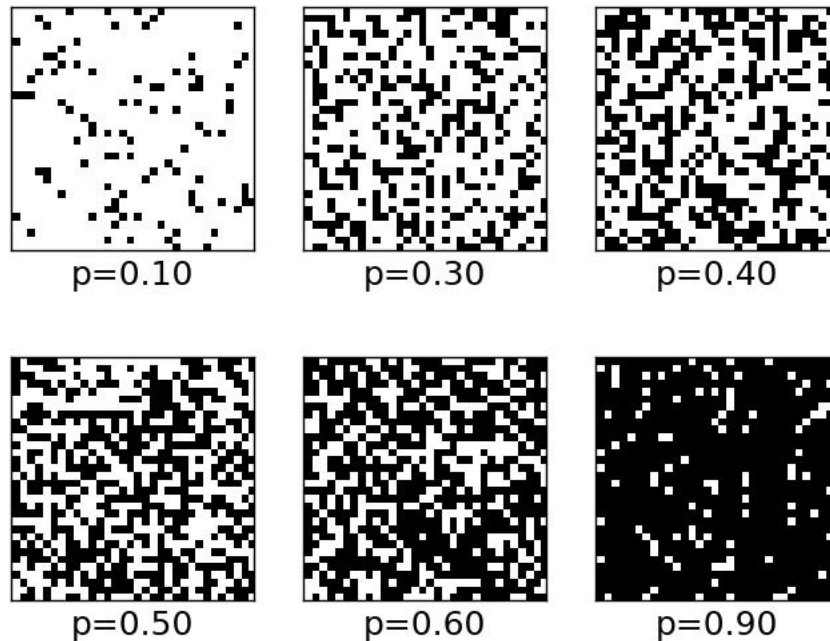


FIGURE 5.1 – Une démonstration de percolation par site sur une grille à deux dimensions pour différentes valeurs de p . Sous le seuil de percolation le système est composé de petits amas, après un certain point critique p_c un amas “infini” occupe la grille.

5.2.1 Percolation dans le réseau aléatoire d'Erdős-Rényi

Un réseau aléatoire peut être 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 [10, 11, 12], il est souvent désigné par son nom mathématique $G(n, E)$, avec n est le nombre de nœuds et E le nombre de liens. Dans la cas où n est très grand, plusieurs propriétés de l'ensemble des réseaux aléatoires ont été exprimées analytiquement et parfois exactement.

Les réseaux aléatoires ER sont les plus étudiés parmi les modèles de réseaux, 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 réseau 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 réseau 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 (voir Fig. 5.2). En particulier, si on considère la fraction de nœuds appartenant à 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. Erdős et Rényi ont étudié la distribution des degrés minimum et maximum dans le graphe aléatoire, mais la distribution des degrés complète a été obtenue plus tard par Bollobás [167]. La probabilité qu'un nœud ait un degré k suit la distribution binomiale :

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

Pour n très grand, cette distribution devient une distribution de Poisson :

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

Le coefficient de clustering, C , est une quantité très simple à calculer pour le graphe aléatoire, rappelons qu'il est défini comme la probabilité que deux voisins d'un nœud du réseau soient

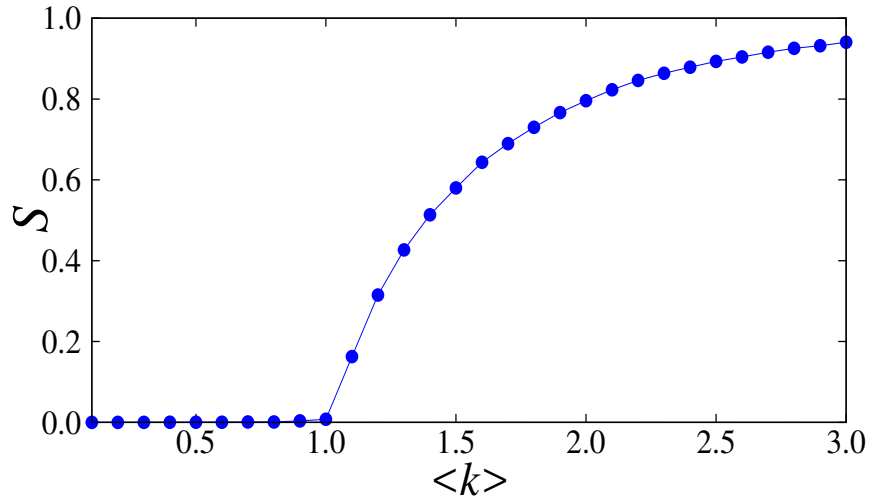


FIGURE 5.2 – Les simulations numériques du seuil de percolation dans le modèle ER en fonction du degré moyen, pour un réseau de taille $n = 10^5$. On observe que le point où la fraction de la composante géante, S , émerge est à $\langle k \rangle = 1$.

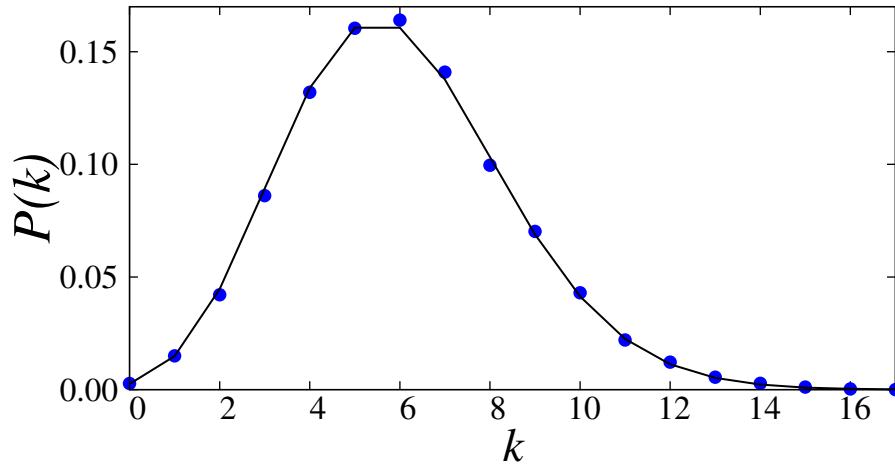


FIGURE 5.3 – Distribution des degrés d'un réseau aléatoire ER de degré moyen $\langle k \rangle = 6$, les cercles représentent les simulations numériques et la ligne noire représente l'Eq. (5.2).

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 :

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

Cette valeur de C 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 clustering assez élevés.

5.2.2 Percolation explosive

Une idée prévalente, pendant assez de temps, est que la plupart des transitions de la percolation classique (ordinaire) semblent être de second ordre (continue) comme la percolation par liens dans les réseaux ER. Dans ce cas, la taille de la composante géante augmente progressivement lorsque le paramètre de contrôle qui est le degré moyen $\langle k \rangle$ dépasse le seuil de percolation. Il convient toutefois de noter qu'il existe également des exemples, bien que rares, de modèles de percolation présentant une transition de phase de premier ordre (discontinue), tels que les percolations bootstrap [168], k-core [169], et brouillage [170].

Les transitions de phase de percolation de premier ordre ont suscité un intérêt considérable depuis le travail de Achlioptas et al. [2]. Dans ce travail, les auteurs fournissent des preuves solides, mais finalement trompeuses, que la connectivité à grande échelle émerge dans une transition de phase discontinue qu'ils ont appelés "percolation explosive".

La percolation explosive (PE) est un phénomène qui résulte souvent de petites interventions conçues pour retarder la transition de phase de percolation. Le déclenchement peut en effet être considérablement retardé, mais une fois la transition de percolation atteinte inévitablement, une connectivité à grande échelle apparaît de manière brusque [171, 172, 173]. L'idée principale du processus Achlioptas (PA) [2] est de modifier la règle pour générer des graphes ER. Au lieu d'ajouter des liens aléatoires, dans le PA 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.

Dans [2], les auteurs se sont demandés s'il était possible de déplacer le point critique de la transition de phase de percolation 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 $\{e_1, e_2\}$ (voir Fig. 5.4). Cette règle a été suggérée dans [174] comme la plus susceptible de retarder le

point critique. Une autre règle est la règle de la somme, où la taille de la nouvelle composante formée est minimisée. Considérons la Fig. 5.4 : (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. Suivant 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 choisit e_2 au lieu de e_1 . (C) Évolution de la fraction de nœud dans la composante géante S pour : ER, BF (une règle de taille bornée, voir [175]), et RP. Le réseau étudié est de taille $n = 512000$.

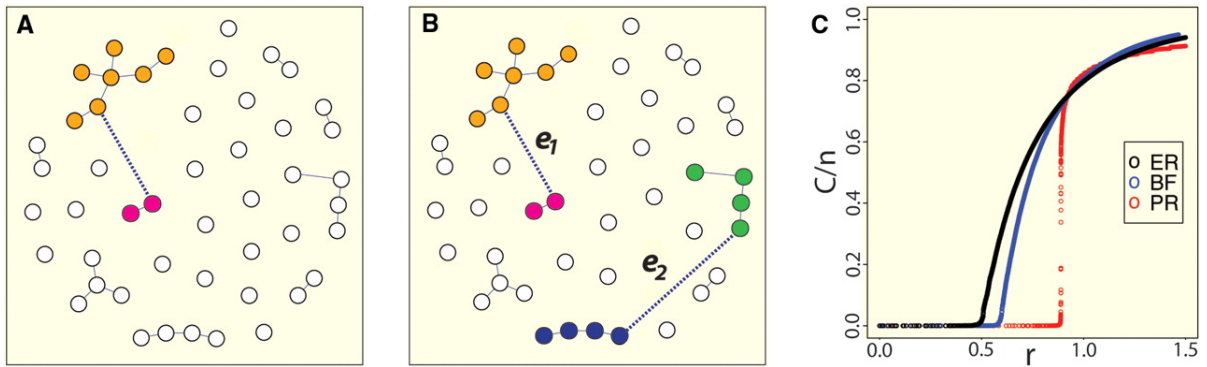


FIGURE 5.4 – Comparaison entre la percolation aléatoire et PA. (A) Percolation ER classique, où les liens sont ajoutés au hasard dans le réseau. (B) Percolation PA, où à chaque étape deux liens sont en compétition pour être établis. (C) Paramètre d'ordre : la taille relative de la composante géante par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

La Fig. 5.4.(c) peut suggérer que la transition de percolation RP est discontinue. Sur la base de l'analyse de l'intervalle de transition, il a été avancé que la transition de percolation sous AP est explosive et discontinue [2]. Cependant, cette conjecture a été rapidement et rigoureusement désapprouvée par Riordan et Warnke [176, 177], 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 [177]. 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 [178] 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 [177] 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 PE 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.

5.2.3 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 la PE. Dans ce qui suit, nous examinons brièvement certains de ces modèles qui sont détaillés entre autres dans la référence [179] :

5.2.3.1 Le modèle de probabilité

Ce modèle consiste à donner pour tout amas i un poids δ_i , à chaque pas de temps un lien entre une paire d'amas (i, j) est occupé par une probabilité proportionnelle à $(\delta_i \delta_j)^\alpha$. La somme de tous les poids est déterminée comme une constante de normalisation $w = \sum_i \sum_j (\delta_i \delta_j)^\alpha$, alors la probabilité pour que l'amas i soit lié à un autre amas j est $\frac{(\delta_i \delta_j)^\alpha}{w}$ [180].

Un autre exemple est celui appelé le modèle hamiltonien de la référence [181], 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 amas et du nombre de liens ajoutés à ces amas. 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 ΔH_e est le changement d'énergie après l'ajout d'un lien.

5.2.3.2 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 [182] la RP est utilisée, dans la référence [183] un mélange de $k = 2$ et $k = 3$ -core sur le réseau aléatoire et 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 [183, 182] et des réseaux ER [183, 182, 184].

5.2.3.3 Les modèles bootstrap

Dans le cadre de la PE, certains modèles bootstrap ont été proposés [185]. Le processus de percolation bootstrap standard sur un réseau suppose 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 i voisins les plus proches sont actifs (avec $i = 2, 3, \dots$). 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.

5.3 Nouvelle approche pour prédire le type de la transition dans les systèmes percolatifs

5.3.1 État de l'art

La prédiction du type de la transition de phase est un sujet assez compliqué, et il n'existe pas encore de formalisme théorique qui traite la question d'une façon rigoureuse. Autrement dit, 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. Et c'est le cas aussi pour la transition de percolation qui est l'un des modèles les plus simples et les plus anciens dans la physique statistique. Cependant pour quelques modèles spécifiques, par

5.3. Nouvelle approche pour prédire le type de la transition dans les systèmes percolatifs

exemple dans [180], une équation cinétique a été utilisée pour déterminer le type de la transition de la percolation. Dans l'évolution des réseaux ER, par exemple, l'équation différentielle est écrite à la limite thermodynamique comme :

$$\frac{\partial n(k_a, t)}{\partial t} = \sum_{i+j=k_a} \frac{\delta_i n(i, t)}{c(t)} \frac{\delta_j n(j, t)}{c(t)} - 2 \frac{\delta_{k_a} n(k_a, t)}{c(t)}, \quad (5.4)$$

où $c(t) = \sum_s \delta_s n(k_a)$ est la constante de normalisation, $n(k_a, t)$ est le nombre d'amas de taille k_a à l'instant t , $\frac{\delta_i n(i, t)}{c(t)} \frac{\delta_j n(j, t)}{c(t)}$ est la probabilité qu'un amas de taille i et de poids δ_i soit connecté à un amas de taille j et de poids δ_j à l'instant t , et $2 \frac{\delta_{k_a} n(k_a, t)}{c(t)}$ est la probabilité qu'un amas de taille k_a et de poids δ_{k_a} soit connecté à n'importe quel amas dans le réseau.

Dans [180, 186] les auteurs ont généralisé l'Eq. (5.4) en prenant $\delta_i = i^\omega$, le cas $\omega = 1$ est exactement le cas ER. En utilisant la technique de la fonction génératrice, ils ont trouvé l'expression de la distribution des amas au point critique, $n(k_a, t_c) \sim k_a^{-\beta}$, où β est tel que :

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

En se basant sur des simulations numériques, ils ont trouvé que pour $\beta > 2$ la transition de phase est continue, et pour $\beta \leq 2$ la transition est discontinue (dans le paragraphe suivant, nous proposons que ce critère sur β peut être généralisé à tout autre modèle). En 1983, un modèle unidimensionnel de percolation a été introduit [187] dans lequel les sites i et j sont reliés par la probabilité $p_{ij} = \frac{p}{|i-j|^\theta}$, où p est un paramètre défini dans l'intervalle $0 \leq p \leq 1$, et θ est un autre paramètre. Dans la référence [188] les auteurs ont prouvé que la transition est discontinue pour $1 < \theta \leq 2$. Pour plus de détails sur la percolation à longue distance voir [189].

5.3.2 Nouvelle approche

Dans les modèles du paragraphe précédent le critère de la caractérisation des transitions de phase dépend de certains paramètres intrinsèques, et ne peut être généralisé aux autres modèles de percolation. La question d'importance majeure qui se pose : Est-il possible d'unifier le critère pour déterminer le type de la transition ? Pour répondre à cette question, il faut d'abord trouver les paramètres ou les grandeurs pertinents pour cette unification.

On propose, dans ce sens, une méthode pour prévoir le type de la transition de phase qui se base :

- Sur le fait que la distribution des amas dans un réseau au voisinage du point critique suit toujours une loi de puissance.
- Sur la façon de choisir les amas.

En général la distribution des amas, $n(k_a)$, dans un réseau est calculée en choisissant les noeuds aléatoirement et après on déduit la taille des amas auxquels ils appartiennent, dans ce cas $n(k_a) = ck_a^{-\beta}$ au voisinage du point de la transition de phase. Si par contre le choix des noeuds n'est pas aléatoire en introduisant une corrélation implicite comme dans le cas de la règle RP, la distribution des amas est telle que $n'(k_a) = c'k_a^{-\beta'}$. Dans la Fig. 5.5, on observe que dans le cas de la règle RP du processus Achlioptas, l'exposant de la distribution des amas β' est différent de la valeur de $\beta = 2.5$ qui est calculée avec un choix aléatoire des noeuds.

L'idée intuitive derrière l'introduction de β' est que lorsque le choix des noeuds est sélectif, il y a un changement non réel de la distribution des amas que nous appelons distribution effective et qu'on désigne par $n'(k_a)$, avec un exposant aussi effectif β' .

5.3.2.1 Composante géante :

Dans cette partie nous cherchons une expression de la composante géante plus générale que celle de ER au voisinage du point critique et en fonction des paramètres β et β' . Nos calculs sont réalisés au voisinage du point critique, qui correspond à $n(k_a) = ck_a^{-\beta}$, puis nous ajoutons des liens au réseau pour permettre l'émergence de la composante géante. 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 qui n'est pas connecté à 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 réseau est

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

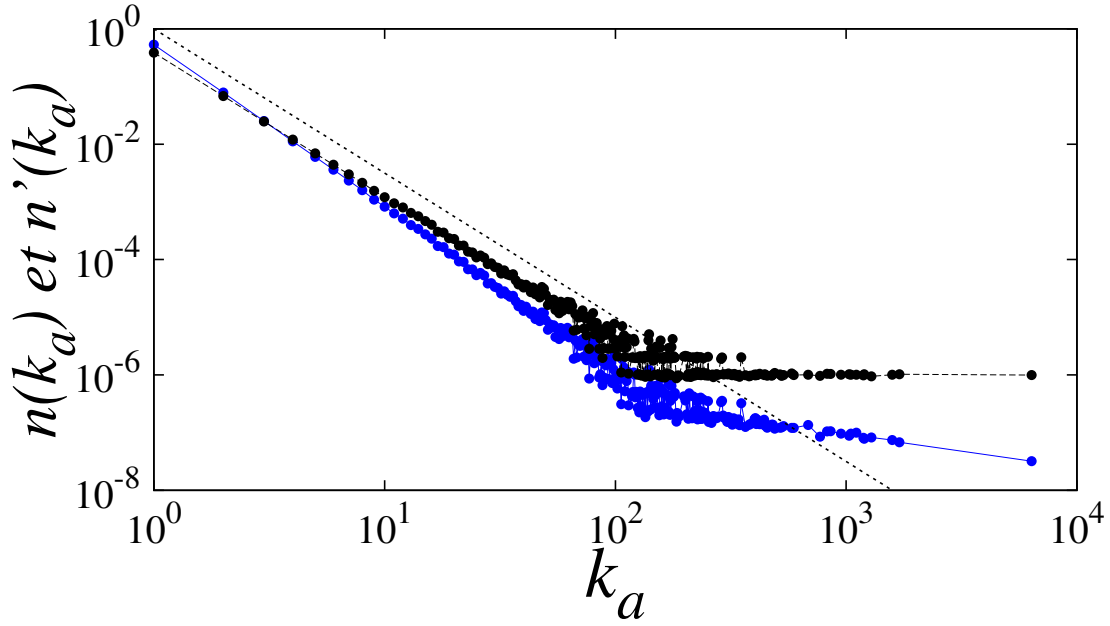


FIGURE 5.5 – La distribution des amas selon deux méthodes différentes pour choisir les noeuds dans le même réseau, la première méthode est aléatoire comme le processus ER (cercles noirs), la deuxième est le processus Achlioptas [2] (cercles bleus). La ligne en pointillée montre l'exposant $\beta = 2.5$ qu'on trouve lorsque le choix des noeuds est aléatoire. La taille du réseau est $n = 10^6$, chaque simulation est une moyenne sur n réalisations.

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

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

où $K_a = n_a^{\frac{1}{\beta-1}}$ est l'amas maximal au voisinage du point critique.

D'où $\sum_{j=1}^{K_a} j n'(j) = \langle k'_a \rangle$ est la taille moyenne effective des amas. Puisque $\langle k_a \rangle$ est la taille moyenne des amas alors $n = n_a \langle k_a \rangle$ et sachant 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 d'amas de taille i qui n'appartiennent pas à la composante géante est

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

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

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

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

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

De l'Eq. (5.8), et l'Eq. (5.9) et sachant que $n = n_a < k_a >$ et $S = 1 - u$ on obtient

$$\begin{aligned} S &= 1 - \frac{1}{< k_a >} \sum_{i=1}^{K_a} in(i) e^{\frac{-iS < k > < k'_a >}{< k_a >}} \\ &= 1 - \frac{1}{< k_a >} \int_{i=1}^{K_a} in(i) e^{\frac{-iS < k > < k'_a >}{< k_a >}} di. \end{aligned} \quad (5.12)$$

Cette formule est une expression plus générale que celle de modèle ER [10]. En effet, dans le cas où on ajoute aléatoirement des liens sans corrélations, $\beta = \beta'$, et si au temps initial tout les amas ont la taille 1, on trouve exactement l'expression de modèle ER, $S = 1 - e^{-S < k >}$.

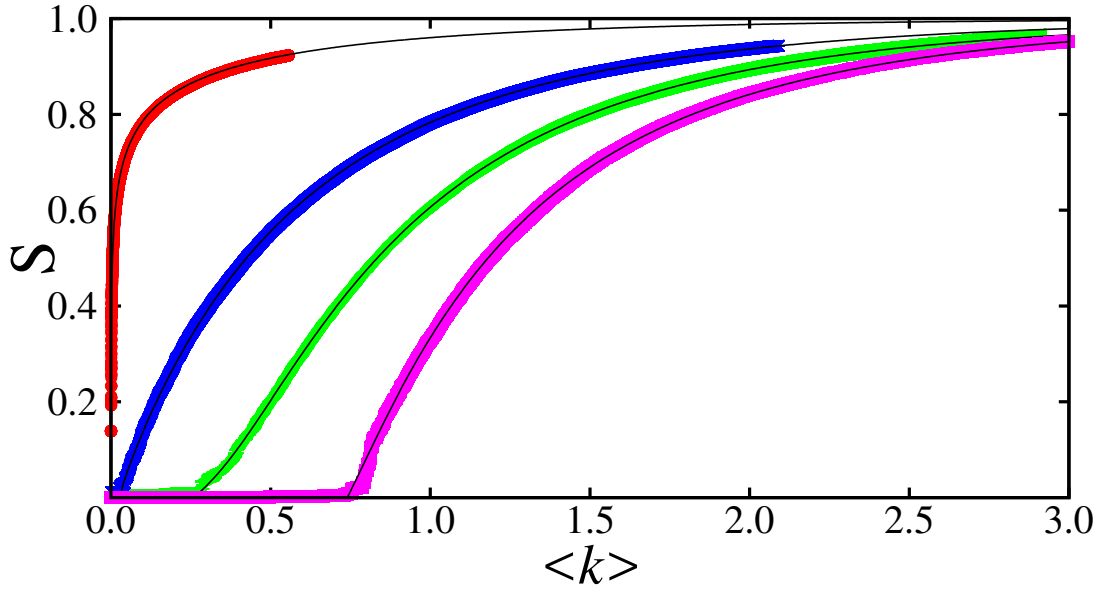


FIGURE 5.6 – Fraction de la composante géante S en fonction de la densité des liens $< k >$ pour différentes valeurs de β , les couleurs de gauche à droite représentent les simulations numériques respectivement pour $\beta = 2, 2.5, 3, 4$, les lignes noire représentent la solution numérique de l'Eq.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

5.3.2.2 Détermination du point de la transition :

A partir de Eq. (5.12) et en utilisant la méthode graphique utilisée dans [190] on obtient facilement le point critique ($\langle k \rangle_c$) où la composante géante apparaît, cette méthode s'appuie sur la condition que le point critique $\langle k \rangle$ est donné par

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

la solution de cette équation donne l'expression suivante

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

dans le cas où il n'y a pas de corrélations $\beta = \beta'$, 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 différente de celle donnée dans la référence [180] où les auteurs ont trouvé que $\langle k \rangle_c = \frac{1}{\langle k_a^2 \rangle}$. Cette expression prévoit des valeurs inférieures à celles prévues par notre équation car $\langle k \rangle_c > 1$. La comparaison de l'Eq. (5.15) avec les simulations numériques (voir Fig. 5.7) montre que notre approche est beaucoup plus précise que celle de la référence [180].

Il convient de mentionner que la détermination du point de la transition par les simulations n'est pas une opération facile, il est assez compliqué et délicat de trouver le avec une grande précision. Dans nos simulations présentées dans la Fig. 5.7, nous avons estimé le point de la transition par la détermination du point où la valeur moyenne des petits amas dans le réseau est maximale, cependant cette estimation est devenue plus précise avec l'augmentation du nombre de nœuds du réseau, car les effets d'échelle diminuent avec la taille. Alors dans la limite thermodynamique l'expression théorique sera souvent exacte.

5.3.2.3 Détermination du type de la transition de phase :

Pour déterminer le type de la transition de phase on détermine la valeur de $\langle k \rangle$ qui fait que S atteigne une valeur finie ($O(1)$) et on nomme Δk . Si Δk tend vers zéro on dit que la transition est de premier ordre sinon on dit qu'elle est de second ordre¹ (voir Fig 5.8 et Fig 5.9). Autrement dit Δk représente la valeur minimale de la densité des liens qu'on doit ajouter au

1. La condition nécessaire de cette méthode est que le point de la transition doit être tendre vers 0. Autrement dit l'origine du point de la transition doit être au point $\langle k \rangle = 0$.

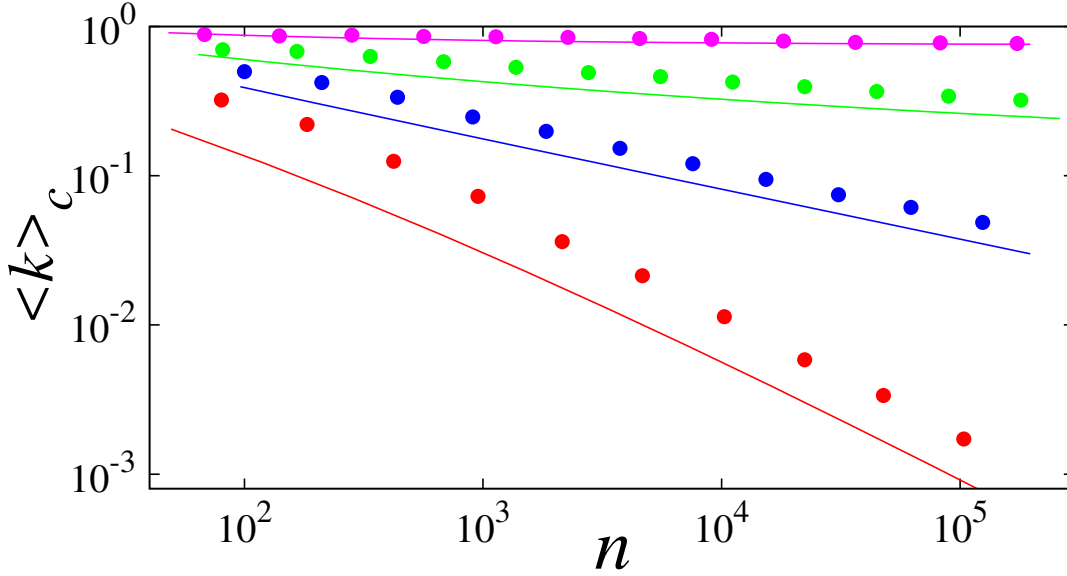


FIGURE 5.7 – Estimation de point critique par les simulations numériques (cercles) et par l'Eq. (5.15) (ligne). Les valeurs de β de haut en bas sont respectivement 4, 3, 2.5, 2. Chaque simulation est moyennée sur 500 réalisations.

réseau pour faire apparaître la composante géante.

A partir de l'Eq. (5.12) et sachant que $n(i) = ci^\beta$, on obtient

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

l'intégrale dans cette expression n'a pas de solution explicite connue, ce qui rend impossible l'extraction de $\langle k \rangle$ en fonction de S . Pour déterminer le type de la transition, on construit la composante géante de la façon suivante : On commence par un réseau dont la distribution des amas suit une loi de puissance, après et successivement, on ajoute un lien entre les deux plus grands amas. Cette construction garantit l'apparition de la composante géante le plus rapidement possible, ce qui implique que si Δk est finie (non nulle) alors la transition est sûrement continue, car toute autre construction donnera un Δk plus grand. Mathématiquement ceci est traduit par l'expression suivante :

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

5.3. Nouvelle approche pour prédire le type de la transition dans les systèmes percolatifs

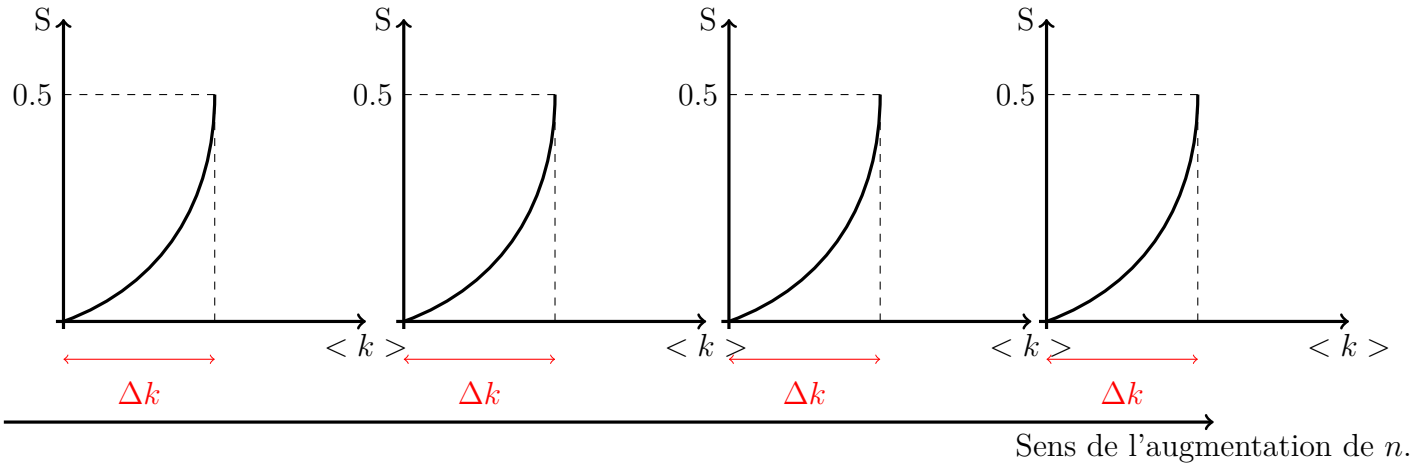


FIGURE 5.8 – Illustration de la variation de Δk en fonction de n lorsque la transition est continue. Δk est la valeur de $\langle k \rangle$ quand $S = 0.5$.

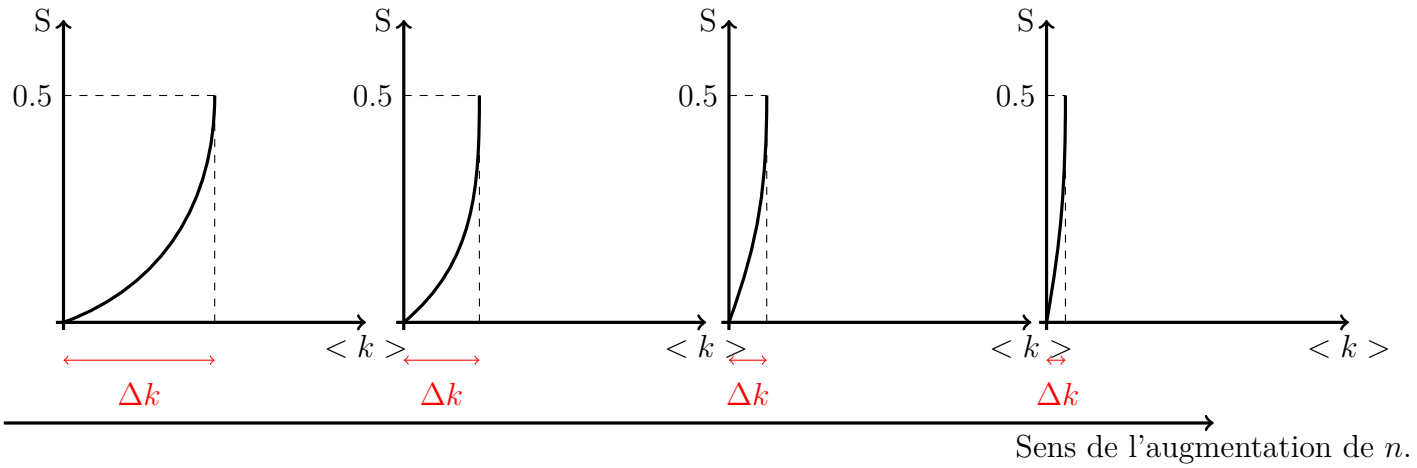


FIGURE 5.9 – Illustration de la variation de Δk en fonction de n lorsque la transition est discontinue. Δk est la valeur de $\langle k \rangle$ quand $S = 0.5$.

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

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

$$\sum_{i=k_a(t)}^{K_a} n_a n(i) = \int_{i=k_a(t)}^{K_a} n_a n(i) di = t, \quad (5.18)$$

on obtient

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

des Eq. (5.17) et Eq. (5.19) on trouve

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

sachant que $t = \frac{n \langle k \rangle}{2}$ et $n = n_a \langle k_a \rangle$, l'Eq. (5.20) devient $\langle k_a \rangle = \frac{\beta-1}{\beta-2}$ et

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

d'où on obtient

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

Pour $\beta > 2$ on a $\langle k_a \rangle = \frac{\beta-1}{\beta-2}$, alors pour une valeur de S finie on peut écrire à la limite thermodynamique que $\langle k \rangle = \frac{2(\beta-2)}{\beta-1} \left(S \right)^{\frac{1-\beta}{2-\beta}}$. Sous ces conditions, on peut dire que Δk ne peut jamais tendre vers zéro, ce qui signifie que la transition de phase est sûrement continue pour $\beta > 2$.

On a démontré, donc, qu'il est impossible que la transition de phase soit de premier ordre lorsque l'exposant $\beta > 2$ indépendamment de l'exposant effectif β' , c'est-à-dire indépendamment des corrélations dans le réseau.

Lorsque $\beta = 2$, l'intégrale de l'Eq. (5.16) peut se calculer 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

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

5.4. Conclusion

en remplaçant l'Eq. (5.23) dans l'Eq. (5.16) on obtient

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

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 $\beta = 2$ on a $K_a = n_a$ 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

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

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

Nous avons montré, alors, que le type de la transition de phase dépend seulement de la distribution des amas au voisinage du point critique (l'exposant β), et ne dépend pas des corrélations entre les amas (l'exposant effectif β').

5.4 Conclusion

Dans ce chapitre nous avons étudié quelques propriétés du modèle de percolation dans les réseaux complexes aléatoires. En particulier, on a trouvé l'expression exacte de la composante géante valable pour toutes les distributions de degrés. Ensuite, on a pu établir l'expression qui donne la valeur du point de la transition de phase. Cette équation est plus précise que celle dont on dispose actuellement [180]. Finalement on a introduit une nouvelle méthode pour prévoir le type de la transition de phase dans le phénomène de percolation. Contrairement aux autres travaux cités dans la Section 5.3.1, où les auteurs ont essayé de déterminer le type de la transition dans des modèles spécifiques, notre calcul est général et s'appuie sur les paramètres universels de la percolation. Notre méthode est basée sur deux paramètres, le premier est l'exposant de la distribution des amas au voisinage du point de la transition β et le deuxième

est l'exposant de la distribution d'amas effectif β' qui est dû aux corrélations dans le réseau. Nous avons trouvé qu'une transition de phase est de premier ordre si $\beta \leq 2$, sinon la transition est de second ordre, indépendamment de la valeur de β' . On en déduit que les corrélations entre les amas au voisinage du point de transition n'ont aucune influence sur le type de la transition de phase.

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 sans 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_ℓ . Nous avons obtenu également la description précise de la forme de la distribution. Profitant de la forme de la distribution n_ℓ , nous avons pu déduire l'expression explicite du plus court chemin. Les expressions obtenues reproduisent les formes de mise à l'échelle connues pour différentes plages de γ . 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_ℓ , mais avec une petite déviation dans la position du maximum. Cette différence n'affecte pas la forme de mise à l'échelle du plus court chemin pour cette valeur de γ , 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 plus court chemin plus

perfectionnée par rapport à l'existant. 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 $\frac{\ell}{\hat{n}} = f(x)$ se révèle plus ou moins tandis que la nouvelle fonction se comporte inversement.

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 $\beta > 2$, quel que soit le type de comportement collectif entre les amas dans le réseau, en revanche elle est du premier ordre si $\beta \leq 2$. De façon générale, il s'avère qu'il suffit de connaître β au voisinage de point critique pour en déduire le type de transition de phase.

Bibliographie

- [1] A. Fronczak, P. Fronczak, and J. A. Holyst. Average path length in random networks. *Phys. Rev. E*, 70 :056110, (2004).
- [2] D. Achlioptas, R. M. D’Souza, and J. Spencer. Explosive percolation in random networks. *Science*, 323 (5920) :1453–1455, (2009).
- [3] 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).
- [4] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *Computer Communication Review*, 29 :251, (1999).
- [5] A.-L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286 :509–512, (1999).
- [6] Cohen, J. E., Briand F., and Newman, C. M. Community food webs : Data and theory. *Biomathematics*, 20, (1990).
- [7] Pimm, S. L. *Food Webs*. University of Chicago Press, Chicago, 2nd ed. (2002).
- [8] M. E. J. Newman and D. J. Watts. Renormalization group analysis of the small-world network model. *Phys. Lett.A*, 263 :341, (1999).
- [9] P. J. Flory. Molecular size distribution in three-dimensional polymers. i. gelation. *Journal of the American Chemical Society*, 63 :3083–3090, (1941).
- [10] P. Erdos and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6 :290–297, (1959).
- [11] 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).

-
- [12] P. Erdos and A. Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Scientia Hungary*, 12 :261–267, (1961).
 - [13] D. J. Watts and S. H. Strogatz. Collective dynamics of “small-world” networks. *Nature*, 393 :440–442, (1998).
 - [14] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *edings of the National Academy of Science*, 99 :7821–7826, (2002).
 - [15] F. R. Pitts. A graph theoretic approach to historical geography. *The Professional Geographer*, 17 :15–20, (1965).
 - [16] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. Classes of small-world networks. *Proc. Natl. Acad. Sci, USA* 97 :11149–11152, (2000).
 - [17] 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).
 - [18] 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).
 - [19] 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).
 - [20] 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).
 - [21] S. N. Dorogovtsev and J. F. F. Mendes. *Evolution of Networks : From Biological Nets to the Internet and WWW*. Oxford Univ Press, Oxford, (2003).
 - [22] R. Cohen and S. Havlin. Scale-free networks are ultrasmall. *Phys. Rev. Lett*, 90 :058701, (2003).
 - [23] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Metric structure of random networks. *Nucl. Phys. B*, 653 :307, (2003).
 - [24] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Math*, 1 :91–113, (2003).
 - [25] 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).
 - [26] J. Scott. *Social Network Analysis : A Handbook*. Sage Publications, London, 2nd ed. (2000).

- [27] W. W. Zachary. An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33 :452–473, (1977).
- [28] J. L. Moreno. *Who Shall Survive ?* Beacon House, Beacon, (1934).
- [29] J. Galaskiewicz and P. V. Marsdeny. Interorganizational resource networks : Formal patterns of overlap. *Social Science Research*, 7 :89–107, (1978).
- [30] Krioukov, D., Kitsak, M., Sinkovits, R. S., Rideout, D., Meyer, D., and Boguna, M. Network cosmology. *Scientific Reports*, 2 :793, (2012).
- [31] E. Bullmore and O. Sporns. Complex brain networks : Graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, 10 :168–198, (2009).
- [32] D. Lazer and al. Computational social science. *Science*, 323 :721–723, (2009).
- [33] Y.Y. Liu, J.J. Slotine, and A.L. Barabasi. Controllability of complex networks. *Nature*, 473 :167–173, (2011).
- [34] L. Smolin. *The life of the cosmos*. Oxford University Press, Oxford, (1997).
- [35] C. A. Trugenberger. Quantum gravity as an information network : Self-organization of a 4d universe. *Phys.Rev. D*, 92 :084014, (2015).
- [36] Bianconi, G., Rahmede, C., and Wu, Z. Complex quantum network geometries : Evolution and phase transitions. *Phys. Rev. E*, 92 :022815, (2015).
- [37] B. Bollobas and O. Riordan. *Percolation*. Cambridge University Press, Cambridge, (2006).
- [38] A. Acin, J. I. Cirac, and M. Lewenstein. Entanglement percolation in quantum networks. *Nature Physics*, 3 :256, (2007).
- [39] S. Sachdev. *Quantum Phase Transitions*. Cambridge University Press, Cambridge, (1999).
- [40] J. Novotny, G. Alber, and I. Jex. Universality in random quantum networks. *Phys. Rev. A*, 92 :062335, December (2015).
- [41] D. Bruß and G. Leuchs (eds.). *Lectures on Quantum Information*. Weinheim, Wiley-VCH, (2007).
- [42] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, (2000).
- [43] M. Mitchell. Complex systems : Network thinking. *Artificial Intelligence*, 170 :1194–1212, (2006).
- [44] I.A. Basheer M. Hajmeer. Artificial neural networks : fundamentals, computing, design and application. *Journal of Microbiological Methods*, 43 :3–31, (2000).

-
- [45] 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).
 - [46] F. Harary. *Graph Theory*. Perseus, Cambridge, MA (1995).
 - [47] D. B. West. *Introduction to Graph Theory*. Prentice Hall, Upper Saddle River, NJ (1996).
 - [48] S. Milgram. The small world problem. *Psychology Today*, page 56–67, May (1967).
 - [49] J. Travers and S. Milgram. *Sociometry*, 32 :425, (1969).
 - [50] P. Dodds, R. Muhamad, and D. Watts. An experimental study of search in global social networks. *Science*, 301(5634) :827–829, (2003).
 - [51] R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet : A Statistical Physics Approach*. Cambridge University Press, Cambridge, (2004).
 - [52] S. Wasserman and K. Faust. *Social Network Analysis : Methods and Applications*. Cambridge University Press, Cambridge, (1994).
 - [53] Boccaletti, Stefano, and al. Complex networks : Structure and dynamics. *Physics reports*, 424.4 :175–308, (2006).
 - [54] 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).
 - [55] V. Latora and M. Marchiori. Efficient behavior of small-world networks. *Phys. Rev. Lett*, 87(19) :198701, Oct (2001).
 - [56] 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).
 - [57] 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).
 - [58] R. Albert, H. Jeong, and A. L. Barabási. Error and attack tolerance of complex networks. *Nature (London)*, 406 :378, (2000).
 - [59] 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).

- [60] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee, and S. S. Manna. *Small-world properties of the indian railway network*. (2002).
- [61] A. Wagner and D. Fell. *The small world inside large metabolic networks*. (August 21 2000).
- [62] S. Bilke and C. Peterson. Topological properties of citation and metabolic networks. *Rev. E*, 64 :036106, (2001).
- [63] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev*45, 2 :167–256, (2003).
- [64] G. Bianconi and A. Capocci. Number of loops of size h in growing scale-free networks. *Phys. Rev. Lett*, 90 :art. no. 078701, (2003).
- [65] A. Fronczak, J. A. Holyst, M. Jedynak, and J. Sienkiewicz. Higher order clustering coefficients in barabasi-albert networks. *Phys. A*, 316 :688–694, (2002).
- [66] 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).
- [67] M. Boguñá and R. Pastor-Satorras. *Phys. Rev. E*, 66 :047104, (2002).
- [68] S. Maslov and K. Sneppen. Specificity and stability in topology of protein networks. *Science*, 296 :910–913, (2002).
- [69] M. Dehmer. *Structural Analysis of Complex Networks*. Springer, (2011).
- [70] P. Hage and F. Harary. Eccentricity and centrality in networks. *Soc Networks*, 17 :57–63, (1995).
- [71] S. Wasserman and K. Faust. *Social network analysis : methods and applications. Structural analysis in the social sciences*. Cambridge University Press, Cambridge, (1994).
- [72] L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40 :35–41, (1977).
- [73] Famous trails to Paul Erdos. Relevant cycles in chemical reaction networks. *Math. Intelligencer*, 51 :51–63, (1999).
- [74] M. E. J. Newman. Power laws, pareto distributions and zipf’s law. *Contemporary Physics*, 46 :323–351, (2005).
- [75] 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).

-
- [76] B. Gutenberg and R. F. Richter. Energy ratio of reflected and refracted seismic waves. *Bull. Seismol. Soc. Am*, 34 :85–102, (1944).
 - [77] 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.
 - [78] G.K. Zipf. *Human Behaviour and the Principle of Least Effort*. Addison-Wesley, Reading, MA, 1949.
 - [79] J.B. Estoup. *Gammes Stenographiques*. Institut Stenographique de France, Paris, (1916).
 - [80] D.H. Zanette and S.C. Manrubia. Vertical transmission of culture and the distribution of family names. *Physica A*, 295 :1, (2001).
 - [81] A.J. Lotka. The frequency distribution of scientific productivity. *Acad. Sci*, 16 :317–323, (1926).
 - [82] D.J. de S. Price. Networks of scientific papers. *Science*, 149 :510–515, (1965).
 - [83] L. A. Adamic and B. A. Huberman. Q. j. electron. *Commerce*, 1 :512, (2000).
 - [84] 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).
 - [85] 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).
 - [86] V. Pareto. *Cours d'Economie Politique*. Droz, Geneva, (1896).
 - [87] G. Bianconi and C. Rahmede. Complex quantum network manifolds in dimension $d > 2$ are scale-free. *Scientific Reports*, 5 :13979, (2015).
 - [88] 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).
 - [89] Flake, G. W., S. Lawrence, C. Lee Giles, and F. M. Coetzee. Self-organization and identification of web communities. *IEEE Computer*, 35 :66, (2002).
 - [90] J. Chen and B. Yuan. Detecting functional modules in the yeast protein–protein interaction network. *Bioinformatics*, 22(18) :2283, (2006).
 - [91] A. W. Rives and T. Galitski. Modular organization of cellular networks. *Proc. Natl. Acad. Sci. USA*, 100(3) :1128, (2003).

- [92] 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).
- [93] M. Steenstrup. Cluster-based networks. *Addison Wesley, Reading, USA*, chapter 4 :75–138, (2001).
- [94] P. Csermely. Cluster-based networks. *Trends Biochem. Sci*, 33(12) :569, (2008).
- [95] D. Chowdhury and D. Stauffer. *Principles of equilibrium statistical mechanics*. Wiley-VCH, Weinheim, (2000).
- [96] H-P. Hsu, V. Mehra, and P. Grassberger. Structure optimization in an off-lattice protein model. *Phys Rev E*, 68 :037703, (2003).
- [97] D.J. Watts and S.H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393 :440–442, (1998).
- [98] Nicholas Rescher. *Complexity : A philosophical overview*. Transaction Publishers, 1998.
- [99] Mitchell M Waldrop. *The emerging science at the edge of order and chaos*. Simon and Schuster, 1993.
- [100] Roger Highfield. *Frontiers of complexity : The search for order in a chaotic world*. Ballantine Books, 1996.
- [101] Marc Mezard and Andrea Montanari. *Information, physics, and computation*. Oxford University Press, 2009.
- [102] Neil F Johnson. *Two's company, three is complexity : A simple guide to the science of all sciences*. Oneworld Pubns Ltd, 2007.
- [103] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys*, 74 :47, (2002).
- [104] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Adv. Phys*, 51 :1079, (2002).
- [105] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev*, 45 :167, (2003).
- [106] V. Pareto. *Cours d'Économie Politique*. (Lausanne, 1897).
- [107] 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).

-
- [108] R. Albert, I. Albert, and G. L. Nakarado. Structural vulnerability of the north american power grid. *Phys. Rev. E*, 69 :025103, (2004).
 - [109] T. B. Achacoso and W. S. Yamamoto. *AY's Neuroanatomy of C. elegans for Computation*. CRC Press, Boca Raton, FL, (1992).
 - [110] A.-L. Barabási, H. Jeong, and R. Albert. Statistical mechanics and its applications. *Physica A*, 272 :173, (1999).
 - [111] Y.-B. Xie, T. Zhou, and B.-H. Wang. Scale-free networks without growth. *Physica A*, 387 :1683, (2008).
 - [112] V. K. Samalam. *arXiv :1202.1498*.
 - [113] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Physical Review Letters*, 85 :4629–4632, (2000).
 - [114] 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).
 - [115] W. H. Furry. On fluctuation phenomena in the passage of high energy electrons through lead. *Phys. Rev*, 52(6) :569–581, (1937).
 - [116] W. Feller. Die grundlagen der volterraschen theorie des kampfes ums dasein in wahrscheinlichkeitstheoretischer behandlung. *Acta Biotheor*, 52(6) :11–40, (1939).
 - [117] M. Delbrück. Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys*, 8(1) :120–124, (1940).
 - [118] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library. Elsevier, Amsterdam, 3rd edition, 2007.
 - [119] 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.
 - [120] Weber M F and Frey E. Master equations and the theory of stochastic path integrals. *Rep. Prog. Phys*, 80 :046601, (2017).
 - [121] R. J. Glauber. Time-dependent statistics of the ising model. *J. Math. Phys*, 4(2) :294–307, (1963).
 - [122] K. Kawasaki. Diffusion constants near the critical point for time-dependent ising models. i. *Phys. Rev*, 145(1) :224–230, (1966).

- [123] K. Kawasaki. Diffusion constants near the critical point for time-dependent ising models. ii. *Phys. Rev*, 148(1) :375–381, (1966).
- [124] K. Kawasaki. Diffusion constants near the critical point for time-dependent ising models. iii. *Phys. Rev*, 150 :285–290, (1966).
- [125] 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).
- [126] C. V. Rao, D. M. Wolf, , and A. P. Arkin. Control, exploitation and tolerance of intracellular noise. *Nature*, 420 :231–237, (2002).
- [127] L. S. Tsimring. Noise in biology. *Rep. Prog. Phys*, 77 :026601, (2014).
- [128] N. T. J. Bailey. A simple stochastic epidemic. *Biometrika*, 37 :193–202, (1950).
- [129] K. Rock, S. Brand, J. Moir, and M. J. Keeling. Dynamics of infectious diseases. *Rep. Prog. Phys*, 77 :026602, (2014).
- [130] 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).
- [131] W. Weidlich and M. Braun. The master equation approach to nonlinear economics. *J. Evol. Econ*, 2 :233–265, (1992).
- [132] 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).
- [133] A. L. Barabási. *Network Science*. Cambridge University Press, (2015).
- [134] 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).
- [135] B. Bollobás and O. Riordan. A polynomial of graphs on surfaces. *Math. Ann*, 323 :1–35, (2002).
- [136] A. L. Barabási. *Linked : The New Science of Networks*. Perseus, Cambridge, 2002.
- [137] A. Barrat et al. *Dynamical Processes on Complex Networks*. (Cambridge University Press, 2008).
- [138] R. Cohen and S. Havlin. *Complex Networks : Structure, Robustness and Function*. (Cambridge University Press, 2010).
- [139] 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).

-
- [140] M. E. J. Newman. *Networks : An Introduction*. University of Michigan and Santa Fe Institute Press, (2010)p.456.
 - [141] R. Cohen and S. Havlin. *Complex Networks : Structure, Robustness and Function*. (Cambridge University Press, 2010).pp :72-77.
 - [142] T. Kalisky, R. Cohen, O. Mokryn, D. Dolev, Y. Shavitt, and S. Havlin. Tomography of scale-free networks and shortest path trees. *Phys. Rev. E*, 74 :066108, (2006).
 - [143] B. Bollobás. *Random Graphs*. Academic Press, London, (1985).
 - [144] M. Nitzan, E. Katzav, R. Khn, and O. Biham. *Phys. Rev. E*, 93 :062309, (2016).
 - [145] E. Katzav, M. Nitzan, D. ben Avraham, P. L. Krapivsky, R. Khn, N. Ross, and O. Biham. *EPL*, 111 :26006, (2015).
 - [146] E. Katzav, O. Biham, and A. K. Hartmann. *Phys. Rev. E*, 98 :012301, (2018).
 - [147] 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 Dezso Miklos*. (Bolyai Society of Mathematical Studies and Springer, 2009).
 - [148] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24 :5–34, (2004).
 - [149] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *PNAS*, 99 :15879–15882, (2002).
 - [150] M. Akbarzadeh and E. Estrada. *Nature Human Behaviour*, 2 :645652, (2018).
 - [151] 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).
 - [152] 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).
 - [153] 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).
 - [154] 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).
 - [155] B. Bollobás and O. Riordan. *Handbook of graphs and networks, eds. S. Bornholdt and H. G. Schuster*. Wiley-VCH, Berlin, (2002).

- [156] M. E. J. Newman. *Networks : An Introduction*. University of Michigan and Santa Fe Institute Press, (2010).
- [157] M. Newman and D. Watts. Scaling and percolation in the small-world network model. *Phys. Lett. A*, 341, (1999).
- [158] A. Barrat and M. Weigt. On the properties of small-world network models. *Eur. Phys. J. B*, 13 :547, (2000).
- [159] 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).
- [160] M. Barthélemy and L. A. N. Amaral. Small-world networks : Evidence for a crossover. *Phys. Rev. Lett*, 82 :3180, (1999).
- [161] A. Barrat. Comment on ‘Small-world networks : crossover picture’.
- [162] M. E. J. Newman and D. J. Watts. Scaling and percolation in the small-world network model. *Phys. Rev.E*, 60 :7332, (1999).
- [163] M. Argollo de Menezes, C. Moukarzel, and T.J.P. Penna. *Europhys. Lett*, 50 :574, (2000).
- [164] 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).
- [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] B. Bollobás. Modern graph theory. *Springer*, (1998,New York).
- [168] J. Adler. Bootstrap percolation. *Physica A*, 453 :171, (1991).
- [169] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. k-core organization of complex networks. *Phys. Rev. Lett*, 96 :040601, (2006).
- [170] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno. Dynamics of jamming transitions in complex networks. *EPL (Europhysics Letters)*, 71 (2) :325–331, (2005).
- [171] 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).
- [172] 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).

-
- [173] Y. S. Cho and B. Kahng. Suppression effect on explosive percolation. *Phys. Rev. Lett*, 107 :275703, (2011).
 - [174] B. Bollobas. The evolution of random graphs. *Trans. Amer. Math. Soc*, 286 :257—274, (1984).
 - [175] T. Bohman and A.M. Frieze. Avoiding a giant component. *Random Structures and Algorithms*, 19 :75–85, (2001).
 - [176] O. Riordan and L. Warnke. Explosive percolation is continuous. *Science*, 333 :322, (2011).
 - [177] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *The Ann. Appl.Prob*, 22 :1450–1464, (2012).
 - [178] J. Nagler, T. Tiessen, and H. W. Gutch. Continuous percolation with discontinuities. *Phys. Rev. X*, 2 :031009, (2012).
 - [179] 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).
 - [180] Y. S. Cho, B. Kahng, and D. Kim. Cluster aggregation model for discontinuous percolation transitions. *Phys. Rev. E*, 81 :030103(R), (2010).
 - [181] 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).
 - [182] 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).
 - [183] L. Cao and J. M. Schwarz. Correlated percolation and tricriticality. *Phys. Rev. E*, 86 :061131, (2012).
 - [184] 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).
 - [185] 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).
 - [186] Y. S. Cho and B. Kahng. Discontinuous percolation transitions in real physical systems. *Phys. Rev. E*, 84 :050102(R), (2011).
 - [187] L. S. Schulman. Long range percolation in one dimension. *Journal of Physics A : Mathematical and General*, 16 :L639, (1983).

- [188] 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).
- [189] D. Lee, Y.S. Cho, and B. Kahng. Diverse types of percolation transitions. *J. Stat. Mech. Theor. Exp*, 107 :124002, (2016).
- [190] M.E.J. Newman. *Networks : An Introduction*. University of Michigan and Santa Fe Institute Press, (2010)p.403-408.