



## UNIVERSITÉ PARIS 6 - PIERRE ET MARIE CURIE

et

# INSTITUT DE PHYSIQUE THÉORIQUE - CEA SACLAY

Thèse de doctorat

Spécialité

Physique Théorique

Sujet de la thèse :

# Effets de taille finie, extrêmes et propagation de défaillances dans des réseaux hors d'équilibre

présentée par

# Hélène GRANDCLAUDE

pour obtenir le grade de

# Docteur de l'Université Paris 6

Soutenue le 10 novembre 2011 devant le jury composé de :<br/>Monsieur Alain BARRATRapporteurMonsieur Alain COMTETExaminateurMonsieur Claude GODRÈCHEDirecteur de thèseMonsieur Malte HENKELExaminateurMonsieur Jörg LEHMANNRapporteur

# Effets de taille finie, extrêmes et propagation de défaillances dans des réseaux hors d'équilibre

#### Résumé :

Cette thèse porte sur deux aspects de la statistique des réseaux hors d'équilibre. Dans un premier temps, nous étudions les effets de taille et de temps fini dans des modèles simples de réseaux en croissance. Nous montrons comment la distribution du degré à temps fini est décrite par une fonction d'échelle et présente une transition entre un régime stationnaire et un régime de grandes déviations de la distribution du degré. La fonction d'échelle décroît exponentiellement vite pour des degrés de l'ordre du degré de coupure  $k_{\star}$ . Nous donnons la formule explicite de cette fonction d'échelle pour trois modèles de réseaux en croissance et deux symétries possibles de la condition initiale.

À partir de ces distributions à temps fini, nous étudions la statistique des extrêmes dans des réseaux en croissance considérés à un temps fini. Nous montrons que la dynamique des noeuds les plus connectés est très différente selon que le réseau croît avec une règle préférentielle ou uniforme. La dynamique des meneurs dans un réseau invariant d'échelle est gelée, dans le sens où seuls quelques uns des premiers noeuds du réseau dominent la connectivité du système. La dynamique des extrêmes se limite alors à l'alternance de quelques noeuds en position de meneur.

Dans un second temps, nous étudions un modèle simple de propagation d'une défaillance dans les réseaux : le modèle de Lehmann et Bernasconi. Grâce à l'analogie entre ce modèle et un processus de branchement généralisé, nous présentons des résultats analytiques et numériques qui montrent l'apparition dans ce modèle de deux seuils de transition vers un régime dans lequel des cascades globales se produisent avec probabilité finie. Nous explorons les propriétés de la cascade dans ce cadre simplifié et montrons la nature différente de chacune des transitions. L'une est dûe aux propriétés de connectivité du réseau de défaillances tandis que l'autre est liée à la fragilité locale des éléments du réseau.

Mots clés : Systèmes Complexes, Statistique Hors d'Équilibre, Réseaux en Croissance, Statistique des Extrêmes, Propagation de Défaillances

# Finite-size effects, extreme-value statistics and failure propagation in nonequilibrium networks

#### Abstract:

This thesis deals with two aspects of the statistics of nonequilibrium networks. In a first part, we present a study of finite-size and finite-time effects in simple models of growing networks. We demonstrate that the finite-time degree distribution is described in terms of a scaling function, and shows a crossover between a stationary regime and a large deviations regime. This scaling function decreases exponentially for degrees of the order of the crossover degree  $k_{\star}$ . The explicit analytical formula is given for three models of growing networks and two possible symmetries of the initial conditions.

Using the finite-time formulation we study the extreme-value statistics in growing networks. We show that the dynamics of the most connected nodes differs strongly, depending on the rule of construction, preferential attachment or uniform attachment. The dynamics of the lead change in a scale-free network is frozen, because only some of the first nodes of the network dominate the connectivity of the system. The extreme value dynamics is just the successive exchanges of the lead between a few nodes.

In the second part, we study a simple model of failure propagation in networks: the model introduced by Lehmann and Bernasconi. By means of an analogy between this model and a generalized branching process, we obtain analytical and numerical results. These results show the existence of two transition thresholds between a regime without global cascades and a regime where the probability to trigger a global cascade with a single failure is finite. We explore the properties of the cascade in this simple formalism and show that the two thresholds behave differently. The first one is connectivity-dependent and the second one depends on the local fragility of the network.

**Key words:** Complex Systems, Nonequilibrium Dynamics, Growing Networks, Extreme-value Statistics, Failure Propagation

# Articles de références

## Publiés dans des revues à comité de lecture

- 1. Finite-Time Fluctuations in the Degree Statistics of Growing Networks, C. Godrèche, H. Grandclaude, J.M Luck, J. Stat. Phys. **137**, 1117 (2009). [arXiv:0907.1470]
- 2. Statistics of leaders and lead change in growing networks, C. Godrèche, H. Grandclaude, J.M Luck, J. Stat. Mech. P02001 (2010). [arXiv:0911.3057]

# Remerciements

# Table des matières

$\mathbf{A}$	In	troduct	ion générale	7
Ι	Ι	ntroduc	tion sur les réseaux	9
	I.1	Théori	e des graphes	9
		I.1.1	Définitions	9
		I.1.2	Graphes aléatoires	15
		I.1.3	Lexique	16
	I.2	Graphe	es aléatoires	17
		I.2.1	Observables dans les graphes aléatoires	17
		I.2.2	Graphes aléatoires à l'équilibre	18
		I.2.3	Graphes en croissance	19
		I.2.4	Exemples de distributions	20
		I.2.5	Modèle de petit-monde	21
	I.3	Observ	vations empiriques	22
		I.3.1	Graphes d'Erdös-Rényi	23
		I.3.2	Graphes exponentiels	23
		I.3.3	Réseaux invariants d'échelle	23
	I.4	Organi	sation de la thèse	28
Б	П	• • • •		90
в	$\mathbf{Pr}$	oprietes	s statistiques des réseaux en croissance	30
Π	Í	État de l	art sur les réseaux invariants d'échelle	33
	II.1	Modèle	e de Barabási-Albert	33
	II.2	Attrac	tivité initiale des sites	35
	II.3	Modèle	$e \text{ avec redirection}  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots  $	36
TT	T T	Effets de	taille finie dans les réseaux en croissance	39
<b>.</b>		Descrii	ption des modèles	39
	111.1	III 1 1	Conditions initiales	40
		III 1 2	Modèle d'attachement uniforme	40
		III 1 3	Modèle de Barabási-Albert	41
		III 1 4	Modèle d'attachement préférentiel généralisé	41
	$\mathbb{H}^2$	Modèle	e d'attachement uniforme	42
		UUUU		

	III.2.2	Distribution du degré pour un noeud quelconque
	III.2.3	Distribution du degré aux temps finis
	III.2.4	Comportement de la fonction d'échelle à la transition
	III.2.5	Grandes déviations
III.3	Modèle	e de Barabási-Albert
	III.3.1	Equation maîtresse
	III.3.2	Fonction génératrice
	III.3.3	Résolution des équations de récurrence
	III.3.4	Distribution d'échelle du degré d'un noeud $i$
	III.3.5	Distribution à temps fini du degré global
	III.3.6	Grandes déviations
III.4	Modèle	e d'attachement préférentiel généralisé $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $.$ 72
	III.4.1	Équation maîtresse $\ldots \ldots 72$
	III.4.2	Fonction génératrice du degré du noeud $i$ à temps fini $\ldots \ldots \ldots 74$
	III.4.3	Fonction génératrice à temps fini du degré global
	III.4.4	Distribution du degré d'un noeud $i$
	III.4.5	Distribution du degré global
	III.4.6	Comportement de la fonction d'échelle
	III.4.7	Grandes déviations
III.5	$\operatorname{Conclu}$	sion
	Statistiqu	ieue des extrêmes dans les reseaux en croissance 87
1 V . 1	$\mathbf{T}_{V} 1 1$	Distribution du plus grand dogré
	1V.1.1	Distribution du pombre de composure
IV 9	Dográ	$\begin{array}{c} \text{Distribution du nombre de co-meneurs} \\ \text{maximum} \\ 02 \end{array}$
1 V.2	IV 9.1	1114 X 11111111 37
	1 V . 4 . 1	Modèle AU 93
	W 2 2	Modèle AU
	IV.2.2 IV.2.3	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 90
IV 3	IV.2.2 IV.2.3	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   e de prépondérances 101
IV.3	IV.2.2 IV.2.3 Nombr	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   e de prépondérances 101   Attachement uniforme 101
IV.3	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   re de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103
IV.3	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2 IV.3.3	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   re de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103   Attachement préférentiel généralisé 105
IV.3 IV 4	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2 IV.3.3 Indice	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   re de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103   Attachement préférentiel généralisé 105   du meneur 105
IV.3 IV.4	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2 IV.3.3 Indice IV.4.1	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   re de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103   Attachement préférentiel généralisé 105   du meneur 105   Attachement uniforme 107
IV.3 IV.4	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2 IV.3.3 Indice IV.4.1 IV.4.2	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   e de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103   Attachement préférentiel généralisé 105   du meneur 107   Barabási-Albert 107   Barabási-Albert 107
IV.3 IV.4	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2 IV.3.3 Indice IV.4.1 IV.4.2 IV.4.3	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   re de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103   Attachement préférentiel généralisé 105   du meneur 105   Attachement uniforme 107   Barabási-Albert 109   Attachement préférentiel généralisé 109
IV.3 IV.4	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2 IV.3.3 Indice IV.4.1 IV.4.2 IV.4.3 Persist	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   e de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103   Attachement préférentiel généralisé 105   du meneur 105   Attachement uniforme 107   Barabási-Albert 105   du meneur 105   Attachement uniforme 107   Barabási-Albert 109   Attachement préférentiel généralisé 109   Attachement préférentiel généralisé 109   Attachement préférentiel généralisé 109   Attachement préférentiel généralisé 109
IV.3 IV.4 IV.5	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2 IV.3.3 Indice IV.4.1 IV.4.2 IV.4.3 Persist IV.5 1	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   re de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103   Attachement préférentiel généralisé 105   du meneur 105   Attachement uniforme 107   Barabási-Albert 109   Attachement préférentiel généralisé 101   Attachement préférentiel généralisé 101   Attachement uniforme 111   Attachement uniforme 111
IV.3 IV.4 IV.5	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2 IV.3.3 Indice IV.4.1 IV.4.2 IV.4.3 Persist IV.5.1 IV.5.2	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   e de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103   Attachement préférentiel généralisé 105   du meneur 105   Attachement uniforme 107   Barabási-Albert 105   du meneur 105   Attachement uniforme 107   Barabási-Albert 109   Attachement préférentiel généralisé 109   Attachement préférentiel généralisé 101   Attachement préférentiel généralisé 111   Attachement préférentiel généralisé 111   Attachement uniforme 111   Attachement uniforme 111   Attachement uniforme 112   Attachement préférentiel 114
IV.3 IV.4 IV.5	IV.2.2 IV.2.3 Nombr IV.3.1 IV.3.2 IV.3.3 Indice IV.4.1 IV.4.2 IV.4.3 Persist IV.5.1 IV.5.2 Conclu	Modèle AU 93   Modèle BA 94   Attachement préférentiel généralisé 99   e de prépondérances 101   Attachement uniforme 101   Barabási-Albert 103   Attachement préférentiel généralisé 105   du meneur 105   Attachement uniforme 107   Barabási-Albert 107   Barabási-Albert 109   Attachement préférentiel généralisé 109   Attachement préférentiel généralisé 109   Attachement préférentiel généralisé 109   Attachement préférentiel généralisé 111   Attachement uniforme 112   Attachement préférentiel 114   sion 118

С	Cascades	de défaillances dans les réseaux	120
$\mathbf{V}$	Propagat	tion de défaillances dans les réseaux	123
	V.1 Modèle	e de Motter-Lai	123
	V.2 Modèle	e de cascade	124
	V.3 Modèle	es de faisceaux de fibres	125
vi	[ Modèles		129
• 1	VI 1 Définit	ion du modèle L&B	120
	VI 1 1	Définition générale	129
	VI 1 2	Bedistribution bimodale	131
	VI 2 Phénor	ménologie du modèle L&B	132
	VI 2 1	Probabilité de non-cascade	132
	VI 2 2	Distribution du nombre total de défaillances dans la cascade	132
	VI 2 3	Probabilité de rupture	133
	VI.2.4	Valeurs de transition	135
	VI.2.5	Nombre de descendants directs d'une défaillance	135
	VI.2.6	Nombre d'ancêtres d'une défaillance	. 138
	VI.2.7	Nombre de membres d'une génération	138
	VI.2.8	Temps d'arrêt de la cascade	141
	VI.2.9	Charge défaillante	141
	VI.3 Du mo	dèle Lehmann & Bernasconi au processus de branchement généralisé	141
	VI.3.1	Approximation pour un système de taille infinie	144
	VI.3.2	Nombre de descendants d'une défaillance	144
	VI.3.3	Distribution conditionnelle des charges défaillantes	145
	VI.4 Proces	sus de branchement généralisé de Harris	146
	VI.5 Validit	é des approximations	147
	VI.5.1	Distribution conditionnelle de la charge défaillante	147
	VI.5.2	Moyenne et distribution du nombre de défaillances	147
	VI.6 Conclu	ision	150
VI	I Etude dı	ı processus de branchement généralisé	153
	VII.1 Définit	$\mathcal{L}$ ion du modèle $H$	153
	VII.2 Probab	bilité de non-cascade	156
	VII.3 Nombr	e total de défaillances	156
	VII.3.1	Pour un processus de branchement simple	157
	VII.3.2	Dans le modèle $H$	161
	VII.3.3	Calcul exact pour les cas simples	167
	VII.4 Nombr	e de défaillances par génération	169
	VII.4.1	Processus de branchement simple	169
	VII.4.2	Processus de branchement généralisé	171
	VII.4.3	Cas simple $\alpha = 0$	172
	VII.5 Probab	pilité de rupture	172
	VII.5.1	Probabilité d'extinction	174

VII.5.2	Probabilité de rupture
VII.5.3	Intervalle de définition de $P_b$
VII.5.4	Résolution numérique de l'équation intégrale
VII.6 Valeur	s de transition $\ldots \ldots \ldots$
VII.7 Temps	d'arrêt de la cascade
VII.8 Distrib	oution des charges défaillantes
VII.8.1	Distribution à la première génération
VII.8.2	Distribution à la s-ième génération
VII.8.3	Distribution limite
VII.8.4	Exemple de cas simples
VII.9 Nomb	e typique de descendants par individu
VII.9.1	Processus de Galton-Watson
VII.9.2	Processus de branchement généralisé
VII.9.3	Cas simples
VII.10 Conclu	$1 sion \ldots \ldots$

# **D** Conclusion

 $\mathbf{204}$ 

V	III R	<b>l</b> ésultat	s et perspectives	207	
VIII.1 Effets de taille finie					
	VIII.	2 Propa	gation de défaillance	209	
A	É	quation	ns intégrales	213	
	A.1	$\operatorname{Relati}$	on de récurrence dans le processus de branchement généralisé	213	
		A.1.1	Distribution de points et intégrales aléatoires	213	
		A.1.2	Relation de récurrence sur les fonctionnelles génératrices des moment	s. 214	
	A.2	Modèl	le de cascade de Harris	215	
		A.2.1	Relation de récurrence pour la FGM	215	
		A.2.2	Probabilité de rupture	217	
в	N	lotatior	18	219	
	B.1	Notat	ions générales	219	
	B.2	Notat	ions spécifiques au modèle LB	220	
С	Cascade avec extension spatiale sur l'axe 1D 22				
	C.1	Modèl	le Lehmann & Bernasconi	221	
	C.2	Modèl	le Recuit	224	
	C.3	Propri	iétés spatiales du modèle recuit	227	

# Partie A Introduction générale

# CHAPITRE I Introduction sur les réseaux

## Qu'est-ce-qu'un réseau?

Un réseau, ou graphe, est un ensemble de noeuds reliés par des connexions. Tout système composé d'éléments en interactions les uns avec les autres peut se modéliser comme un graphe. Cette modélisation est particulièrement utile pour la représentation de systèmes complexes lorsque les processus dynamiques étudiés reposent sur les interactions entre un grand nombre d'éléments distincts. De récents progrès ont été réalisés dans l'étude et la compréhension de systèmes complexes, tels le World Wide Web (www) [1], l'Internet [2], les réseaux de communication, d'énergie [3] et de transports [4], les interactions sociales [5] ou les processus biologiques [6], grâce à la modélisation en réseau de ces systèmes.

Les concepts clés de l'étude des réseaux ont été introduits dans le domaine de la théorie des graphes en mathématiques, dont nous allons faire un court résumé ci-dessous.

## 1.1 Théorie des graphes

L'origine de la théorie des graphes est généralement attribuée à Leonard Euler et à son article [7] sur le problème des sept ponts de Königsberg. La ville est alors construite sur deux îles et les berges environnantes, qui sont reliées par sept ponts. Euler se demande si il existe une promenade à travers la ville qui permet de passer une fois et une seule sur chaque pont et de revenir à son point de départ. Pour résoudre ce problème, Euler introduit les premiers éléments de la théorie des graphes et montre qu'un tel chemin n'existe pas car ce graphe comprend un trop grand nombre de sommets incidents à un nombre impair d'arêtes. Les mathématiciens se sont ensuite largement intéressés aux graphes réguliers, et à des problèmes statiques sur graphes, comme le problème de la coloration de graphes, avant que le développement des lois de l'électricité n'attire leur attention sur des problèmes de flots sur réseaux. Ce n'est finalement qu'au milieu du XX<sup>e</sup> siècle qu'Erdös et Rényi [8, 9] introduisirent la notion de probabilités dans la construction de graphes.

#### I.1.1 Définitions

Un graphe est composé d'un ensemble  $\mathcal{V}$  dénombrable non vide de *noeuds*, ou *sommets*, ou *sites*, et d'un ensemble  $\mathcal{E}$  de paires de noeuds, qui représentent les *connexions*, ou *arêtes*. Deux

sommets connectés par une arête sont dits adjacents. La représentation mathématique d'un graphe  $\mathcal{G}$  est donc la paire des deux ensembles :

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}). \tag{I.1}$$

Un exemple de représentation graphique d'un réseau est donné figure I.1.



FIGURE I.1 – Représentation graphique d'un réseau connexe non dirigé.

#### a Types de graphes

Si les arêtes ne sont pas orientées, *i.e.* les paires de sommets de l'ensemble  $\mathcal{E}$  ne sont pas ordonnées, le réseau est dit *non dirigé*. Si les arêtes sont orientées, le graphe est dirigé. La représentation graphique d'une arête orientée comporte une flèche dans le sens de l'interaction. Les arêtes peuvent également porter un poids, qui modélise la force de l'interaction entre les sommets qu'elles relient. Les réseaux sont alors appelés graphes pondérés, et la représentation graphique habituelle consiste à faire varier l'épaisseur de l'arête. On peut définir toute une zoologie de graphes pour modéliser un système en particulier, en faisant varier les propriétés des sommets et des arêtes, par exemple en distinguant plusieurs types de sommets ou d'arêtes, en autorisant une arête à connecter plus de deux sommets (hypergraphes), en attribuant un poids aux sommets... On représente quelques uns de ces types de graphes sur la figure I.2.

#### b Matrice d'adjacence

La taille du réseau est donnée par le nombre N de sommets, *i.e.* le cardinal de l'ensemble  $\mathcal{V}^{a}$ . Le nombre maximum d'arêtes distinctes dans un graphe de taille N est  $\binom{N}{2}$ , et un graphe dont le nombre d'arêtes E est maximum est appelé graphe complet.

Un graphe peut être représenté mathématiquement par sa matrice d'adjacence  $\mathbb{M} = m_{ij}$ . Cette matrice de taille  $N \times N$  code les arêtes qui existent entre les N sommets du graphe. Pour

a. En théorie des graphes, la taille d'un réseau est généralement donnée par le nombre total E d'arêtes.



FIGURE I.2 – Représentations graphiques de différentes espèces de graphes.

un graphe non pondéré, les éléments de matrice prennent les valeurs :

$$m_{ij} = \begin{cases} 1 & \text{si } (i,j) \in \mathcal{E} \\ 0 & \text{sinon} \end{cases}$$
(I.2)

La matrice d'adjacence d'un graphe non dirigé est symétrique, car alors l'arête (i, j) est égale à l'arête (j, i). Elle est en général non symétrique dans un graphe dirigé, où l'arête (i, j) correspond à la connexion du sommet i vers le sommet j. Dans un graphe pondéré, les éléments de matrice sont des réels qui indiquent le poids de chaque arête. Un élément non nul dans la diagonale correspond à une boucle unitaire, c'est-à-dire une arête qui connecte un sommet à lui-même. Cette structure est souvent appelée boucle d'auto-rétroaction dans les réseaux biologiques et correspond par exemple à l'inhibation de la synthèse d'une protéine lorsque sa concentration est suffisante dans une cellule. On ne considérera par la suite que des réseaux sans boucle unitaire.



FIGURE I.3 – Représentation d'un arbre. On vérifie bien sur la figure que le nombre d'arêtes E = N - 1.

#### c Densité d'un graphe

Un graphe est dit *connecté* ou connexe si tous ses sommets peuvent être atteints par au moins un chemin d'arêtes à partir de n'importe quel sommet du graphe, c'est-à-dire qu'il n'y a aucun sous-ensemble déconnecté du reste du réseau. On ne considérera dans la suite que des graphes connectés. Le nombre minimum d'arêtes que doit contenir un graphe connecté de taille N est  $E_{min} = N - 1$ , et dans un graphe sans boucle unitaires, le nombre maximum possible d'arêtes est  $E_{max} = N(N-1)/2$ . Lorsque le nombre d'arêtes du réseau est de l'ordre de  $E \sim E_{max}$ , le réseau est dit dense, tandis que si  $E \sim N^{\delta}$ , pour  $1 < \delta < 2$  le graphe est *clairsemé*.

Une idée de la structure générale d'un graphe peut donc être donnée par la densité  $\mathcal{D}$  d'arêtes qu'il contient. On définit la densité par :

$$\mathcal{D} = \frac{E}{E_{max}}.\tag{I.3}$$

Les réseaux que l'on considérera par la suite sont des *arbres*, c'est-à-dire des graphes hiérarchiques non dirigés dans lesquels chaque sommet ne peut être atteint à partir d'un autre sommet que par un unique chemin. Dans ces arbres, le nombre d'arêtes est donc le plus petit possible : E = N - 1, et le réseau est clairsemé. On représente un exemple d'arbre sur la figure I.3

#### d Chemins et connectivité

Un graphe connecté dans lequel le nombre d'arêtes est plus grand que  $E_{min}$  peut contenir des boucles, ou cycles, c'est-à-dire des ensembles d'arêtes distinctes qui permettent, en partant d'un sommet *i* de revenir à ce même sommet. On souligne de telles boucles en rouge sur la figure I.4.

Le nombre I de boucles irréductibles dans un graphe se calcule par :

$$I = E + 1 - N, \tag{I.4}$$



FIGURE I.4 – Représentation d'un réseau quelconque dans lequel les arêtes appartenant à des boucles sont tracées en rouge.



FIGURE I.5 – Représentation du plus court chemin entre les deux noeuds rouges.

où par irréductible on entend que la boucle ne peut être obtenue par combinaison de boucles plus petites. Le graphe de la figure I.4 compte par exemple N = 12 sommets, E = 14 arêtes et I = 3 boucles irréductibles. Dans un arbre, le nombre de boucles est évidemment I = 0.

La notion de chemin sur un réseau est liée à la notion de distance entre deux noeuds. On définit la longueur d'un chemin comme le nombre n d'arêtes distinctes que ce chemin traverse. Un chemin qui emprunte une seule arête a donc longueur n = 1. La distance  $d_{i,j}$  entre deux sommets respectivement étiquetés i et j est la taille du plus petit chemin qui permet de rejoindre j en partant de  $i^{b}$ . On souligne sur la figure I.5 le plus court chemin entre les deux sommets rouges, et la distance entre ces deux noeuds est donc d = 3.

À partir de la définition de chemin, on peut déterminer l'importance d'un noeud dans la connectivité du graphe à travers la notion de *centralité*. Il existe plusieurs types de centralité. La centralité de proximité  $g_i$  du sommet i est définie comme l'inverse de la distance d'un noeud

b. Équivalent au chemin qui rejoint i en partant de j dans un réseau non orienté.



FIGURE I.6 – Centralité des sommets d'un graphe quelconque : le sommet bleu a une centralité de proximité  $g_b = 1/19$  et d'intermédiation  $b_b = 154/6$ , pour le sommet vert  $g_v = 1/20$  et  $b_v = 149/6$ .

d'indice i à tous les autres sommets j du graphe :

$$g_i = \frac{1}{\sum_{j \neq i} d_{i,j}},\tag{I.5}$$

et elle est d'autant plus grande que le sommet est à plus petite distance de tous les autres sommets du graphe. Sur la figure I.6, le sommet vert possède par exemple une centralité de proximité de  $g_v = 1/20$ , tandis que le sommet bleu est l'élément le plus central selon cette définition, avec  $g_b = 1/19$ .

La centralité  $b_i$  d'intermédiation (souvent appelée uniquement centralité) reflète le poids du sommet i dans la liaison entre deux parties du graphes peu connectées entre elles. Elle est définie comme le rapport du nombre de plus courts chemins entre toutes les paires de sommets possibles passant par i:

$$b_i = \sum_{k \neq j \neq i} \frac{B(k, i, j)}{B(k, j)},\tag{I.6}$$

où B(k, i, j) est le nombre de plus courts chemins entre les sommets k et j passant par i, et B(k, j) le nombre total de plus courts chemins entre k et j. Cette centralité détermine si le sommet joue un rôle important dans le trafic sur le réseau. Elle est parfois aussi appelée *charge*. Sur la figure I.6, le sommet vert a une centralité  $b_v = 149/6$  et le sommet bleu  $b_b = 154/6$ .

#### e Degré

Nous nous intéresserons dans la suite aux propriétés locales des réseaux, contenues dans leur proche voisinage. En particulier, nous étudierons en grand détail le nombre de plus proches voisins d'un noeud *i*, c'est-à-dire son *degré*. Le degré  $k_i$  du sommet *i* est le nombre d'arêtes qui connectent directement le sommet *i* à d'autres sommets du graphe. Sur la figure I.6, le degré des sommets bleu et vert est identique et vaut  $k_v = k_b = 4$ . Dans un réseau dirigé, on distingue le degré entrant  $k_i$ , c'est-à-dire le nombre d'arêtes qui pointent vers le sommet considéré, et le degré sortant  $k_o$  qui compte le nombre d'arêtes issues du sommet. Le degré total du sommet est alors  $k = k_i + k_o$ .

Dans un graphe *complet*, tous les sommets sont connectés entre eux, et le degré d'un noeud quelconque du graphe est k = N-1. C'est le degré maximum possible dans un graphe sans boucle unitaire. Nous étudions des graphes connectés, donc le degré minimum d'un noeud quelconque est  $k_{min} = 1$ .

Une quantité proche du degré est le nombre de plus proches *m*-ièmes voisins  $z_m$  d'un noeud, c'est-à-dire le nombre de sommets situés à une distance *m* du sommet considéré. Cette quantité contient les propriétés de voisinage étendu du sommet considéré.

#### I.1.2 Graphes aléatoires

La notion de probabilité dans la construction des graphes fut introduite par Erdös et Rényi [8, 9]. Dans leur modèle, le nombre total N de sommets est fixé et l'existence d'une arête entre deux noeuds dépend d'une probabilité p. Chaque configuration de la position des arêtes est un élément de l'ensemble statistique des réalisations possibles, et son poids statistique dépend du modèle considéré. Les quantités étudiées sur le réseau sont alors statistiquement distribuées. Un graphe aléatoire est souvent construit à partir de la distribution de probabilité du degré d'un noeud quelconque P(k). Cette distribution fixe les propriétés locales du graphe, et on peut calculer la plupart des autres quantités à partir de ces propriétés.

La probabilité qu'un sommet *i* appartenant à un graphe de taille N soit de degré k est donnée par la distribution du degré p(k, i, N).

La probabilité qu'un sommet tiré au hasard parmi tous les sommets de toutes les réalisations possibles d'un graphe de N sommets soit de degré k est donnée par la distribution totale du degré

$$P(k,N) = \frac{1}{N} \sum_{i=1}^{N} p(k,i,N).$$
 (I.7)

Si tous les sommets sont statistiquement équivalents, la distribution totale du degré est aussi la distribution du degré de chaque sommet. La distribution du degré correspond à la distribution d'une quantité liée à un sommet, et donc aux propriétés locales du réseau.

Le degré moyen  $\langle k \rangle$  d'un graphe est le premier moment de la distribution totale du degré :

$$\langle k \rangle = \sum_{k} k P(k, N), \tag{I.8}$$

et le nombre d'arêtes E est donné par :

$$E = \frac{\langle k \rangle N}{2},\tag{I.9}$$

pour un réseau de taille  $N \gg 1$ . Un graphe clairsemé a un degré moyen  $\langle k \rangle \ll N$ .

Pour un réseau dirigé, on peut distinguer la distribution des degrés entrants  $p_i(k_i, s, N)$  et sortants  $p_o(k_o, s, N)$  d'un sommet s. La distribution totale respective des degrés entrants et

sortants est donnée par :

$$P_i(k_i, N) = \frac{1}{N} \sum_{i=1}^{N} p_i(k_i, i, N)$$
(I.10)

$$P_o(k_o, N) = \frac{1}{N} \sum_{i=1}^{N} p_o(k_o, i, N)$$
(I.11)

Pour un réseau sans connexions à l'extérieur, on doit avoir :

$$\langle k_i \rangle = \langle k_o \rangle = \frac{\langle k \rangle}{2}.$$
 (I.12)

Du point de vue de la physique statistique, on distingue deux types de graphes aléatoires : les réseaux à l'équilibre et les réseaux hors équilibre.

- Les réseaux à l'équilibre contiennent en général un nombre fixé de noeuds. Ces noeuds sont statistiquement indépendants et équivalents. On peut citer comme exemple le modèle du réseau aléatoire classique d'Erdös et Rényi.[8]
- Les réseaux hors équilibre sont caractérisés par une distribution du degré non homogène sur l'ensemble des noeuds du réseau. Cette non-homogénéité peut être due à l'addition de nouveaux noeuds dans le réseau au cours du temps, et on parle alors de réseaux en croissance. Dans le cas de résaux hors équilibre, on doit considérer la variable temporelle tdans l'étude des quantités, et la distribution d'une variable quelconque X est une fonction du temps : P(X, t). La variation temporelle des distributions s'exprime dans le formalisme de l'équation maîtresse [10, 11].

On présente section I.2 différents modèles de graphes aléatoires, après un court résumé des termes utilisés en théorie des graphes

#### I.1.3 Lexique

- Sommet = Noeud = Site : élement du réseau,
- Arête = Lien = Connexion : relation entre les éléments du réseau,
- Réseau connecté : réseau sans sous-parties distinctes,
- Réseau non dirigé : réseau dont les arêtes ne sont pas orientées, il n'y a pas de sens de parcours d'une connexion,
- Réseau dirigé : réseau dont les arêtes sont orientées,
- Réseau pondéré : réseau dont les arêtes portent un poids,
- Degré k d'un noeud : nombre de connexions du noeud,
- Pour les réseaux orientés : Degré entrant : nombre de connexions vers le sommet, Degré sortant : nombre de connexions sortant du sommet,
- Arbre : réseau sans boucles
- Matrice d'adjacence : matrice carrée  $N \times N$  où N est la taille du graphe. Les éléments de matrice  $m_{ij}$  indiquent la connexion *i.e.* l'adjacence de deux sommets

$$\begin{array}{ll} m_{ij} = 1 & \text{s'il existe une connexion entre le sommet } i \text{ et le sommet j} \\ m_{ij} = m_{ji} & \text{si le réseau est non-dirigé} \\ m_{ii} = 0 & \text{si le réseau ne contient pas de boucle unitaire} \end{array}$$
(I.13)

- Taille linéaire d'un graphe : nombre minimal de connexions à utiliser pour passer d'un noeud quelconque à un autre noeud du réseau
- Distance  $d_{i,j}$ : Nombre minimum d'arêtes que l'on doit emprunter sur le graphe pour aller d'un sommet i à un sommet j.
- Nombre de plus proches *m*-ièmes voisins  $z_m$  : nombre de sommets situés à une distance *m* du sommet considéré.
- Centralité  $b_i$ : nombre total de plus courts chemins entre toutes les paires de sommets possibles dans le réseau passant par le sommet i, parfois appelée charge.
- Réseau aléatoire : réseau construit avec un arrangement aléatoire des connexions. Un réseau particulier est un membre d'un ensemble statistique de toutes les réalisations possibles.
- Réseau en équilibre : réseau aléatoire dans lequel tous les noeuds sont équivalents et statistiquement indépendants.
- Réseau hors équilibre : réseau aléatoire dont les noeuds ont une distribution du degré inhomogène.

## I.2 Graphes aléatoires

#### I.2.1 Observables dans les graphes aléatoires

Les observables, définies exactement dans les graphes combinatoires, sont dans les graphes aléatoires définies par le biais de distributions. Le degré k d'un noeud quelconque est ainsi distribué selon la distribution du degré P(k), qui caractérise le modèle de graphe utilisé.

Le nombre de *m*-ièmes plus proches voisins  $z_m$  est distribué selon  $P_m(z)$  sur l'ensemble du graphe.

Les distances d entre deux noeuds quelconques du réseau sont distribuées selon la fonction de distribution  $P_d(d)$ . Si la distribution décroît assez vite, on peut définir un diamètre [12] :

$$\langle d \rangle = \sum_{d} dP_d(d)$$
 (I.14)

ou une distance maximale  $d_{max}$  qui correspond à la taille du plus long des plus courts chemins entre n'importe quelle paire de noeuds du réseau.

La centralité d'intermédiation b est distribuée selon  $P_b(b)$ .

Ces différentes distributions sont reliées entre elles par les relations suivantes :

• Les distributions des distances et des *m*-ièmes plus proches voisins dans un graphe sont reliées par :

$$P_d(d) = \frac{1}{N} \sum_{z} z P_d(z) = \frac{\langle z_d \rangle}{\sum_d \langle z_d \rangle},$$
(I.15)

où  $\langle z_d \rangle$  est le nombre moyen de noeuds à la distance d d'un sommet que l'onque. La relation inverse est donnée par :

$$P_d(z) = \binom{N-1}{z} \left( P(d) \right)^z \left( 1 - P(d) \right)^{N-1-z}.$$
 (I.16)



FIGURE I.7 – Une des réalisations possibles d'un réseau d'Erdös-Rényi.

• La centralité moyenne d'un réseau est liée à la longueur moyenne du plus court chemin dans le réseau de taille N par :

$$\langle b \rangle = (N-1) (\langle d \rangle - 1) \tag{I.17}$$

#### I.2.2 Graphes aléatoires à l'équilibre

Le premier modèle de graphe aléatoire à l'équilibre fut défini par Erdös et Rényi [8] comme suit :

- Le graphe contient N sommets et E arêtes.
- Les E arêtes sont choisies au hasard parmi les  $\binom{N}{2}$  possibles.
- Ce graphe a donc

$$\binom{\binom{N}{2}}{E} = \binom{\frac{N(N-1)}{2}}{E}$$
(I.18)

réalisations possibles.

On trace sur la figure I.7 l'une des 6435 réalisations possibles d'un réseau d'Erdös-Rényi pour N = 6 et E = 8. Une définition alternative de ce modèle consiste, pour chaque paire de sommets du graphe, à les connecter avec probabilité p, et à ne pas tracer d'arête avec probabilité 1-p. Le nombre d'arêtes E de la définition précédente est relié à la probabilité p par le nombre moyen d'arêtes obtenues selon la seconde méhode de construction :

$$E = \frac{N(N-1)}{2}p.$$
 (I.19)

La distribution du degré d'un tel réseau est donnée par la relation bimodale :

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}.$$
 (I.20)

Un graphe d'Erdös-Rényi a pour degré moyen :

 $\langle k \rangle = p(N-1). \tag{I.21}$ 



FIGURE I.8 – Construction d'une réalisation d'un réseau aléatoire à partir de la distribution du degré de ses noeuds. Sur la figure de gauche, on représente les noeuds munis de leur degré tiré dans P(k), et on relie ces ébauches d'arêtes au hasard sur la figure de droite.

Ce modèle montre une transition de percolation selon la valeur de p, qui a été largement étudiée par Erdös et Rényi [9], Bollobás [13], Kolchin [14] ou Luczak [15].

On peut construire un réseau aléatoire à partir de la distribution du degré de ses noeuds de la façon suivante :

- 1. Pour chacun des N sommets du graphe, on tire un degré k de la distribution P(k).
- 2. Un noeud possède donc k ébauches d'arêtes, que l'on prolonge jusqu'au crochet libre d'un autre noeud, jusqu'à ce que tous les crochets soient associés à une arête.

On représente sur la figure I.8 ces deux étapes de la construction d'un réseau dont le degré est distribué selon une distribution P(k).

#### I.2.3 Graphes en croissance

Un graphe en croissance peut être considéré comme hors équilibre [16], tous ses noeuds n'ayant pas la même distribution du degré, et les distributions de probabilité dépendant du temps. L'origine de l'inhomogénéité des noeuds provient de l'évolution temporelle du réseaux. En effet, si un réseau croît par ajout de nouveaux noeuds, les noeuds les plus anciens auront statistiquement un degré plus élevé que les noeuds les plus récents. Si l'addition de nouveaux noeuds est stoppée après un certain laps de temps, la distribution du degré de chaque noeud pourra tendre, par addition continue de connexions, vers l'état d'équilibre correspondant, mais les noeuds ne seront jamais tout à fait indiscernables [12]. On représente en exemple les premières étapes d'un graphe en croissance sur la figure I.9. Les distributions des variables X qui apparaissent dans les réseaux hors d'équilibre sont en général notées P(X, t), où t est la variable temporelle du système.

Les réseaux aléatoires sont souvent étudiés grâce aux outils de la physique statistique, et on considère donc en général la limite  $N \to \infty$  de la taille des réseaux. Les modèles de réseaux aléatoires les plus étudiés sont caractérisés par les distributions du degré décrites ci-dessous.



FIGURE I.9 – Premières étapes d'un réseau en croissance suivant le modèle de citation : chaque nouveau noeud arrive dans le réseau porteur d'un lien qu'il connecte à l'un des noeuds déjà présent.

#### I.2.4 Exemples de distributions

Le degré des sommets d'un réseau peut être distribué selon de nombreuses distributions différentes, et les plus fréquentes appartiennent aux catégories suivantes :

1. Distribution de Poisson :

$$P(k) \propto \frac{\mathrm{e}^{-\langle k \rangle} \langle k \rangle^k}{k!}.$$
 (I.22)

Cette distribution est la limite  $N \to \infty$ ,  $pN \to cste$  de la distribution bimodale du réseau aléatoire classique d'Erdös-Rényi [8], et sa valeur moyenne  $\langle k \rangle$  est définie et vérifie :  $\langle k \rangle = p(N-1)$ .

2. Distribution exponentielle :

$$P(k) \propto e^{-k/\langle k \rangle}$$
 (I.23)

Cette distribution correspond à la distribution des graphes simples en croissance, comme par exemple le modèle dont on a présenté les premières étapes figure I.9. L'échelle qui émerge naturellement de cette distribution est de l'ordre du degré moyen. On reviendra plus en détail dans les chapitres suivants sur un modèle présentant ce type de distribution.

3. Distribution en loi de puissance :

$$P(k) \propto k^{-\gamma}, \ k \neq 0, \tag{I.24}$$

où  $\gamma > 1$  est l'exposant de la distribution. Pour ce type de distribution du degré dans le réseau, aucune échelle naturelle n'apparaît. On dit alors que le réseau est *invariant d'échelle*. À la limite asymptotique  $N \to \infty$ , tous les moments d'ordre  $m \ge \gamma - 1$  divergent. En particulier, si le degré moyen  $\langle k \rangle$  est fini, alors  $\gamma > 2$ , et les moments d'ordre m > 1 ne sont pas définis si  $\gamma \le 3$ . Si le degré moyen  $\langle k \rangle$  diverge également, pour une distribution du degré stationnaire, alors  $1 < \gamma \le 2$ . Nous étudions par la suite le modèle d'attachement

préférentiel de Barabási-Albert, dans lequel la distribution en loi de puissance apparaît naturellement.

 Il existe encore d'autres types de distributions du degré, comme les distributions multifracales ou discrètes, pour lesquelles nous renverrons le lecteur au livre de Dorogotsev et Mendes [16].

#### I.2.5 Modèle de petit-monde

Au-delà de la mesure locale représentée par la distribution du degré des noeuds, on peut également s'intéresser au regroupement des sommets à plus grande échelle. Le concept de *petit monde* fait référence aux expériences de sociologues et psychologues comme Milgram [17]. Celui-ci observa que deux individus pris au hasard dans la population nord-américaine peuvent être reliés par une chaîne de relations sociales remarquablement courte devant la taille totale du système. En 1998, Watts et Strogatz [18] proposèrent un modèle de graphe aléatoire qui permet une interpolation entre les graphes réguliers et les graphes complétement aléatoires. Ils montrent que ce modèle présente les caractéristiques de courts chemins et de grand coefficient de regroupement observées dans maints réseaux réels.

Watts et Strogatz considèrent un réseau de taille N clairsemé, donc de degré typique  $N \gg k$ , mais cependant connecté :  $k \gg \ln N$ . Dans un réseau régulier présentant ces caractéristiques, la longueur typique des plus courts chemins est  $L \sim N/2k \gg 1$  et le coefficient de regroupement vaut  $C \sim 3/4$ . Ce coefficient de regroupement est défini comme la moyenne de la fraction d'arêtes qui existent dans un réseau entre les n voisins d'un noeud sur le nombre total n(n-1)/2d'arêtes qui pourraient exister. Un réseau aléatoire classique, à la Erdös-Rényi, présente les valeurs suivantes :  $C \sim k/N \ll 1$  et  $L \sim \ln N/\ln k$ .

En effet, dans un réseau aléatoire dont le nombre d'éléments N est grand, quelle que soit la distribution du degré, pourvu qu'elle décroisse assez vite, la probabilité d'existence de boucles dans le voisinage d'un noeud quelconque peut être considérée comme négligeable. Le réseau présente une structure locale en arbre. Le nombre de sommets situés à une distance plus petite ou égale à l d'un sommet quelconque est de l'ordre de  $k^l$ , où k est le nombre moyen de voisins d'un sommet. La taille du réseau N et le nombre moyen de noeuds situés à une distance plus petite ou égale à la distance moyenne L entre deux sommets sont du même ordre de grandeur :  $N \sim k^L$ . La distance moyenne entre deux sommets est donc de l'ordre de :

$$L \sim \frac{\ln N}{\ln k}.\tag{I.25}$$

La dépendance de la distance typique dans le réseau à la taille du réseau n'est donc que logarithmique dans un graphe aléatoire, et la distance moyenne entre deux sommets reste donc petite même pour de grands réseaux. Les réseaux aléatoires montrent bien un effet de petit monde, mais ils sont beaucoup moins groupés qu'observés dans les réseaux réels [19].

Pour réconcilier ces deux aspects, Watts et Strogatz développèrent donc un modèle simple dans lequel la taille moyenne des plus courts chemins et le coefficient de regroupement peuvent reproduire les observations. On part d'un graphe régulier de N sommets connectés exactement à k voisins, représentés sur un cercle par simplicité (figure I.10). Pour chaque noeud, avec probabilité p, l'arête qui le lie à son plus proche voisin dans le sens des aiguilles d'une montre



FIGURE I.10 – Schéma d'un graphe de N = 15 noeuds et E = 28 arêtes selon la valeur du paramètre p de Watts et Strogatz.

est reconnectée à un noeud pris au hasard dans le réseau, les doublons n'étant pas autorisés. On reprend la même routine pour les arêtes qui lient chaque noeud à son deuxième plus proche voisin dans le sens des aiguilles d'une montre. On recommence jusqu'à ce que toutes les connexions avec les voisins dans le sens anti-trigonométrique aient été testées.

Ce modèle permet l'interpolation entre les réseaux réguliers pour lesquels p = 0 et les graphes aléatoires, pour lesquels p = 1. Lorsque p augmente, la distance moyenne entre deux noeuds diminue très vite, tandis que le coefficient de regroupement reste de l'ordre de 1. À  $p \approx 0, 1$ ,  $L(p)/L(0) \approx 0, 1$  et  $C(p)/C(0) \approx 0.6$ . Ce phénomène est expliqué par Watts [20] et Pandit et Amritkar [21] par l'apparition de raccourcis entre les noeuds. De nombreuses études de ce modèle ont suivi, en particulier sur les effets de la taille du réseau sur la valeur de transition, parmi lesquelles on peut citer Barthélemy et Amaral [22], Barrat [23], Newman et Watts [24], Argollo de Menezes *et al.* [25] ou Barrat et Weigt [26]

Cet article de Watts et Strogatz, mettant en regard des observations empiriques et un modèle simple qui en explique les caractéristiques, a marqué le début du vaste intérêt qu'ont suscités les réseaux et les systèmes complexes.

# **I.3** Observations empiriques

Les premiers articles présentant une modélisation en réseaux de systèmes réels complexes et qui en étudient les propriétés, portaient sur l'étude des citations scientifiques [27], bien connues des intéressés. Des études comme celle de Travers et Milgram [17] ont porté sur le réseau de connaissances, qui montre des propriétés de petit-monde, et de Sola Pool et Kochen [28] se sont intéressés à la topologie et aux strucures récurrentes des réseaux de contacts sociaux. Le développement d'Internet, tant en taille du système qu'en nombre d'utilisateurs, a, en grande partie, réactivé l'intérêt des recherches sur l'organisation des réseaux complexes. Il a également été contemporain de l'accroissement de la taille des bases de données, qui permet l'étude et la visualisation de systèmes de grande taille. Ce dernier phénomène a permis de mettre en évidence la fréquence des distributions en loi de puissance dans des réseaux de multiples origines [29].

#### I.3.1 Graphes d'Erdös-Rényi

Le premier modèle de graphe aléatoire, d'Erdös-Rényi [8, 9], décrit un graphe de N sommets connectés par n arêtes choisies au hasard parmi les N(N-1)/2 possibles. Le nombre de voisins, *i.e.* le degré, d'un sommet pris au hasard est donné dans ce modèle par une loi de Poisson. Pour de grands réseaux, le degré d'un noeud est typiquement de l'ordre du degré moyen. C'est le plus simple des modèles intégrant une composante stochastique dans sa construction, dans lequel les noeuds sont statistiquement homogènes. On le trouve donc comme première approximation ou comme référence dans nombre de modélisations de systèmes complexes. Les études de résistance de réseaux comparent ainsi souvent le même processus dynamique sur des graphes aléatoires classiques d'Erdös-Rényi et sur des graphes plus complexes. Voir par exemple l'article d'Albert, Jeong et Barabási [30] sur la résistance à des attaques aléatoires ou ciblées de réseaux, ou celui de Watts [20] sur la propagation en cascade dans les réseaux.

Les réseaux réels ont cependant un degré distribué selon des lois plus complexes, dont nous présentons ici un rapide aperçu.

#### I.3.2 Graphes exponentiels

La distribution du degré d'un réseau en croissance dans lequel les nouveaux noeuds se connectent au hasard à l'un des noeuds déjà présent suit une loi de probabilité exponentielle. Cette distribution du degré correspond donc au modèle de graphe en croissance le plus simple possible.

Les réseaux de transmission d'électricité présentent typiquement une distribution exponentielle du nombre de connexions par élément, comme l'ont étudié Solé et al. [3] pour les réseaux électriques italien, anglais et d'Europe continentale centre-ouest. La distribution cumulative du degré de ces différents réseaux électriques, reportée sur la figure I.11, montre une distribution exponentielle dont le degré moyen  $\langle k \rangle$  est de l'ordre de 2,70 [3]. Le réseau nord-américain suit également une telle distribution du nombre de connexion par sous-station [31], et son degré moyen est plus proche de 2 [32].

#### I.3.3 Réseaux invariants d'échelle

La grande majorité des systèmes étudiés a montré que la distribution du degré des éléments suit une loi de puissance dont l'exposant  $\gamma$  est supérieur à 2. Cette propriété peut révéler un phénomène d'auto-organisation des éléments au sein du réseau [12]. Barabási et Albert ont montré [33] que la loi de puissance du degré émerge naturellement dans les réseaux en croissance avec attachement préférentiel. Nous détaillons ce modèle dans la partie suivante. Nous présentons ici quelques exemples de réseaux réels invariants d'échelle.

#### a Réseaux scientifiques

Le premier système dont l'étude par de Solla Price [27] a montré une distribution invariante d'échelle du degré est le réseau des citations scientifiques. On observe en effet dans la littérature



FIGURE I.11 – Distribution du degré cumulative des réseaux électrique européens, pour un système composé de 3000 générateurs et sous-stations électriques. Emprunt à Solé, Rosas-Casals, Corominas-Murtra, Valverde [3].



FIGURE I.12 – Distribution du nombre de collaborateurs par auteur, pour différentes bases de données regroupant les publications scientifiques. Emprunt à Newman [34].

scientifique que les papiers fondateurs d'un domaine sont cités quelques milliers de fois, tandis que de nombreux articles ne sont cités que par des spécialistes d'une branche spécifique et d'autres, plus nombreux encore, ne sont jamais cités.

Dans le même domaine, Newman a étudié en 2001 [34, 35] les réseaux de collaborations entre scientifiques, étant entendu que des chercheurs ont collaboré s'ils ont écrit un article ensemble. Il a observé que le nombre d'articles par auteur, le nombre d'auteurs par article et le nombre de collaborateurs par auteur sont distribués en lois de puissances, avec des corrections qui peuvent être dues à des effets de taille finie. On présente sur la figure I.12 les distributions du nombre de collaborateurs par auteur issues de différentes bases de données. Barabási *et al.* expliquent ce comportement par les règles de construction du réseau de collaborations [36].

#### b World Wide Web

Un grand nombre d'articles parus sur les réseaux complexes concernent les propriétés du World Wide Web. Les données sur ce système sont en effet facilement accessibles, actuelles et nombreuses. C'est de plus le réseau artificiel le plus étendu, le nombre de pages référencées par Google aujourd'hui est de l'ordre de 45 milliards<sup>c</sup>. Les noeuds de ce réseau sont les pages web, reliées entre elles par des hyperliens. Le degré d'une page est donné par le nombre d'autres pages qui redirigent l'internaute vers elle. Les données sont en général obtenues par des programmes qui naviguent aléatoirement d'une page à la suivante, mais restent parcellaires du fait de la vitesse d'évolution du World Wide Web. Parmi les études menées sur le sujet, on peut citer les travaux d'Albert, Barabási et Jeong [1], de Kumar *et al.* [37], ainsi que les travaux de Broder

c. données fournies par le site www.worldwidewebsize.com



FIGURE I.13 – Distribution du nombre d'hyperliens entrants (symboles vides) et sortants (symboles pleins), pour les pages de Wikipédia en anglais (cercles) et en portugais (triangles). Les lignes correspondent à des simulations des auteurs. Emprunt à Capocci *et al.* [40].

*et al.* [38] et de Adamic et Huberman [39], et surtout l'article dans lequel Barabási et Albert introduisent le modèle d'attachement préférentiel [33]. Comme exemple de la distribution du degré typique d'une page web, nous présentons les résultats plus récents de Capocci *et al.* sur la connectivité des pages de l'encyclopédie Wikipédia [40], figure I.13.

#### c Internet

La structure physique qui contient le réseau virtuel du World Wide Web est un réseau informatique qui utilise un protocole de transfert de données. On peut étudier la structure d'Internet au niveau des routeurs ou au niveau des systèmes autonomes. Dans les deux cas, le degré de ce réseau est distribué selon une loi de puissance, comme vérifié par Faloutsos *et al.* [41] ou Govindan et Tangmunarunkit [42]. Jeong, Néda et Barabási [43] et Pastor-Satorras *et al.* [44] ont montré que la distribution du degré est de plus linéaire :  $P(k) \propto k$ . On présente figure I.14 une carte partielle de la structure d'Internet à un instant donné. Internet étant composé d'un grand nombre de réseaux locaux mis en commun, on s'attend à ce que la distribution de son degré suive une loi de puissance. On représente la distribution du degré de ce réseau, tel que mesuré par Chang *et al.* [45] sur la figure I.15.



FIGURE I.14 – Carte partielle d'Internet, basée sur les données du 15 juin 2005 situées à opte.org. Les éléments sont les adresses IP et les arêtes les connexions entre deux adresses. Les arêtes sont colorées suivant l'origine géographique de l'adresse IP. CC The Opte Project



FIGURE I.15 – Distribution du degré des éléments d'internet, au niveau des systèmes autonomes. Emprunt à Chang *et al.* [45].

#### d Réseaux de transports

Les réseaux de transport, et en particulier les lignes aériennes, sont très largement caractérisés par une structure dominée par de grands centres qui drainent et redistribuent les flux de voyageurs. Ces systèmes sont en général représentés par des réseaux pondérés, exprimant le nombre de voyageurs fréquantant les différentes lignes. On observe dans ces réseaux une distribution du poids des arêtes en loi de puissance, comme l'ont montré Barrat *et al.* [46].

Les exemples de réseaux invariants d'échelle répertoriés ci-dessus montrent des distributions du degré en loi de puissance dont l'exposant est généralement compris entre 2 et 3. Certaines distributions n'ont pas le même exposant pour les noeuds de faible degré et les noeuds de plus grand degré. Cet effet est souvent attribué à la taille finie des réseaux réels qui coupe la queue des distributions [31, 47, 34, 35].

## 1.4 Organisation de la thèse

Nous allons étudier en détail dans la première partie de ce travail les effets de taille finie dans les réseaux invariants d'échelle, après une courte introduction sur le modèle d'attachement préférentiel de Barabási et Albert et quelques variantes. Nous nous intéresserons ensuite à la statistique des extrêmes dans ces réseaux en croissance, modifiée autant que possible par les résultats de temps fini du premier chapitre.

Dans la seconde partie, nous étudierons le modèle de Lehmann et Bernasconi qui décrit un

réseau de défaillances, et sa propagation, analogue à la percolation d'un sous-réseau dans un graphe aléatoire.

# Partie B

# Propriétés statistiques des réseaux en croissance
# CHAPITRE |

# État de l'art sur les réseaux invariants d'échelle

Les réseaux invariants d'échelle apparaissent largement dans la littérature décrivant les phénomènes complexes par le biais des réseaux. De nombreux modèles ont été proposés pour décrire les propriétés des réseaux rééls, et en particulier cette organisation auto-similaire caractérisée par des distributions de probabilités en loi de puissance. Les premiers modèles, comme celui d'Erdös-Rényi ou de Watts-Strogatz, étudient des réseaux dont la taille est une constante du problème. Barabási et Albert remarquent que la majorité des réseaux réels sont des systèmes ouverts qui se sont construits par adjonctions successives d'éléments. Ils font également l'observation, à la suite de Redner [5] que la popularité est attractive. Ils en déduisent un modèle simple dans lequel l'invariance d'échelle apparaît naturellement, et que nous détaillons dans cette partie. Nous présentons également des modèles qui généralisent le propos de Barabási et Albert.

# II.1 Modèle de Barabási-Albert

Barabási et Albert ont proposé en 1999 [33] un modèle dans lequel la distribution du degré en loi de puissance émerge naturellement de deux propriétés :

- la croissance du réseau,
- l'attachement préférentiel des nouveaux noeuds aux noeuds les plus connectés.

Le réseau se construit par l'ajout successif de nouveaux noeuds tout au long de son histoire. À l'instant initial t = 0, le réseau est composé d'un petit nombre  $m_0$  de noeuds. Pour chaque intervalle de temps, on ajoute un nouveau noeud qui se connecte par m liens au reste du réseau. Le nouveau noeud se lie à m noeuds pré-existants distincts. Le nombre m de liens avec lequel un nouveau noeud arrive dans le système est inférieur à la taille du noyau initial  $m \leq m_0$ .

L'attachement est défini comme préférentiel car la probabilité d'attachement d'un nouveau noeud au noeud i de degré  $k_i$  dépend du degré  $k_i$  comme :

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}.$$
(II.1)

Après t unités de temps, le réseau compte donc  $t + m_0$  noeuds et mt liens. Le taux d'acquisition d'un nouveau lien par le noeud i est donné par :

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_j k_j},\tag{II.2}$$

où la somme  $\sum_j k_j = 2mt$  est le nombre total d'extrémités de liens dans le réseau. L'équation différentielle pour le degré  $k_i(t)$  du noeud *i* au temps *t* est donc :

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t},\tag{II.3}$$

Et le degré typique  $k_i(t)$  du noeud labellisé *i*, au temps *t* est donc :

$$k_i(t) = m\sqrt{\frac{t}{t_i}},\tag{II.4}$$

où  $t_i$  est le temps d'apparition du noeud i et m le degré initial du noeud i, qui entre dans le système porteur de m liens.

La distribution cumulative  $\mathcal{P}(k_i(t) < k)$  du degré d'un noeud *i* au temps *t* est donc donnée par :

$$\mathcal{P}(k_i(t) < k) = \mathcal{P}\left(t_i > m^2 \frac{t}{k^2}\right) \tag{II.5}$$

$$= 1 - \mathcal{P}\left(t_i \le m^2 \frac{t}{k^2}\right) \tag{II.6}$$

$$= 1 - \frac{m^2}{t + m_0} \frac{t}{k^2},\tag{II.7}$$

où la probabilité de tirer un noeud au hasard parmi  $t + m_0$  et que ce noeud soit apparu avant le temps  $m^2 \frac{t}{k^2}$  est  $\frac{m^2}{t+m_0} \frac{t}{k^2}$ . La distribution du degré d'un noeud quelconque du réseau est donc donnée par :

$$P_{k(t)}(k) = \frac{\partial \mathcal{P}(k_i(t) < k)}{\partial k} = 2\frac{m^2}{t + m_0}\frac{t}{k^3},\tag{II.8}$$

et pour des temps longs  $t \gg 1$ , et en particulier significativement plus longs que la taille du noyau initial  $t \gg m_0$ , la distribution du degré se comporte en loi de puissance :

$$P(k) \underset{t\gg1}{\to} \frac{2m^2}{k^3} \propto k^{-3}, \tag{II.9}$$

avec un exposant  $\gamma = 3$ , indépendant du nombre *m* de nouveaux liens introduits par unité de temps.

Ce modèle a également été résolu par le biais d'équations maîtresse [48, 49] ou d'équations cinétiques [50].

Bien que ce modèle ne reproduise pas tout l'intervalle des valeurs possibles de l'exposant mesuré dans les réseaux réels, l'exposant  $\gamma = 3$  se situe dans cet intervalle. De plus, le comportement en loi de puissance émerge naturellement des deux simples propriétés de croissance et d'attachement préférentiel.

Pour vérifier que ces deux propriétés sont nécessaires simultanément, on considère un modèle simplifié d'un réseau en croissance dans lequel la probabilité d'attachement est uniforme, et donc non-préférentielle,  $\Pi(k) = a$ , où a est une constante quelconque, identique pour tous les noeuds du réseau. La distribution du degré décroit alors exponentiellement vite :  $P(k) \sim e^{-\beta k}$ .

Respectivement, on considère un modèle avec attachement préférentiel dans un réseau composé d'un nombre constant N de noeuds. Au temps initial, ce réseau ne contient aucun lien. À chaque intervalle de temps, on sélectionne aléatoirement un noeud i et on lui ajoute un lien avec probabilité  $\Pi(k_i) = k_i / \sum_j k_j$ . Au cours des premières étapes, ce modèle correspond au modèle de Barabási-Albert, car les noeuds non encore connectés peuvent être vus comme un réservoir des noeuds à ajouter dans le réseau. Dans les premiers temps, ce modèle présente donc une distribution du degré en loi de puissance. Ce comportement n'est cependant pas stationnaire, car dès que tous les noeuds sont connectés, la distribution du degré explose.

Les deux propriétés de croissance et d'attachement préférentiel sont donc nécessaires à l'émergence d'un comportement invariant d'échelle, mais ce modèle simple ne permet cependant pas de couvrir la gamme des valeurs possibles de l'exposant dans les réseaux réels. On considère donc dans la suite quelques modifications possibles du modèle qui influent sur la valeur de l'exposant de la distribution.

# II.2 Attractivité initiale des sites

Dorogovtsev, Mendes et Samukhin [48] proposent de modifier le modèle originel de Barabási-Albert en ajoutant aux sites une attractivité initiale C constante. Dans ce modèle, le degré  $k_i$ d'un noeud *i* correspond au nombre de connexions entrantes. La probabilité d'attachement  $\Pi(k_i)$ d'un nouveau lien au site *i* de degré  $k_i$  est proportionnelle à :

$$\Pi(k_i) \propto C + k_i, \tag{II.10}$$

où  $C \ge 0$  est l'attractivité initiale de chaque noeud. On s'intéresse à la distribution du nombre de liens pointant vers un noeud. Dans le modèle de Dorogovtsev et al., à chaque intervalle de temps, un nouveau noeud et m connexions dirigées apparaissent dans le réseau. Ces connexions sont issues d'un noeud quelconque, et pointent vers le noeud i avec probabilité :

$$\Pi(k_i) = \frac{C+k_i}{\sum_j C+k_i} = \frac{C+k_i}{(C+m)t}.$$
(II.11)

La probabilité  $p_k(k_i(t))$  que k des m nouvelles connexions se connectent pointant vers le noeud i de degré  $k_i(t)$  à l'unité de temps t est donné par la probabilité binomiale :

$$p_k(k_i(t)) = \binom{m}{k} \left( \Pi(k_i(t)) \right)^k \left( 1 - \Pi(k_i(t)) \right)^{m-k}.$$
 (II.12)

La probabilité  $P(k_i, t+1)$  qu'un site *i* soit de degré  $k_i$  au temps t+1 peut être obtenue à partir de cette même probabilité au temps précédent t:

$$P(k_i, t+1) = \sum_{k=0}^{m} p_k(k_i - k) P(k_i - k, t),$$
(II.13)

et la distribution du degré pour un noeud quelconque  $P(k,t) = \sum_{i=1}^{t} P(k_i,t)/t$  au temps t vérifie la relation asymptotique à la limite  $t \to \infty$ :

$$\left(1 + \frac{C}{m}\right)P(k) + (k+C)P(k) - (k-1+C)P(k-1) = \left(1 + \frac{C}{m}\right)\delta(k),$$
 (II.14)

où le terme  $\delta(k)$  correspond aux nouveaux noeuds arrivant dans le système avec degré 0. La solution de cette équation se comporte, pour  $k + C \gg 1$  en :

$$P(k) \approx \left(1 + \frac{C}{m}\right) \frac{\Gamma\left(1 + C + \frac{C}{m}\right)}{\Gamma(C)} (k + C)^{-\left(2 + \frac{C}{m}\right)},\tag{II.15}$$

avec une loi de puissance d'exposant  $\gamma = 2 + \frac{C}{m} \ge 2$ .

Dorogovtsev et al. ont également déterminé le degré typique  $\langle k_i \rangle$  (t) d'un noeud né au temps  $t_i$  au temps t :

$$\langle k_i(t) \rangle = C\left(\left(\frac{t_i}{t}\right)^{-\frac{m}{m+C}} - 1\right),$$
 (II.16)

ce qui donne, pour des temps grand devant le temps de naissance du noeud  $t \gg t_i$ :

$$\langle k_i(t) \rangle \sim t_i^{-\nu}, \ \nu = \frac{m}{m+C},$$
 (II.17)

et les deux exposants  $\gamma$  et  $\nu$  des lois de puissance de la distribution du degré et du degré typique d'un noeud né au temps  $t_i$  vérifient la relation :

$$\nu(\gamma - 1) = 1. \tag{II.18}$$

Ce modèle se ramène au modèle de Barabási-Albert pour une attractivité initiale C = m, qui est équivalente à un noeud arrivant dans le réseau avec m connexions, et on retrouve alors les exposants du modèle :  $\gamma = 3$  et  $\nu = \frac{1}{2}$ . Les deux cas limites de ce modèle sont le cas C = 0, dans lequel seul le premier noeud du réseau peut recevoir de nouvelles connexions, et  $C \to \infty$  dans lequel tous les sites ont même probabilité d'attachement, et pour lequel l'invariance disparaît :  $\gamma \to \infty$  et  $\nu = 0$ .

Ce modèle présente une distribution du degré en loi de puissance, dont l'exposant  $\gamma$  varie entre 2 et l'infini. Par une généralisation simple du modèle de Barabási-Albert, il permet donc d'obtenir un exposant qui couvre le spectre des mesures dans les réseaux réels invariants d'échelle.

# II.3 Modèle avec redirection

Krapivsky et Redner ont proposé un modèle [11] inspiré d'une proposition d'organisation du World Wide Web par un processus de copie entre les liens des différentes pages web [51]. Dans ce modèle de réseau en croissance, un nouveau noeud apparaît à chaque unité de temps, et s'attache à l'un des noeuds déjà existant avec une probabilité uniforme sur tous les noeuds du réseau. Les connexions sont dirigées, sortant du nouveau noeud et pointant vers le noeud préexistant. Avec probabilité r, la connexion glisse le long de la connexion sortante du noeud initialement choisi, vers celui auquel il est connecté. Avec probabilité 1 - r, il reste connecté au noeud initialement choisi. La distribution f(k) du degré k d'un noeud pris au hasard dans le réseau au temps n, et donc de taille N = n, est donnée par l'équation maîtresse :

$$\frac{\mathrm{d}f(n,k)}{\mathrm{d}n} = \frac{1-r}{n-1}f(n-1,k-1) - \frac{1-r}{n-1}f(n-1,k) + r\frac{k-2}{n-1}f(n-1,k-1) - r\frac{k-1}{n-1}f(n-1,k) + \delta(k,1),$$
(II.19)

où  $\delta(k, 1)$  correspond au degré initial du noeud *i* arrivant dans le réseau au temps *i*. On peut réécrire cette équation sous une forme proche de celle obtenue pour le modèle de Barabási-Albert :

$$(n-1)\frac{\mathrm{d}f(n,k)}{\mathrm{d}n} = r\left((k-2) + \frac{1-r}{r}\right)f(n-1,k-1) - r\left((k-1) + \frac{1-r}{r}\right)f(n-1,k),$$
(II.20)

et on identifie alors la probabilité d'attachement  $\Pi(k)$  en fonction de la probabilité r de redirection :

$$\Pi(k) = r\left(k - 1 + \frac{1 - r}{r}\right).$$
(II.21)

On retrouve les probabilités d'attachement préalablement connues pour :

- r = 0, il n'y a jamais de redirection, alors  $\Pi(k) = cste$  correspond à la probabilité d'attachement uniforme,
- r = 1/2, alors  $\Pi(k) \propto k$  et on retrouve le modèle de Barabási-Albert,
- r = 1, alors  $\Pi(k) \propto k 1$ , et, aucun des noeuds à part le premier, ne peut recevoir de nouveau lien,
- pour r quelconque,  $\Pi(k) \propto k+C$ , avec  $C = \frac{1}{r} 1$  l'attractivité initiale d'un site, introduite dans la section ci-dessus.

Ce modèle présente un intérêt algorithmique pour la simulation des modèles d'attachement préférentiel. En effet, un tirage uniforme puis un test de redirection sont plus rapide qu'un tirage préférentiel. On peut, de plus, faire varier la valeur de l'attachement préférentiel sur son intervalle de définition en variant la probabilité de redirection r. Ce modèle présente donc également une distribution du degré en loi de puissance, dont l'exposant  $\gamma$  couvre les valeurs mesurées dans les réseaux réels,  $\gamma > 2$ . Il est mathématiquement équivalent au modèle de construction par copie de Kumar *et al.* [37].

Nous présentons dans le chapitre suivant une étude méthodique des effets de taille finie dans les modèles décrits ici.

# CHAPITRE **[]]**

# Effets de taille finie dans les réseaux en croissance

Les réseaux invariants d'échelle sont caractérisés par une distribution du degré qui se comporte en loi de puissance, pour les noeuds de degré suffisament grand. La distribution du degré k suit, dans un réseau de taille infinie, la loi :

$$f(k) \sim k^{-\gamma},$$
 (III.1)

dite stationnaire. Dans un réseau de taille finie N, un noeud ne peut avoir un degré k > N. On observe, de fait, une coupure naturelle dans la distribution du degré des réseaux réels, de taille finie. Dans un système invariant d'échelle, caractérisé par des distributions en loi de puissance, les effets de taille finie entraînent des corrections importantes sur les distributions statistiques. Il existe un degré critique  $k_{\star}$  qui définit trois régimes de la distribution du degré :

- stationnaire si le degré est assez petit devant  $k_{\star}$
- exponentiellement décroissant pour un degré significativement plus grand que  $k_{\star}$
- critique à la transition entre ces deux régimes.

Le degré de coupure  $k_{\star}$  représente le degré à partir duquel un noeud commence à sentir la taille finie du réseau.

Dans ce chapitre, nous nous intéresserons à la définition du degré de coupure, et à la description de chacun des trois régimes pour différents modèles de réseaux en croissance. Dans un premier temps, on posera un certain nombre de définitions et de techniques dans le cadre du modèle le plus simple : le modèle avec attachement uniforme, dans lequel un noeud entrant s'attache avec égale probabilité à l'un des noeuds déjà présent. Nous considérerons ensuite les effets de taille finie qui apparaissent dans le modèle d'attachement préférentiel de Barabási-Albert [33], puis nous concluerons par le modèle d'attachement préférentiel généralisé, introduit par Dorogovtsev, Mendes et Samukhin [48] comme généralisation du modèle de Barabási-Albert.

# III.1 Description des modèles

On considère dans ce chapitre des réseaux en croissance, c'est-à-dire des réseaux dont le nombre de noeuds augmente avec le temps. Dans les modèles étudiés ici, un nouveau noeud apparaît à chaque unité de temps n > 0, que l'on étiquette par sa date de naissance n. Ce nouveau noeud n s'attache à l'un des noeuds  $i = 1, \ldots, n-1$  déjà présents, avec une probabilité  $p_{n,i}$  qui dépend du noeud-cible *i*, et du modèle considéré. Le modèle est entièrement décrit par la règle d'attachement des nouveaux noeuds et les conditions initiales du système.

# **III.1.1** Conditions initiales

On cherche à déterminer dans notre étude comment les différentes quantités du système dépendent des conditions initiales. On considère les conditions initiales les plus simples, auxquelles peuvent se ramener par un changement d'échelle tous les autres choix. Les deux conditions initiales que nous présentons ici diffèrent par le choix de la symétrie initiale.

La condition initiale (A) consiste en un premier noeud qui apparaît au temps n = 1 sans connexion, c'est-à-dire avec degré  $k_1(1) = 0$ . Tous les noeuds successifs *i* apparaissent avec degré  $k_i(i) = 1$ . Le second noeud qui apparaît au temps n = 2 ne peut se connecter qu'au seul noeud déjà présent, et les degrés des deux noeuds valent alors :  $k_1(2) = k_2(2) = 1$ . On parle de cette configuration comme de la configuration de dimère, utilisée dans les travaux de Krapivsky et Redner [11, 52]. Au temps suivant n = 3, les deux premiers noeuds sont donc, dans tous les modèles considérés, équivalents, et ont même probabilité de recevoir la connexion du noeud n = 3 entrant. L'emsemble des noeuds des réseaux issus de cette configuration ont un nombre total de degrés :

$$K^{(A)}(n) = \sum_{i} k_i(n) = 2n - 2$$
 (III.2)

au temps n.

Pour la condition initiale (B), le premier noeud apparaît au temps n = 1 porteur d'une connexion qui ne se lie à aucun autre noeud, mais qui peut être interprétée comme une racine de l'arbre formé par le réseau. Le premier noeud a donc degré  $k_1(1) = 1$  au temps initial. Le second noeud qui apparaît au temps n = 2 porteur d'une seule connexion se lie au noeud déjà présent, et les degrés des noeuds au temps n = 2 sont :  $k_1(2) = 2$  et  $k_2(2) = 1$ . Ce réseau se construit en arbre enraciné, et la somme des degrés de tous les noeuds à un temps n quelconque vérifie :

$$K^{(B)}(n) = \sum_{i} k_i(n) = 2n - 1.$$
 (III.3)

On représente sur la figure III.1 les trois premières étapes de construction du réseau pour chaque condition initiale. À l'étape n = 3 le nouveau noeud entrant dans le système peut se connecter à un seul des noeuds présents i, avec une probabilité  $p_{3,i}$  qui dépend du modèle considéré et des conditions initiales.

On distinguera dans la suite les quantités X spécifiquement mesurées dans les réseaux nés d'une condition initiale particulière par la notation du cas considéré en exposant :  $X^{(A)}$ , respectivement  $X^{(B)}$ .

## III.1.2 Modèle d'attachement uniforme

Dans le modèle avec attachement uniforme, la probabilité que le nouveau noeud entrant dans le système au temps n se connecte à l'un des noeuds i = 1, ..., n - 1 ne dépend que du nombre n-1 de noeuds dejà présents. Tous les noeuds ont donc même probabilité de recevoir une nouvelle connexion à chaque instant et le degré d'un noeud i au temps n ne dépend typiquement que du rapport entre les temps de naissance i et d'observation n.



FIGURE III.1 – Les trois premières étapes de construction du réseau pour les deux conditions initiales (A) et (B).

# III.1.3 Modèle de Barabási-Albert

Ce modèle de réseau en croissance présente naturellement une distribution stationnaire du degré en loi de puissance, dont l'émergence est rappelée dans le chapitre précédent II.1. Il repose sur l'attachement préférentiel des nouveaux noeuds aux noeuds les plus connectés. La probabilité que le noeud  $i = 1, \ldots, n-1$  reçoive la connexion du nouvel élément n qui entre dans le système au temps n dépend du degré  $k_i(n-1)$  de ce noeud i à la fin de l'étape précédente. La probabilité qu'un noeud i soit de degré  $k_i(n)$  dépend donc de toute son histoire, et, en particulier, la date de naissance i du noeud doit avoir un fort impact sur le degré  $k_i(n)$  considéré.

# III.1.4 Modèle d'attachement préférentiel généralisé

Ce modèle est une extension du modèle de Barabási-Albert dans lequel les nouveaux noeuds entrant dans le système avec degré  $k_i(i) = 1$  ont une attractivité initiale additionnelle c qui tempère ou amplifie l'effet de l'attachement préférentiel. En particulier, lorsque  $c \to \infty$ , tous les noeuds ont même attractivité, quelque soit leur degré, et on retrouve les propriétés du modèle avec attachement uniforme, tandis que pour c = 0, on retrouve le modèle de Barabási-Albert.

# III.2 Modèle d'attachement uniforme

Le modèle d'attachement uniforme est le modèle le plus simple de graphe en croissance. Un noeud entre dans le système avec une arête dont le bout libre s'attache à l'un des noeuds déjà présents, choisi uniformément.

## III.2.1 Distribution du degré pour un noeud né au temps i

#### a Définition du degré

On étiquette chaque noeud par sa date d'entrée dans le réseau. Un noeud né au temps n se connecte au moment de son apparition dans le réseau à un noeud i parmi les n-1 noeuds déjà présents avec probabilité :

$$p_{n,i} = \frac{1}{n-1}.$$
 (III.4)

Le degré  $k_i$  du noeud *i* au moment de son apparition est donc :  $k_i(i) = 1$ . À chaque étape *n* subséquente, il peut recevoir une connexion du nouveau noeud entrant dans le système, avec la même probabilité pour les n - 1 noeuds.

Par analogie avec les modèles de marcheurs aléatoires, on peut définir l'incrément du degré d'un noeud i à chaque étape n > i par :

$$I_i(n) = k_i(n) - k_i(n-1) = \begin{cases} 1 & \text{avec probabilité } p_{n,i} \\ 0 & \text{sinon.} \end{cases}$$
(III.5)

Sous cette analogie, le réseau peut être vu comme un ensemble de taille croissante de marcheurs aléatoires 1D totalement biaisés, en ce qu'ils ne peuvent que rester à coordonnée constante ou augmenter leur coordonnée d'une unité par unité de temps, ou comme un processus de condensation d'un gaz sur un média dont la taille croît sans discontinuer.

Le degré  $k_i(n)$  du noeud *i* pour un temps postérieur à sa date d'apparition *i* est donc :

$$k_i(n) = k_i(i) + \sum_{j=i+1}^n I_i(j),$$
 (III.6)

où le degré du noeud au moment de son apparition  $k_i(i)$  est 1 pour i > 1. Le degré initial du noeud 1 dépend de la condition initiale :

$$k_1^{(A)}(1) = 0,$$
  
 $k_1^{(B)}(1) = 1.$ 
(III.7)

La définition bimodale (III.5) de l'incrément  $I_i(n)$  du degré du noeud i au temps n permet d'écrire directement le degré moyen  $\langle k_i(n) \rangle$  au temps n d'un noeud né au temps i:

$$\langle k_i(n) \rangle = 1 + \sum_{j=i+1}^{n} \frac{1}{j-1} = H_{n-1} - H_{i-1} + 1,$$
 (III.8)

où  $H_n$  est le nombre harmonique défini par :

$$H_n = \sum_{i=1}^n \frac{1}{i},\tag{III.9}$$

dont la limite continue donne :

$$H_n \approx \int_1^n \frac{\mathrm{d}i}{i} = \ln n. \tag{III.10}$$

Le comportement logarithmique du nombre harmonique de premier ordre pour  $n \gg 1$  sera détaillé plus bas (III.17).

Le degré moyen du noeud i au temps n peut donc, pour n et i grands, s'approximer par :

$$\langle k_i(n) \rangle \approx \ln \frac{n}{i} + 1.$$
 (III.11)

La distribution de probabilité du degré  $k_i(n)$  au temps n est notée  $f_i(n,k) = P(k_i(n) = k)$ . Les conditions aux limites sont données par :

$$f_i(i,k) = \delta_{k,1} \tag{III.12}$$

$$f_i(n,k) = \delta_{k,0}, \ n < i, \tag{III.13}$$

et assurent que le noeud i arrive dans le système au temps i avec degré 1. Les conditions aux limites de la distribution du degré sont différentes pour le noeud i = 1 selon les conditions initiales choisies pour le système. On distinguera par la suite les fonctions correspondant à des réseaux issus respectivement des conditions initiales (A) et (B) par la notation correspondante en exposant. Les conditions aux limites de la distribution du degré du noeud i = 1 dépendent de la condition initiale :

$$f_1^{(A)}(1,k) = \delta_{k,0}$$
(III.14)  
$$f_1^{(B)}(1,k) = \delta_{k,1}.$$

Dans le cas de l'attachement uniforme, on pourrait calculer la distribution du degré d'un noeud né à l'étape i par un raisonnement combinatoire, mais on va ici introduire le cadre technique qui sera nécessaire au calcul de la distribution pour les autres modèles considérés.

#### b Nombres harmoniques

Les nombres harmoniques de premier ordre  $H_n$  sont définis à partir de la fonction digamma pour un argument n entier :

$$\psi_0(n) = \frac{\Gamma'(n)}{\Gamma(n)} = H_{n-1} - \gamma, \qquad (\text{III.15})$$

et le développement asymptotique de la fonction digamma est donné par :

$$\psi_0(z+1) \underset{z \to \infty}{\approx} \ln z + \frac{1}{2z} - \sum_{i=1}^{\infty} \frac{B_{2i}}{2iz^{2i}} \underset{z \to \infty}{\to} \ln z, \qquad \text{(III.16)}$$

où les  $B_i$  sont les nombres de Bernoulli. La limite asymptotique des nombres harmoniques  $H_n$  est donc :

$$H_n \approx \ln n - \gamma.$$
 (III.17)

On définit les nombres harmoniques d'ordre supérieur à partir des fonctions polygamma d'ordre m:

$$\psi_m(z) = \left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{m+1} \ln \Gamma(z), \qquad (\text{III.18})$$

dont la relation avec les nombres harmoniques est donnée par :

$$\psi_m(n) = (-1)^{m-1} m! \left( \zeta(m+1) - H_{n-1}^{(m+1)} \right), \qquad \text{(III.19)}$$

où  $\zeta$  est la fonction zeta de Riemann et  $H_n^{(m)}$  les nombres harmoniques d'ordre m:

$$H_n^{(m)} = \sum_{i=1}^n \frac{1}{i^m}.$$
 (III.20)

En particulier, le nombre harmonique de deuxième ordre  $H_n^{(2)}$  peut s'approximer, pour  $n \gg 1$  par :

$$H_n^{(2)} \approx \frac{\pi^2}{6} - \frac{1}{n}.$$
 (III.21)

### c Fonction génératrice

La fonction génératrice  $F_{n,i}$  de la distribution du degré  $k_i(n)$  du noeud *i* au temps *n* est définie par :

$$F_{n,i} = \left\langle x^{k_i(n)} \right\rangle = \sum_{k=1}^n f_i(n,k) x^k.$$
 (III.22)

On peut également exprimer cette fonction génératrice par la définition incrémentale (III.6) du degré du noeud i au temps n, et elle devient alors :

$$F_{n,i} = \left\langle x^{k_i(n)} \right\rangle = \left\langle x^{k_i(i)} x^{\sum_{j=i+1}^n I_i(j)} \right\rangle.$$
(III.23)

L'incrément du degré du noeud i au temps j est, pour le modèle avec attachement uniforme, indépendant de ce même incrément à un autre temps j', et on peut donc écrire la fonction génératrice du degré du noeud i au temps n comme :

$$F_{n,i}(x) = x \prod_{j=i+1}^{n} \left\langle x^{I_i(j)} \right\rangle.$$
(III.24)

La fonction génératrice de l'incrément  $I_i(j)$  du degré du noeud *i* au temps *j* vérifie :

$$\langle x^{I_i(j)} \rangle = p_{i,j} x^1 + (1 - p_{i,j}) x^0 = 1 + (x - 1) p_{i,j} = \frac{x + j - 2}{j - 1},$$
 (III.25)

et ne dépend donc pas de l'identité du noeud i. La dépendance en l'étape de naissance i de la fonction génératrice du degré du noeud i est contenue dans les bornes du produit des fonctions génératrices de l'incrément du degré du noeud i de l'équation (III.24). La fonction génératrice du degré du noeud i s'écrit donc :

$$F_{n,i}(x) = \frac{x(i-1)!\Gamma(x+n-1)}{(n-1)!\Gamma(x+i-1)},$$
(III.26)

pour tout noeud quelconque  $i \ge 2$ . La forme de cette fonction ne diffère du régime général que pour le premier noeud du réseau i = 1, selon la condition initiale considérée. La fonction génératrice initiale du degré du noeud i = 1 est en effet donnée par :

$$F_{1,1}^{(A)}(x) = 1 \tag{III.27}$$

$$F_{1,1}^{(B)}(x) = x, (III.28)$$

et à un temps ultérieur n, la fonction génératrice du degré du noeud i = 1 est définie en fonction de la condition initiale considérée :

$$F_{n,1}^{(A)}(x) = \frac{\Gamma(x+n-1)}{(n-1)!\Gamma(x)}$$
(III.29)

$$F_{n,1}^{(B)}(x) = \frac{x\Gamma(x+n-1)}{(n-1)!\Gamma(x)}.$$
(III.30)

#### d Équation maîtresse

L'équation maîtresse pour la distribution du degré  $k_i(n)$  du noeud *i* au temps *n* s'écrit :

$$\frac{\mathrm{d}f_i(n,k)}{\mathrm{d}n} = p_{n,i}f_i(n-1,k-1) - p_{n,i}f_i(n-1,k), \qquad (\mathrm{III.31})$$

et le temps étant discret dans ce problème, on obtient l'équation de récurrence suivante pour la distribution du degré du noeud i:

$$f_i(n,k) = \frac{1}{n-1} f_i(n-1,k-1) + \left(1 - \frac{1}{n-1}\right) f_i(n-1,k).$$
(III.32)

La fonction génératrice  $F_{n,i}$  du degré du noeud i au temps n vérifie donc elle-même la récurrence :

$$F_{n,i}(x) = \frac{x}{n-1} F_{n-1,i}(x) + \left(1 - \frac{1}{n-1}\right) F_{n-1,i}(x)$$
  
=  $\frac{x+n-2}{n-1} F_{n-1,i}(x),$  (III.33)

dans laquelle on peut reconnaître la fonction génératrice de l'incrément (III.25). Cette relation de récurrence peut en effet également s'obtenir directement à partir de la définition (III.24) de la fonction génératrice  $F_{n,i}$  comme produit des fonctions génératrices de l'incrément à tout

temps j > i. La fonction génératrice du degré du noeud i au temps n - 1 est en effet donnée, selon cette formule, par :

$$F_{n-1,i}(x) = x^{k_i(i)} \prod_{j=i+1}^{n-1} \left\langle x^{I_i(j)} \right\rangle,$$
(III.34)

et la relation de récurrence s'obtient directement :

$$F_{n,i}(x) = \left\langle x^{I_i(n)} \right\rangle F_{n-1,i}(x). \tag{III.35}$$

Cette approche par le biais de l'incrément du degré du noeud i au cours du temps n'est cependant applicable que dans le cas du modèle de croissance par attachement uniforme, dans lequel il ne dépend pas de l'histoire préalable du noeud i considéré, au contraire des modèles avec attachement préférentiel.

## e Moments du degré

On obtient à partir de la formule de la fonction génératrice (III.26) du degré du noeud i tous les moments d'ordre m de la distribution  $f_i(n, k)$  du degré du noeud i au temps n, par le développement en série entière pour  $x \to 1$  de la fonction génératrice :

$$\left\langle x^{k_i(n)} \right\rangle = \underset{x \to 1}{=} 1 + (x-1) \sum_{k=1}^n k f_i(n,k) + \frac{(x-1)^2}{2!} \sum_{k=1}^n k(k-1) f_i(n,k) + \dots$$
 (III.36)

$$= 1 + (x - 1) \langle k_i(n) \rangle + \frac{(x - 1)^2}{2!} \left( \left\langle k_i(n)^2 \right\rangle - \left\langle k_i(n) \right\rangle \right) + \dots$$
(III.37)

$$= 1 + (x - 1)(\psi_0(n) - \psi_0(i) + 1) + \frac{(x - 1)^2}{2!} \left( (\psi_0(n) - \psi_0(i))^2 + 2(\psi_0(n) - \psi_0(i)) + \psi_1(n) - \psi_1(i) \right) + \dots,$$
(III.38)

où  $\psi_i$  est la fonction polygamma, qui peut s'exprimer en fonction des nombres harmoniques (III.19). La valeur moyenne et la variance du degré du noeud *i* sont donc données par les expressions suivantes, indépendantes des conditions initiales :

$$\langle k_i(n) \rangle = H_{n-1} - H_{i-1} + 1,$$
 (III.39)

var 
$$(k_i(n)) = \langle k_i(n)^2 \rangle - \langle k_i(n) \rangle^2 = H_{n-1} - H_{i-1} - H_{n-1}^{(2)} + H_{i-1}^{(2)}.$$
 (III.40)

On retrouve la valeur moyenne (III.8) du degré du noeud i au temps n obtenue directement à partir du produit des moyennes des incréments successifs  $I_i$  du degré  $k_i$ .

#### f Régime d'échelle

On considère le régime d'échelle, pour lequel le temps de naissance i et le temps courant n sont grands et de même ordre, et on définit leur rapport comme :

$$z = \frac{n}{i} \ge 1. \tag{III.41}$$

Dans ce régime, les fonctions  $\Gamma$  qui entrent dans la formule de la fonction génératrice du degré  $k_i(n)$  se simplifient grâce au développement de Stirling :

$$\frac{\Gamma(i)\Gamma(x+n-1)}{\Gamma(n)\Gamma(x+i-1)} \approx z^{1/2} \left(1 + \frac{x-1}{n}\right)^n \left(1 + \frac{x-1}{i}\right)^{-i} \left(\frac{x+n-1}{x+i-1}\right)^{x-3/2}$$
(III.42)

$$\approx z^{x-1},$$
 (III.43)

et la fonction génératrice a, dans le régime d'échelle, la forme suivante :

2

$$F_{n,i}(x) \approx x \mathrm{e}^{(x-1)\ln z}.$$
 (III.44)

On reconnait ici la fonction génératrice d'une distribution de Poisson pour  $k \ge 1$ , de paramètre  $\mu = \ln z$ . La forme d'échelle de la distribution du degré du noeud *i* au temps *n* est donc donnée par la formule de la distribution de Poisson avec un degré minimum  $k_{min} = 1$ :

$$f_i(n,k) \approx \frac{(\ln z)^{k-1} \mathrm{e}^{-\ln z}}{(k-1)!}.$$
 (III.45)

### III.2.2 Distribution du degré pour un noeud quelconque

On considère maintenant le réseau pris dans son ensemble, c'est-à-dire qu'on s'intéresse au degré k(n) d'un noeud que l'on sélectionne au hasard dans le réseau au temps n sans rien connaître de son étape d'apparition. La distribution f du degré d'un noeud quelconque peut être obtenue en moyennant toutes les distributions  $f_i$  des noeuds i sur tous les noeuds  $i \leq n$  du réseau :

$$f(n,k) = \frac{1}{n} \sum_{i=1}^{n} f_i(n,k),$$
 (III.46)

où tous les noeuds peuvent être tirés avec même probabilité. Les conditions initiales, *i.e.* les distributions du degré au temps n = 1, sont respectivement :

$$f^{(A)}(1,k) = \delta_{k,0} \tag{III.47}$$

$$f^{(B)}(1,k) = \delta_{k,1}.$$
 (III.48)

Quelques valeurs particulières peuvent être extraites de considérations combinatoires simples. La probabilité qu'un noeud tiré au hasard soit de degré k(n) = 1 au temps n se calcule à partir de la distribution du degré d'un noeud d'indice i. Les noeuds d'indice i < n ont une probabilité d'être de degré 1 donnée par :

$$P(k_i(n) = 1) = \prod_{j=i}^{n} \frac{(j-1)}{j}.$$
(III.49)

Pour les conditions initiales (A), le noeud d'indice i = 1 peut être de degré 1, et la relation est valable pour  $i \ge 1$ . Pour les conditions initiales (B), le noeud d'indice i = 1 est de degré au moins égal à 2, et la relation est valable pour i > 1. Le noeud d'indice n est de degré 1 avec probabilité 1 quelle que soit la condition intiale. La probabilité de tirer un noeud de degré 1 au hasard dans le réseau est donc donnée par :

$$f^{(A)}(n,1) = \frac{1}{2} + \frac{1}{n(n-1)}, \quad f^{(B)}(n,1) = \frac{1}{2}.$$
 (III.50)

La probabilité qu'un noeud tiré au hasard soit de degré maximum est la probabilité jointe de tirer l'un des noeuds initiaux, et que ce noeud ait attiré toutes les connexions entrant dans le système. Cette probabilité dépend donc des conditions initiales :

$$f^{(A)}(n, n-1) = \frac{2}{n!}, \quad f^{(B)}(n, n) = \frac{1}{n!}.$$
 (III.51)

# a Fonction génératrice

La fonction génératrice  $F_n(x)$  du degré d'un noeud tiré au hasard dans le réseau à un temps fini n peut également s'exprimer en fonction de la fonction génératrice du degré d'un noeud ipar :

$$F_n(x) = \left\langle x^{k(n)} \right\rangle$$
  
=  $\sum_{k=1}^n f(n,k) x^k$   
=  $\frac{1}{n} \sum_{k=1}^n \sum_{i=1}^n f_i(n,k) x^k$   
=  $\frac{1}{n} \sum_{i=1}^n F_{n,i}(x).$  (III.52)

La fonction génératrice  $F_{n,i}$  du degré d'un noeud *i* obéit à la relation de récurrence définie en (III.33), et la fonction génératrice  $F_n$  du degré vérifie donc la relation de récurrence suivante :

$$\frac{1}{n}\sum_{i=1}^{n}F_{n,i}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{x+n-2}{n-1}F_{n-1,i}(x)$$
(III.53)

$$\Leftrightarrow nF_n(x) = (x+n-2)F_{n-1}(x) + \frac{x+n-2}{n-1}F_{n-1,n}(x)$$
(III.54)

$$\Leftrightarrow nF_n(x) = (x+n-2)F_{n-1}(x) + x. \tag{III.55}$$

Cette relation de récurrence peut aussi être obtenue à partir de la relation de récurrence de la distribution f(n,k) du degré.

# b Équation maîtresse

La distribution du degré d'un noeud quelconque obéit à l'équation maîtresse suivante :

$$\frac{\mathrm{d}f(n,k)}{\mathrm{d}n} = \frac{1}{n} \Big( f(n-1,k-1) - 2f(n-1,k) + \delta_{k,1} \Big), \tag{III.56}$$

qui exprime le fait qu'à chaque instant, deux unités de degré sont ajoutées au réseau, et qu'un nouveau noeud de degré k = 1 arrive dans le réseau. Le temps est discret, et l'équation maîtresse peut donc se simplifier en :

$$nf(n,k) = f(n-1,k-1) + (n-2)f(n-1,k) + \delta_{k,1}.$$
 (III.57)

Cette relation de récurrence peut encore s'obtenir à partir de la relation de récurrence de la distribution du degré du noeud d'indice i, ou à partir de la relation de récurrence (III.55) de la fonction génératrice du degré.

#### c Distribution stationnaire

La solution indépendante du temps de l'équation maîtresse donne la distribution stationnaire du degré :

$$2f_{stat}(k) - f_{stat}(k-1) = \delta_{k,1}$$
 (III.58)

$$\Leftrightarrow f_{stat}(k) = \frac{1}{2^k},\tag{III.59}$$

où  $k \ge 1$ . On reconnaît une distribution exponentielle du degré, en particulier sous la forme :

$$f_{stat}(k) = e^{-k\ln 2}.$$
 (III.60)

La fonction génératrice associée à la distribution stationnaire est donnée par :

$$nF_{stat}(x) - (x+n-2)F_{stat}(x) = x$$
(III.61)

$$\Leftrightarrow F_{stat}(x) = \frac{x}{x-2}.$$
 (III.62)

# d Temps fini

La solution générale de l'équation de récurrence (III.55) apporte une correction de temps fini à la fonction génératrice de la distribution stationnaire. Elle s'obtient par résolution de l'équation de récurrence, dans laquelle la solution particulière est donnée par la fonction génératrice stationnaire  $F_{stat}(x)$ . La solution de l'équation homogène :

$$F_{n+1}(x) = \frac{x+n-1}{n+1}F_n(x)$$
(III.63)

est donnée modulo un facteur de proportionnalité  $\lambda$  :

$$F_n^{(H)}(x) = \lambda \frac{\Gamma(x+n-1)}{n!\Gamma(x)},$$
(III.64)

dont la valeur dépend des conditions initiales du système :

$$F_1^{(A)}(x) = 1 (III.65)$$

$$F_1^{(B)}(x) = x.$$
 (III.66)

La fonction génératrice du degré d'un noeud quelconque du réseau à un temps fini n est donc définie par les formules suivantes pour chacune des deux conditions initiales A et B étudiées :

$$F_n^{(A)}(x) = \frac{x}{2-x} + \frac{2(1-x)}{2-x} \frac{\Gamma(x+n-1)}{n!\Gamma(x)},$$
(III.67)

$$F_n^{(B)}(x) = \frac{x}{2-x} + \frac{x(1-x)}{2-x} \frac{\Gamma(x+n-1)}{n!\Gamma(x)}.$$
 (III.68)

La parité de ces fonctions génératrices diffère selon les conditions initiales considérées. Les rapports des fonctions  $\Gamma$  peuvent être approximés par un polynôme de degré n-1, et les fonctions génératrices pour les conditions initiales A et B sont respectivement des polynômes de degré n-1 et n.

La distribution du degré d'un noeud quelconque au temps n s'obtient à partir de l'expansion en série entière de la fonction génératrice autour de x = 0 et la correction au comportement stationnaire s'écrit comme une somme de nombres harmoniques  $H_n^{(m)}$  d'ordres  $m = 1, \ldots, k-1$ .

# III.2.3 Distribution du degré aux temps finis

On s'intéresse ici au comportement de la distribution du degré quand le temps n est grand. On cherche alors le comportement de la queue de la distribution, pour les degrés k grands. Pour un degré k fini et un temps n grand, on suppose que la distribution du degré se comporte selon la distribution stationnaire :  $f(n,k) \approx f_{stat}(k)$ . On suppose également qu'il existe un degré limite  $k_{\star}(n)$  tel que pour des degrés de l'ordre de  $k_{\star}$  les effets de taille finie ne sont plus négligeables, et tel que pour  $k \ll k_{\star}$  la distribution est stationnaire :  $f(k,n) \approx f_{stat}(k)$ . Dans les sytèmes de taille finie, la probabilité qu'un noeud ait un degré de l'ordre de la taille du système est négligeable. La queue de la distribution, pour  $k \gg k_{\star}$ , est donc tronquée par rapport à la distribution stationnaire :  $f(n,k) \ll f_{stat}(k)$ . Le degré de coupure doit donc être petit devant la taille du système :  $k_{\star} \ll n$ .

Les noeuds de plus grand degré sont presque surement des noeuds apparus au cours des premières étapes de croissance du système lorsque le taux de croissance de chaque noeud n'est pas petit devant 1, c'est-à-dire des noeuds dont l'indice est de l'ordre de  $i \sim 1$ . En première approximation, on peut considérer que l'échelle à partir de laquelle la distribution du degré ressent la taille finie du système est de l'ordre de la valeur typique du degré des premiers noeuds du système :

$$k_{\star}(n) \sim \langle k_1(n) \rangle \approx \ln n + 1.$$
 (III.69)

On confirme cette estimation par des arguments de statistique des extrêmes. Le degré de coupure est de l'ordre du degré maximal atteint par le système en n étapes, c'est-à-dire qu'il n'y a typiquement qu'un seul noeud dont le degré est plus grand que  $k_{\star}$ , sur l'ensemble du système :

$$\int_{k_{\star}(n)}^{\infty} f_{stat}(k) \sim \frac{1}{n}.$$
(III.70)

Le degré de coupure se comporte donc comme :

$$k_{\star}(n) \sim \frac{\ln n}{\ln 2}.$$
 (III.71)

Ces deux approches valident donc un comportement en  $\ln n$  du degré de coupure  $k_{\star}(n)$ .

si l'on s'intéresse aux corrections de taille finie sur la distribution du degré, la quantité qui permet de voir ce phénomène au mieux est le rapport de la distribution à taille finie et de la distribution stationnaire :

$$R(n,k) = \frac{f(n,k)}{f_{stat}(k)} = 2^k f(n,k).$$
 (III.72)

Vue l'analyse heuristique précédente, la fonction R est supposée varier entre deux plateaux : 1 pour  $k \ll k_{\star}(n)$  et 0 pour  $k \gg k_{\star}(n)$ . Le graphe du rapport R(n,k) des distributions en fonction de la variable redimensionnée  $k/\ln n$  pour différentes tailles n représenté figure III.2 confirme le comportement en  $\ln n$  du degré de coupure. On s'intéresse alors aux propriétés de la transition.



FIGURE III.2 – Rapport des distributions du degré à temps fini et stationnaire, pour un réseau issu des conditions initiales (A). On représente ce rapport pour deux tailles  $N = 10^3$  et  $N = 10^6$ .

#### III.2.4 Comportement de la fonction d'échelle à la transition

On introduit la dérivée discrète du rapport R(n,k) des distributions à temps fini et stationnaires pour analyser la transition entre les deux plateaux pour  $k \ll k_{\star}(n)$  et  $k \gg k_{\star}(n)$ . Cette transition étant décroissante, on utilisera la quantité :

$$d(n,k) = R(n,k-1) - R(n,k),$$
(III.73)

pour  $k \ge 2$ , avec R(n,0) = 1, et la première valeur de cette fonction est donc d(n,1) = 1 - R(n,1). À la limite  $k \to \infty$ , R(n,k) = 0, et la somme des d(n,k) est simplement

$$\sum_{k} d(n,k) = 1. \tag{III.74}$$

D'après la forme de la transition des rapports R(n, k) figure III.2, on s'attend à ce que la distribution d(n, k) soit une distribution non nulle dans la région de transition et centrée autour de la valeur de transition  $k_{\star}$ . La fonction génératrice  $D_n(x)$  de cette distribution est donnée par :

$$D_n(x) = \sum_{k \ge 1} d(n,k) x^k = (x-1)F_n(2x) + x.$$
(III.75)

À partir de cette fonction génératrice, on peut calculer tous les moments de la distribution d(n,k), et en particulier le premier moment, que l'on définit comme la valeur de transition  $k_{\star}$ :

$$k_{\star}(n) = \sum_{k \ge 1} k d(n,k) = \left. \frac{\mathrm{d}D_n(x)}{\mathrm{d}x} \right|_{x=1}.$$
 (III.76)

La formule de  $D_n(x)$  fait intervenir la fonction génératrice du degré global  $F_n(x)$ , calculée en (III.26), et la valeur de transition est donc :

$$k_{\star}(n) = \lim_{x \to 1} F_n(2x) + 1.$$
(III.77)

La valeur de transition dépend donc des conditions initiales selon :

$$k_{\star}^{(A)}(n) = 2H_n \approx 2\ln n + 2\gamma$$
  

$$k_{\star}^{(B)}(n) = 2H_n + 1 \approx 2\ln n + 2\gamma + 1.$$
(III.78)

On vérifie bien, sur la figure III.3 du rapport R(n, k) en fonction des conditions initiales, la dépendance de la position de la transition aux conditions initiales. Dans la limite des n grands, la transition se comporte donc bien en  $2 \ln n$ , quelque soit la condition initiale considérée.

La largeur  $\sigma$  du front de transition se détermine à partir du second moment de la distribution d(n,k):

$$\sigma^{2}(n) = \sum_{k \ge 1} k^{2} d(n,k) - \left(k_{\star}(n)\right)^{2} = \left.\frac{\mathrm{d}^{2} D_{n}(x)}{\mathrm{d}x^{2}}\right|_{x=1} + k_{\star}(n) - \left(k_{\star}(n)\right)^{2},\tag{III.79}$$



FIGURE III.3 – Rapport des distributions du degré à temps fini et stationnaire, pour un réseau issu des conditions initiales (A) et (B). On représente ce rapport pour un système de taille  $N = 10^6$ .

soit encore :

$$\sigma^{2}(n) = 4 \lim_{x \to 1} \frac{\mathrm{d}F_{n}(2x)}{\mathrm{d}x} + k_{\star}(n) - \left(k_{\star}(n)\right)^{2} = 2H_{n} + 4H_{n}^{(2)} \approx 2\left(\ln n + 2\gamma - \frac{\pi^{2}}{3}\right), \quad (\mathrm{III.80})$$

et ne dépend pas des conditions initiales. La largeur du front de transition croît donc avec le temps comme  $\sigma(n) \approx (2 \ln n)^{1/2}$  pour  $n \gg 1$ . Ces résultats sont représentés figure III.4, sur laquelle on a tracé la distribution d(n,k) en fonction de la variable réduite  $k/\ln n$  pour deux tailles finies différentes. On observe que cette distribution est bien centrée autour de  $2 \ln n$ , comme attendu pour la valeur de transition. Le front de transition s'élargit bien en fonction de la taille n du système, mais sa largeur relative décroît, car la distribution est représentée en fonction de la variable redimensionnée. Cette décroissance est cependant très lente, en  $(\ln n)^{-1/2}$ .



FIGURE III.4 – Distribution de transition d(N, k) en fonction de la variable réduite  $k/\ln N$ , pour un réseau issu des conditions initiales (A). On représente ce rapport pour deux tailles  $N = 10^3$ et  $N = 10^6$ .

## III.2.5 Grandes déviations

Les résultats obtenus pour la distribution de la transition indiquent donc que la probabilité de trouver un noeud de degré  $k \gg 2 \ln n$  décroît en 1/n. La distribution est confinée autour de  $k_{\star}(n)$ , et de largeur typique  $(2 \ln n)^{1/2}$ . On cherche à déterminer la probabilité qu'un noeud ait un degré proche du degré maximum  $k \sim n$ , c'est-à-dire bien plus grand que le degré de coupure  $k_{\star}(n) \sim \ln n \ll n$ .

On détaillera les calculs pour le cas des conditions initiales (A). La fonction génératrice de la distribution du degré global  $F_n(x)$  peut être définie à partir de la distribution du degré, ou comme une série entière autour de x = 0 :

$$F_n(x) = \sum_{k \ge 1} f(n,k) x^k, \qquad (\text{III.81})$$

$$= \sum_{k \ge 1} \frac{1}{k!} \left. \frac{\mathrm{d}^{k} F_{n}(x)}{\mathrm{d}x^{k}} \right|_{x=0} x^{k}.$$
 (III.82)

La formule intégrale de Cauchy pour la dérivée k-ième d'une fonction vérifie :

$$\left. \frac{\mathrm{d}^k F_n(x)}{\mathrm{d}x^k} \right|_{x=0} = \frac{k!}{2\mathrm{i}\pi} \oint \frac{F_n(y)}{y^{k+1}} \mathrm{d}y,\tag{III.83}$$

et en identifiant, on obtient une formule intégrale explicite pour la distribution du degré :

$$f(n,k) = \oint \frac{\mathrm{d}y}{2\mathrm{i}\pi y^{k+1}} F_n(y).$$
(III.84)

Or la fonction génératrice  $F_n(x)$  contient un rapport de fonctions  $\Gamma$ , d'expansion exponentielle. Le premier terme de la fonction génératrice s'annule sous l'intégration de contour, tandis que l'intégration du deuxième terme peut être estimée par la méthode du col. Le point col  $y_s$  vérifie

$$\frac{\mathrm{d}}{\mathrm{d}y} \left[ \left( y + n - \frac{3}{2} \right) \ln(y + n - 1) - \left( n + \frac{1}{2} \right) \ln(n) - \left( y - \frac{1}{2} \right) \ln(y) \right]_{y = y_s} = 0, \quad (\mathrm{III.85})$$

soit, après quelques approximations dues au régime des grandes déviations  $k \gg 1$  et  $n \gg 1$  :

$$y_s \approx \frac{n}{1 + \frac{n}{k}W\left(-\frac{k}{n}e^{-\frac{k}{n}}\right)},\tag{III.86}$$

où W est la fonction de Lambert, définie comme la fonction inverse de :

$$W \mapsto W e^W.$$
 (III.87)

Par souci de simplicité, on introduit la variable  $\xi = k/n$ , avec laquelle le dénominateur prend la forme :

$$\nu = 1 + \frac{W\left(-\xi e^{-\xi}\right)}{\xi},\tag{III.88}$$

et les deux nouvelles variables  $\xi$  et  $\nu$  sont liées par la relation plus commune :

$$\xi = -\frac{\ln(1-\nu)}{\nu}.$$
 (III.89)

La variable  $\xi$  est comprise dans l'intervalle (0, 1), et la variable  $\nu$  est une fonction décroissante de  $\xi$  définie sur l'intervalle  $(0, \infty)$ .

On introduit ensuite cette valeur du point col  $y_s = n/\nu$  dans l'intégrale de la partie singulière de la fonction génératrice, et la distribution du degré dans le régime de grandes déviations est de la forme :

$$f(n,k) \sim \exp\left(-n\left(\xi \ln n + S(\xi)\right)\right),$$
 (III.90)

où la fonction de grande déviation  $S(\xi)$  prend la forme, simplifiée grâce à l'introduction de la variable  $\nu$ :

$$S(\xi) = \frac{1}{\nu} \Big( \nu \ln \nu + \ln(1-\nu) \ln \nu - (1+\nu) \ln(1+\nu) \Big).$$
(III.91)

Cette fonction varie entre les valeurs extrêmes S(0) = 0 et S(1) = -1. La probabilité d'obtenir un noeud de degré  $k \approx n$ , *i.e*  $\xi \approx 1$ , est donc de la forme :

$$f(n,n) \sim \exp\left(-n(\ln n - 1)\right) \sim \frac{1}{n!},\tag{III.92}$$

et on retrouve bien le résulat attendu par les considérations combinatoires (III.51).

La probabilité d'obtenir un noeud de degré proche du degré maximum atteignable décroît donc plus vite qu'exponentiellement en n.

# III.3 Modèle de Barabási-Albert

Ce modèle de réseau en croissance, décrit dans le chapitre précédent II.1, a été défini par Barabási et Albert [33]. Un nouveau noeud qui arrive dans le système est porteur d'une connexion libre et s'attache préférentiellement aux noeuds déjà très connectés.

# III.3.1 Equation maîtresse

#### a Degré d'un noeud i

La probabilité qu'un noeud, apparu à l'étape n, se connecte au noeud i déjà présent dans le réseau dépend du degré  $k_i(n-1)$  de ce noeud à l'étape précédente n-1, selon :

$$p_{n,i} = \frac{k_i(n-1)}{Z(n-1)},$$
(III.93)

où Z(n) est la fonction de partition qui assure la normalisation des probabilités de chaque noeud à l'étape n + 1:

$$Z(n) = \sum_{i=1}^{n} k_i(n) = K(n).$$
 (III.94)

K(n) est la somme des degrés de tous les noeuds du réseau au temps n. Elle dépend des conditions initiales (III.2) et (III.3), et pour  $n \ge 2$ , elle vérifie la récurrence K(n) = K(n-1)+2, car à chaque étape, un noeud entre dans le système porteur d'un lien qui se connecte à un autre noeud.

On définit comme dans le cas de l'attachement uniforme la distribution du degré  $f_i(n,k)$  au temps n d'un noeud né au temps i, de fonction génératrice  $F_{n,i}(x)$ .

La distribution vérifie l'équation maîtresse :

$$\frac{\mathrm{d}f_i(n,k)}{\mathrm{d}n} = p_{n,i}f_i(n-1,k-1) - p_{n,i}f_i(n-1,k), \qquad (\text{III.95})$$

soit encore, à temps discret :

$$f_i(n,k) = \frac{k-1}{Z(n-1)} f_i(n-1,k-1) + \left(1 - \frac{k}{Z(n-1)}\right) f_i(n-1,k).$$
(III.96)

Les conditions initiales (A) et (B) donnent respectivement les conditions aux limites suivantes pour le noeud i = 1:

$$f_1^{(A)}(1,k) = \delta_{k,0} \tag{III.97}$$

$$f_1^{(B)}(1,k) = \delta_{k,1}, \tag{III.98}$$

et pour un noeud quelconque  $i \ge 2$ ,

$$f_i(i,k) = \delta_{k,1}.\tag{III.99}$$

# b Degré global

La distribution f(n,k) du degré d'un noeud pris au hasard dans le réseau au temps n peut s'obtenir par la moyenne de la distribution du degré de tous les noeuds :

$$f(n,k) = \frac{1}{n} \sum_{i=1}^{n} f_i(n,k).$$
 (III.100)

La probabilité qu'un noeud de degré k reçoive le degré entrant est donnée par le taux d'attachement  $\pi_{n,k} = k/Z(n-1)$ . L'équation maîtresse s'obtient directement à partir de l'équation pour le degré d'un noeud *i*. On peut également par un raisonnement simple obtenir l'équation maîtresse pour le nombre typique nf(n,k) de noeuds de degré k au temps n:

$$\frac{\mathrm{d}\left(nf(n,k)\right)}{\mathrm{d}n} = \pi_{n,k-1}(n-1)f(n-1,k-1) - \pi_{n,k}(n-1)f(n-1,k) + \delta_{k,1},\qquad(\mathrm{III}.101)$$

où  $\delta_{k,1}$  correspond à l'ajout un noeud de degré 1 au réseau. En temps discret, cette équation devient :

$$nf(n,k) = \frac{k-1}{Z(n-1)}(n-1)f(n-1,k-1) + \left(1 - \frac{k}{Z(n-1)}\right)(n-1)f(n-1,k) + \delta_{k,1}, \quad \text{(III.102)}$$

avec les conditions aux limites suivantes :

$$f^{(A)}(1,k) = \delta_{k,0}, \tag{III.103}$$

$$f^{(B)}(1,k) = \delta_{k,1}.$$
 (III.104)

La fonction génératrice du degré global au temps n est notée :

$$F_n(x) = \left\langle x^{k(n)} \right\rangle. \tag{III.105}$$

On peut obtenir directement à partir de l'équation maîtresse une relation de récurrence pour la distribution stationnaire  $f_{stat}(k)$  du degré d'un noeud quelconque :

$$(k-1)f_{stat}(k-1) - (k+2)f_{stat}(k) = 2\delta_{k,1},$$
(III.106)

avec  $Z(n) \approx 2n$ . La distribution stationnaire du degré d'un noeud quelconque est donc :

$$f_{stat}(k) = \frac{2}{k(k+1)(k+2)} \approx \frac{2}{k^3}.$$
 (III.107)

On retrouve bien ici le résultat obtenu par Barabási-Albert d'une distribution du degré en loi de puissance d'exposant  $\gamma = 3$ .

# III.3.2 Fonction génératrice

## a Du degré d'un noeud i

Pour obtenir une solution à temps fini des distributions du degré, on suit le raisonnement développé pour le modèle avec attachement uniforme, par le biais de la fonction génératrice du degré  $F_{n,i}(x)$ :

$$F_{n,i}(x) = \left\langle x^{k_i(n)} \right\rangle \tag{III.108}$$

$$=\sum_{k=1}^{n} f_i(n,k) x^k.$$
 (III.109)

On peut définir, comme pour le modèle AU, une relation de récurrence pour la fonction génératrice du degré à partir de la relation de récurrence du degré du noeud i:

$$k_i(n) = k_i(n-1) + I_i(n),$$
 (III.110)

où l'incrément du degré du noeud i au temps n est donné par :

$$I_i(n) = \begin{cases} 1 & \text{avec probabilité } p_{n,i} = \frac{k_i(n-1)}{Z(n-1)} \\ 0 & \text{sinon} \end{cases}$$
(III.111)

La fonction génératrice du degré du noeud i s'écrit donc :

$$F_{n,i}(x) = \left\langle x^{k_i(n)} \right\rangle \tag{III.112}$$

$$= \left\langle x^{I_i(n)} x^{k_i(n-1)} \right\rangle \tag{III.113}$$

$$= \left\langle \left( (1 - p_{n,i}) + x p_{n,i} \right) x^{k_i(n-1)} \right\rangle$$
(III.114)

$$= \left\langle x^{k_i(n-1)} + \left(\frac{x-1}{Z(n-1)}k_i(n-1)\right) x^{k_i(n-1)} \right\rangle.$$
(III.115)

Le terme en  $kx^k$  peut s'écrire comme :

$$kx^k = x\frac{\mathrm{d}x^k}{\mathrm{d}x},\tag{III.116}$$

et la fonction génératrice du degré du noeud i vérifie la relation de récurrence :

$$F_{n,i}(x) = F_{n-1,i}(x) + \frac{x(x-1)}{Z(n-1)} \frac{\mathrm{d}F_{n-1,i}(x)}{\mathrm{d}x},$$
 (III.117)

où Z(n) est le facteur de normalisation défini en (III.94). Les conditions initiales sont :

$$F_{i,i}(x) = x, \quad i \ge 2, \tag{III.118}$$

$$F_{1,1}^{(A)}(x) = 1,$$
 (III.119)

$$F_{1,1}^{(B)}(x) = x. (III.120)$$

La fonction génératrice à taille finie s'obtient par résolution de l'équation de récurrence (III.117). Pour résoudre cette équation, on utilise le changement de variables suivant :

$$u = \frac{x}{1-x} \quad \Leftrightarrow \quad x = \frac{u}{u+1}$$
 (III.121)

d'où 
$$x(x-1)\frac{\mathrm{d}}{\mathrm{d}x} = -u\frac{\mathrm{d}}{\mathrm{d}u},$$
 (III.122)

grâce auquel l'équation de récurrence de la fonction génératrice  $F_{n,i}(x) = F_{n,i}\left(\frac{u}{u+1}\right)$  du degré du noeud *i* s'écrit :

$$F_{n,i}\left(\frac{u}{u+1}\right) = F_{n-1,i}\left(\frac{u}{u+1}\right) - \frac{u}{Z(n-1)}\frac{\mathrm{d}F_{n-1,i}\left(\frac{u}{u+1}\right)}{\mathrm{d}u}.$$
 (III.123)

Pour simplifier les équations, on notera par la suite :

$$\widehat{F}_{n,i}(u) = F_{n,i}\left(\frac{u}{u+1}\right) = F_{n,i}(x), \qquad (\text{III.124})$$

et l'équation (III.117) devient :

$$\widehat{F}_{n,i}(u) = \widehat{F}_{n-1,i}(u) - \frac{u}{Z(n-1)} \frac{\mathrm{d}\widehat{F}_{n-1,i}(u)}{\mathrm{d}u}.$$
(III.125)

# b Du degré global

À partir de la fonction génératrice  $F_{n,i}$  du degré  $k_i(n)$  du noeud i, on peut définir la fonction génératrice  $F_n$  du degré k(n) d'un noeud quelconque :

$$F_n(x) = \left\langle x^{k(n)} \right\rangle \tag{III.126}$$

$$=\sum_{k=1}^{n} f(n,k)x^{k}$$
(III.127)

$$= \frac{1}{n} \sum_{i=1}^{n} F_{n,i}(x).$$
(III.128)

La fonction génératrice  ${\cal F}_n$  du degré global vérifie donc la relation :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n F_{n,i}(x)$$
(III.129)

$$= \frac{1}{n} \sum_{i=1}^{n} \left( F_{n-1,i}(x) + \frac{x(x-1)}{Z(n-1)} \frac{\mathrm{d}F_{n-1,i}(x)}{\mathrm{d}x} \right),$$
(III.130)

qui donne la relation de récurrence suivante pour la fonction génératrice du degré global :

$$nF_n(x) = (n-1)F_{n-1}(x) + (n-1)\frac{x(x-1)}{Z(n-1)}\frac{\mathrm{d}F_{n-1}(x)}{\mathrm{d}x} + x.$$
 (III.131)

Les conditions initiales sont définies comme dans le modèle AU pour n = 1 (III.66).

On peut également obtenir les relations de récurrence pour cette fonction génératrice à partir de la relation de récurrence pour la distribution du degré d'un noeud quelconque (III.102).

La fonction génératrice  $F_n$  du degré global converge asymptotiquement vers la fonction génératrice  $F_{stat}$  de la distribution stationnaire, qui vérifie l'équation différentielle :

$$2F_{stat}(x) - x(x-1)\frac{\mathrm{d}F_{stat}(x)}{\mathrm{d}x} = 2x.$$
 (III.132)

Cette équation a pour solution :

$$F_{stat}(x) = 3 - \frac{2}{x} - \frac{2(x-1)^2}{x^2} \ln(1-x), \qquad (\text{III.133})$$

dont l'expansion en série entière autour de  $x \to 0$  donne tous les termes de la distribution stationnaire :

$$F_{stat}(x) = \frac{4}{6}x + \frac{4}{24}\frac{x^2}{2} + \dots = \sum_{k=1}^{\infty} \frac{4}{k(k+1)(k+2)}\frac{x^k}{k!},$$
 (III.134)

et on retrouve bien la solution stationnaire de la distribution du degré global (III.107).

Sous le changement de variables (III.122), la fonction génératrice du degré global  $F_n(u) = F_n(x)$  vérifie la relation de récurrence suivante :

$$n\widehat{F}_{n}(u) = \widehat{F}_{n-1}(u) - \frac{u}{Z(n-1)}\frac{\mathrm{d}F_{n-1}(u)}{\mathrm{d}u} + \frac{u}{u+1},$$
(III.135)

que l'on obtient à partir de l'équation (III.131).

## III.3.3 Résolution des équations de récurrence

## a Transformée de Mellin pour le degré du noeud i

Les équations se simplifient encore en utilisant la transformée de Mellin. La transformée  $M_{n,i}$  de la fonction  $\widehat{F}_{n,i}$  est définie par :

$$M_{n,i}(s) = \int_0^\infty \widehat{F}_{n,i}(u) u^{-(s+1)} du,$$
 (III.136)

de transformation inverse

$$\widehat{F}_{n,i}(u) = \oint_C \frac{\mathrm{d}s}{2\pi \mathrm{i}} M_{n,i}(s) u^s, \qquad (\mathrm{III.137})$$

où l'intégrale de contour est définie dans le plan complexe, le long du contour C, dont la nature est précisée ci-après. La transformée de Mellin de la fonction génératrice  $F_{i,i}$  du degré du noeud

i au temps i est en particulier donnée par :

$$M_{i,i}(s) = \int_0^\infty \widehat{F}_{i,i}(u) u^{-(s+1)} du$$
  
=  $\int_0^\infty x(u) u^{-(s+1)} du$   
=  $\int_0^1 x^{-s} (1-x)^{s-1} dx$ , (III.138)

dont il existe une solution analytique pour  $0 < \operatorname{Re}(s) < 1$ , et :

$$M_{i,i}(s) = \frac{\pi}{\sin(\pi s)}.$$
(III.139)

Le contour C est donc défini comme le périmètre de la bande 0 < Re(s) < 1 du plan complexe. La transformée de Mellin est liée à la dérivée de la fonction  $\widehat{F}_{n,i}$  par :

$$M_{n,i}(s) = \int_0^\infty \widehat{F}_{n,i}(u) u^{-(s+1)} du = \frac{1}{s} \int_0^\infty \widehat{F}'_{n,i}(u) u^{-s} du, \qquad (\text{III.140})$$

soit encore

$$sM_{n,i}(s) = \int_0^\infty u\widehat{F}'_{n,i}(u)u^{-(s+1)}\mathrm{d}u.$$
 (III.141)

La relation de récurrence (III.117) après changement de variable (III.122) et transformation de Mellin se simplifie en :

$$M_{n,i}(s) = \left(1 - \frac{s}{Z(n-1)}\right) M_{n-1,i}(s).$$
(III.142)

Pour un noeud que l<br/>conque,  $i\geq 2,$  la transformée de Mellin des conditions initiales est donnée par :

$$M_{i,i}(s) = m(s) = \frac{\pi}{\sin(\pi s)},$$
 (III.143)

et la solution de l'équation de récurrence dépend de la condition initiale :

$$M_{n,i}^{(A)}(s) = \frac{\Gamma(n - \frac{s}{2} - 1)\Gamma(i - 1)}{\Gamma(i - \frac{s}{2} - 1)\Gamma(n - 1)}m(s).$$

$$M_{n,i}^{(B)}(s) = \frac{\Gamma\left(n - \frac{s}{2} - \frac{1}{2}\right)\Gamma\left(i - \frac{1}{2}\right)}{\Gamma\left(i - \frac{s}{2} - \frac{1}{2}\right)\Gamma\left(n - \frac{1}{2}\right)}m(s)$$
(III.144)

La fonction génératrice du degré du noeud i peut être développée en série pour  $x \to 1$  i.e.  $u \to \infty$  :

$$\widehat{F}_{n,i}(u) = \sum_{k=0}^{\infty} \frac{1}{k! u^k} \left. \frac{\mathrm{d}^k \widehat{F}_{n,i}(u)}{\mathrm{d} u^k} \right|_{u \to \infty},\tag{III.145}$$

dont les coefficients  $\frac{d^k \widehat{F}_{n,i}(u)}{du^k} \Big|_{u \to \infty}$  peuvent être calculés à partir des résidus négatifs de la transformée de Mellin. En effet, la fonction génératrice s'obtient à partir de la transformée de Mellin par :

$$\widehat{F}_{n,i}(u) = \oint_C \frac{\mathrm{d}s}{2\pi \mathrm{i}} M_{n,i}(s) u^s, \qquad (\mathrm{III.146})$$

$$= \sum_{j \in \mathbb{N}} \operatorname{Res} \left[ M_{n,i}(s) u^s, s = -j \right],$$
(III.147)

et comme la transformée de Mellin ne dépend pas du paramètre u, on peut comparer un à un les termes de la série des résidus de la transformée de Mellin et les termes du développement en série entière de la fonction génératrice du degré du noeud i autour de  $u \to \infty$ . Les termes de cette dernière série contiennent, à leur tour, tous les moments de la distribution du degré du noeud i au temps fini n.

En particulier, les deux premiers coefficients de l'expansion de  $\widehat{F}_{n,i}$  en série entière correspondent à la valeur moyenne et à la variance du degré  $k_i(n)$  du noeud *i* au temps *n*, qui prennent respectivement les valeurs :

$$\left\langle k_i^{(A)}(n) \right\rangle = \frac{\Gamma(n-\frac{1}{2})\Gamma(i-1)}{\Gamma(i-\frac{1}{2})\Gamma(n-1)},\tag{III.148}$$

$$\left\langle k_i^{(B)}(n) \right\rangle = \frac{\Gamma(n)\Gamma\left(i - \frac{1}{2}\right)}{\Gamma(i)\Gamma\left(n - \frac{1}{2}\right)},$$
 (III.149)

$$\operatorname{var}k_{i}^{(A)}(n) = 2\frac{n-1}{i-1} - \left\langle k_{i}^{(A)}(n) \right\rangle^{2} - \left\langle k_{i}^{(A)}(n) \right\rangle, \qquad (\text{III.150})$$

$$\operatorname{var}k_{i}^{(B)}(n) = 2\frac{2n-1}{2i-1} - \left\langle k_{i}^{(B)}(n) \right\rangle^{2} - \left\langle k_{i}^{(B)}(n) \right\rangle.$$
(III.151)

# b Transformée de Mellin pour le degré global

On définit également la transformée de Mellin de la fonction génératrice du degré global, et on la note  $M_n(s)$ . La valeur initiale est définie dans le cas de la condition initiale (A) pour n = 2 et pour la condition initiale (B) au temps n = 1, et les fonctions génératrices du degré  $F_2^{(A)}(x) = F_1^{(B)}(x) = x$  ont pour transformée de Mellin :

$$M_2^{(A)}(s) = M_1^{(B)}(s) = \frac{\pi}{\sin(\pi s)},$$
 (III.152)

que l'on notera par simplicité m(s). On reproduit le même raisonnement que pour la fonction génératrice  $F_{n,i}(x)$  du degré d'un noeud *i*. La relation de récurrence (III.131) devient, sous transformation de Mellin :

$$nM_n(s) = (n-1)\left(1 - \frac{s}{Z(n-1)}\right)M_{n-1}(s) + m(s), \qquad (\text{III.153})$$

dont les conditions initiales sont données par (III.152). Cette équation de récurrence se résout comme une somme d'une solution homogène et d'une solution particulière. L'équation homogène

$$M_n^{(H)}(s) = M_{n-1}^{(H)}(s) \frac{n-1}{n} \left(1 - \frac{s}{Z(n-1)}\right)$$
(III.154)

a pour solution

$$M_n^{(H)}(s) = \lambda(s) \frac{2\Gamma(n-1-\frac{s}{2})}{n\Gamma(1-\frac{s}{2})\Gamma(n-1)},$$
 (III.155)

où  $\lambda$  est une fonction de s que l'on détermine à partir des conditions initiales. La solution particulière est la solution indépendante du temps de l'équation :

$$nM(s) = (n-1)\left(1 - \frac{s}{Z(n-1)}\right)M(s) + m(s).$$
 (III.156)

La transformée de Mellin de la fonction génératrice du degré d'un no eud quelconque au temps  $n\ {\rm est}\ {\rm donc}$  :

$$M_n^{(A)}(s) = \frac{2m(s)}{n(s+2)} \left( n - 1 + (s+1) \frac{\Gamma\left(n - 1 - \frac{s}{2}\right)}{\Gamma\left(1 - \frac{s}{2}\right)\Gamma(n-1)} \right)$$

$$M_n^{(B)}(s) = \frac{m(s)}{n(s+2)} \left( 2n - 1 + (s+1) \frac{\sqrt{2\pi}\Gamma\left(n - \frac{s}{2} - \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2} - \frac{s}{2}\right)\Gamma\left(n - \frac{1}{2}\right)} \right),$$
(III.157)

pour chacune des conditions initiales considérées.

La limite stationnaire  $n \to \infty$  de ce résultat vaut :

$$M_{stat}(s) = \frac{2m(s)}{s+2},\tag{III.158}$$

à partir de laquelle, par transformation inverse, on retrouve bien la solution stationnaire (III.133):

$$\widehat{F}_{stat}(u) = \frac{1}{2i\pi} \oint_C ds \frac{2m(s)}{s+2} u^s$$
(III.159)

$$= \sum_{j \in \mathbb{N}} \operatorname{Res}\left[\frac{2\pi}{\sin(\pi s)} \frac{u^s}{(s+2)}, s = -j\right]$$
(III.160)

$$= 1 - \frac{2}{u} + 2\frac{\ln u}{u^2} + \frac{2}{u^2} \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{ju^j}$$
(III.161)

$$=1 - \frac{2}{u} + 2\frac{\ln(1+u)}{u^2} \tag{III.162}$$

$$= 3 - \frac{2}{x} - 2\frac{(1-x)^2}{x^2}\ln(1-x)$$
(III.163)

$$=F_{stat}(x).$$
 (III.164)

On considère ensuite les solutions à temps fini, en particulier dans le régime d'échelle.

# III.3.4 Distribution d'échelle du degré d'un noeud i

Pour la distribution du degré d'un noeud i, le régime d'échelle est déterminé dans les cas où le temps de naissance i du noeud et le temps d'observation n sont grands et de même ordre. On utilise le rapport z des temps, défini pour le modèle AU :

$$z = \frac{n}{i} \ge 1. \tag{III.165}$$

Dans ce régime, le degré moyen du noeud i au temps n et sa variance, définies respectivement en (III.148) et (III.150), sont, pour  $i \sim n \gg 1$ , de la forme :

$$\langle k_i(n) \rangle \approx z^{1/2} \tag{III.166}$$

$$\operatorname{var}(k_i(n)) \approx z - z^{1/2} = z^{1/2}(z^{1/2} - 1).$$
 (III.167)

On peut retrouver ces résultats à partir de la distribution complète du degré du noeud i au temps n. Elle s'obtient, dans le régime d'échelle, par transformations inverses successives, à partir de la transformée de Mellin  $M_{n,i}$  de la fonction génératrice du degré du noeud i au temps n. Les formules (III.144) deviennent, dans le régime d'échelle :

$$M_{n,i}(s) \approx z^{-s/2} m(s)$$
  
$$\approx z^{-s/2} \frac{\pi}{\sin(\pi s)}.$$
 (III.168)

Dans ce régime, par transformation de Mellin inverse, la fonction génératrice du degré du noeud i au temps n s'obtient explicitement :

$$F_{n,i}(x) \approx \frac{1}{2i\pi} \oint_C z^{-s/2} \frac{\pi}{\sin(\pi s)} \left(\frac{x}{1-x}\right)^s \tag{III.169}$$

$$\approx \sum_{j \in \mathbb{N}} \frac{(x-1)^j z^{j/2}}{x^j} \tag{III.170}$$

$$\approx \frac{x}{x+z^{1/2}(1-x)}.$$
 (III.171)

Et on obtient finalement la distribution du degré du noeud i au temps n dans le régime d'échelle par expansion de la fonction génératrice autour de  $x \to 0$ :

$$F_{n,i}(x) \approx \frac{x}{x + z^{1/2}(1-x)} = \sum_{k=1}^{\infty} z^{-1/2} \left(1 - z^{-1/2}\right)^{k-1} x^k, \qquad (\text{III.172})$$

et donc :

$$f_i(n,k) \approx z^{-1/2} \left(1 - z^{-1/2}\right)^{k-1},$$
 (III.173)

où z = n/i.

Dans le régime d'échelle, la distribution du degré du noeud i au temps n suit une loi géométrique qui ne dépend que du rapport des temps de naissance et d'observation, et est indépendante des conditions initiales.

### III.3.5 Distribution à temps fini du degré global

Les résidus positifs de la transformation de Mellin correspondent aux termes de l'expansion de la fonction génératrice en série autour de u = 0. On peut donc calculer à partir des expressions obtenues en (III.157) les termes de la distribution du degré global à un temps fini n. Par exemple, les premiers termes de l'expansion permettent d'obtenir la distribution des premières valeurs du degré d'un noeud quelconque. Ainsi, la probabilité que le noeud tiré au hasard soit de degré k = 1, dans un réseau issu des conditions initiales (A), est donnée par :

$$f^{(A)}(n,1) = \frac{2(n-1)}{3n} + \frac{4\Gamma\left(n-\frac{3}{2}\right)}{3\sqrt{\pi}n\Gamma(n-1)} = \frac{2}{3} - \frac{2}{2n} + \frac{4}{3\sqrt{\pi}n^{3/2}} + \dots$$
(III.174)

On retrouve bien la probabilité stationnaire  $f_{stat}(1) = 2/3$ , et un terme correctif. Pour la condition initiale (A), toutes les probabilités f(n, k) présentent cette correction en  $n^{-3/2}$ . Pour les réseaux issus des conditions initiales (B), toutes les probabilités sont des fractions rationnelles de n. On a en particulier pour les plus petits degrés :

$$f^{(B)}(n,1) = \frac{2n-1}{3n} = \frac{2}{3} - \frac{1}{3n}$$
  
$$f^{(B)}(n,2) = \frac{n^2 - 2n + 3}{3n(2n-3)} = \frac{1}{6} - \frac{1}{12n} + \frac{3}{8n^2} + \dots$$
 (III.175)

Toute la distribution du degré est contenue dans les transformées de Mellin (III.157).

#### a Valeur de transition

Le comportement de la distribution du degré global f(n, k) au temps n se divise en trois régimes pour n grand. Pour les degrés petits,  $k \sim 1$ , les effets de taille finie n'influencent pas la distribution du degré, qui est analogue à la distribution stationnaire  $f(n, k) \sim f_{stat}(k)$ . Pour les degrés grands, de l'ordre de la taille du système  $k \sim n$ , la probabilité de trouver un noeud de ce degré k, donc auquel tous les autres noeuds ou presque se sont attachés, est très faible, et nulle dès que k > n. La valeur k = n est donc une coupure naturelle de la distribution du degré à temps fini. On s'intéresse au comportement de la distribution du degré lorsque celui-ci prend une valeur intermédiaire  $1 \ll k \ll n$ . La valeur de transition, notée  $k_*(n)$  est telle que, en moyenne, seule une fraction finie des noeuds du réseau ont un degré supérieur :

$$n \int_{k_{\star}(n)}^{\infty} \mathrm{d}k f_{stat}(k) \sim 1, \qquad (\text{III.176})$$

 $\operatorname{soit}$ 

$$k_{\star}(n) \approx (2n)^{1/2}.$$
 (III.177)

On retrouve un résultat de cet ordre si l'on considère que le plus grand degré typique se comporte comme le degré typique du premier noeud, en effet le noeud de plus grand degré est le plus souvent parmi les premiers, et le degré de transition  $k_{\star}$  est de l'ordre du plus grand degré typique :

$$k_{\star}(n) \sim \langle k_1(n) \rangle \approx n^{1/2}.$$
 (III.178)

#### b Fonction d'échelle

Pour étudier le régime d'échelle, on redimensionne le degré par le degré de transition. On note la variable redimensionnée :  $y = k/k_{\star}(n)$ . On suppose que la distribution du degré à taille finie se comporte selon la fonction d'échelle  $\Phi(y)$  :

$$f(n,k) \approx f_{stat}(k)\Phi(y), \quad y = \frac{k}{n^{1/2}}, \tag{III.179}$$

où cette fonction  $\Phi$ , selon le raisonnement précédent, varie de  $\Phi(y) \sim 1$  pour  $y \ll 1$  à  $\Phi(y) \sim 0$  pour  $y \gg 1$ .

On calcule d'abord la fonction d'échelle  $\Phi^{(A)}(y)$  dans le cas de la condition initiale (A). La transformée de Mellin correspondant à la distribution du degré global est donnée par la formule (III.157), dont l'approximation, pour n grand, est :

$$M_n^{(A)}(s) \approx \frac{2m(s)}{(s+2)} \left( 1 + \frac{s+1}{\Gamma\left(1 - \frac{s}{2}\right)} n^{-s/2 - 1} \right).$$
(III.180)

On va s'intéresser au terme de droite qui seul dépend de n et porte la correction à temps fini, et on le note :

$$M_{n,corr}(s) \approx \frac{2(s+1)m(s)}{(s+2)\Gamma\left(1-\frac{s}{2}\right)} n^{-s/2-1}.$$
 (III.181)

La notation de la condition initiale en exposant est ici supprimée des formules par soucis de lisibilité, et il est sous-entendu que les formules présentées s'appliquent à la condition initiale (A). Le terme correctif de la fonction génératrice du degré global est déterminé par la transformation de Mellin inverse :

$$\widehat{F}_{n,corr}(u) \approx \frac{1}{n} \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} \frac{2(s+1)m(s)}{(s+2)\Gamma\left(1-\frac{s}{2}\right)} \left(\frac{u}{n^{1/2}}\right)^s.$$
(III.182)

Or, on a vu dans la section précédente que la distribution du degré se calcule à partir de sa fonction génératrice par le biais d'une intégrale de contour (III.84). Pour la fonction géneratrice  $\widehat{F}$ , redimensionnée selon les formules de changement de variable (III.122), on obtient donc :

$$f(n,k) = \oint \frac{\mathrm{d}x}{2\mathrm{i}\pi} \frac{F_n(x)}{x^{k+1}} = \oint \frac{\mathrm{d}u}{2\mathrm{i}\pi} \frac{\widehat{F}_n(u)(u+1)^{k-1}}{u^{k+1}}.$$
 (III.183)

La partie constante en n de la transformée de Mellin se transforme en la partie régulière de la distribution du degré, après une double intégration :

$$f_{reg}(n,k) = \oint \frac{\mathrm{d}u}{2\mathrm{i}\pi} \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} \frac{2m(s)}{(s+2)} \frac{(u+1)^{k-1}}{u^{k-s+1}},$$
(III.184)

et cette formule se résout après inversion des intégrales et utilisation de la relation [53] :

$$\oint \frac{\mathrm{d}u}{2\mathrm{i}\pi} \frac{(u+1)^{k-1}}{u^{k-s+1}} = \frac{\Gamma(k)}{\Gamma(s)\Gamma(k-s+1)}.$$
(III.185)
La valeur de l'intégrale de contour en s est donnée par :

$$f_{reg}(n,k) = \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} \frac{2m(s)}{(s+2)} \frac{\Gamma(k)}{\Gamma(s)\Gamma(k-s+1)}$$
(III.186)

$$= \operatorname{Res}\left[\frac{2m(s)}{(s+2)}\frac{\Gamma(k)}{\Gamma(s)\Gamma(k-s+1)}, s = -2\right]$$
(III.187)  
4 (III.187)

$$=\frac{1}{k(k+1)(k+2)}$$
(III.188)

$$= f_{stat}(k). \tag{III.189}$$

La correction de temps fini est donc contenue dans la partie corrective de la fonction génératrice, dont on extrait la partie corrective de la distribution du degré :

$$f_{corr}(n,k) = \oint \frac{\mathrm{d}u}{2\mathrm{i}\pi} \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} \frac{2(s+1)m(s)}{(s+2)\Gamma\left(1-\frac{s}{2}\right)n^{s/2}} \frac{(u+1)^{k-1}}{u^{k-s+1}}.$$
 (III.190)

On utilise encore la relation (III.185) après échange des intégrales, puis les relations pour les fonctions  $\Gamma$  suivantes :

$$\Gamma(1-z) = -z\Gamma(-z) \tag{III.191}$$

$$\Gamma(z) = \frac{2^{z-1/2}}{\sqrt{2\pi}} \Gamma\left(\frac{z}{2}\right) \Gamma\left(\frac{z+1}{2}\right), \qquad (\text{III.192})$$

pour  $z \in \mathbb{C}$  quelconque, ainsi que la formule asymptotique de Stirling. La partie corrective de la distribution du degré global se met sous la forme :

$$f_{corr}(n,k) = \frac{2}{\sqrt{\pi}} \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} \frac{s+1}{s+2} \Gamma\left(\frac{1-s}{2}\right) f_{stat}(k) \frac{\Gamma(k+3)}{\Gamma(k+1-s)} \frac{1}{\left(2n^{1/2}\right)^{s+2}},\tag{III.193}$$

$$\approx f_{stat}(k) \frac{2}{\sqrt{\pi}} \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} \frac{s+1}{s+2} \Gamma\left(\frac{1-s}{2}\right) \left(\frac{k}{2n^{1/2}}\right)^{s+2}.$$
 (III.194)

La fonction d'échelle, définie par la formule (III.179), est donc donnée par la formule suivante, pour la condition initiale (A):

$$\Phi^{(A)}(y) = 1 + \frac{2}{\sqrt{\pi}} \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} \frac{s+1}{s+2} \Gamma\left(\frac{1-s}{2}\right) \left(\frac{y}{2}\right)^{s+2},\tag{III.195}$$

où  $y = k/n^{1/2}$ .

Pour la condition initiale (B), on suit les mêmes lignes de raisonnement, et la partie régulière est égale à la distribution stationnaire :

$$f_{reg}^{(B)}(n,k) = f_{stat}(k).$$
 (III.196)

La partie corrective dérive de la partie corrective de la transformée de Mellin, qui prend la forme d'échelle suivante :

$$M_{n,corr}^{(B)}(s) \approx \frac{\sqrt{\pi}(s+1)m(s)}{(s+2)\Gamma\left(\frac{1-s}{2}\right)} n^{-s/2-1}.$$
 (III.197)

La fonction d'échelle pour la condition initiale (B) vérifie donc la formule :

$$\Phi^{(B)}(y) = 1 + \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} \frac{s+1}{s+2} \Gamma\left(1-\frac{s}{2}\right) \left(\frac{y}{2}\right)^{s+2}.$$
 (III.198)

Les intégrales de contour obtenues pour les fonctions d'échelle  $\Phi^{(A)}$  et  $\Phi^{(B)}$  peuvent être évaluées par utilisation du théorème des résidus, pour un contour contenant les pôles des fonctions  $\Gamma$ , dans le demi-plan des réels positifs. La somme s'effectue sur les entiers impairs pour la condition initiale (A) et sur les entiers pairs pour la condition initiale (B). Les fonctions d'échelle pour les deux conditions initiales s'écrivent donc respectivement sous la forme des sommes :

$$\Phi^{(A)}(y) = 1 + \frac{8}{\sqrt{\pi}} \sum_{j \ge 0} \frac{(-1)^j (j+1)}{(2j+3)j!} \left(\frac{y}{2}\right)^{2j+3}$$
(III.199)

$$\Phi^{(B)}(y) = 1 + \sum_{j \ge 0} \frac{(-1)^j (j+1)(2j+3)}{(j+2)!} \left(\frac{y}{2}\right)^{2j+4},$$
 (III.200)

qui se calculent analytiquement et donnent les résultats finaux suivants :

$$\Phi^{(A)}(y) = \operatorname{erfc}\left(\frac{y}{2}\right) + \frac{e^{-\frac{y^2}{4}}}{\sqrt{\pi}}\left(y + \frac{y^3}{2}\right)$$
(III.201)

$$\Phi^{(B)}(y) = e^{-\frac{y^2}{4}} \left( 1 + \frac{y^2}{4} + \frac{y^4}{8} \right).$$
(III.202)

La fonction  $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$  est la fonction erreur complémentaire. Ces expressions ont été obtenues respectivement par Waclaw et Sokolov [54] et Krapivsky et Redner [52], et Dorogotsev et al. [10]

On observe sur la figure III.5 la convergence du rapport de la distribution du degré à temps fini et de la distribution stationnaire vers la courbe asymptotique  $\Phi$  pour une grande taille du système  $n \gg 1$ . On compare sur la figure III.6 les fonctions d'échelle  $\Phi$  pour les conditions initiales (A) et (B). Ces deux fonctions d'échelle ont un comportement semblable : comme attendu, pour  $y \ll 1$ , *i.e*  $k \ll k_{\star}$ ,  $\Phi(y) = 1$ , et pour  $y \gg 1$ ,  $\Phi(y) = 0$ . Elles atteignent cette valeur asymptotique par une décroissance en  $\exp(-y^2/4)$  après avoir atteint un maximum à une valeur finie de y. Cependant, la parité du polynôme associé au terme exponentiel dépend de la condition initiale, et il en est de même pour la forme du terme dominant pour y petit. Le développement des fonctions d'échelle autour de ces deux limites est, pour y petit :

$$\Phi^{(A)}(y) = 1 + \frac{y^3}{3\sqrt{\pi}} + \dots$$
(III.203)

$$\Phi^{(B)}(y) = 1 + \frac{3y^4}{32} + \dots, \qquad (\text{III.204})$$

et pour y grand :

$$\Phi^{(A)}(y) \approx \frac{y^3}{2\sqrt{\pi}} e^{-\frac{y^2}{4}}$$
(III.205)

$$\Phi^{(B)}(y) \approx \frac{y^4}{8} e^{-\frac{y^2}{4}}.$$
(III.206)



FIGURE III.5 – Rapport de la distribution à temps fini et de la distribution stationnaire en fonction de la variable redimensionnée  $y = k/N^{1/2}$  pour deux tailles du réseau N = 100 et  $N = 10^4$ . On présente ici comme exemple les résultats obtenus pour la condition initiale (A).

Le terme correctif, pour les y petits, se comporte donc en  $y^3$  pour les conditions initiales (A) et en  $y^4$  pour (B). La correction est donc moins sensible pour les conditions initiales (B) pour les petites valeurs de y, mais la correction est plus forte à l'approche de la transition. Le maximum de  $\Phi^{(B)}$  est donc plus grand et décalé vers de plus grandes valeurs de y que le pic de  $\Phi^{(A)}$ .

#### III.3.6 Grandes déviations

La probabilité de trouver dans le système un noeud de degré  $k \approx n \gg 1$  décroît comme une loi de puissance d'exposant  $\gamma = 3$  dans le régime stationnaire :

$$f_{stat}(n) \approx \frac{4}{n^3},$$
 (III.207)



FIGURE III.6 – Fonctions d'échelle pour les conditions initiales (A) et (B) en fonction de la variable redimensionnée  $y = k/N^{1/2}$ , à temps fini  $N = 10^4$  et asymptotique  $\Phi(y)$ .

et la correction de temps fini est donnée par la partie singulière de la distribution. Elle dérive de la partie singulière de la transformée de Mellin pour le degré global, et vaut :

$$M_{n,sing}^{(A)}(s) = M_n^{(A)}(s) - \frac{n-1}{n} M_{stat}(s) = \frac{2}{n} \frac{(s+1)m(s)}{s+2} \frac{\Gamma\left(n-\frac{s}{2}-1\right)}{\Gamma\left(1-\frac{s}{2}\right)\Gamma(n-1)}.$$
 (III.208)

Cette expression croît exponentiellement avec n, et peut être approximée par :

$$M_{n,sing}(s) \approx \frac{1}{n\sqrt{2\pi}} \frac{2(s+1)m(s)}{s+2} \\ \exp\left[\left(n-\frac{s}{2}\right)\ln\left(n-\frac{s}{2}\right) - \left(\frac{1}{2}-\frac{s}{2}\right)\ln\left(1-\frac{s}{2}\right) - n\ln n\right]$$
(III.209)

On supprime la notation de la condition initiale par souci de clarté, étant sous-entendu ici que l'on considère la condition initiale (A), mais le système issu des conditions initiales (B) se

comporte similairement dans le régime des grandes déviations. La partie singulière de la fonction génératrice est donnée par l'intégrale de contour :

$$\widehat{F}_{n,sing}(u) = \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} M_{n,sing}(s) u^s, \qquad (\text{III.210})$$

qui s'obtient en bonne approximation par la méthode du col. Le point col vaut ici :

$$s_s \approx \frac{2n}{(1-u^2)},\tag{III.211}$$

et la partie singulière de la fonction génératrice est donc de l'ordre de :

$$\widehat{F}_{n,sing}(u) \sim \left(1 - \frac{1}{u^2}\right)^{-n}.$$
(III.212)

#### a Position des racines

La position des racines du polynôme  $F_n(x)$  peut être calculée ici, comme les points du plan complexe pour lequel l'estimation ci-dessus reste asymptotiquement constante, soit :

$$\left|1 - \frac{1}{u_0^2}\right| = 1.$$
 (III.213)

La valeur de ces racines est donnée par la paramétrisation par un angle  $0 \le \theta \le 2\pi$  dans le plan complexe :

$$u_0 = \left(1 - e^{-i\theta}\right)^{-1/2},$$
 (III.214)

soit en termes de la variable x :

$$x_0 = \frac{1}{1 - (1 - e^{-i\theta})^{-1/2}}.$$
 (III.215)

#### b Fonction de grande déviation

La partie singulière de la distribution du degré peut s'estimer grâce à l'approximation faite ci-dessus (III.212), et elle est donnée par l'intégrale suivante :

$$f_{sing}(n,k) \sim \oint \frac{\mathrm{d}u}{2\mathrm{i}\pi} \left(\frac{u+1}{u}\right)^k \left(1 - \frac{1}{u^2}\right)^{-n}.$$
 (III.216)

Cette expression se résout à son tour par la méthode du col, où le point-col vaut :

$$u_s = \frac{k - 2n}{k},\tag{III.217}$$

et se met sous la forme :

$$f_{sing}(n,k) \sim \exp\left(-nS(\xi)\right),\tag{III.218}$$

où  $\xi$  est la variable réduite  $\xi = k/n$ , et la fonction de grande déviation prend la forme :

$$S(\xi) = (1 - \xi)\ln(1 - \xi) - (2 - \xi)\ln\left(\frac{2 - \xi}{2}\right).$$
 (III.219)

La fonction de grande déviation présente bien une croissance quadratique pour  $\xi$  petit, et correspond bien à la décroissance en  $\exp(-y^2/4)$  des fonctions d'échelle pour les conditions initiales (A) et (B). La limite  $\xi \to 1$  redonne la probabilité  $f(n, k_{max}(n)) \sim 2^{-n}$  d'obtenir un noeud de degré maximum.

# III.4 Modèle d'attachement préférentiel généralisé

Ce modèle est une généralisation du modèle d'attachement préférentiel de Barabási-Albert. Dans ce modèle, l'intensité de l'attachement préférentiel est modulée par un paramètre c qui amplifie ou estompe son effet.

# III.4.1 Équation maîtresse

A un temps n donné, la probabilité qu'un noeud i précédemment apparu reçoive la connexion du noeud n est proportionnelle à la somme  $k_i(n) + c$ , où  $k_i(n)$  est le degré du noeud i au temps n et c une constante réelle : l'attractivité initiale des noeuds. Selon la valeur de la constante additive c on peut retrouver le modèle avec attachement uniforme décrit section III.2, pour  $c \to \infty$ , ou le modèle de Barabási-Albert décrit section III.3, pour c = 0. La constante c peut également prendre des valeurs négatives, car les noeuds arrivent dans le réseau avec degré 1, et cette attractivité initiale prend ses valeurs dans :

$$-1 < c < \infty. \tag{III.220}$$

La probabilité que le noeud i reçoive la connexion du noeud n au temps n est donc donnée par :

$$p_{n,i} = \frac{k_i(n-1) + c}{Z(n-1)},$$
(III.221)

où  $k_i(n-1)$  est le degré du noeud *i* au moment de l'apparition du noeud *n*, et la fonction de partition Z(n-1) assure la normalisation des probabilités  $p_{n,i}$ :

$$Z(n) = \sum_{i=1}^{n} (k_i(n) + c) = K(n) + cn.$$
(III.222)

On rappelle que K(n) est le nombre de terminaisons de connexions présentes dans le système au temps n, et elle varie selon la condition initiale (III.2) et (III.3). On a donc deux cas :

$$Z^{(A)} = (c+2)n - 2 \tag{III.223}$$

$$Z^{(B)} = (c+2)n - 1. \tag{III.224}$$

#### a Degré du noeud i

La distribution  $f_i(n,k)$  du degré du noeud *i* au temps *n* vérifie l'équation maîtresse :

$$\frac{\mathrm{d}f_i(n,k)}{\mathrm{d}n} = p_{n,i}f_i(n-1,k-1) - p_{n,i}f_i(n-1,k), \qquad (\text{III.225})$$

 $\operatorname{soit}$ 

$$f_i(n,k) = \frac{k-1+c}{Z(n-1)} f_i(n-1,k-1) + \left(1 - \frac{k+c}{Z(n-1)}\right) f_i(n-1,k), \quad (\text{III.226})$$

dont les conditions initiales dépendent du cas considéré, de la même façon que pour les modèles précédents (III.14).

#### b Degré global

La distribution f(n,k) du degré d'un noeud pris au hasard dans le réseau au temps n est obtenue par la moyenne de toutes les distributions du degré de tous les noeuds du réseau :

$$f(n,k) = \frac{1}{n} \sum_{i=1}^{n} f_i(n,k),$$
 (III.227)

et la distribution du degré d'un noeud quelconque vérifie donc la relation de récurrence suivante, obtenue à partir de la relation de récurrence (III.226) :

$$nf(n,k) = (n-1)\frac{k+c-1}{Z(n-1)}f(n-1,k-1) + (n-1)\left(1 - \frac{k+c}{Z(n-1)}\right)f(n-1,k) + \delta_{k,1}, \quad \text{(III.228)}$$

où le  $\delta$  correspond à l'apparition dans le système d'un nouveau noeud, de degré 1, et où les conditions initiales sont données par (III.66).

La distribution asymptotique  $f_{stat}(k)$  pour  $n \to \infty$  du degré d'un noeud quelconque du réseau est la solution indépendante de n de l'équation de récurrence (III.228) et vérifie :

$$\frac{k+2c+2}{c+2}f_{stat}(k) = \frac{k+c-1}{c+2}f_{stat}(k-1) + \delta_{k,1},$$
 (III.229)

et donc :

$$f_{stat}(k) = \frac{(c+1)\Gamma(2c+3)\Gamma(k+c)}{\Gamma(c+1)\Gamma(k+2c+3)}.$$
 (III.230)

Pour de grandes valeurs de k, la distribution stationnaire décroît en loi de puissance :

$$f_{stat}(k) \approx \frac{(c+2)\Gamma(2c+3)}{\Gamma(c+1)}k^{-(c+3)},$$
 (III.231)

d'exposant  $\gamma = c + 3$ . L'attractivité initiale c d'un site est définie telle que c > -1, donc dans le modèle d'attachement préférentiel généralisé l'exposant  $\gamma$  couvre toute la gamme des exposants pour un réseau invariant d'échelle dont la valeur moyenne est définie [1, 55, 56, 57].

#### III.4.2 Fonction génératrice du degré du noeud *i* à temps fini

La distribution  $f_i(n,k)$  du degré  $k_i(n)$  du noeud i est encodée dans la fonction génératrice :

$$F_{n,i}(x) = \sum_{k=1}^{n} f_i(n,k) x^k = \left\langle x^{k_i(n)} \right\rangle.$$
 (III.232)

Le degré du noeud i au temps n peut être défini par récurrence :

$$k_i(n) = k_i(n-1) + I_i(n), \quad n > i,$$
 (III.233)

où  $I_i(n)$  est l'incrément du degré du noeud *i* au temps *n* et prend la valeur :

$$I_i(n) = \begin{cases} 1 & \text{avec probabilité } \frac{k_i(n-1)+c}{Z(n-1)} \\ 0 & \text{sinon} \end{cases}$$
(III.234)

La fonction génératrice  $F_{n,i}$  peut donc s'écrire en fonction de celle de l'étape précédente  $F_{n-1,i}$ :

$$F_{n,i}(x) = \left\langle x^{I_i(n)} x^{k_i(n-1)} \right\rangle$$
  
=  $\left\langle \left( 1 + (x-1) \frac{k_i(n-1) + c}{Z(n-1)} \right) x^{k_i(n-1)} \right\rangle$ , (III.235)

soit :

$$F_{n,i}(x) = F_{n-1,i}(x) + \frac{x-1}{Z(n-1)} \left( cF_{n-1,i}(x) + x \frac{\mathrm{d}F_{n-1,i}(x)}{\mathrm{d}x} \right),$$
(III.236)

où Z est le facteur de normalisation défini en (III.222).

#### a Résolution par transformée de Mellin

Cette équation de récurrence (III.236) peut être résolue par transformation de Mellin (III.136) après un changement de variable (III.122), défini dans la section du modèle de Barabási-Albert. L'équation de récurrence pour la fonction génératrice du degré du noeud i en la variable  $u = \frac{x}{1-x}$  s'écrit :

$$F_{n,i}\left(\frac{u}{u+1}\right) = \left(1 - \frac{c}{Z(n-1)(u+1)}\right)F_{n-1,i}\left(\frac{u}{u+1}\right) - \frac{u}{Z(n-1)}\frac{\mathrm{d}F_{n-1,i}}{\mathrm{d}u}\left(\frac{u}{u+1}\right),\tag{III.237}$$

et pour simplifier les équations par la suite, on pose :

$$\widehat{F}_{n,i}(u) = (1-x)^c F_{n,i}(x).$$
 (III.238)

L'équation de récurrence pour la fonction  $\widehat{F}_{n,i}$  prend la forme :

$$\widehat{F}_{n,i}(u) = \left(1 - \frac{c}{Z(n-1)}\right)\widehat{F}_{n-1,i}(u) - \frac{u}{Z(n-1)}\frac{\mathrm{d}\widehat{F}_{n-1,i}(u)}{\mathrm{d}u}.$$
(III.239)

La transformée de Mellin  $M_{n,i}(s)$  de la fonction génératrice du degré du noeud *i* est définie par :

$$M_{n,i}(s) = \int_0^\infty \widehat{F}_{n,i} u^{-(s+1)} \mathrm{d}u, \qquad (\text{III.240})$$

et vérifie donc l'équation de récurrence :

$$M_{n,i}(s) = \left(1 - \frac{s+c}{Z(n-1)}\right) M_{n-1,i}(s), \quad n \le i$$
(III.241)

L'équation de récurrence présente une dépendance simple en n, contenue uniquement dans la fonction de partition  $Z(n) \propto n$ , et a donc pour solution dépendante des conditions initiales :

$$M_{n,i}^{(A)}(s) = \frac{\Gamma\left(n - \frac{s+c+2}{c+2}\right)\Gamma\left(i - \frac{2}{c+2}\right)}{\Gamma\left(i - \frac{s+c+2}{c+2}\right)\Gamma\left(n - \frac{2}{c+2}\right)}m_{c}(s),$$

$$M_{n,i}^{(B)}(s) = \frac{\Gamma\left(n - \frac{s+c+1}{c+2}\right)\Gamma\left(i - \frac{1}{c+2}\right)}{\Gamma\left(i - \frac{s+c+1}{c+2}\right)\Gamma\left(n - \frac{1}{c+2}\right)}m_{c}(s),$$
(III.242)
(III.243)

Pour c = 0, on retrouve la transformée de Mellin de la fonction génératrice du degré d'un noeud *i* obtenue pour le modèle de Barabási-Albert (III.144).

#### III.4.3 Fonction génératrice à temps fini du degré global

On définit la fonction génératrice  $F_n(x)$  du degré k(n) d'un noeud pris au hasard dans le réseau au temps n, qui peut s'obtenir à partir de la fonction génératrice  $F_{n,i}$ :

$$F_n(x) = \sum_{k=1}^n f(n,k) x^k = \frac{1}{n} \sum_{k=1}^n \sum_{i=1}^n f_i(n,k) x^k = \frac{1}{n} \sum_{i=1}^n F_{n,i}(x).$$
(III.244)

Elle vérifie donc la relation de récurrence suivante, tirée de la relation de récurrence de la fonction génératrice du degré du noeud i (III.236) :

$$nF_n(x) = (n-1)F_{n-1}(x) + (x-1)\frac{n-1}{Z(n-1)}\left(cF_{n-1}(x) + x\frac{\mathrm{d}F_{n-1}(x)}{\mathrm{d}x}\right) + x,\qquad(\mathrm{III.245})$$

le dernier terme x correspondant au noeud n de degré 1 qui apparaît dans le système à l'étape n. Cette relation peut s'obtenir aussi à partir de la relation de récurrence (III.228) pour la distribution du degré f(n, k).

La fonction génératrice  $F_{stat}(x)$  de la distribution stationnaire du degré global  $f_{stat}(k)$  vérifie l'équation différentielle indépendante du temps n:

$$x(1-x)F'_{stat}(x) + (2+2c-cx)F_{stat}(x) = (2+c)x.$$
 (III.246)

La solution cette équation est donnée par :

$$F_{stat}(x) = \lambda(x) \frac{(1-x)^{2+c}}{x^{2c+2}},$$
(III.247)

où  $\lambda(x)$  s'obtient par calcul de l'intégrale suivante :

$$\lambda(x) = (c+2) \int_0^x \mathrm{d}y \frac{y^{2c+2}}{(1-y)^{c+3}},\tag{III.248}$$

qui aboutit à une fonction hypergéométrique. On laisse donc la fonction génératrice du degré stationnaire sous cette forme intégrale :

$$F_{stat}(x) = \frac{(c+2)(1-x)^{2+c}}{x^{2c+2}} \int_0^x \mathrm{d}y \frac{y^{2c+2}}{(1-y)^{c+3}}.$$
 (III.249)

#### a Résolution par transformée de Mellin

On transforme la relation de récurrence (III.245) par le changement de variables (III.122), et elle vérifie donc, pour la fonction  $\widehat{F}_n(u) = (1-x)^c F_n(x)$ :

$$n\widehat{F}_{n}(u) = (n-1)\left(1 - \frac{c}{Z(n-1)}\right)\widehat{F}_{n-1}(u) - (n-1)\frac{u}{Z(n-1)}\frac{\mathrm{d}\widehat{F}_{n-1}(u)}{\mathrm{d}u} + u(u+1)^{c-1}.$$
(III.250)

Cette expression se simplifie encore par transformation de Mellin, où la transformée  $M_n(s)$  de la fonction génératrice du degré global est définie par :

$$M_n(s) = \int_0^\infty \widehat{F}_n u^{-(s+1)} \mathrm{d}u, \qquad (\text{III.251})$$

et vérifie donc la relation de récurrence :

$$nM_n(s) = (n-1)\left(1 - \frac{s+c}{Z(n-1)}\right)M_{n-1}(s) + m_c(s), \quad n \le 1,$$
(III.252)

où  $m_c(s)$  est la transformée de Mellin de la condition initiale  $\widehat{F}_{i,i}(u) = x(1-x)^c$ :

$$m_{c}(s) = \int_{0}^{\infty} \frac{u^{-s}}{(1+u)^{c+1}} du$$
  
=  $\int_{0}^{1} (1-x)^{s+c-1} x^{-s} dx$  (III.253)  
=  $\frac{\Gamma(1-s)\Gamma(s+c)}{\Gamma(c+1)}$ .

Ces intégrales ne sont définies que si -c < Re(s) < 1. Le contour sur lequel s'intègre la transformée de Mellin inverse, définie en (III.137) est donc contraint à cette bande de l'espace complexe.

La transformée de Mellin pour le degré global se présente sous la forme d'une équation différentielle discrète en n non homogène. On suit les lignes de raisonnement présentées pour le modèle de Barabási-Albert. La solution homogène est ici :

$$M_n^{(H)}(s) = \lambda(s) \frac{2\Gamma\left(n - 1 - \frac{s+c}{2}\right)}{n\Gamma\left(1 - \frac{s+c}{2}\right)},$$
(III.254)

tandis que la solution particulière est :

$$M(s) = \frac{Z(n)m_c(s)}{(s+2c+2)n}.$$
 (III.255)

La solution finale dépend des conditions initiales qui interviennent dans le calcul de la constante  $\lambda$  :

$$M_2^{(A)}(s) = M_1^{(B)}(s) = m_c(s)$$
(III.256)

$$\Rightarrow \lambda^{(A)}(s) = m_c(s) \frac{1+s+c}{2+s+2c} \tag{III.257}$$

$$\Rightarrow \lambda^{(B)}(s) = m_c(s) \frac{(1+s+c)(c+s)}{4(2+s+2c)},$$
 (III.258)

et la transformée de Mellin de la distribution du degré global au temps n s'écrit, pour la condition initiale (A):

$$M_{n}^{(A)}(s) = \frac{m_{c}(s)}{(2+2c+s)n} \left( (c+2)n - 2 + 2(1+c+s) \frac{\Gamma\left(n - \frac{s+c+2}{c+2}\right)\Gamma\left(\frac{2c+2}{c+2}\right)}{\Gamma\left(1 - \frac{s}{c+2}\right)\Gamma\left(n - \frac{2}{c+2}\right)} \right),$$
(III.259)

et pour la condition initiale (B) :

$$M_{n}^{(B)}(s) = \frac{m_{c}(s)}{(2+2c+s)n} \left( (c+2)n - 1 + (1+c+s) \frac{\Gamma\left(n - \frac{s+c+1}{c+2}\right)\Gamma\left(\frac{c+1}{c+2}\right)}{\Gamma\left(\frac{1-s}{c+2}\right)\Gamma\left(n - \frac{1}{c+2}\right)} \right).$$
(III.260)

Ces équations se simplifient en (III.153) pour c = 0, *i.e.* pour le modèle BA.

# III.4.4 Distribution du degré d'un noeud *i*

#### a Premiers moments

À partir de la transformée de Mellin (III.242) pour le degré d'un noeud i, on peut obtenir tous les coefficients de l'expansion en série de la fonction génératrice des moments du degré du noeud i à un temps fini n.

En particulier, on obtient, pour la valeur moyenne du degré du noeud i:

$$\left\langle k_i^{(A)}(n) \right\rangle = (c+1) \frac{\Gamma\left(n - \frac{1}{c+2}\right) \Gamma\left(i - \frac{2}{c+2}\right)}{\Gamma\left(i - \frac{1}{c+2}\right) \Gamma\left(n - \frac{2}{c+2}\right)} - c, \qquad (\text{III.261})$$

$$\left\langle k_i^{(B)}(n) \right\rangle = (c+1) \frac{\Gamma(n) \Gamma\left(i - \frac{1}{c+2}\right)}{\Gamma(i) \Gamma\left(n - \frac{1}{c+2}\right)} - c \tag{III.262}$$

et pour la variance autour de la moyenne :

$$\operatorname{var}(k_{i}^{(A)}(n)) = (c+1)(c+2) \frac{\Gamma(n) \Gamma\left(i - \frac{2}{c+2}\right)}{\Gamma(i) \Gamma\left(n - \frac{2}{c+2}\right)} - \left\langle k_{i}^{(A)}(n) \right\rangle^{2} - (2c+1) \left\langle k_{i}^{(A)}(n) \right\rangle - c(c+1),$$
(III.263)  
$$\operatorname{var}(k_{i}^{(B)}(n)) = (c+1)(c+2) \frac{\Gamma\left(n + \frac{1}{c+2}\right) \Gamma\left(i - \frac{1}{c+2}\right)}{\Gamma\left(i + \frac{1}{c+2}\right) \Gamma\left(n - \frac{1}{c+2}\right)} - \left\langle k_{i}^{(B)}(n) \right\rangle^{2} - (2c+1) \left\langle k_{i}^{(B)}(n) \right\rangle - c(c+1).$$

#### b Régime d'échelle

On s'intéresse plus particulièrement au régime d'échelle, lorsque n et i, les temps d'observation et de naissance, sont tous les deux grands et de même ordre, et on développe les résultats précédents selon le rapport z = n/i:

$$\langle k_i(n) \rangle \approx (c+1)z^{1/(c+2)} - c \tag{III.264}$$

$$\operatorname{var}(k_i(n) \approx (c+1)z^{1/(c+2)} \left(z^{1/(c+2)} - 1\right)$$
 (III.265)

On peut également obtenir, dans ce régime, une formule explicite pour la distribution du degré du noeud i. La transformée de Mellin pour le degré du noeud i prend la forme compacte :

$$M_{n,i}(s) \approx z^{-\frac{s+c}{c+2}} m_c(s). \tag{III.266}$$

La fonction génératrice du degré du noeud i s'obtient par la transformée de Mellin inverse :

$$\widehat{F}_{n,i}(u) \approx \oint \frac{\mathrm{d}s}{2\mathrm{i}\pi} z^{-\frac{s+c}{c+2}} m_c(s) u^s, \qquad (\text{III.267})$$

que l'on obtient par le calcul des résidus sur les pôles entiers positifs de l'intégrande :

$$\widehat{F}_{n,i}(u) \approx \sum_{j>0} (-1)^j u^j z^{-\frac{c+j}{c+2}} \frac{\Gamma(j+c)}{\Gamma(j)\Gamma(1+c)},$$
(III.268)

qui donne finalement :

$$\widehat{F}_{n,i}(u) \approx \frac{x}{\left(x + z^{1/(c+2)}(1-x)\right)^{c+1}}.$$
 (III.269)

Le développement en série entière de la fonction génératrice autour de x = 0 a pour coefficients la distribution du degré  $k_i(n)$  du noeud *i* au temps *n* en fonction du rapport z = n/i dans le régime d'échelle :

$$f_i(n,k) \approx \frac{\Gamma(k+c)}{\Gamma(k)\Gamma(1+c)} z^{-\frac{c+1}{c+2}} \left(1 - z^{-1/(2+c)}\right)^{k-1}.$$
 (III.270)

Dans ce régime, on retrouve bien les résultats précédents, à savoir que le degré du noeud i est distribué selon une loi de Poisson dans le modèle AU  $c \to \infty$  et selon une série géométrique pour le modèle BA c = 0.

#### III.4.5 Distribution du degré global

À partir de la transformée de Mellin de la fonction génératrice du degré global pour les conditions initiales (A) (III.259) et (B) (III.260), on peut calculer toute la distribution de probabilité du degré global. La probabilité qu'un noeud pris au hasard soit de degré 1 est ainsi le résidu de  $M_n(s)$  à s = 1. Cette probabilité dépend des conditions initiales, par exemple on a pour les conditions initiales (A) :

$$f^{(A)}(n,1) = \frac{1}{(2c+3)n} \left( (c+2)n - 2 + 2(2+c) \frac{\Gamma\left(n - \frac{c+3}{c+2}\right)\Gamma\left(\frac{2c+2}{c+2}\right)}{\Gamma\left(\frac{c+1}{c+2}\right)\Gamma\left(n - \frac{2}{c+2}\right)} \right)$$
  
$$= \frac{1}{(2c+3)} \left( (c+2) - \frac{2}{n} + 2(2+c) \frac{\Gamma\left(\frac{2c+2}{c+2}\right)}{\Gamma\left(\frac{c+1}{c+2}\right)} n^{-\frac{2c+3}{c+2}} + \dots \right),$$
 (III.271)

pour  $n \ge 2$ , et pour les conditions initiales (B) :

$$f^{(B)}(n,1) = \frac{(c-2)n-1}{(2c+3)n},$$
(III.272)

pour  $n \ge 2$ . Les formules en fractions rationnelles à k = 1 et k = 2 pour la condition initiale (B) sont une exception dans la distribution du degré. Les termes généraux présentent une correction en  $n^{-(2c+3)/(c+2)}$ , quelque soit la condition initiale. On présente donc dans la suite la distribution du degré à temps fini en fonction de la distribution stationnaire et de la fonction d'échelle.

#### a Valeur de transition

Dans ce modèle d'attachement préférentiel généralisé, l'échelle de transition  $k_{\star}(n)$  d'un réseau de taille *n* donnée par l'argument de la valeur typique des premiers noeuds se comporte en :

$$k_{\star}(n) \sim \langle k_1(n) \rangle \approx (c+1) \frac{\Gamma\left(\frac{c}{c+2}\right)}{\Gamma\left(\frac{c+1}{c+2}\right)} n^{\frac{1}{c+2}}.$$
 (III.273)

L'argument de la statistique des extrêmes donne :

$$n \int_{k_{\star}(n)}^{\infty} \mathrm{d}k f_{stat}(k) \sim 1$$

$$k_{\star}(n) \sim n^{\frac{1}{c+2}}.$$
(III.274)

Les deux résultats sont bien cohérents, et on étudie donc la transition de la distribution à temps fini du degré global de part et d'autre de  $k_{\star}$ .

#### b Fonction d'échelle

La distribution du degré global à temps fini est obtenue par correction de la distribution stationnaire  $f_{stat}$  par la fonction d'échelle multiplicative  $\Phi_c$ :

$$f(n,k) \approx f_{stat}(k)\Phi_c(y), \quad y = \frac{k}{n^{1/(c+2)}}.$$
 (III.275)

On reproduit, pour obtenir la fonction d'échelle, les différentes étapes du raisonnement suivi pour le modèle de Barabási-Albert, section III.3. Les fonctions d'échelle sont explicitement données par le développement en série sur les résidus des transformées de Mellin inverses :

$$\Phi_c^{(A)}(y) = 1 + \frac{2\Gamma\left(\frac{2c+2}{c+2}\right)}{(c+2)\Gamma(2c+3)}y^{2c+3}\sum_{j\geq 0}\frac{(-1)^j(j+c+2)y^j}{(j+2c+3)j!\Gamma\left(1-\frac{j+1}{c+2}\right)}$$
(III.276)

$$\Phi_c^{(B)}(y) = 1 + \frac{\Gamma\left(\frac{c+1}{c+2}\right)}{(c+2)\Gamma(2c+3)} y^{2c+3} \sum_{j\geq 1} \frac{(-1)^j (j+c+2)y^j}{(j+2c+3)j!\Gamma\left(\frac{j}{c+2}\right)}.$$
 (III.277)

Ces expressions sont complexes à calculer analytiquement en l'état. On s'intéresse donc plutôt à leurs dérivées  $\Phi'_c(y)$ , dans lesquelles le facteur 1/(j + 2c + 3) se simplifie, et où l'on peut appliquer les identités suivantes [58] :

$$\sum_{j\geq 0} \frac{(-1)^j y^j}{j! \Gamma\left(1 - \frac{j+1}{c+2}\right)} = (c+2) \oint \frac{\mathrm{d}z}{2\mathrm{i}\pi} \exp\left(-yz + z^{c+2}\right)$$
(III.278)

$$\sum_{j\geq 1} \frac{(-1)^j y^j}{j! \Gamma\left(-\frac{j}{c+2}\right)} = y \oint \frac{\mathrm{d}z}{2\mathrm{i}\pi} \exp\left(-yz + z^{c+2}\right).$$
(III.279)

Les dérivées de la fonction d'échelle s'expriment donc en termes d'intégrales de contour :

$$\Phi_c^{\prime(A)}(y) = \frac{2\Gamma\left(\frac{2c+2}{c+2}\right)}{\Gamma(2c+3)} y^{2c+2} \oint \frac{\mathrm{d}z}{2\mathrm{i}\pi} (c+2-yz) \exp\left(-yz+z^{c+2}\right)$$
(III.280)

$$\Phi_c^{\prime(B)}(y) = \frac{\Gamma\left(\frac{c+1}{c+2}\right)}{(c+2)\Gamma(2c+3)} y^{2c+3} \oint \frac{\mathrm{d}z}{2\mathrm{i}\pi} (c+3-yz) \exp\left(-yz+z^{c+2}\right).$$
(III.281)

À partir de ces expressions, on peut déterminer le comportement asymptotique de la fonction d'échelle. Pour les y petits, lorsque  $\Phi_c(y)$  est proche de 1 :

$$\Phi_c^{(A)}(y) = 1 + \frac{2\Gamma\left(\frac{2c+2}{c+2}\right)}{\Gamma\left(\frac{c+1}{c+2}\right)\Gamma(2c+4)}y^{2c+3} + \dots$$
(III.282)

$$\Phi_c^{(B)}(y) = 1 + \frac{(c+3)}{2(c+2)^3\Gamma(2c+3)}y^{2c+4} + \dots$$
(III.283)

Pour les y grands la fonction d'échelle décroît vers 0 selon :

$$\Phi_c^{(A)}(y) \approx 2(c+2)\lambda(c)\Gamma\left(\frac{2c+2}{c+2}\right)\Psi(y)$$
(III.284)

$$\Phi_c^{(B)}(y) \approx \lambda(c) \Gamma\left(\frac{c+1}{c+2}\right) \Psi(y) y, \qquad (\text{III.285})$$

où la dépendance en y est contenue dans la fonction

$$\Psi(y) = y^{2c+3} - c/2(c+1) \exp\left[-(c+1)\left(\frac{y}{c+2}\right)^{\frac{c+2}{c+1}}\right],$$
 (III.286)

et la constante  $\lambda$  dépend de la valeur de c :

$$\lambda(c)^{-1} = \sqrt{2\pi(c+1)}(c+2)^{\frac{2c+3}{2(c+1)}}\Gamma(2c+3).$$
(III.287)

On représente, figure III.7, la fonction d'échelle tracée pour différentes valeurs de c et pour chaque condition initiale. On retrouve les comportements attendus pour les fonctions d'échelle. Un résultat similaire a été établi pour la conditions initiale (A) par Waclaw et Sokolov [54]. Pour des valeurs finies de c, la fonction passe par un maximum entre les deux plateaux, qui croît si c décroît. La position du maximum et de la transition subséquente se rapproche de l'axe des ordonnées à mesure que  $c \to -1$ . On retrouve bien, en particulier, le comportement obtenu pour le modèle de Barabási-Albert à c = 0. Pour  $c \to \infty$ , le pic s'estompe et la transition est directe entre un plateau à 1 pour  $y \ll 1$  et un à 0 pour  $y \gg 1$ , comportement attendu pour le modèle AU.

#### III.4.6 Comportement de la fonction d'échelle

#### a Valeurs typiques

On s'intéresse à la dépendance en c de la position du maximum et de sa largeur. On considère donc la fonction d'échelle comme une distribution de probabilité à partir de laquelle on peut définir des pseudo-moments d'ordre p:

$$\mu_p = \int_0^\infty p \Phi_c(y) y^{p-1} dy = -\int_0^\infty \Phi'_c(y) y^p dy.$$
 (III.288)

Après introduction de l'expression obtenue pour la dérivée de la fonction d'échelle et inversion des intégrales, on obtient les expressions suivantes, respectivement pour les conditions initiales (A) et (B):

$$\mu_p^{(A)} = 2(p+c+1) \frac{\Gamma\left(\frac{2c+2}{c+2}\right) \Gamma(p+2c+3)}{\Gamma(2c+3)} \oint \frac{\mathrm{d}z}{2\mathrm{i}\pi} \exp\left(z^{c+2}\right) z^{-(p+2c+3)}, \quad (\text{III.289})$$

$$\mu_p^{(B)} = (p+c+1) \frac{\Gamma\left(\frac{c+1}{c+2}\right) \Gamma(p+2c+4)}{(c+2)\Gamma(2c+3)} \oint \frac{\mathrm{d}z}{2\mathrm{i}\pi} \exp\left(z^{c+2}\right) z^{-(p+2c+4)}.$$
 (III.290)

Ces expressions se résolvent par la méthode du col, aux points

$$z_s^{(A)} = \left(\frac{p+2c+3}{2+c}\right)^{\frac{1}{c+2}},\tag{III.291}$$

$$z_s^{(B)} = \left(\frac{p+2c+4}{2+c}\right)^{\frac{1}{c+2}},\tag{III.292}$$

et on obtient donc une forme explicite pour les pseudo-moments des fonctions d'échelle :

$$\mu_p^{(A)} = \frac{(p+c+1)}{(c+1)} \frac{\Gamma\left(\frac{3c+4}{c+2}\right)\Gamma(p+2c+3)}{\Gamma(2c+3)\Gamma\left(\frac{p+3c+4}{c+2}\right)}$$
(III.293)

$$\mu_p^{(B)} = \frac{\Gamma\left(\frac{c+1}{c+2}\right)\Gamma(p+2c+3)}{\Gamma(2c+3)\Gamma\left(\frac{p+c+1}{c+2}\right)}.$$
(III.294)



FIGURE III.7 – Fonction d'échelle  $\Phi(y)$  en fonction de la variable redimensionnée  $y = k/n^{1/(c+2)}$ pour différentes valeurs du paramètre c = -1/2, c = 0 et c = 2, pour chacune des conditions initiales. Les fonctions sont obtenues par calcul numérique de la relation de récurrence pour la distribution du degré f(n, k).

Ces résultats permettent d'analyser la transition entre les valeurs 1 à  $y \sim 0$  et 0 pour  $y \to \infty$ . Sur la figure III.7 on observe une transition continue entre le comportement de la fonction d'échelle à de petites valeurs de c et son comportement à de grandes valeurs de c. Pour c petit, la fonction d'échelle présente un maximum qui se déplace vers l'axe des ordonnées et une fonction très étroite autour de ce maximum. Pour c grand, la transition est brusque entre les deux valeurs extrêmes de  $\Phi(y)$ . On va donc s'intéresser au comportement de la fonction d'échelle pour des modèles proches du modèle d'attachement uniforme  $(c \to \infty)$  et pour les petites valeurs de  $c (c \to -1)$ .

#### *b* Grandes valeurs de *c*

La valeur de transition  $y_{\star}$  de y pour laquelle la fonction d'échelle varie brusquement entre 1 et 0 peut être estimée par le pseudo-moment  $\mu_1$ , car pour les grandes valeurs de c, la fonction d'échelle est pratiquement  $\Phi(y) = 1$  pour  $y < y_{\star}$  et  $\Phi(y) = 0$  pour  $y > y_{\star}$ . La valeur de la transition vérifie  $y_{\star} \sim \mu_1$ , où :

$$\mu_1^{(A)} = 2c + 2(\gamma + 1) + \dots \tag{III.295}$$

$$\mu_1^{(B)} = 2c + 2(\gamma + 1) + 1 + \dots$$
(III.296)

Avec ces résultats et le changement de variable  $k \approx y \left(1 + \frac{\ln n}{c}\right)$ , on retrouve bien pour  $c \to \infty$  l'estimation de la valeur de transition obtenue pour le modèle d'attachement uniforme (III.71). La position de la valeur de la transition croît en fonction de c comme 2c, pour des valeurs assez grandes de c.

On cherche également à estimer la largeur de la région de transition, dont le carré peut être donné par :

$$\sigma^2 = \mu_2 - \mu_1^2, \tag{III.297}$$

qui a, au voisinage de  $c \to \infty$ , le comportement :

$$\sigma^{(A)2} = 2c + 2(2\gamma + 1) - \frac{2\pi^2}{3}) + \dots$$
(III.298)

$$\sigma^{(B)2} = 2c + 2(2\gamma + 1) + 1 - \frac{2\pi^2}{3}) + \dots, \qquad \text{(III.299)}$$

avec lequel on retrouve bien le comportement en  $2 \ln n$  du degré de transition (III.78) pour  $c \to \infty$ . Cette quantité croît avec c comme 2c, et confirme l'impression que la transition s'adoucit pour de plus grandes valeurs de c sur les représentations en fonction de y comme sur le graphe représenté figure III.7.

#### c Petites valeurs de c

La fonction d'échelle présente un pic qui s'élève, s'étrécit et se décale vers la gauche pour c décroissant. La position  $\langle y \rangle$  de ce pic est donnée, pour  $c \to -1$ , par le rapport des pseudomoments :

$$\langle y \rangle = \frac{\mu_2}{2\mu_1},\tag{III.300}$$

dont le développement en série autour de c = -1 donne :

$$\langle y \rangle^{(A)} = 1 + \left(\frac{3}{2} - \gamma\right)(c+1) + \dots$$
 (III.301)

$$\langle y \rangle^{(A)} = 1 + (2 - \gamma) (c + 1) + \dots$$
 (III.302)

On vérifie sur la figure III.7 que la position du maximum de la fonction d'échelle se trouve bien au voisinage des valeurs calculées ici. La position du pic converge donc vers la valeur finie  $\langle y \rangle_{max} = 1$ , qui correspond à k = n.

On s'intéresse également à la largeur de ce pic, et comment elle décroît pour  $c \to -1$ . La largeur du pic est donnée par la racine carrée de la pseudo-variance :

$$\tilde{\sigma}^2 = \frac{4\mu_1\mu_3 - 3\mu_2^2}{12\mu_1^2},\tag{III.303}$$

dont les valeurs, en fonction des conditions initiales (A) et (B), sont données pour les petites valeurs de c par :

$$\tilde{\sigma}^{(A)2} = \frac{1}{6} \left( 5(c+1) + (19 - 10\gamma - \pi^2)(c+1)^2 \right) + \dots$$
(III.304)

$$\tilde{\sigma}^{(B)2} = \frac{1}{6} \left( 4(c+1) + (22 - 8\gamma - \pi^2)(c+1)^2 \right) + \dots$$
(III.305)

L'aire sous le pic se comporte comme  $\mu_1$ , soit :

$$\mu_1^{(A)} = \frac{1}{(c+1)} + 2 - \gamma + \dots$$
(III.306)

$$\mu_1^{(B)} = \frac{1}{(c+1)} + 3 - \gamma + \dots, \qquad (\text{III.307})$$

c'est-à-dire que le pic s'étrécit comme  $\sqrt{c+1}$  lorsque  $c \to 1$ , et sa hauteur croît comme  $\mu_1/\tilde{\sigma} \sim (1+c)^{-3/2}$ .

Le modèle correspondant à c = -1 est en effet un réseau dans lequel seul le premier noeud peut recevoir des connexions au cours de toute l'histoire du réseau, et où seules deux valeurs du degré existent : k = 1 pour tous les noeuds, sauf le premier pour lequel k = n(-1).

#### III.4.7 Grandes déviations

On répète le raisonnement mené dans le cadre du modèle de Barabási-Albert III.3 pour déterminer la probabilité de trouver un noeud dont le degré est de l'ordre de la taille du réseau.

La partie régulière de la distribution de probabilité décroît comme :

$$f_{stat}(n) \approx \frac{(c+2)\Gamma(2c+3)}{\Gamma c+1} n^{-(c+3)},$$
 (III.308)

et la partie singulière est calculée à partir de la partie singulière de la transformée de Mellin :

$$M_{n,sing}(s) = \frac{2}{n} \frac{(s+c+1)m_c(s)}{(s+c+2)} \frac{\Gamma\left(n - \frac{s+c+2}{c+2}\right)\Gamma\left(\frac{2c+2}{c+2}\right)}{\Gamma\left(1 - \frac{s}{c+2}\right)\Gamma\left(n - \frac{2}{c+2}\right)},$$
(III.309)

pour la condition initiale (A), sous-entendue jusqu'à mention du contraire. On généralise l'approximation (III.212) effectuée pour le modèle de Barabási-Albert de l'intégrale de contour contenue dans la transformée de Mellin inverse, et la fonction génératrice prend ici la forme :

$$\widehat{F}_{n,sing}(u) \sim \left(1 - \frac{1}{(-u)^{c+2}}\right)^{-n}.$$
 (III.310)

Cette forme de la fonction génératrice permet de déterminer la position des racines complexes de la fonction génératrice, vérifiant la condition :

$$\left|1 - \frac{1}{(-u_0)^{c+2}}\right| = 1,\tag{III.311}$$

d'où l'on obtient les positions des racines de la fonction génératrice, paramétrisées par l'angle  $0 \le \theta \le 2\pi$ :

$$x_0 = \frac{1}{1 - (1 - e^{-i\theta})^{1/(c+2)}}.$$
 (III.312)

L'estimation de la fonction génératrice permet de donner la forme de la fonction de grande déviation de la distribution du degré, qui se comporte dans ce régime  $k \sim n$  comme ;

$$f(n,k) \sim \exp(-nS(\xi)), \tag{III.313}$$

avec  $\xi$  la variable redimensionnée :  $\xi = k/n$ . La fonction de grande déviation prend donc la forme :

$$S(\xi) = \ln\left(\nu^{c+2} - 1\right) - \frac{c+2}{\nu^{c+2} - 1}\left((\nu - 1)\ln(\nu - 1) + \left(\nu^{c+1} - 1\right)\nu\ln\nu\right),\tag{III.314}$$

où on a introduit une seconde variable pour simplifier les expressions, qui vérifie :

$$\xi = \frac{(c+2)(\nu-1)}{\nu^{c+2} - 1}, \quad -1 < \nu < \infty.$$
(III.315)

La fonction de grande déviation varie entre un comportement en loi de puissance lorsque  $\xi \to 0$ , *i.e.*  $\nu \to \infty$ :

$$S(\xi) \approx (c+1) \left(\frac{\xi}{c+2}\right)^{\frac{c+2}{c+1}},$$
 (III.316)

et un maximum, atteint pour  $\xi = 1$ , *i.e.*  $\nu = 1$ :

$$S(\xi = 1) = \ln(c+2).$$
 (III.317)

Ces deux comportements correspondent aux résultats obtenus de décroissance superexponentielle des fonctions d'échelle  $\Phi(y)$  lorsque  $k \ll n$ , c'est-à-dire à  $\xi$  petit, tandis que la valeur de S(1) assure une décroissance exponentielle de la probabilité qu'un noeud ait le degré maximum :

$$f(n, k_{max}) \sim (c+2)^{-n},$$
 (III.318)

où  $k_{max} = n - 1$  pour la condition initiale (A) et  $k_{max} = n$  pour la condition initiale (B).

# III.5 Conclusion

Nous avons étudié méthodiquement dans ce chapitre les effets de taille finie dans des réseaux en croissance avec des règles de construction différentes. Les effets de taille finie et de coupure des distributions avaient déjà été étudiées par de nombreux auteurs [10, 11, 52, 54] pour chaque modèle.

Nous avons traité le modèle d'attachement uniforme, le modèle d'attachement préférentiel de Barabási-Albert et le modèle d'attachement préférentiel généralisé avec la même méthodologie. Dans chaque cas, nous avons calculé la distribution à temps fini du degré d'un noeud donné et d'un noeud pris au hasard dans le réseau.

Les caractéristiques de ces distributions ont été déterminées dans le régime d'échelle pour deux conditions initiales (A) et (B) qui diffèrent par la symétrie des premiers noeuds. Nous avons précisé l'impact du choix de la condition initiale dans la statistique des réseaux en croissance à temps fini.

Dans cette étude nous avons mis en évidence et déterminé analytiquement la valeur  $k_{\star}(n)$ du degré de transition entre le régime stationnaire et le régime des grandes déviations dans les réseaux finis. Cette valeur divise l'intervalle de définition des distributions en trois régimes :

- Le régime stationnaire, lorsque  $k \ll k_{\star}$ , pour lequel la distribution du degré est donnée en bonne approximation par la distribution stationnaire  $f_{stat}$ . Un noeud, dont le degré se trouve dans cet intervalle de valeurs, ne ressent pas la taille finie du réseau.
- Le régime de transition, lorsque k ~ k<sub>\*</sub>. La distribution du degré à temps fini, pour ces valeurs, suit une fonction d'échelle Φ multiplicative, qui dépend des conditions initiales du réseau. Cette fonction assure la continuité de la distribution du degré à temps fini entre le régime stationnaire pour lequel elle vaut Φ(0) = 1 et le régime des grandes déviations. Aucun noeud d'un réseau de taille n ne peut atteindre des valeurs du degré k > n, et donc Φ(∞) = 0. Entre ces deux valeurs limites, la fonction d'échelle atteint un maximum et prend des valeurs supérieures à 1, dans lesquelles s'accumule le poids de la distribution du degré qui, dans un réseau de taille infini, correspondrait à la probabilité qu'un noeud soit de degré k > k<sub>\*</sub>.
- Le régime de grandes déviations pour  $k \gg k_{\star}$ , dans lequel la distribution du degré décroît exponentiellement vers 0. Nous avons étudié la forme de la fonction des grandes déviations, et montré qu'elle ne dépend pas de la condition initiale.

L'analyse de ces trois régime a été menée sur trois modèles de réseaux aléatoires en croissance, dont le dernier, le modèle d'attachement préférential généralisé, caractérisé par une attractivité initiale c des noeuds, permet d'extrapoler les résultats entre le modèle d'attachement uniforme pour  $c \to \infty$  et au-delà du modèle de Barabási-Albert (c = 0). La dépendance de l'exposant de la loi de puissance de la distribution stationnaire dans le paramètre, et la dépendance des distributions de temps fini dans les conditions initiales confirment que le comportement d'échelle des réseaux en croissance n'est pas universel.

Nous utilisons dans le chapitre suivant les résultats obtenus ici pour étudier les caractéristiques des extrêmes dans les réseaux en croissance.

# Chapitre |V|

# Statistique des extrêmes dans les réseaux en croissance

Dans l'étude à temps fini des réseaux en croissance, on a constaté que les effets de taille jouent principalement au voisinage d'une échelle de transition, qui est de l'ordre du plus grand degré typique du système. Nous avons également observé que ce plus grand degré est de l'ordre du degré moyen des premiers noeuds. Une question naturelle se pose quand à la validité de cette hypothèse. On s'intéresse, donc dans ce chapitre, à la statistique des extrêmes dans les réseaux en croissance à temps fini.

Dans un réseau en croissance, le nombre de connexions entrant dans le système à chaque unité de temps est fini, et un seul des noeuds déjà présent peut voir son degré augmenter. La croissance du degré d'un noeud donné se fait donc aux dépends du reste du réseau et on peut donc, à la suite de Luczak et Erdös [59] considérer l'acquisition de nouvelles connexions comme une course au degré, dans laquelle le noeud le plus connecté est le meneur. On s'intéresse à la statistique de ce meneur.

Comment son identité change-t-elle au cours du temps? Quelle est la probabilité pour un meneur de rester dominant au cours du temps?

Nous avons étudié ces questions pour les trois modèles décrits dans le chapitre précédent :

• Attachement uniforme : la probabilité d'attachement du nouveau noeud ne dépend pas de l'identité du noeud-cible, au temps n elle est équivalente pour tous les n-1 noeuds du réseau déjà présents. La probabilité  $p_{i,n}$  que le noeud n s'attache à un noeud i précédent est :

$$p_{n,i} = \frac{1}{n-1}.\tag{IV.1}$$

• Barabási-Albert : la probabilité d'attachement  $p_{n,i}$  du nouveau noeud n au noeud i déjà présent dépend du degré  $k_i(n-1)$  du noeud i avant ajout du noeud n :

$$p_{n,i} = \frac{k_i(n-1)}{Z(n-1)},$$
 (IV.2)

où  $Z(n-1) = \sum_{i=1}^{n-1} k_i(n-1)$  assure la normalisation des probabilités.

• Attachement préférentiel généralisé : la probabilité d'attachement est préférentielle comme dans le modèle BA, et chaque noeud possède une attractivité initiale  $-1 < c < \infty$ . La

probabilité  $p_{n,i}$  d'attachement du nouveau noeud n au noeud i est donnée par :

$$p_{n,i} = \frac{k_i(n-1) + c}{Z(n-1)},$$
(IV.3)

où le facteur de normalisation est  $Z(n-1) = \sum_{i=1}^{n-1} (k_i(n-1) + c)$  pour ce modèle.

Pour chacun des modèles considérés, nous allons, après un court rappel sur la statistique des extrêmes de variables aléatoires discrètes, considérer un certain nombre de quantités typiques des extrêmes :

- le degré maximum  $k_{max}(n)$  du réseau au temps n,
- le nombre de co-meneurs C(n), c'est-à-dire le nombre de noeuds portant le degré maximum à un temps n,
- le nombre de périodes de prépondérance  $\mathcal{L}(n)$  d'un meneur sur la connectivité du réseau jusqu'au temps n,
- l'indice I(n) du meneur au temps n, c'est-à-dire l'indice du noeud de degré maximum,
- la persistance du meneur, *i.e.* la probabilité S(n) que le premier meneur soit encore meneur au temps n et
- le nombre de meneurs distincts D(n) pendant l'histoire du résau jusqu'au temps n.

Nous étudierons pour chaque modèle l'impact des conditions initiales sur les quantités étudiées :

- la condition initiale (A) correspond à la condition initiale de dimère, dans laquelle les deux premiers noeuds du réseau sont équivalents. Ce cas a été étudié par Krapivsky et Redner dans le cadre d'un travail préliminaire [60].
- la condition initiale (B) aboutit à la construction d'un arbre enraciné. Pour cette condition initiale, tous les noeuds sont statistiquement différents.

Les conditions initiales ont été définies dans le chapitre précédent et sont résumées figure III.1. On précise ici que la condition initiale (A) est modifiée pour les simulations, de telle sorte que la symétrie entre les deux premiers noeuds est brisée, c'est-dire que l'on impose le premier noeud comme le meneur, de degré  $k_1(3) = 2$  à la troisième étape.

# IV.1 Statistique des extrêmes pour les variables entières

Dans un réseau classique, tous les noeuds ont un degré tiré indépendemment d'une même distribution  $f_k$ . L'étude de la statistique des plus grands degrés et du comportement des meneurs repose donc sur la statistique des extrêmes d'un ensemble de n variables aléatoires entières indépendantes, tirées d'une même distribution. Pour introduire les quantités étudiées dans ce chapitre, nous allons considérer ici la statistique des extrêmes d'un ensemble de n variables aléatoires tirées d'une distribution géométrique discrète :

$$f(k) = (1-a)a^{k-1}, \quad k \ge 1,$$
 (IV.4)

où le paramètre a prend ses valeurs dans l'intervalle (0, 1). On s'intéresse au degré maximum  $k_{max}(n) = \max(k_1, \ldots, k_n)$  et au nombre C(n) de co-meneurs, c'est-à-dire le nombre de noeuds ayant le degré maximum.

#### IV.1.1 Distribution du plus grand degré

On introduit la distribution cumulative du degré :

$$\mathcal{F}(k) = P\left(k_i \ge k\right) = \sum_{j=k}^{\infty} f(j).$$
(IV.5)

Les variables  $k_i$  étant entières, la distribution du degré vérifie, en termes de la distribution cumulative :

$$f(k) = \mathcal{F}(k) - \mathcal{F}(k+1), \qquad (\text{IV.6})$$

et la distribution cumulative du degré maximum est donc donnée par :

$$P(k_{max}(n) < k) = (1 - \mathcal{F}(k))^n.$$
 (IV.7)

Les variables sont entières, et la distribution  $\phi_n(k) = \operatorname{Prob}(k_{max}(n) = k)$  du degré maximum parmi n noeuds est donc la dérivée discrète de la distribution cumulative :

$$\phi_n(k) = (1 - \mathcal{F}(k+1))^n - (1 - \mathcal{F}(k))^n,$$
 (IV.8)

qui peut s'exprimer en fonction de la distribution du degré f(k) :

$$\phi_n(k) = \sum_{c=1}^n \binom{n}{c} f(k)^c (1 - \mathcal{F}(k))^{n-c}.$$
 (IV.9)

L'entier c, sur lequel la somme est effectuée, est le nombre de noeuds de degré maximum, *i.e.* le nombre de co-meneurs dans le réseau.

Les formules ci-dessus s'appliquent à toute distribution f(k). Dans le cas de la distribution géométrique (IV.4), la distribution cumulative du degré s'écrit :

$$\mathcal{F}(k) = a^{k-1},\tag{IV.10}$$

et la distribution du degré maximum est alors donnée par :

$$\phi_n(k) = (1 - a^k)^n - (1 - a^{k-1})^n.$$
 (IV.11)

On focalise la suite notre étude au cas des réseaux de taille  $n \gg 1$ . Dans cette limite, chaque terme de la distribution du degré maximum peut être approximée par un terme exponentiel :

$$\phi_n(k) \approx e^{-na^k} - e^{-na^{k-1}}.$$
 (IV.12)

La valeur moyenne du degré maximum s'obtient directement à partir de cette distribution, dans l'approximation  $n \gg 1$ :

$$\langle k_{max}(n) \rangle = \sum_{k=1}^{\infty} k \phi_n(k)$$
 (IV.13)

$$\approx \sum_{k=1}^{\infty} k \left( e^{-na^k} - e^{-na^{k-1}} \right), \qquad (IV.14)$$

où l'essentiel des valeurs de la somme (IV.14) est concentré dans les termes qui correspondent aux valeurs de k telles que :

$$na^k \sim 1$$
 (IV.15)

$$\Rightarrow k \sim -\frac{\ln n}{\ln a} \tag{IV.16}$$

$$\Rightarrow k \sim \frac{\ln n}{|\ln a|},\tag{IV.17}$$

comme 0 < a < 1. Pour  $na^k \sim 1$ , le terme entre parenthèses de la somme (IV.14) est aussi d'ordre 1, et on obtient donc une valeur approchée pour la valeur moyenne du degré maximum :

$$\langle k_{max}(n) \rangle \approx \frac{\ln n}{|\ln a|}.$$
 (IV.18)

Pour obtenir tous les moments de la distribution du degré maximum, on définit la fonction génératrice des moments :

$$G_n(s) = \left\langle e^{sk_{max}(n)} \right\rangle = \sum_{k=1}^{\infty} e^{sk} \phi_n(k), \qquad (\text{IV.19})$$

so it, dans l'approximation  $n\gg 1$  :

$$G_n(s) \approx \sum_{k=1}^{\infty} e^{sk} \left( e^{-na^k} - e^{-na^{k-1}} \right).$$
 (IV.20)

Comme on a approximé le premier moment du degré maximum par la formule (IV.18), on développe la fonction génératrice des moments autour de cette estimation. Les variables aléatoires sont ici entières, et on va donc distinguer, dans l'approximation (IV.18), la partie entière j et le reste  $\xi$ . On pose :

$$\frac{\ln n}{|\ln a|} = j + \xi \Leftrightarrow n = a^{-(j+\xi)}, \tag{IV.21}$$

où  $j \in \mathbb{N}$  et  $0 < \xi < 1$ . L'entier j est donc entièrement défini, et on exprime k en fonction de j en posant :

$$k = j + m = \frac{\ln n}{|\ln a|} + m - \xi,$$
 (IV.22)

où  $m \in \mathbb{Z}$ .

Avec ces nouvelles notations, la fonction génératrice des moments du degré maximum s'écrit comme une série :

$$G_n(s) \approx \sum_{m=-\infty}^{\infty} \exp\left\{s\left(\frac{\ln n}{|\ln a|} + m - \xi\right)\right\} \left(e^{-a^{m-\xi}} - e^{-a^{m-1-\xi}}\right).$$
(IV.23)

La partie indépendante de n est notée comme une amplitude :

$$A(s,\xi) = \sum_{m=-\infty}^{\infty} e^{s(m-\xi)} \left( e^{-a^{m-\xi}} - e^{-a^{m-1-\xi}} \right),$$
 (IV.24)

de telle sorte que la fonction génératrice se met sous la forme :

$$G_n(s) \approx A(s,\xi) n^{s/|\ln a|}.$$
 (IV.25)

L'amplitude A est une fonction de  $\xi$  périodique de période 1, qui s'écrit également comme la série :

$$A(s,\xi) = \sum_{j=-\infty}^{\infty} \alpha_n(s) e^{2i\pi j\xi},$$
 (IV.26)

dont les coefficients ont la forme :

$$\alpha_j(s) = \frac{\mathrm{e}^s - 1}{\ln a} \Gamma\left(\frac{s + 2\mathrm{i}\pi j}{\ln a}\right). \tag{IV.27}$$

Pour  $j \neq 0$ , les coefficients  $\alpha_j$  décroissent exponentiellement vite en j, donc, en pratique, seuls les premiers termes comptent dans la somme.

Les expressions asymptotiques pour les moments du degré maximum peuvent être obtenues par l'expansion de la génératrice des moments autour de s = 0. Les oscillations en  $\xi$  de l'amplitude sont négligeables, et la fonction amplitude  $A(s,\xi)$  est donnée, en bonne approximation, par le coefficient constant  $\alpha_0(s)$ . Les deux premiers moments sont donc, en bonne approximation :

$$\langle k_{max}(n) \rangle \approx \frac{\ln n + \gamma_E}{|\ln a|}$$
 (IV.28)

$$\operatorname{var}(k_{max}(n)) \approx \frac{\pi^2}{6\ln(a)^2} + \frac{1}{12},$$
 (IV.29)

où  $\gamma_E$  est la constante d'Euler.

Les oscillations de l'amplitude de la distribution du degré se retrouvent dans la forme de la distribution du maximum, et sont visibles autour du maximum de la distribution. On en observera des exemples plus loin dans notre étude.

La statistique des extrêmes a été largement étudiée dans le cas des variables aléatoires continues, et on a pu définir quelques distributions limites vers lesquelles convergent les distributions d'extrêmes. On cherche ici à quelle classe se rattache notre problème. Pour ce faire, on approxime la distribution géométrique IV.4 par une densité de probabilité exponentielle :

$$f(x) = b e^{-bx}, (IV.30)$$

dans laquelle  $b = |\ln a|$ . La valeur maximale se comporte comme :

$$x_{max}(n) = \frac{\ln n + \xi}{b},\tag{IV.31}$$

où la variable réduite  $\xi$  est distribuée selon une loi de Gumbel :

$$\rho_{\xi}(\xi) = \mathrm{e}^{-\xi - \mathrm{e}^{-\xi}},\tag{IV.32}$$

grâce à laquelle on retrouve des résultats semblables à ceux obtenus pour une variables discrète (IV.29):

$$\langle x_{max}(n) \rangle \approx \frac{\ln n + \gamma_E}{b}, \quad \text{var } x_{max}(n) \approx \frac{\pi^2}{6b^2}.$$
 (IV.33)

On peut donc considérer la distribution du degré maximum d'un ensemble tiré d'une distribution géométrique comme une loi de Gumbel discrète.

#### IV.1.2 Distribution du nombre de co-meneurs

Il peut y avoir plusieurs noeuds de degré maximum  $k_{max}(n)$  au temps n. On appelle comeneurs tous les noeuds de degré maximum. Le meneur est, parmi eux, celui qui a atteint le degré maximum le premier. Le nombre de co-meneurs permet de calculer par la suite la probabilité que le meneur change, c'est-à-dire la probabilité que l'un des co-meneurs, différent du meneur courant, voie son degré augmenter avant celui du meneur. La distribution du nombre de co-meneurs C(n) au temps n s'obtient à partir de la distribution du degré maximum. En effet, la somme de la distribution (IV.9) est prise sur le nombre de co-meneurs c au temps n, et la probabilité jointe du degré maximum  $k_{max}(n)$  et du nombre de co-meneurs C(n) s'écrit simplement :

$$P(k_{max}(n) = k, C(n) = c) = {\binom{n}{c}} f(k)^c (1 - \mathcal{F}(k))^{n-c}, \qquad (IV.34)$$

où f(k) est la distribution et  $\mathcal{F}(k)$  la distribution cumulative du degré global. La distribution  $\rho_C(c) = \operatorname{Prob}(C(n) = c)$  du nombre de co-meneurs est donc :

$$\rho_C(c) = \binom{n}{c} \sum_{k=1}^{\infty} f(k)^c \left(1 - \mathcal{F}(k)\right)^{n-c}.$$
 (IV.35)

On applique ce cadre général à la distribution géométrique définie en (IV.4), avec l'approximation par une exponentielle, pour  $n \gg 1$ , utilisée précédemment. La distribution du nombre de co-meneurs est donnée en bonne approximation par :

$$\rho_C(c) \approx \frac{(1-a)^c}{c!} \sum_{m=-\infty}^{\infty} a^{c(m-1-\xi) \exp\left(-a^{m-1-\xi}\right)}.$$
 (IV.36)

La somme est une fonction périodique en  $\xi$ , dont les oscillations autour de la valeur constante sont ici aussi petites. On les néglige en bonne approximation, et la distribution asymptotique du nombre de co-meneurs prend la forme :

$$\rho_C(c) \approx \frac{1}{|\ln a|} \frac{(1-a)^c}{c},\tag{IV.37}$$

qui est connue comme une distribution logarithmique. Le nombre moyen de co-meneurs, pour  $n \gg 1$  est donc :

$$\langle C(n) \rangle \approx \frac{1-a}{a \left| \ln a \right|}.$$
 (IV.38)

On a fait l'approximation, dans cette partie, de variables alétoires indépendantes et identiquement distribuées. On vérifiera dans la suite, pour chaque modèle, la validité de cette approximation et l'effet sur les quantités observées.

# IV.2 Degré maximum

La valeur typique du plus grand degré  $k_{max}(n)$ , atteint par un noeud quelconque au temps n, s'obtient par le biais de la statistique des extrêmes. En effet, le plus grand degré du réseau

est tel que seul un noeud parmi les n éléments du réseau peut être d'un degré plus grand que  $k_{max}(n)$ , et ce dernier est défini par :

$$\int_{k_{max(n)}}^{\infty} nf(n,k) \mathrm{d}k \approx \frac{1}{n},$$
 (IV.39)

où f(n,k) est la distribution du degré d'un noeud quelconque du réseau au temps n.

Les premières approches par la statistique stationnaire des extrêmes reposent sur l'hypothèse que  $k_{max}(n)$  est la plus grande variable aléatoire parmi un tirage de n variables tirées indépendament, selon la même distribution stationnaire  $f_{stat}(k)$ . Des effets correctifs potentiellement importants, suggérés par Moreira *et al.* [61], peuvent apparaître, qui doivent être pris en compte dans les modèles étudiés :

- Les degrés  $k_i(n)$  ne sont pas des variables indépendantes. Elles sont liées les unes aux autres par le facteur de normalisation qui assure qu'un degré et un seul est ajouté aux noeuds déjà présents dans le réseau. Le degré  $k_i(n)$  dépend donc de quel noeud parmi les n-1 éléments reçoit la nouvelle connexion. Pour les grands temps cependant, la contrainte s'affaiblit car le nombre de degrés de liberté croît avec la taille n du réseau.
- On a montré dans le chapitre précédent que les effets de taille et temps finis font apparaître une correction à la distribution stationnaire  $f_{stat}(k)$  du degré, d'autant plus grande que le degré k considéré est proche de la valeur de transition  $k_{\star}(n) \sim k_{max}(n)$ . Les effets de taille finie peuvent donc avoir un impact fort sur la statistique du plus grand degré.
- Enfin, la distribution  $f_i(n,k)$  des degrés  $k_i(n)$  des noeuds dépend de l'étape *i* d'arrivée dans le système du noeud *i*. En particulier, pour *n* grand, les degrés des noeuds les plus anciens et des noeuds les plus récents sont différemment distribués. Cette remarque aura un impact sur l'identité du meneur.

Pour chaque modèle, l'effet et l'intensité de ces corrections peuvent être très différents. On détaille donc le calcul des degrés maximaux pour chaque modèle.

## IV.2.1 Modèle AU

En première approximation, on va considérer la valeur typique du degré maximal donnée par la statistique des extrêmes sur la distribution stationnaire du degré, pour n grand. La distribution stationnaire  $f_{stat}$  du degré global vérifie la formule calculée au chapitre précédent (III.59) :

$$f_{stat}(k) = 2^{-k}.$$
 (IV.40)

Le plus grand degré typique est donc, en première approximation :

$$\frac{2^{-k_{max}(n)}}{\ln 2} = \frac{1}{n}$$
(IV.41)

$$\Rightarrow k_{max(n)} \approx \frac{\ln n}{\ln 2}.$$
 (IV.42)

Dans le cas du modèle AU, le comportement logarithmique du degré maximum est préservé après les corrections détaillées ci-dessus. En effet, si on ajoute la dépendance en l'indice i du noeud, la dépendance de la distribution du degré du noeud i en son indice n'apparaît qu'en correction logarithmique. Le degré typique d'un noeud d'indice i est en effet donné par :

$$\langle k_i(n) \rangle \approx \ln \frac{n}{i} + 1.$$
 (IV.43)

Pour les effets de temps finis, la transition entre le régime stationnaire et le régime d'échelle est très étroite autour de  $k_{\star}(n) \approx 2 \ln n$ . On a vu dans le chapitre précédent que la variance de la fonction d'échelle de temps fini se comporte en  $2 \ln n$ . Or, la première approximation pour la valeur moyenne du degré maximum (IV.42) donne un préfacteur de l'ordre de  $1/\ln 2$ , qui est nettement inférieur à 2. Dans ce modèle, la distribution stationnaire est donc bien justifiée pour le calcul du degré maximum typique. On retrouve donc la distribution calculée pour des variables discrètes indépendantes et identiquement distribuées, avec a = 1/2. La distribution  $\phi$  du degré maximum s'écrit donc, sous les approximations de variables indépendantes et identiquement distribuées, comme :

$$\phi(k) = \left(1 - \frac{1}{2^k}\right)^n - \left(1 - \frac{1}{2^{k-1}}\right)^n,$$
 (IV.44)

qui devient en bonne approximation pour  $n \to \infty$  :

$$\phi(k) = \exp\left(-\frac{n}{2^k}\right) - \exp\left(-\frac{n}{2^{k-1}}\right).$$
(IV.45)

On compare sur la figure IV.1 le résultat des simulations et la formule stationnaire pour la distribution du degré maximum dans un réseau, pour différentes tailles et pour les deux conditions initiales. On observe, pour la distribution stationnaire, une oscillation dans la forme du maximum de la distribution, dûe aux oscillations de l'amplitude de la distribution (IV.24).

En suivant le raisonnement décrit pour les variables distribuées géométriquement, on retrouve bien l'estimation attendue pour le degré maximum moyen :

$$\langle k_{max}(n) \rangle \approx \frac{\ln n}{\ln 2}.$$
 (IV.46)

On compare les résultats obtenus par l'approche stationnaire aux résultats de la simulation d'un réseau en croissance avec attachement uniforme figure IV.2. La valeur moyenne obtenue par la distribution stationnaire du plus grand degré converge très vite vers sa valeur asymptotique :

$$\langle k_{max}(n) \rangle \rightarrow \frac{\ln n + \gamma_E}{\ln 2} \approx 1.332746.$$
 (IV.47)

Les valeurs obtenues par simulation à temps fini du modèle AU convergent très lentement vers une limite qui semble être de l'ordre de la valeur stationnaire.

## IV.2.2 Modèle BA

La distribution stationnaire du degré d'un noeud quelconque dans le réseau est donnée par la formule (III.107) du chapitre précédent. Pour de grandes valeurs de k, la distribution stationnaire du degré se comporte en :

$$f_{stat}(k) \approx \frac{4}{k^3},$$
 (IV.48)



FIGURE IV.1 – Distribution du degré maximum pour différentes tailles du système, selon la condition initiale. Les courbes jaune et magenta sont obtenues à partir de la distribution IV.44 du degré maximum.

et la distribution cumulative :

$$\mathcal{F}_{stat}(k) \approx \frac{2}{k^2}.$$
 (IV.49)

On obtient donc une première approximation pour le degré maximum, par la statistique des extrêmes, et on attend une valeur typique en :

$$k_{max}(n) \sim n^{1/2}.$$
 (IV.50)

La distribution complète du degré maximum dans l'approximation de variables aléatoires indépendantes et identiquement distribuées selon  $f_{stat}$  est donnée par (IV.9) :

$$\phi_n(k) \approx \sum_{c=1}^n \binom{n}{c} f_{stat}(k)^c (1 - \mathcal{F}_{stat}(k))^{n-c}, \qquad (\text{IV.51})$$



FIGURE IV.2 – Valeur moyenne du degré maximum en fonction de la taille N du système. La droite correspond à la limite asymptotique de la distribution stationnaire  $\gamma_E/\ln 2 + 1/2$ .

dans laquelle la valeur du premier terme de la somme, c = 1, domine tous les autres, pour  $k \gg 1$ , *i.e.*  $f_{stat}(k) \ll 1$ :

$$\phi_{stat}(k) \approx n f_{stat}(k) (1 - \mathcal{F}_{stat}(k))^n \tag{IV.52}$$

$$\approx n f_{stat}(k) \exp(-n \mathcal{F}_{stat}(k))$$
 (IV.53)

$$\approx \frac{4n}{k^3} \mathrm{e}^{-2n/k^2}.\tag{IV.54}$$

Les arguments utilisés pour estimer la valeur du degré de transition dans le chapitre précédent établissaient que  $k_{max}(n) \sim k_{\star}(n) \sim n^{1/2}$ . On peut donc définir une variable aléatoire continue :

$$Y(n) = \frac{k_{max}(n)}{n^{1/2}},$$
 (IV.55)

dont la distribution asymptotique  $\varphi_{stat}$  pour  $n\gg 1$  peut être déterminée dans le régime sta-

tionnaire à partir de la distribution stationnaire  $\phi_{stat}$  du degré maximum :

$$\varphi_{stat}(Y) \approx \frac{4}{Y^3} \mathrm{e}^{-2/Y^2},\tag{IV.56}$$

où l'on reconnait une distribution de Fréchet. Pour ce type de distribution, seul le premier moment est défini, et vaut :

$$\langle Y \rangle_{stat} = (2\pi)^{1/2} \approx 2.506628.$$
 (IV.57)

Or les moments d'ordres supérieurs de la distribution du degré maximum doivent être définis. La distribution de Fréchet est donc incorrecte. En particulier, les effets de taille finie introduisent une coupure dans les valeurs du degré et donc dans la distribution du degré maximum. Les moments d'ordres supérieurs sont alors définis.

En seconde approximation, on va donc introduire les effets de taille finie dans les considérations de statistique des extrêmes. Dans le chapitre précédent, on a obtenu la distribution du degré global à taille finie :

$$f(n,k) \approx \frac{4}{k^3} \Phi(y), \quad y = \frac{k}{n^{1/2}},$$
 (IV.58)

dans laquelle la fonction d'échelle dépend de la condition initiale considérée :

$$\Phi^{(A)}(y) = \operatorname{erfc}\left(\frac{y}{2}\right) + \frac{y}{\sqrt{\pi}}\left(1 + \frac{y^2}{2}\right) e^{-y^2/4},$$
 (IV.59)

$$\Phi^{(B)}(y) = \frac{y}{\sqrt{\pi}} e^{-y^2/4} + \left(1 + \frac{y^2}{2}\right) \operatorname{erfc}\left(\frac{y}{2}\right).$$
(IV.60)

La distribution cumulative du degré  $\mathcal{F}(n,k) = \sum_{j=k}^{\infty} f(j)$  peut aussi être définie comme le produit de la distribution stationnaire  $\mathcal{F}_{stat}(k) = 2/k^2$  et d'une fonction d'échelle  $\Psi(y)$ , avec la même variable réduite  $y = \frac{k}{n^{1/2}}$ :

$$\mathcal{F}(n,k) \approx \frac{2}{k^2} \Psi(y),$$
 (IV.61)

où la fonction d'échelle  $\Psi(y)$  vérifie la relation :

$$\Phi(y) = \Psi(y) - \frac{y}{2}\Psi'(y).$$
 (IV.62)

En effet, en approximant la somme par une intégrale, on retrouve bien :

$$\sum_{j=k}^{\infty} f(j) \approx \frac{4}{n} \int_{y}^{\infty} \frac{\Psi(z) - \frac{z}{2} \Psi'(z)}{z^{3}} dz$$

$$\approx \frac{2}{ny^{2}} \Psi(y)$$

$$\approx \mathcal{F}(n,k).$$
(IV.63)

La fonction d'échelle de la distribution cumulative dépend donc aussi de la condition initiale selon :

$$\Psi^{(A)}(y) = \frac{y}{\sqrt{\pi}} e^{-y^2/4} + \left(1 + \frac{y^2}{2}\right) \operatorname{erfc}\left(\frac{y}{2}\right), \qquad (IV.64)$$

$$\Psi^{(B)}(y) = \left(1 + \frac{y^2}{2}\right) e^{-y^2/4}.$$
 (IV.65)

Dans cette approximation, on considère toujours que le degré de chaque noeud est identiquement distribué, selon la distribution f(n,k), et on insère donc les distributions à temps fini f(n,k) et  $\mathcal{F}(n,k)$  dans la formule (IV.9). On peut ici aussi faire l'approximation que seul le terme en c = 1 compte dans la somme, et on obtient une approximation de temps fini pour la distribution du degré maximum :

$$\phi(n,k) \approx \frac{4n}{k^3} \Phi(y) \exp\left(-\frac{2n}{k^2} \Psi(y)\right)$$
(IV.66)

$$\varphi_Y(Y) \approx \frac{4\Phi(Y)}{Y^3} e^{-2\Psi(Y)/Y^2}, \qquad (IV.67)$$

où  $Y(n) = k_{max}(n)/n^{1/2}$ . L'influence en les conditions initiales est contenue dans les fonctions d'échelle  $\Phi$  et  $\Psi$ . Cette expression décroît exponentiellement vite, en  $\varphi_Y(Y) \propto \exp(-Y^2/4)$ , quelle que soit la condition initiale considérée, et tous les moments de Y(n) sont finis. Les préfacteurs dépendent cependant de la condition initiale, et les premiers moments prennent les valeurs :

$$\left\langle Y^{(A)}(n)\right\rangle \approx 1,842146\tag{IV.68}$$

$$\langle Y^{(B)}(n) \rangle \approx 1,966678,$$
 (IV.69)

pour valeur moyenne, tandis que la variance réduite :

$$v(n) = \frac{\operatorname{var}(k_{max}(n))}{\langle k_{max}(n) \rangle^2} = \frac{\operatorname{var}(Y(n))}{\langle Y(n) \rangle^2},$$
 (IV.70)

prend, selon les conditions initiales, les valeurs

$$v^{(A)}(n) \approx 0,161205, \quad v^{(B)}(n) \approx 0,179513.$$
 (IV.71)

Cette seconde approximation ne tient cependant pas compte de la dépendance du degré des noeuds à leur date d'apparition dans le réseau. Or la dépendance en la date d'apparition i du noeud est multiplicative dans le régime d'échelle  $n \sim i \gg 1$  (III.173) et l'approximation de variables identiquement distribuées néglige ce point.

On observe donc sur la figure IV.3 que les résultats des simulations confirment le comportement en  $n^{1/2}$  de la valeur moyenne du plus grand degré. Le préfacteur est cependant différent de notre estimation. La valeur moyenne du degré maximum converge vers :

$$\left\langle Y_{\infty}^{(A)}(n) \right\rangle \approx 2,00 \quad \left\langle Y_{\infty}^{(B)}(n) \right\rangle \approx 2,16.$$
 (IV.72)



FIGURE IV.3 – Valeur moyenne du degré maximum en fonction de la taille n du système.

Cette différence est dûe à l'approximation de variables identiquement distribuées. Lors de la construction du réseau, les premiers noeuds reçoivent plus de connexions que les noeuds les plus récents, et le degré maximum sera donc en pratique plus élevé qu'avec l'estimation de variables aléatoires iid.

On représente également, figure IV.4 la distribution du degré maximum redimensionné. On observe bien que la distribution est plus étroite pour les conditions initiales (A) que pour les conditions initiales (B).

#### IV.2.3 Attachement préférentiel généralisé

La distribution du degré dans le modèle d'attachement préférentiel généralisé décroît en loi de puissance pour de grandes valeurs de k (III.231) :

$$f_{stat}(k) \approx \frac{(c+2)\Gamma(2c+3)}{\Gamma(c+1)}k^{-(c+3)},$$
 (IV.73)



FIGURE IV.4 – Distribution du degré maximum redimensionné  $Y(n) = k_{max}(n)/n^{1/2}$  pour un système de taille  $n = 10^4$ .

et dans le régime  $k \sim c \gg 1$ , la forme d'échelle de la distribution stationnaire s'écrit :

$$f_{stat}(k) \sim \exp\left(-c\Phi(\kappa)\right),$$
 (IV.74)

où  $\kappa$  est le rapport  $\kappa = k/c$ , et où la fonction d'échelle prend la forme :

$$\Phi(\kappa) = (\kappa + 2)\ln(\kappa + 2) - (\kappa + 1)\ln(\kappa + 1) - 2\ln(2).$$
 (IV.75)

On rappelle les deux limites de ce modèle :

- pour  $c \to \infty$ , le rapport  $\kappa$  tend vers 0 et on retrouve bien la distribution stationnaire du modèle AU avec  $\Phi(\kappa) \approx \kappa \ln(2)$ .
- pour  $c \to 0$  correspondant au modèle BA,  $\kappa \to \infty$ , la fonction d'échelle croît logarithmiquement comme :

$$\phi(\kappa) \approx \ln(\kappa) + 1 - 2\ln(2), \qquad (IV.76)$$

et la distribution stationnaire décroît en loi de puissance d'exposant 3.

Pour les valeurs finies de c > -1, le modèle APG se comporte de façon semblable au modèle BA. Comme étudié dans le chapitre précédent, le degré maximum typique se comporte comme le degré de transition (III.274) :

$$k_{max}(n) \sim n^{\frac{1}{c+2}},\tag{IV.77}$$

où l'on retrouve bien le comportement du modèle BA pour c = 0. Par simplicité de notation, on pose :

$$\nu = \frac{1}{c+2},\tag{IV.78}$$

et le rapport entre le comportement général de  $k_{max}(n)$  et sa valeur moyenne est noté comme précédemment :

$$Y(n) = \frac{k_{max}(n)}{n^{\nu}}.$$
 (IV.79)

Cette variable redimensionnée est distribuée selon la distribution non-triviale  $\varphi_Y$ , qui dépend de  $\nu$ , et donc du paramètre d'attractivité initiale c. Les premiers moments de la variable rescalée  $\langle Y(n) \rangle$  et var(Y(n)) convergent vers des valeurs asymptotiques  $\langle Y \rangle_{\infty}$  et  $v_{\infty}$ , où v(n) est la variance redimensionnée.

Les corrections de taille finie à cette convergence sont proportionnelles à  $1/k_{\star}(n) \sim n^{-\nu}$ .

# IV.3 Nombre de prépondérances

On appelle une prépondérance une période de temps au cours de laquelle l'identité du meneur ne change pas. La distribution du nombre de changements de meneurs est fortement liée à la seconde quantité considérée pour les variables aléatoires indépendantes et indentiquement distribuées : le nombre de co-meneurs. En effet, un changement de meneur à un certain temps n est dû à la présence au temps n-1 de co-meneurs susceptibles de gagner un degré au temps n et donc de devenir meneurs. La probabilité qu'un changement de meneur survienne au temps n+1 est donc proportionnel au nombre C(n)-1 de noeuds de plus grand degré qui ne sont pas meneur au temps n. Le nouveau noeud n+1 s'attache à un co-meneur i avec probabilité  $p_{n+1,i}$ , et la différence du nombre moyen de prépondérances  $\langle \mathcal{L}(n+1) \rangle$  entre l'étape n et l'étape n+1s'écrit :

$$\langle \mathcal{L}(n+1) \rangle - \langle \mathcal{L}(n) \rangle = p_{n+1,i} \left( \langle C(n) \rangle - 1 \right).$$
 (IV.80)

La probabilité d'attachement  $p_{n+1,i}$  à l'un des co-meneurs dépend du modèle considéré.

#### IV.3.1 Attachement uniforme

Dans le modèle d'attachement uniforme  $p_{n+1,i} = 1/n$ , la distribution géométrique obtenue par l'approximation de variables aléatoires indépendantes et identiquement distribuées donne des résultats cohérents avec les simulations. On reprend donc le résultat asymptotique (IV.37) pour la distribution du nombre de co-meneurs de variables aléatoires distribuées selon la distribution géométrique (IV.4) avec a = 1/2. Dans cette approximation, le nombre moyen de co-meneurs, dans la limite stationnaire  $n \gg 1$  est donné par :

$$\langle C \rangle \approx \frac{1-a}{a \left| \ln a \right|} = \frac{1}{\ln 2}.$$
 (IV.81)

À partir de la relation de récurrence pour le nombre moyen de changements de meneur (IV.80), on obtient le résultat asymptotique suivant pour le nombre de changements de meneur jusqu'au temps  $n \gg 1$ :

$$\langle \mathcal{L}(n) \rangle \approx \left(\frac{1}{\ln(2)} - 1\right) \ln n.$$
 (IV.82)

Les effets de temps fini expliquent la différence entre les résultats de simulation et la courbe théorique du nombre moyen de changements de meneur. La statistique des changements de meneur est particulièrement sensible aux effets de taille finie dans les réseaux avec condition initiale (A). En effet, dans ces réseaux la variation du nombre de changements de meneur dépend fortement des premières étapes de la construction du réseau, et pour un réseau à temps fini, l'effet de ces premières étapes se fait encore sentir.



FIGURE IV.5 – Nombre moyen de meneurs jusqu'au temps n en fonction de la taille n du système.
## IV.3.2 Barabási-Albert

Dans le modèle de Barbási-Albert, la probabilité d'attachement du nouveau noeud n + 1 à l'un des co-meneurs *i* dépend du degré maximum  $k_{max}(n)$ :

$$p_{n+1,i} = \frac{k_{max}(n)}{Z(n)},$$
 (IV.83)

où Z(n), défini dans le chapitre précédent, assure la normalisation de la probabilité d'attachement sur tous les noeuds existants dans le réseau.

Dans l'approximation de la distribution stationnaire des degrés, indépendants et identiquement distribués, le nombre de co-meneurs pour le modèle BA, de distribution stationnaire  $f_{stat}(k) \approx 4k^{-3}$ , est distribué selon (IV.37) :

$$\rho_C(c) = \binom{n}{c} \sum_{k=1}^{\infty} \left(\frac{4}{k^3}\right)^c \left(1 - \frac{2}{k^2}\right)^{n-c}.$$
 (IV.84)

Dans le régime stationnaire, le degré maximum est grand  $k_{max}(n) \gg 1$ , et, comme la distribution du degré décroît en loi de puissance, le nombre de noeuds de degré  $k_{max}(n)$  est petit devant la taille du réseau. La probabilité que le nombre de co-meneurs soit c = 2 est faible, mais elle est encore plus faible pour  $c \ge 2$  et on peut donc négliger les cas où  $C(n) \ge 2$ . La probabilité qu'il y ait deux co-meneurs au temps  $n \gg 1$  est donc donnée par :

$$\rho_C(2) \approx \sum_{k=1}^{\infty} \binom{n}{2} f_{stat}^2(k) \mathrm{e}^{-n\mathcal{F}_{stat}(k)}.$$
 (IV.85)

L'incrément du nombre de changements de meneur au temps n + 1 est donnné par le produit de la probabilité d'avoir deux co-meneurs et la probabilité que la nouvelle connexion s'attache au co-meneur, soit  $k_{max(n)}/Z(n)$ :

$$\langle \mathcal{L}(n+1) \rangle - \langle \mathcal{L}(n) \rangle \approx \sum_{k=1}^{\infty} \frac{k}{Z(n)} {n \choose 2} f_{stat}^2(k) \mathrm{e}^{-n\mathcal{F}_{stat}(n)}.$$
 (IV.86)

La limite asymptotique  $n \gg 1$  donne :

$$\langle \mathcal{L}(n+1) \rangle - \langle \mathcal{L}(n) \rangle \approx \int_{1}^{\infty} \mathrm{d}k \frac{k}{Z(n)} {n \choose 2} f_{stat}^{2}(k) \mathrm{e}^{-n\mathcal{F}_{stat}(n)}$$
(IV.87)

$$\approx \frac{1 - e^{-2n}(1+2n)}{2n}$$
 (IV.88)

$$\approx \frac{1}{2n},$$
 (IV.89)

où on a posé  $Z(n) \approx 2n$ . On assimile la différence des nombres de changements de meneur à une dérivée discrète, et le nombre de changements de meneur jusqu'au temps n présente un comportement asymptotique logarithmique, comme pour le modèle AU :

$$\langle \mathcal{L}(n) \rangle \approx \frac{\ln n}{2}.$$
 (IV.90)

Cette approximation repose sur la limite stationnaire de la statistique des extrêmes, et on a vu dans la section précédente que cette approximation donne pour le modèle BA le bon comportement, mais que le préfacteur varie. Le nombre de changements de meneurs dans le modèle BA prend donc la forme :

$$\langle \mathcal{L}(n) \rangle \approx A_{\mathcal{L}} \ln n,$$
 (IV.91)

où l'amplitude est déterminée à partir des résultats de simulation figure IV.6. L'amplitude dépend de la condition initiale considérée, et on obtient :

$$A_{L}^{(A)} \approx 0.45, \quad A_{L}^{(B)} \approx 0.39.$$
 (IV.92)

On retrouve, pour la condition (A), les résultats précédents de Krapivsky et Redner [60].



FIGURE IV.6 – Nombre moyen de meneurs jusqu'au temps n en fonction de la taille n du système.

Le système subit donc en pratique moins de changements de meneurs qu'attendu par l'analyse stationnaire. On a en effet vu que la distribution du degré à temps fini décroît plus vite vers 0 pour les grands degrés que la distribution stationnaire. La probabilité que plusieurs noeuds soient de grand degré est donc plus faible dans un réseau à temps fini que dans un réseau stationnaire, et la probabilité de changer de meneur au cours du temps est donc réduite par rapport au régime stationnaire.

### IV.3.3 Attachement préférentiel généralisé

On pour suit pour ce modèle l'analyse par le biais des distributions stationnaires. Dans ce modèle, la probabilité d'attachement à l'un des co-meneurs i est donnée par :

$$p_{n+1,i} = \frac{k_{max}(n) + c}{Z(n)},$$
 (IV.93)

où  $Z(n) \approx (c+2)n$ , et la distribution stationnaire du degré se comporte, pour k grand, en :

$$f_{stat}(k) \approx \frac{(c+2)\Gamma(2c+3)}{\Gamma(c+1)}k^{-(c+3)}.$$
 (IV.94)

La distribution cumulative stationnaire du degré est donc proportionnelle à :

$$\mathcal{F}_{stat}(k) \propto k^{-(c+2)}.$$
 (IV.95)

Comme pour le modèle BA, la probabilité que plus de deux noeuds soient co-meneurs peut être négligée, et l'incrément du nombre de changements de meneur vérifie asymptotiquement :

$$\langle \mathcal{L}(n+1) \rangle - \langle \mathcal{L}(n) \rangle \approx \sum_{k=1}^{\infty} \frac{k+c}{Z(n)} {n \choose 2} f_{stat}^2(k) \mathrm{e}^{-n\mathcal{F}_{stat}(k)}$$
 (IV.96)

$$\approx \int_{1}^{\infty} \frac{k+c}{Z(n)} {n \choose 2} f_{stat}^2(k) \mathrm{e}^{-n\mathcal{F}_{stat}(k)} \mathrm{d}k \qquad (\mathrm{IV.97})$$

$$\propto \frac{1}{n}.$$
 (IV.98)

Le nombre moyen de changements de meneur se comporte donc asymptotiquement en :

$$\langle \mathcal{L}(n) \rangle \propto \ln n,$$
 (IV.99)

et quel que soit la valeur de c, comme on peut le voir sur la figure IV.7 pour c = -1/2, 0, 2. Seuls les préfacteurs varient avec c.

## IV.4 Indice du meneur

Krapivsky et Redner ont introduit en [60] le calcul de l'identité du meneur. Le meneur est le noeud de degré  $k_{max}$ . Pour déterminer son identité, on calcule d'abord l'indice moyen d'un noeud de degré k quelconque. On le définit par :

$$\langle i_k(n) \rangle = \frac{g_k(n)}{f(n,k)},\tag{IV.100}$$



FIGURE IV.7 – Nombre moyen de meneurs jusqu'au temps n en fonction de la taille n du système, pour différentes valeurs du paramètre c.

où f(n,k) est la distribution du degré d'un noeud quelconque du réseau au temps n, et

$$g_k(n) = \frac{1}{n} \sum_{i=1}^n i f_i(n,k), \qquad (\text{IV.101})$$

la distribution de l'indice d'un noeud de degré k, avec  $f_i(n, k)$  la distribution du degré du noeud *i* au temps n. On peut donc définir cette distribution de l'indice par une relation de récurrence tirée de l'équation de récurrence pour la distribution du degré du noeud i, pour chaque modèle (III.32), (III.96) et (III.226).

## IV.4.1 Attachement uniforme

La formule de récurrence pour la distribution de l'indice d'un no eud de degré k est donnée dans le modèle AU par :

$$ng_k(n) = g_{k-1}(n-1) + (n-2)g_k(n-1) + n\delta_{k,1},$$
(IV.102)

où le dernier terme provient du nouveau noeud n qui apparaît dans le système avec degré k = 1. On s'intéresse au régime stationnaire  $n \to \infty$  pour k fixé. On suppose que la distribution de l'indice  $g_k$  d'un noeud de degré k croît linéairement en temps *i.e.* en taille n, et on pose donc :

$$g_k(n) \approx n\gamma(k),$$
 (IV.103)

où la fonction  $\gamma$  ne dépend que du degré k, et vérifie l'équation de récurrence :

$$3\gamma(k) = \gamma(k-1) + \delta_{k,1}.$$
 (IV.104)

La solution de cette équation est simplement :

$$\gamma(k) = 3^{-k}.\tag{IV.105}$$

On considère également la distribution stationnaire du degré global  $f_{stat}(k) \approx 2^{-k}$  dans la formule du degré moyen d'un noeud de degré k, et on obtient donc, dans le régime stationnaire :

$$\langle i_k \rangle \approx n \left(\frac{2}{3}\right)^k.$$
 (IV.106)

On a discuté de la validité de l'approche stationnaire dans le modèle AU pour le degré maximum IV.2.1.

Une approximation de l'indice du meneur I s'obtient en insérant la forme logarithmique pour le degré maximum (IV.46) :

$$\langle I(n) \rangle = \left\langle i_{k_{max}(n)} \right\rangle$$
 (IV.107)

$$\approx n \exp\left(\ln(n)\left(1 - \frac{\ln 3}{\ln 2}\right)\right)$$
 (IV.108)

$$\approx n^{\sigma}$$
. (IV.109)

avec

$$\sigma = 2 - \frac{\ln 3}{\ln 2}.\tag{IV.110}$$

On utlise l'analyse de la statistique des extrêmes de la section IV.1 pour définir le préfacteur de la loi de puissance de l'indice moyen du meneur. On moyenne l'indice moyen d'un noeud de degré k (IV.106) sur la distribution de  $k_{max}(n)$ :

$$\langle I(n) \rangle \approx \sum_{k=1}^{\infty} n\left(\frac{2}{3}\right)^k \phi_n(k),$$
 (IV.111)

et on retrouve la fonction génératrice (IV.19) du degré maximum  $G_n(s = \ln(2/3))$ , avec a = 1/2. On poursuit les calculs effectués section IV.1, et on obtient :

$$\langle I(n) \rangle \approx n G_n \left( \ln(2/3) \right) \approx A_{I,stat} n^{\sigma},$$
 (IV.112)

où l'amplitude de cette approche stationnaire est donnée par le coefficient j = 0 de (IV.27) :

$$A_{I,stat} = \frac{1}{3\ln 2} \Gamma\left(\frac{\ln 3}{\ln 2} - 1\right) \approx 0,733240.$$
 (IV.113)

On représente donc figure IV.8 l'indice moyen du meneur redimensionné par l'estimation stationnaire  $n^{\sigma}$ . On observe que les simulations initialisées par les conditions initiales (A) et (B) convergent vers une valeur asymptotique  $A_I \approx 0.68$ , proche de la valeur  $A_{I,stat}$  calculée par le raisonnement sur les variables stationnaires.



FIGURE IV.8 – Indice moyen du meneur au temps n en fonction de la taille n du système.

## IV.4.2 Barabási-Albert

Dans le cadre du modèle avec attachement préférentiel, la relation de récurrence pour la distribution de l'indice d'un noeud de degré k (IV.101) vérifie la relation suivante :

$$nZ(n-1)g_k(n) = (n-1)(k-1)g_{k-1}(n) + (n-1)(Z(n-1)-k)g_k(n-1) + nZ(n-1)\delta_{k,1}.$$
 (IV.114)

On fait la même hypothèse que pour le modèle AU, de linéarité en n de la distribution de l'indice dans le régime stationnaire :  $g_k(n) \approx n\gamma(k)$ , où la fonction  $\gamma$  vérifie, dans le modèle BA, la relation de récurrence :

$$(k+4)\gamma(k) = (k-1)\gamma(k-1) + 2\delta_{k,1},$$
 (IV.115)

et donc

$$\gamma(k) = \frac{48}{k(k+1)(k+2)(k+3)(k+4)}.$$
(IV.116)

L'indice typique d'un noeud de degré k a donc, dans l'approximation stationnaire où

$$f_{stat}(k) = \frac{4}{k(k+1)(k+2)},$$
 (IV.117)

la forme suivante :

$$\langle i_k(n) \rangle \approx \frac{12n}{(k+3)(k+4)}.$$
 (IV.118)

Et l'indice typique du meneur  $\langle I(n) \rangle = \langle i_k(n) \rangle$  pour  $k = k_{max}(n) \sim n^{1/2}$  est donc bien constant dans le régime stationnaire, comme l'annonçaient Krapivsky et Redner [60], avec une correction de l'ordre de  $n^{1/2}$ . On cherche à estimer la valeur de cette constante dans le régime stationnaire. La distribution du degré maximum est donnée pour le modèle BA par (IV.54), et l'indice moyen du noeud de degré maximum est donné par l'approximation suivante dans le régime stationnaire :

$$\langle I(n) \rangle \approx \sum_{k=1}^{\infty} \left(\frac{12n}{k^2}\right) \phi_{stat}(k) \approx 6.$$
 (IV.119)

Les corrections de taille finie sont liées aux termes de l'expansion pour k grand de l'indice moyen d'un noeud de degré k. Les premiers termes correctifs se comportent en  $n/k^3$ . La correction à l'indice moyen du meneur est donc de l'ordre de  $1/k_{max}(n) \sim n^{-1/2}$ .

On lit sur le graphe IV.9 la valeur asymptotique  $\langle I\rangle_\infty$  de l'indice du meneur, dépendante des conditions initiales :

$$\langle I \rangle_{\infty}^{(A)} \approx 3.40, \quad \langle I \rangle_{\infty}^{(B)} \approx 2.67,$$
 (IV.120)

et les principales corrections de temps fini se comportent en  $n^{-1/2}$ .

#### IV.4.3 Attachement préférentiel généralisé

Dans ce modèle, la formule de récurrence pour la distribution de l'indice du degré s'écrit :

$$nZ(n-1)g_k(n) = (n-1)(k-1+c)g_{k-1}(n) + (n-1)(Z(n-1)-k-c)g_k(n-1) + nZ(n-1)\delta_{k,1}.$$
(IV.121)



FIGURE IV.9 – Indice moyen du meneur au temps n en fonction de la taille n du système.

On répète l'approximation de linéarité effectuée pour les modèles précédents en posant  $g_k(n) \approx n\gamma(k)$ , et la fonction  $\gamma$  vérifie dans ce modèle :

$$(k+2c+2)\gamma(k) = (k-1+c)\gamma(k-1) + (c+2)\delta_{k,1}.$$
 (IV.122)

La fonction  $\gamma$  s'exprime donc comme un rapport de fonctions gamma. L'indice moyen du meneur dépend donc non trivialement de la valeur du paramètre c, et converge vers une valeur asymptotique  $\langle I \rangle_{\infty}$ . Cette valeur est finie pour des valeurs finies du paramètre c, *i.e.*  $\nu$ . Lorsque  $c \to \infty$ , le modèle converge vers le modèle AU, dans lequel l'indice moyen du meneur croît indéfiniment. Il y a donc un saut dans la fonction  $\langle I \rangle_{\infty} (\nu)$ , et en pratique, on ne peut représenter une approximation de la valeur asymptotique que pour des valeurs de  $\nu > 0, 1$ .

La dépendance en  $\nu$  de la valeur asymptotique de l'indice moyen du meneur est estimée pour des valeurs finies de la taille du réseau, sur la figure IV.10. On retrouve bien pour  $\nu = 1/2$ , *i.e.* pour le modèle BA un indice du meneur de l'ordre de 2.



FIGURE IV.10 – Indice moyen du meneur au temps  $N = 10^4$  en fonction du paramètre  $\nu$ .

## IV.5 Persistance du meneur et nombre de meneurs distincts

La probabilité de persistance du meneur S(n) est la probabilité que le premier noeud, d'indice i = 1, du réseau soit resté meneur jusqu'au temps n, c'est-à-dire qu'aucun changement de meneur n'ait eu lieu jusqu'au temps n. En première estimation [60], on peut considérer que S(n) est égale à la probabilité que le degré  $k_1(n)$  soit égal au degré maximum  $k_{max}(n)$  au temps n. On a déterminé au chapitre précédent, pour chaque modèle, la distribution du degré de chaque noeud i. La probabilité de persistance dépend cependant de toute l'histoire du système et de chaque étape m précédent le temps d'observation n. Nous proposons donc un résultat plus précis en étudiant la probabilité que le degré du premier noeud  $k_1(m)$  ait été égal au degré maximum pour toute étape  $m \leq n$ .

On va tout d'abord s'intéresser aux grandes déviations, *i.e* à la probabilité que  $k_1$  ait été très différent de sa valeur typique  $\langle k_1(n) \rangle$ . On a défini dans le chapitre précédent la fonction génératrice  $F_{n,i}(x)$  des degrés du noeud *i* (III.26), (III.144) et (III.242), et on peut donc exprimer

la distribution du degré du noeud i en tant qu'intégrale de contour de la fonction génératrice :

$$f_i(n,k) = \oint \frac{\mathrm{d}x}{2\mathrm{i}\pi} \frac{F_{n,i}(x)}{x^{k+1}}.$$
 (IV.123)

La probabilité que le noeud *i* ait un degré très différent de son degré typique  $\langle k_i(n) \rangle$  peut être estimée par la méthode du col appliquée à l'intégrale de contour ci-dessus. On insère ensuite dans le résultat obtenu le degré maximum typique du modèle au temps *n* pour obtenir la probabilité que le noeud *i* ait degré maximum. On détaille les résultats pour les différents modèles étudiés.

Le nombre typique de meneurs distincts  $\langle D(n) \rangle$  jusqu'au temps *n* est une quantité complémentaire à la probabilité de persistence  $\langle S(n) \rangle$  du premier noeud comme meneur. En effet, si le nombre typique de meneurs distincts est de l'ordre de l'unité, le premier noeud reste meneur avec une probabilité non négligeable. Mais le nombre de meneurs distincts renvoie aussi à l'indice typique du meneur. Dans le modèle d'attachement uniforme, cet indice croît continuement, ce qui est peu compatible avec un nombre asymptotique fini de meneurs distincts, tandis que pour les modèles d'attachement préférentiel, dans lesquels l'indice du meneur converge vers une valeur asymptotique finie il ne peut y avoir qu'un nombre fini de noeuds ayant pris la prépondérance à un moment de l'histoire du réseau.

### IV.5.1 Attachement uniforme

#### a Persistance

Dans ce modèle, la probabilité que le degré  $k_1(n)$  du premier noeud soit de l'ordre du degré maximum  $k_1(n) \approx b \ln n, b \neq 1$  est donnée par :

$$f_1(n, b \ln n) \sim (\ln n)^{-1/2} n^{-\tau_b},$$
 (IV.124)

où l'exposant dépend de la valeur de la constante b:

$$\tau_b = 1 - b + b \ln b. \tag{IV.125}$$

Dans le modèle AU, le degré maximum vaut  $k_{max}(n) \approx \ln n / \ln 2$ , et on a donc  $b = 1/\ln 2$ . L'exposant de la probabilité que le premier noeud soit de degré  $k_{max}(n)$  vaut donc  $\tau_{max} = \tau_{b=1/\ln 2} \approx 0.08671$ , résultat obtenu par Krapivsky et Redner [60]. La probabilité de survie du premier meneur dépend cependant aussi de si le premier noeud a été de degré  $k \approx b \ln m$  pour toute étape  $m \leq n$ . On peut faire ici une analogie avec la probabilité de survie d'un marcheur aléatoire sur réseau 1D avec bords absorbants. On ne considère ici que le degré du noeud 1, indépendemment de la dynamique du degré des autres noeuds, avec des murs absorbants dont la position croît au cours du temps, qui se ramène à l'étude de Bauer, Godrèche et Luck en [62]. Cette analyse entraine que la probabilité de survie se met sous la forme :

$$S(n) \sim (\ln n)^{-3/2} n^{-\tau_{max}}.$$
 (IV.126)

Le graphe IV.11 représente la probabilité de survie en fonction de la taille du système obtenue par simulation pour les deux conditions initiales et compare ces résultats avec l'estimation de Krapivsky et Redner (ligne discontinue) et notre estimation, obtenue par analogie avec les marcheurs aléatoires (ligne continue). La décroissance de la probabilité de survie du premier meneur est beaucoup plus rapide que la probabilité que le noeud 1 soit de degré maximum au temps n.



FIGURE IV.11 – Probabilité de survie du premier meneur jusqu'au temps n.

#### b Nombre de meneurs distincts

Le nombre moyen de meneurs distincts  $\langle D(n) \rangle$  dépend de résultats obtenus précédemment : • le nombre de changements de meneur au cours du temps croît logarithmiquement :

$$\langle \mathcal{L}(n) \rangle \sim \ln n,$$
 (IV.127)

• l'indice typique du meneur croît en loi de puissance d'exposant  $\sigma \approx 0.415$  avec le temps :

$$\langle I(n) \rangle \sim n^{\sigma}.$$
 (IV.128)

Le nombre de changements de meneur croît donc typiquement moins vite que l'indice du meneur. On peut donc raisonnablement supposer que lors d'un changement de meneur, le nouveau meneur est un noeud plus récent, qui n'a encore jamais été meneur. De plus, le degré maximum croît selon une loi logarithmique :  $k_{max}(n) \sim \ln n$ , et le degré typique d'un noeud né à l'instant *i* ne subit qu'une correction additive en ln *i* au degré typique  $\langle k_i(n) \rangle$ . Un noeud plus récent peut donc atteindre le degré maximum avec probabilité non nulle pour  $n \gg 1$ .

On trace figure IV.12 le nombre moyen  $\langle D(n) \rangle$  de meneurs distincts dans l'histoire du réseau jusqu'au temps n, en fonction de n. Le nombre moyen de meneurs distincts croît bien logarithmiquement, avec un préfacteur  $A_D$  lu sur le graphe :

$$A_D \approx 0.29,\tag{IV.129}$$

quelle que soit la conditions initiale.

On estime la probabilité qu'un changement de meneur se fasse au profit d'un noeud qui n'a jamais été meneur par le rapport des nombres moyens de prépondérances et de meneurs distincts :

$$\mathcal{P} \approx \frac{\langle D(n) \rangle}{\langle \mathcal{L}(n) \rangle} = \frac{A_D}{A_{\mathcal{L}}} \approx 0,65.$$
 (IV.130)

Ce résultat confirme notre analyse heuristique : dans le modèle AU, un meneur peut être un noeud récent dans l'histoire du réseau.

## IV.5.2 Attachement préférentiel

Le raisonnement pour les deux modèles d'attachement préférentiel repose sur les résultats obtenus pour l'indice du meneur. En effet, le nombre moyen de meneurs distincts ne peut pas croître plus vite que l'indice du meneur. On s'attend donc à ce qu'il reste de l'ordre de l'unité, et qu'il atteigne une limite asymptotique non triviale  $\langle D \rangle_{\infty}$ . Dans les deux modèles d'attachement préférentiel, l'indice typique du noeud de degré maximum  $\langle I(n) \rangle$  sature asymptotiquement à une valeur finie, tandis que le nombre de changements de meneur croît logarithmiquement. On attend donc un comportement typique à la tête du réseau d'un petit nombre de noeuds dont le degré est proche ou égal au degré maximum, et qui s'echangent la prépondérance à un rythme logarithmique en la taille du système. Ces noeuds attirent la quasi-totalité des nouvelles connexions et gardent, voire creusent leur avantage sur le reste des noeuds du réseau. La distribution complète de D(n) converge vers une distribution  $\rho_D$  dans la limite des temps longs. Sur le graphe de cette distribution pour le modèle BA, représenté figure IV.13, on observe bien



FIGURE IV.12 – Nombre de meneurs distincts jusqu'au temps n. La droite a pour équation  $A_D \ln n + 0.75$ 



FIGURE IV.13 – Distribution du nombre de meneurs distincts dans l'histoire du réseau jusqu'au temps  $n = 10^5$  pour les deux conditions initiales dans le modèle BA.

que la probabilité que la prépondérance soit partagée entre deux noeuds est proche de 1 et décroît ensuite très vite vers 0. On lit sur la figure IV.14 que la valeur moyenne asymptotique du nombre de meneurs distincts au cours de l'histoire du réseau est de l'ordre de 2 quelque soit la condition initiale :

$$\langle D \rangle_{\infty}^{(A)} \approx 2,22, \quad \langle D \rangle_{\infty}^{(B)} \approx 1,94.$$
 (IV.131)

On peut définir la probabilité de persistance S(n) à partir du nombre de meneurs distincts, comme la probabilité que D(n) = 1, soit encore :

$$S_{\infty} = \rho_D(1). \tag{IV.132}$$

Le nombre de meneurs distincts restant fini tout au long de l'histoire du réseau, la probabilité de persistance du premier meneur n'est pas asymptotiquement nulle, et on lit sur la figure IV.15 la probabilité asymptotique de persistance dans le modèle BA :

$$S_{\infty}^{(A)} \approx 0,279, \quad S_{\infty}^{(B)} \approx 0,389.$$
 (IV.133)

Le modèle d'attachement préférentiel généralisé permet une extrapolation entre le cas c = -1, *i.e*  $\nu = 1$ , pour lequel tous les noeuds se connectent en étoile autour du premier noeud, et dans lequel le nombre de meneur distinct, et la probabilité de survie du premier meneur sont 1, et le cas  $c \to \infty$ , *i.e.*  $\nu \to 0$  pour lequel on retrouve le modèle AU. Dans ce modèle, le nombre de meneurs distincts augmente indéfiniment, et la probabilité de survie du premier noeud décroît vers 0. Entre ces deux extrêmes, le modèle de Barabási-Albert, pour  $\nu = 1/2$ 



FIGURE IV.14 – Nombre de meneurs distincts jusqu'au temps n.



FIGURE IV.15 – Probabilité de survie du premier meneur jusqu'au temps n.



(a) Nombre moyen de meneurs distincts, pour N = (b) Probabilité de persistance du premier meneur, 10<sup>4</sup>, en fonction de  $\nu$  pour  $N = 10^4$ , en fonction de  $\nu$ 

FIGURE IV.16 – Valeurs proches des valeurs asymptotiques du nombre moyen  $\langle D \rangle$  de meneurs distincts et de la probabilité de survie S du premier meneur.

présente déjà les caractéristiques du cas  $\nu = 1$  d'un nombre fini de meneurs distincts et d'une probabilité de persistance du premier meneur non nulle. On représente sur la figure IV.16 les valeurs asymptotiques de ces deux quantités en fonction de  $\nu$ . Comme ces quantités n'ont pas de limite asymptotique pour  $\nu \approx 0$  (ou une convergence très lente si  $\nu > 0$ ), on ne représente pas les valeurs pour  $\nu < 0.1$ .

## IV.6 Conclusion

Nous avons étudié la statistique des extrêmes à temps fini dans les réseaux en croissance pour trois modèles typiques : le modèle d'attachement uniforme, le modèle d'attachement préférentiel de Barabási-Albert et le modèle d'attachement généralisé qui permet l'interpolation des quantités entre les deux autres modèles. Notre étude s'est concentrée sur les caractéristiques liées au noeud de plus grand degré dans un réseau en croissance : la valeur du plus grand degré  $k_{max}(n)$  au temps n, l'identité I du meneur, c'est-à-dire le premier porteur de ce degré, ou le nombre C d'autres noeuds qui portent ce degré. À coté de ces quantités ponctuelles, nous avons caractérisé l'histoire de ces noeuds extrêmes, par l'étude du rôle des co-meneurs dans les changements d'identité du meneur et du nombre de périodes L de prépondérance d'un noeud sur les autres, ou du nombre D de meneurs au cours de l'histoire du réseau. Nous avons tiré de ces informations la probabilité S que le premier meneur garde la prépondérance sur les autres noeuds au cours de l'histoire du réseau.

Dans le modèle d'attachement uniforme, nous avons déterminé avec de bonnes approximations la distribution du degré maximum, dont les fluctuations autour de la valeur moyenne sont données par une loi de Gumbel discrète. Dans le modèle de Barabási-Albert, le degré maximum croît comme le degré de transition  $k_{\star}$  déterminé dans le chapitre précédent, et la distribution de la variable redimensionnée  $Y \approx k_{max}/n^{1/2}$  dépend des conditions initiales. Le modèle d'attachement préférentiel généralisé présente un degré maximum qui dépend de l'attractivité initiale c, de telle sorte que le modèle permet une interpolation entre les modèles AU et BA.

Notre étude heuristique et numérique des autres quantités permet de quantifier les résultats d'études précédentes [60], et fournit une analyse des phénomènes à l'oeuvre dans la dynamique des noeuds de plus grand degré. Les réseaux invariants d'échelle présentent une statistique des extrêmes relativement figée. Tandis que le degré maximum continue de croître avec le temps, seuls quelques noeuds se partagent la prépondérance sur la connectivité du réseau. Les valeurs asymptotiques finies de l'indice du meneur et du nombre distinct de meneurs attribuent à quelques uns des premiers noeuds la prépondérance tout au long de la construction du réseau. Le nombre de changements de meneur croît cependant continuement au cours du temps, et le scénario typique se trouve donc être celui d'un partage de la prépondérance entre quelques uns des noeuds les plus vieux, et donc les plus connectés du réseau. On trouve ici un exemple de processus du type *les riches deviennent plus riches*, évoqué dans leur étude par Krapivsky et Redner. Dans le cas de réseaux avec attachement uniforme, la situation n'est pas gelée au sommet et des noeuds plus jeunes peuvent venir prendre la prépondérance sur les meneurs.

Les quantités étudiées ici à temps fini montrent une dépendance sensible aux conditions initiales. Ces systèmes, opérant dans un régime proche du régime critique montrent un effet de mémoire de toute l'histoire de leur construction, et, en particulier, des différences de symétrie des toutes premières étapes de leur croissance.

L'effet de monopole de la prépondérance par les noeuds les plus anciens peut être atténuée par l'introduction d'un paramètre intrinsèque à chaque noeud qui viendrait compenser l'attachement préférentiel pour des noeuds plus jeunes mais intrinséquement plus attractif. On renvoie ici le lecteur aux travaux de Bianconi et Barabási [63] et de Godrèche et Luck [64].

## Partie C

## Cascades de défaillances dans les réseaux

## CHAPITRE V Propagation de défaillances dans les réseaux

La question de la résistance des réseaux à la défaillance de leurs éléments a intéressé une vaste communauté. La résistance de réseaux aléatoires à la suppression ciblée ou aléatoire de noeuds du réseau a été étudiée en premier par Albert, Jeong et Barabási [30]. Ils montrèrent qu'un réseau exponentiel reste globalement bien connecté pour quelques défaillances aléatoires ou ciblées et se morcelle ensuite en petits groupes au-delà d'une certaine fraction de défaillances. Un réseau invariant d'échelle montre un comportement similaire sous attaque ciblée des noeuds les plus connectés, mais une bien meilleure résistance aux défaillances aléatoires [65]. D'autres auteurs, comme Cohen *et al.* [66] ont également découvert une transition dans la connectivité de réseaux aléatoires classiques subissant des défaillances. Cependant ces modèles ne considèrent les défaillances que comme des événements indépendants, c'est-à-dire que la défaillance ou l'attaque d'une arête ou d'un noeud particulier n'influe pas la probabilité de défaillance de ses voisins. Ces modèles peuvent être considérés comme des modèles de percolation inversée, dans lesquels en-dessous de la présence d'une certaine fraction du nombre d'éléments initiaux, le composant géant n'existe plus [67].

Dans un réseau réel, la défaillance d'un élément influe sur les éléments voisins. Si par exemple un tronçon de route est fermé entre deux points, la circulation va se reporter sur les axes voisins, et potentiellement les saturer. Repoussant les flux de circulation vers d'autres axes, la perturbation peut se répandre sur le réseau routier. De nombreuses études, comme celle de Guirema *et al.* ont porté sur la congestion du traffic sur Internet [68].

## V.1 Modèle de Motter-Lai

Sur ces considérations, Motter et Lai [69] ont proposé un modèle de réseau dans lequel les éléments portent une charge, définie comme le nombre de plus courts chemins qui passent par l'élément considéré [34]. Chaque noeud j du réseau a une capacité maximale  $C_j$ , définie à partir de la charge initiale  $L_j$  du noeud comme :

$$C_j = (1+\alpha)L_j, \quad j = 1, ..., N,$$
 (V.1)

où  $\alpha \leq 0$  est le paramètre de tolérance du réseau. La suppression de noeuds dans le réseau change la distribution des plus courts chemins du réseau, et donc la charge des éléments. Lorsque la

charge d'un noeud devient plus grande que sa capacité, alors le noeud est supprimé, et peut provoquer la surcharge d'autres noeuds du réseau. Dans un tel processus, Motter et Lai ont montré qu'une cascade globale de défaillances peut se produire si les degrés des noeuds du réseau sont très hétérogènement distribués et que le premier noeud supprimé aléatoirement est l'un des noeuds les plus centraux du réseau. Les auteurs confirment avec cette étude que les réseaux réels sont très résistants à des erreurs aléatoires, mais très sensibles à des attaques qui ciblent les noeuds de plus grande centralité, comme attendu suite à l'étude de Callaway *et al.* [65].

D'autres modèles de cascades, basés sur la topologie du réseau et sur la résistance interne des éléments, ont été définis, comme par exemple le modèle de Watts [20], de Crucitti *et al.* [70] ou de Sachtjen *et al.* [71].

## V.2 Modèle de cascade

Watts [20] constate qu'un petit choc initial peut parfois se propager dans l'ensemble d'un réseau alors que dans la majorité des cas une petite perturbation n'a pas d'effet notable. Il se propose d'étudier la propagation d'une opinion (ou d'une mode ou d'une rumeur), phénomène appelé cascade d'information [72]. Il part du constat qu'un individu peut faire sienne une opinion si un nombre suffisant des individus avec lesquels il est en contact partagent cette opinion. La propagation de cette opinion est modélisée comme suit.

Un individu possède k voisins qui ont soit l'opinion 0, soit l'opinion 1. Si au moins une fraction-seuil  $\phi$  de ses k voisins ont l'opinion 1, alors il l'adopte, sinon il conserve l'opinion 0. Le modèle est complétement aléatoire, en ce que le nombre de voisins et la fraction-seuil de chaque individu sont distribués. Une fraction-seuil est attribuée à chaque individu au début des temps selon la distribution normalisée  $f(\phi)$ ,  $\phi \in (0, 1)$ . On construit ensuite un réseau de N individus, dans lequel le nombre de voisins de chacun est distribué selon  $p_k$ , et dont la valeur moyenne est  $\langle k \rangle = z$ . La dynamique de la cascade est initialisée comme suit : au temps t = 0, tous les individus sont dans l'état 0. À t = 0, une petite fraction  $\Phi_0 \ll 1$  des individus passe dans l'état 1, il y reste.

Watts étudie la probabilité qu'une cascade globale soit déclenchée par un petit noyau initial dans un réseau clairsemé ( $z \ll N$ ), c'est-à-dire que la cascade ait concerné une fraction finie des éléments du réseau. Dans un réseau suffisamment clairsemé, la probabilité de trouver une boucle dans le voisinage local d'un élément est négligeable. Le noyau initial ne peut donc croître que si l'un des éléments voisins est *vulnérable*, c'est-à-dire qu'il a une fraction seuil  $\phi < 1/k$ , ou un degré  $k \leq 1/\phi$ . Une cascade globale ne peut apparaître que s'il existe un sous-réseau connecté de noeuds vulnérables, de taille comparable à celle du réseau. Watts ramène donc le problème dynamique à un problème statique de percolation sur un graphe aléatoire [73].

La fonction génératrice des degrés de noeuds vulnérables est notée  $G_0(x)$ , et permet de générer tous les moments de la distribution du degré des noeuds vulnérables. La condition de cascade est donnée par :

$$G_0''(1) = z_1$$

et permet de définir deux régimes. Si  $G''_0(1) < z$ , tous les sous-réseaux de noeuds vulnérables

sont petits, et les individus vulnérables reliés au noyau initial sont isolés les uns des autres, et ne peuvent contribuer à propager la cascade. Pour  $G_0''(1) > z$ , la taille typique des sousréseaux de noeuds vulnérables est infinie, et si l'un des éléments d'un tel réseau est activé, tout le sous-réseau s'active.

Watts montre que dans ce modèle, la transition vers une cascade globale peut survenir dans deux régimes distincts :

- un régime de faible connectivité, dans lequel la taille des cascades suit une distribution en loi de puissance, avec une pente 3/2, qui correspond à la distribution de la taille des groupes dans les réseaux aléatoires à la percolation. Dans ce régime,  $z \sim 1$ , et presque tous les noeuds sont vulnérables.
- un régime de haute connectivité. Dans ce régime, la propagation de la cascade est limitée par la stabilité locale des individus. Chaque noeud a en effet assez de voisins pour que l'activation de l'un d'eux ne soit pas suffisant pour l'activer à son tour. La distribution de la taille de la cascade décroît donc exponentiellement vite, avec cependant une composante de très grande taille de faible poids statistique.

Nous montrons par la suite que le modèle de Lehmann et Bernasconi [74] d'une propagation de défaillance dans un réseau complet présente des caractéristiques similaires et analytiquement accessibles. Ce modèle est inspiré de concepts des modèles de faisceaux de fibres que nous présentons succinctement dans la section suivante.

## V.3 Modèles de faisceaux de fibres

Les physiciens étudiant les propriétés des matériaux se sont également intéressés aux questions de propagations de défaillances dans les systèmes, distinctement de la communauté des réseaux. L'histoire des études sur les matériaux composés de multiples brins tissés ensemble peut remonter à Léonard de Vinci et à ses expériences sur la résistance des cordes. Les articles fondateurs de ce domaine sont ceux de Peirce (1962) [75] et de Daniels (1945) [76]. On peut renvoyer au papier de revue récent de Pradhan *et al.* [77].

Un faisceau de fibres est un ensemble d'éléments élastiques connectés en parallèle à une plaque qui transmet la force entre toutes les fibres. Une fibre peut supporter une certaine force avant de casser et de reporter l'effort qu'elle supportait sur les autres fibres du faisceau. La force que chaque fibre peut supporter est tirée dans une distribution qui dépend du problème considéré. Si on tire avec une certaine force sur un faisceau de fibres (une corde par exemple), au-delà d'une certaine intensité de la force, la corde casse. Les études sur les faisceaux de fibres cherchent à déterminer comment la rupture se produit. Si la résistance des éléments est distribuée, toutes les fibres ne cassent pas au même moment. On peut aussi observer que dans certains cas, une partie des fibres cède, mais la corde résiste encore.

Cet effet peut être considéré comme analogue à une avalanche de défaillances, largement étudiée dans les années 1990 [78, 79, 80, 81, 82, 83, 84, 85, 77]. Si pour une charge donnée une fibre cède, la fraction de la charge qu'elle supportait est redistribuée aux autres fibres du faisceau. Sous cette surcharge, une ou plusieurs autres fibres peuvent défaillir et redistribuer à leur tour la charge supportée. Ce processus se poursuit jusqu'à ce que toutes les fibres aient cédé ou que les fibres survivantes supportent l'ensemble de la charge. La redistribution des charges peut se faire selon deux grandes familles de modèles : si la charge est redistribuée également à toutes les fibres survivantes, on parle de *partage global de charge*, dont la dynamique a été résolue par Pradhan *et al.* [77, 86] ou Bhattacharyya *et al.* [87]. Si la charge n'est redistribuée qu'à une fraction des fibres survivantes, spatialement proches de la fibre défaillante, on parle de *partage local de charge* [88, 89].

Lehmann et Bernasconi [74] se sont inspirés des modèles de faisceaux de fibres pour proposer un modèle qui décrit la propagation d'une défaillance dans des réseaux d'infrastructures. Les défaillances dans les réseaux électriques ont été très étudiées, principalement par les ingénieurs, pour comprendre et limiter la portée des pannes d'électricité, comme par exemple la panne d'août 2003 (figure V.1) qui a plongé dans le noir une grande partie de la côte est des États-Unis. La majorité des modèles cherchent donc à reproduire précisement les caractéristiques d'un réseau électrique en particulier [90], et ne permettent pas d'extraire les processus fondamentaux à l'origine d'une cascade de défaillances lors d'une panne.

Motter et Lai [69] décrivent un modèle simplifié de cascades dans les réseaux électriques, qui ne permet cependant pas d'effectuer de calculs analytiques sur les quantités entrant en jeu dans les cascades. Leur modèle contient cependant l'idée que la cascade peut se propager vers des noeuds à l'autre bout du réseau, par le biais de la centralité par exemple. Dobson *et al.* [91] ont généralisé cette idée, et proposent de considérer le réseau électrique comme un graphe complet, dans lequel un noeud défaillant redistribue sa charge selon une distribution de Poisson aux noeuds survivants. Ils décrivent analytiquement la propagation d'une cascade de défaillances comme un processus de branchement. Lehmann et Bernasconi ont prolongé ce raisonnement et décrit la famille de défaillances dans un réseau électrique comme un processus de branchement généralisé [92] en bonne approximation. D'autres études, basées sur la topologie du réseau [32] ou sur la dynamique des flux de courant dans le réseau [93] proposent un autre point de vue sur la propagation de défaillances dans les réseaux électriques.

Dans un premier temps, nous présentons le modèle de Lehmann et Bernasconi et vérifions l'approximation en processus de branchement généralisé de la propagation des défaillances dans le système. Grâce au modèle de branchement généralisé appliqué aux caractéristiques du modèle de Lehmann et Bernasconi, nous décrivons analytiquement et numériquement les quantités à l'oeuvre dans la propagation d'une cascade globale.



August 15, 2003 • 9:14 p.m. EDT • About 7 hours after blackout

FIGURE V.1 – Photos satellites de la côte nord-est des États-Unis avant et après la panne d'électricité, le 23 août 2003 (Air Force Weather Agency)

# Chapitre VI Modèles

## VI.1 Définition du modèle L&B

#### VI.1.1 Définition générale

Le modèle de Lehmann & Bernasconi (L&B) décrit la propagation de défaillances dans un réseau électrique, considéré comme statique pour la durée de la cascade. On modélise le réseau commme un ensemble de N éléments complétement connectés les uns aux autres. Chacun de ces éléments porte une charge, déterminée à l'origine des temps, et l'un de ces éléments défaille. La redistribution de sa charge au reste du réseau peut engendrer une cascade par générations de défaillances successives. Ce modèle dépend de deux paramètres :  $\alpha$  représente la marge de tolérance du réseau, et  $\beta$  la valeur de la fraction redistribuée des charges défaillantes. La cascade de défaillances évolue comme suit.

#### 1. À la génération s = 0:

Chaque élément porte une charge  $u_i(0)$ , i = 1, ..., N. Les  $u_i(0)$  sont des variables aléatoires indépendantes et identiquement distribuées selon la distribution  $f_{L(0)}$ . Les charges initiales de tous les noeuds sont tirées une fois et une seule à cette étape s = 0 et évoluent tout au long de l'histoire du réseau. À chaque noeud i du réseau est associée une charge-seuil  $u_i^{max}$  définie à partir de sa charge initiale aléatoire  $u_i(0)$ :

$$u_i^{max} = (1+\alpha)u_i(0). \tag{VI.1}$$

Cette charge-seuil est ainsi construite qu'aucun des noeuds n'est en surcharge à la génération initiale s = 0. Si à une génération s de la cascade la charge  $u_i(s)$  du noeud i devient supérieure à ce seuil :  $u_i(s) > u_i^{max}$ , alors l'élément i défaille à son tour.

La cascade est initialisée par un élément j choisi au hasard, de charge  $L(0) = \lambda = u_j(0)$ , qui défaille.

Le nombre de défaillances  $N_0$  de la génération s = 0 est donc :  $N_0 = 1$ , et le nombre total d'éléments  $M_0$  qui ont défailli jusqu'à l'étape s = 0 est :  $M_0 = N_0 = 1$ .

2. À la génération s = 1 :

La charge L(0) de l'élément qui a défailli à la génération s = 0 est redistribuée aux  $\nu(1) = N - M_0$  éléments survivants au début de la génération s = 1. Chaque élément i

apartenant aux  $\nu(1)$  survivants reçoit une fraction aléatoire  $\Delta_{i,1}(0)$  de la charge défaillante  $L(0) = \lambda$ , et sa charge à la première génération est donc :

$$u_i(1) = u_i(0) + \Delta_{i,1}(0)\lambda. \tag{VI.2}$$

Les variables aléatoires  $\Delta_{i,1}(0)$  sont définies de telle façon que, en moyenne, la charge défaillante L(0) est entièrement redistribuée sur les  $\nu(1)$  éléments survivants :  $\nu(1) \langle \Delta(0) \rangle =$ 1, soit, pour  $\nu(1) \gg 1$  :

$$\sum_{i} \Delta_{i,1}(0) \approx 1. \tag{VI.3}$$

Parmi les  $\nu(1)$  éléments survivants au début de la génération s = 1, un nombre aléatoire  $N_1 = n$  d'éléments de charges  $L_1(1), \ldots, L_n(1)$  défaillent suite à la surcharge provoquée par la redistribution de la charge L(0). Le nombre total de défaillances jusqu'à l'étape s = 1 est donc  $M_1 = M_0 + N_1$ .

3. À la génération s = 2 :

Les  $N_1 = n$  défaillances de la génération s = 1 redistribuent leur charge à chaque élément *i* parmi les  $\nu(2) = N - M_1$  éléments survivants par fractions aléatoires  $\Delta_{i,1}(1), \ldots, \Delta_{i,n}(1)$ indépendantes et identiquement distribuées. La charge  $u_i(2)$  d'un élément *i* parmi les  $\nu(2)$ survivants à l'étape s = 2 est alors :

$$u_i(2) = u_i(1) + \Delta_{i,1}(1)L_1(1) + \dots + \Delta_{i,n}(1)L_n(1), \qquad (\text{VI.4})$$

où les variables aléatoires  $\Delta_{i,j}(1)$  vérifient la conservation de chaque charge défaillante associée  $L_j(1)$ :

$$\nu(2) \left\langle \Delta_{i,j}(1) \right\rangle = 1, \ j = 1, \dots, n, \tag{VI.5}$$

soit pour  $\nu(2) \gg 1 : \sum_i \Delta_{i,j}(1) \approx 1, \ \forall j$ 

Un nombre aléatoire  $N_2$  d'éléments défaillent suite à cette redistribution de charge et engendreront à leur tour des défaillances subséquentes. Le nombre total de défaillances jusqu'à l'étape s = 2 est donc  $M_2 = M_1 + N_2$ .

4. À la génération s :

Les défaillances de charges  $L_1(s-1), \ldots, L_{N_{s-1}}(s-1)$  de la génération précédente s-1 distribuent leur charges aux  $\nu(s)$  éléments survivants au début de la s-ième génération, de charges  $u_i(s-1)$ , où le nombre de survivants est donné par :

$$\nu(s) = N - M_{s-1} = N - \sum_{t=0}^{s-1} N_t = \nu(s-1) - N_{s-1}$$
(VI.6)

avec  $M_{s-1}$  le nombre total de défaillances jusqu'à la fin de la génération s-1 et  $N_t$  le nombre de défaillances survenues pendant la génération t.

Chaque élément *i* survivant reçoit une fraction aléatoire  $\Delta_{i,j}(s-1)$  de la charge  $L_j(s-1)$  de chacune des  $N_{s-1} = n$  défaillances de la génération précédente s - 1:

$$u_i(s) = u_i(s-1) + \sum_{j=1}^n \Delta_{i,j}(s-1)L_j(s-1).$$
 (VI.7)

Pour assurer la conservation en moyenne des charges défaillantes, la relation suivante sur les  $\Delta(s)$  doit être vérifiée à toute génération s:

$$\nu(s) \left\langle \Delta(s-1) \right\rangle = 1, \tag{VI.8}$$
  
$$j = 1, \dots, N_{s-1}.$$

où la moyenne est prise sur l'ensemble des valeurs possibles pour j donné. Ces variables aléatoires  $\Delta(s-1)$  dépendent du nombre total de défaillances, et donc de l'histoire de la cascade depuis la génération d'origine s = 0.

Suite à cette redistribution,  $N_s$  éléments défaillent à la génération s et redistribueront leur charge aux survivants lors de la génération suivante de la cascade. La taille  $M_s$  de la cascade à la fin de cette génération est donc :  $M_s = M_{s-1} + N_s$ 

On itère ce processus jusqu'à ce qu'à une génération s la cascade n'engendre plus de défaillances :  $N_s = 0$ , ou que tous les noeuds aient défailli :

$$M_s = \sum_{t=0}^{s} N_t = N.$$
 (VI.9)

Ce dernier cas correspond à la rupture du réseau.

Les notations sont résumées en appendice B.

### VI.1.2 Redistribution bimodale

Lehmann et Bernasconi considèrent un mode de redistribution simple : la variable aléatoire  $\Delta$  est donnée par une distribution bimodale. À la génération s, toute charge survivante se voit ajouter une fraction aléatoire  $\Delta(s-1)$  de chaque charge défaillante de la génération précédente, où :

$$\Delta(s-1) = \begin{cases} \beta & \text{avec probabilité } p(s) \\ 0 & \text{avec probabilité } 1-p(s) \end{cases},$$
(VI.10)

où  $\beta \in (0,1)$  est la fraction de charge redistribuée avec probabilité p(s) aux éléments de la génération s. La probabilité p(s) est définie de telle façon que chaque charge défaillante est entièrement redistribuée en moyenne, selon (VI.8). La valeur moyenne de la variable aléatoire distribuée selon cette loi bimodale est donc :

$$\langle \Delta(s-1) \rangle = \beta p(s) = \frac{1}{\nu(s)}.$$
 (VI.11)

Soit  $u_i(s)$  une charge prise au hasard parmi les  $\nu(s)$  charges survivantes au début de la génération s. Pour chacune des  $N_{s-1}$  défaillances de l'étape précédente, la fraction redistribuée à l'élément *i* est tirée dans la distribution bimodale (VI.10). La charge  $u_i(s)$  subit une surcharge  $\beta L_j$  due à la défaillance  $j = 1, \ldots, N_{s-1}$  avec probabilité p(s), et défaille si sa valeur après redistribution de toutes les  $N_{s-1}$  charges défaillantes dépasse sa valeur-seuil  $u_i^{max}$ .

Ce cas n'est traitable analytiquement que sous certaines approximations que nous détaillerons section VI.3.

## VI.2 Phénoménologie du modèle L&B

On peut s'intéresser à de multiples quantités caractéristiques du modèle de cascade, que l'on décrit ici.

#### VI.2.1 Probabilité de non-cascade

L'une des premières quantités que l'on peut considérer est la probabilité que la première défaillance ne soit suivie d'aucune autre, c'est-à-dire qu'elle n'engendre aucune défaillance sub-séquente : la probabilité de non-cascade  $P_{nc}$ . Cette probabilité dépend des deux paramètres du modèle :

- $\beta$  la fraction de la charge qui est redistribuée aux éléments subissant une surcharge suite à la première défaillance et l'inverse du nombre moyen de noeuds subissant surcharge,
- $\alpha$  qui exprime la résistance d'un élément à une surcharge.

La probabilité  $P_{nc}(\lambda)$  que la défaillance d'un élément de charge  $L(0) = \lambda$  à la génération s = 0ne soit suivie d'aucune autre défaillance est bien évidemment dépendante de la charge  $\lambda$ , tirée de la distribution  $f_{L(0)}$  des charges initiales L(0). La probabilité de non-cascade  $P_{nc}$  est la moyenne sur la distribution des charges initiales de la probabilité de non-cascade  $P_{nc}(\lambda)$  suite à une défaillance initiale de charge  $L(0) = \lambda$ .

On représente sur la figure VI.1 la probabilité de non-cascade obtenue par simulation du modèle L&B en fonction de  $\alpha$  pour différentes valeurs du second paramètre  $\beta$ . Pour des valeurs de  $\beta \leq 1$  (triangles), la probabilité de non-cascade est non nulle même pour  $\alpha = 0$ . En effet, la probabilité de subir une surcharge à la première génération est donnée par l'équation (VI.11) :  $p(1) = 1/\beta(N-1)$ , et pour  $\beta$  grand, la probabilité est faible. Dans ce cas, la cascade s'interrompt à la génération s = 0 car aucun noeud n'a subi de surcharge.

## VI.2.2 Distribution du nombre total de défaillances dans la cascade

Si la cascade continue au-delà de la première défaillance, on peut s'intéresser au nombre total d'éléments M ayant défailli durant toute l'histoire de la cascade. Ce nombre dépend de la charge des défaillances tout au long de la cascade, car un élément fortement chargé a une probabilité plus élevée qu'un élément faiblement chargé d'engendrer un grand nombre de défaillances. De plus, la charge des défaillances descendantes dépend de la charge défaillante qui les engendre. Le nombre total de défaillances  $M(\lambda)$  est donc dépendant de la charge  $L(0) = \lambda$  de la première défaillance, distribuée selon la distribution  $f_{L(0)}$ . On ne considère ici que la moyenne sur la distribution de la charge initiale  $f_{L(0)}$  du nombre total de défaillances, noté M.

La forme de la distribution du nombre total de défaillances dépend de la valeur des deux paramètres du modèle  $\alpha$  et  $\beta$ . Pour une valeur de  $\alpha$  donnée on fixe la taille du système, et on fait varier le paramètre  $\beta$ . On observe sur la figure VI.2(a) l'apparition pour certaines valeurs de  $\beta$ d'un pic de probabilité pour la valeur M = N. Il correspond au poids de toutes les simulations du modèle L&B dont la taille de la cascade est supérieure à N. Le seuil d'apparition de ce pic dépend de la taille du système et de l'autre paramètre  $\alpha$ . On note  $\beta_c(\alpha)$  la valeur-seuil d'apparition de ce pic à la limite  $N \to \infty$ . À la limite asymptotique, l'existence d'un pic à  $N \to \infty$  signifie que la probabilité d'obtenir une cascade infinie est non nulle. Dans un système



FIGURE VI.1 – Probabilité de non-cascade en fonction de  $\alpha$  pour différentes valeurs de  $\beta$ .

fini, on observe un pic de probabilité pour la valeur M = N pour de plus petites valeurs de  $\beta$ que  $\beta_c$ . Ce pic correspond à la probabilité que le système présente des cascades de grande taille, et la valeur de  $\beta$  pour lequel il apparaît dépend donc de la taille du système et du paramètre  $\alpha : \beta_{c,N}(\alpha, N)$ .

À  $\beta_c$ , la distribution du nombre total de défaillances tend vers une loi de puissance si la taille du système augmente, comme on le voit sur la figure VI.2(b). Pour des tailles finies, la distribution est une loi de puissance pour les petites valeurs de M et décroit plus rapidement ensuite. Avant l'apparition d'un pic à n = N, la distribution du nombre total de défaillances suit une loi exponentielle (points bleus).

Pour  $\beta = 1$ , le nombre de défaillances est distribué selon une loi de puissance d'exposant 3/2. Ce comportement correspond au régime critique des processus de branchement.

Si le nombre total de défaillances M = N, alors tous les éléments du système ont défailli, et on parle alors de rupture du réseau.

## VI.2.3 Probabilité de rupture

On regarde plus en détail la probabilité de rupture  $P_b$ , c'est-à-dire la probabilité pour qu'une cascade initiée par une défaillance de charge quelconque ne s'achève pas avant que les éléments du réseau aient défailli. Dans un réseau infini, la probabilité de rupture est la probabilité pour que la cascade se poursuive indéfiniment. On note  $P_b(\lambda)$  la probabilité de rupture pour une cascade initiée par une défaillance de charge  $L(0) = \lambda$ . La probabilité de rupture  $P_b$  pour une



(a) À  $\alpha$  fixé pour différentes valeurs de  $\beta$ 



(b) Proche de la valeur critique  $\beta_c$  pour différentes tailles du système.

FIGURE VI.2 – Distribution du nombre de défaillances totales d'une cascade dans un système de taille N. Sur la figure (a) la taille du système N et le paramètre  $\alpha$  sont fixés et  $\beta$  prend ses valeurs sous la valeur seuil, à la valeur-seuil  $\beta_{c,N}$ , à la valeur critique  $\beta_c \approx 0.33$  et au-dessus de la valeur critique. Sur la figure (b), on choisit les paramètres du système proches de la valeur critique  $\beta_c$  pour  $\alpha$  donné et on varie la taille N du système.

charge initiale quelconque est la moyenne de la probabilité  $P_b(\lambda)$  sur la distribution  $f_{L(0)}$  de la charge de la défaillance initiale  $L(0) = \lambda$ .

La valeur de la probabilité de rupture dépend des paramètres du problème  $\alpha$  et  $\beta$ , et de la taille N du système. On remarque sur la figure VI.3 qu'il existe un seuil critique  $\beta_c(N)$ dépendant des deux paramètres et de la taille du système, pour lequel la probabilité de rupture devient non nulle. Si la taille du système augmente, ce seuil se déplace vers les plus grandes valeurs de  $\beta$ , comme on l'observe sur la figure VI.3(b), et tend vers une valeur limite. On s'attend à ce que la valeur-seuil converge vers une limite non triviale, et que cette valeur soit proche de la valeur  $\beta_c$  de la section précédente pour  $N \to \infty$ , pour une valeur de  $\alpha$  donnée.

La probabilité de rupture atteint une valeur maximale pour  $\beta$  entre  $\beta_c$  et 1 avant de décroître vers 0 pour la valeur de transition supérieure  $\beta = 1$ . Cette décroissance est due à la définition de la probabilité de surcharge p(s). En effet, la probabilité de surcharge est définie de telle sorte qu'en moyenne toute la charge est redistribuée, c'est-à-dire que le nombre de noeuds subissant une surcharge est  $1/\beta$ . Or, si  $\beta = 1$ , en moyenne un seul noeud subit une surcharge. Si ce noeud ne défaille pas, cette branche de la cascade s'éteint. De plus, les fluctuations statistiques peuvent entraîner qu'une charge défaillante ne distribue pas de surcharge. Ces deux effets s'additionnent, et la cascade de défaillance s'éteint avec probabilité 1 pour  $\beta = 1$ , si la taille N du système est assez grande.

## VI.2.4 Valeurs de transition

On observe une transition d'une valeur nulle à une valeur non-nulle de la probabilité de rupture pour une valeur donnée de  $\beta < 1$  qui dépend de  $\alpha$ . Dans le modèle de Lehmann & Bernasconi, cette valeur-seuil dépend de la taille du système, mais converge pour  $N \to \infty$  vers une valeur non-triviale  $\beta_c(\alpha)$ . On retrouve la transition entre une cascade de défaillances de taille totale finie et une cascade de taille infinie avec probabilité non nulle dans l'étude de toutes les quantités caractéristiques de la cascade.

On représente sur la figure VI.4 les valeurs de  $\beta_c(N)$  en fonction du second paramètre  $\alpha$ , relevées pour des simulations de réseaux de différentes tailles N. Ces valeurs seuil semblent converger vers une valeur limite pour  $N \to \infty$ .

Pour  $\beta \to 1$ , la probabilité de rupture  $P_b$  converge vers une valeur nulle, quelle que soit la valeur du paramètre  $\alpha$ . On étudiera la distribution des différentes quantités observées à cette valeur de transition supérieure pour en comprendre les propriétés.

## VI.2.5 Nombre de descendants directs d'une défaillance

Une défaillance à une génération s distribue des fractions  $\beta$  de sa charge  $L(s) = \ell$  aux éléments survivants à la génération suivante s+1. Si l'élément surchargé défaille à cette génération, on considère que la défaillance a été engendrée par la défaillance de charge L(s). L'ensemble de tous les éléments subissant une surcharge  $\beta \ell$  et défaillant à la génération s + 1 sont les descendants directs de la défaillance de charge L(s). On note  $N_1(\ell)$  le nombre de descendants directs de l'élément de charge  $L(s) = \ell$ . Le nombre de descendants directs d'un individu dépend des paramètres du problème, car la probabilité de subir une surcharge et l'intensité de cette surcharge dépendent du paramètre  $\beta$ , et la résistance de chaque élément dépend du paramètre



(a) Pour différentes valeurs du paramètre  $\alpha$ .



(b) Pour différentes tailles N du système.

FIGURE VI.3 – Probabilité de rupture dans le modèle Lehmann & Bernasconi en fonction du paramètre  $\beta$ , pour une taille N du système fixée, pour différentes valeurs de  $\alpha$  (a), et pour une valeur de  $\alpha$  fixée, pour différentes tailles du système (b).



FIGURE VI.4 – Estimation numérique de la valeur du paramètre  $\beta_c$  à la transition en fonction du second paramètre  $\alpha$  pour un réseau de taille finie.

 $\alpha$ . Il dépend aussi de la charge de l'ancêtre, car un élément portant une charge importante provoquera avec une plus grande probabilité la défaillance des éléments qu'il surcharge, et aura donc plus de descendants qu'un individu avec une faible charge.

La probabilité de subir une surcharge est donnée par la relation (VI.11), et dépend donc du nombre  $\nu(s+1)$  d'éléments survivants dans le système à la génération s+1. La distribution du nombre de descendants d'un élément de charge L(s) est également liée au nombre d'éléments survivants dans le système à la génération considérée. On trace sur la figure VI.5 la distribution du nombre de descendants d'éléments de différentes charges pour différents nombres de survivants et valeurs de  $\beta$ .

Sur la figure VI.5(a), on fixe la taille du système N, et les paramètres  $\alpha$  et  $\beta$  dans la simulation. On fait varier le nombre  $\nu$  limite de survivants jusqu'auquel on considère la distribution de  $N_1(\lambda)$ , c'est-à-dire que l'on arrête la simulation dès que le nombre de survivants est plus petit qu'un seuil arbitraire. Pour un nombre de survivants proche de la taille du système :  $\nu(s) > N/2$ , on vérifie sur la figure VI.5(b) que les charges défaillantes dont la valeur est plus élevée engendrent plus de défaillances que les plus petites charges. On observe également sur la figure VI.5(c) l'influence du paramètre  $\beta$  sur la distribution du nombre de défaillances engendrées par un élément d'une charge donnée.

### VI.2.6 Nombre d'ancêtres d'une défaillance

La phénoménologie du modèle est différente selon que le système est paramétré par des valeurs de  $\beta$  inférieures à la valeur de transition ou supérieures. On observe sur la figure VI.6(a) que les éléments du système sous les valeurs de transition (carrés bleus) subissent une seule surcharge avec une probabilité proche de 1, alors que pour un système au-dessus des valeurs de transition, la distribution du nombre d'ancêtres d'une défaillance  $n_{anc}$  a une queue de probabilité plus large.

Dans des systèmes de taille infinie, la probabilité pour qu'un individu subisse une surcharge est de l'ordre de

$$\frac{1}{\nu(s)} \approx \frac{1}{N} \ll 1,\tag{VI.12}$$

et donc la probabilité pour qu'il en subisse deux est encore plus faible et peut être négligée. Dans le modèle de Lehmann & Bernasconi, la taille du système est finie, et la probabilité de subir plusieurs surcharges avant défaillance peut être plus importante si le nombre de survivants  $\nu(s)$ à une génération quelconque s est petit devant la taille N du système. On trace sur la figure VI.6(b) la distribution du nombre de surcharges subies par un élément avant défaillance si le nombre de survivants au moment de la défaillance est indifférent (carrés rouge) et si le nombre de survivants  $\nu(s)$  à la génération s de la défaillance reste de l'ordre de la taille du système N. Pour un nombre limite de survivants  $\nu(s)$  quelconque, la probabilité que le nombre d'ancêtres d'une défaillance soit supérieur à 1 est plus élevée que pour un nombre limite de survivants petit devant la taille du système.

## VI.2.7 Nombre de membres d'une génération

Chaque génération de défaillances peut être vue comme un ensemble de charges indépendantes, dont la taille dépend du nombre et de la charge des défaillances de la génération précé-


(a) Pour différentes valeurs du nombre de survi- (b) Pour différentes valeurs de la charge défaillante vants  $\nu(s)$  à toute génération s L



(c) Pour différentes valeurs de  $\beta$ 

FIGURE VI.5 – Distribution du nombre de descendants directs  $N_1(L)$  d'une défaillance de charge L pour un système de taille N et de paramètre  $\alpha$  fixés.



(b) Pour différentes valeurs du nombre de survivants  $\nu(s)$  à toute génération s

FIGURE VI.6 – Distribution du nombre de défaillances des générations précédentes ayant causé la défaillance d'un élément, pour une valeur donnée de  $\alpha$ , si N est fixé et  $\beta$  prend des valeurs entre  $\beta < \beta_c$  et  $\beta > \beta_c$  (a); si  $\beta$  est fixé et que l'on considère le cas où le nombre de survivants reste notablement plus grand que la taille du système N (triangles), ou tous les cas (carrés) (b). dente, ainsi que de l'histoire antérieure de la cascade (surcharges déjà subies par les différents éléments, nombre d'éléments survivants...). On s'intéressera à l'évolution de la taille des générations au cours de l'histoire de la cascade.

## VI.2.8 Temps d'arrêt de la cascade

Une cascade s'arrête lorsque tous les éléments du système ont défailli, ou lorsqu'une génération ne compte aucun membre, puisqu'alors plus aucune défaillance ne sera engendrée dans le système. La génération s à laquelle la cascade s'éteint est appelée le temps d'arrêt de la cascade. On distingue les deux cas, selon que la cascade s'arrête faute d'éléments survivants, et on note alors  $s_r$  la génération d'arrêt, ou qu'elle s'arrête à la génération  $s_a$  car elle n'engendre plus de défaillances. On observe sur la figure VI.7 que la durée mise par la cascade pour s'arrêter, dans chacun des deux cas est plus longue lorsque le système est proche de la valeur de transition  $\beta_c \approx 0.33$ , c'est-à-dire lorsque le nombre total de défaillances est distribué en loi de puissance.

## VI.2.9 Charge défaillante

Chaque élément i porte une charge  $u_i$ ,  $i = 1, \ldots, N$ , et défaille lorsque cette charge dépasse le seuil de tolérance de l'élément. La charge  $L = \ell$  de l'élément qui défaille est ensuite redistribuée aux noeuds survivants du système. La poursuite de la cascade de défaillances dépend donc de la valeur de la surcharge  $\beta \ell$  imposée aux éléments survivants, c'est-à-dire de la valeur de la charge L des éléments défaillants. On s'intéresse donc à la distribution de la charge défaillante à différentes générations. La distribution des charges défaillantes à chaque étape est une distribution de sommes de variables aléatoires. La distribution des charges défaillantes L(s) à une étape s dépend de la distribution des charges défaillantes L(s-1) à l'étape précédente s-1, et de la génération s considérée. Elle dépend également des paramètres  $\alpha$  et  $\beta$  du système, qui donnent les règles de défaillance des éléments. On représente sur la figure VI.8 la distribution de la charge défaillante pour différentes générations, en fonction de la valeur du paramètre  $\beta$ . La distribution de la charge défaillante à la génération s = 0 est toujours la distribution  $f_{L(0)}$ , uniforme sur (0,1). Dès la première génération, la distribution se déforme. Pour des valeurs des paramètres telles que la probabilité de rupture est nulle, on observe que la distribution se concentre très vite autour de petites valeurs de L, tandis que pour  $P_b \neq 0$ , la distribution couvre des valeurs finies de L.

## VI.3 Du modèle Lehmann & Bernasconi au processus de branchement généralisé

Lehmann & Bernasconi ont proposé [74] une série d'approximations pour définir un cadre analytique à l'étude des différentes quantités présentées. Nous rappelons et discutons ici ces approximations. On rappelle que les défaillances d'une génération en engendrent d'autres à la génération suivante. Un noeud *i* défaille à une étape *s* s'il subit une surcharge et si la nouvelle valeur de la charge  $u_i(s)$  dépasse le seuil de tolérance  $u_i^{max}$ .



(a) Distribution de la génération d'arrêt par extinction



(b) Distribution de la génération de la rupture

FIGURE VI.7 – Distribution de la génération d'arrêt de la cascade pour différentes valeurs de  $\beta$ , pour un système de taille N et de paramètre  $\alpha$  fixés. L'arrêt de la cascade est dû à l'extinction du nombre de défaillances à une génération  $s_a$  (a), ou à la rupture du système à une génération  $s_r$  (b).



(c) Au-dessus de la transition  $\beta = 0.5 > \beta_c$ 

FIGURE VI.8 – Distribution des charges défaillantes à une génération donnée s pour différentes valeurs du paramètre  $\beta$ , sous la transition (a), à la transition (b) et au-dessus de la transition (c). Les simulations sont réalisées pour  $\alpha = 0.2$  et un système de taille N = 2000, et les quantités sont moyennées sur 10000 échantillons.

#### VI.3.1 Approximation pour un système de taille infinie

Dans un ensemble infini de charges, on considère qu'un élément ne reçoit qu'une seule surcharge au cours de son histoire, *i.e.* que la distribution des charges survivantes u(s) à l'étape s est identique à la distribution des u quelle que soit l'étape considérée. En effet, un élément subit une surcharge lors d'une génération s quelconque avec probabilité  $p(s) \sim 1/\nu(s)$ . Si le réseau est infini  $N \to \infty$ , le nombre d'éléments survivants  $\nu(s) \sim N$  à chaque étape est infini, et la probabilité pour un élément donné de recevoir une surcharge est très faible :  $p(s) \ll 1$ . La probabilité pour cet élément de recevoir deux surcharges au cours de son histoire est encore plus faible :  $p(s)p(t) \sim 1/N^2 \sim 0$ , et peut donc être négligée dans la limite des systèmes de taille infinie.

La probabilité qu'un noeud i de charge :

$$u_i(s) = u_i(0) + \Delta(s-1)L(s-1),$$
 (VI.13)

pris au hasard à l'étape s, défaille suite à une défaillance de charge  $L(s-1) = \ell$  à la génération précédente est donc :

$$P(u_i(s) > u_i^{max}) = p(s)P(u_i(0) + \beta \ell > (1 + \alpha)u_i(0)) = p(s)P(\alpha u_i(0) < \beta \ell).$$
(VI.14)

Pour une distribution  $f_{L(0)}$  des charges initiales  $u_i(0)$  uniforme sur (0, 1), la probabilité que l'ajout d'une surcharge  $\beta \ell$  suffise à dépasser la charge-seuil  $u_i^{max}$  de l'élément ne dépend que de la valeur des paramètres du problème  $\alpha$  et  $\beta$ , et de la valeur de la charge défaillante  $L(s-1) = \ell$  comme :

$$P(\alpha u_i(0) < \beta \ell | L(s-1) = \ell) = \begin{cases} 1 & \text{si } \alpha < \beta \ell \\ \frac{\beta \ell}{\alpha} & \text{sinon} \end{cases} = \min\left(1, \frac{\beta \ell}{\alpha}\right).$$
(VI.15)

La probabilité que l'élément *i* de charge  $u_i(s)$  défaille suite à la défaillance de charge L(s-1)à la génération s-1 est donc :

$$P(u_i(s) > u_i^{max} | L(s-1) = \ell) = p(s) \min\left(1, \frac{\beta\ell}{\alpha}\right).$$
(VI.16)

### VI.3.2 Nombre de descendants d'une défaillance

Un élément de charge  $L(s-1) = \ell$  qui défaille à l'étape (s-1) engendre  $N_1(\ell)$  défaillances directes à la génération s parmi les  $\nu(s)$  survivants. L'indice 1 souligne que les défaillances sont engendrées avec une génération d'écart. La probabilité d'engendrer  $N_1(\ell) = n$  descendants est donnée par la loi binomiale :

$$P(N_1(\ell) = n) = {\binom{\nu(s)}{n}} p^n \left(1 - p\right)^{\nu(s) - n}, \qquad (\text{VI.17})$$

où  $p = P(u_i(s) > u_i^{max} | L(s-1) = \ell) = p(s) \min(1, \frac{\beta \ell}{\alpha})$ , avec l'approximation que l'élément *i* de charge  $u_i(s)$  n'a subi qu'une seule surcharge au cours de son histoire.

Le nombre moyen de descendants de la défaillance de charge  $L(s-1) = \ell$  à l'étape s est :

$$\mu(\ell) = \sum_{k=0}^{\nu(s)} k P(N_1(\ell) = k)$$
  
=  $\nu(s) p(s) \min\left(1, \frac{\beta \ell}{\alpha}\right).$  (VI.18)

Par définition de la probabilité p(s) de surcharge à l'étape s d'un élément pris au hasard (VI.11), le nombre typique de descendants d'une défaillance de charge  $L(s-1) = \ell$  se simplifie en :

$$\mu(\ell) = \min\left(\frac{1}{\beta}, \frac{\ell}{\alpha}\right). \tag{VI.19}$$

Pour un nombre de survivants  $\nu(s)$  assez grand, la loi binomiale pour le nombre de descendants peut être approximée par une loi de Poisson de paramètre  $\mu(\ell)$ :

$$P(N_1(\ell) = n) = \frac{\mu(\ell)^n}{n!} e^{-\mu(\ell)},$$
 (VI.20)

de fonction génératrice :

$$G(\ell, t) = \left\langle t^{N_1(\ell)} \right\rangle = e^{\mu(\ell)(t-1)}.$$
 (VI.21)

### VI.3.3 Distribution conditionnelle des charges défaillantes

La charge  $L = \ell$  engendre  $N_1(\ell) = n$  défaillances à la génération suivante. Les charges  $L_1, \ldots, L_n$  de ces n descendants sont indépendantes les unes des autres. La charge  $L_i = \ell_i$  de la défaillance i issue de  $L = \ell$  est distribuée selon une densité de probabilité  $f_{L_i|L}(\ell_i|\ell)$  que nous allons déterminer. Les charges des défaillances qui surviennent à la suite d'une défaillance de charge  $\ell$  sont construites de la façon suivante :

- La charge initiale  $u_i(s)$  de l'élément défaillant *i* est tirée de la distribution  $f_{L(0)}$ , uniforme sur (0, 1), pour un système de taille infinie.
- Cette charge initiale u<sub>i</sub>(s), augmentée de la surcharge βℓ doit être supérieure à la valeurseuil u<sub>i</sub><sup>max</sup> de l'élément i pour que l'élément défaille :

$$u_i(s) + \beta \ell > u_i^{max} = u_i(s)(1+\alpha).$$
(VI.22)

- La charge initiale  $u_i(s)$  est donc contrainte à vérifier la condition :  $\alpha u_i(s) < \beta \ell$ .
- Deux cas se présentent :
  - si  $\beta \ell > 1$ , alors toute charge  $u_i(s)$  tirée de  $f_{L(0)}$  vérifie la condition :  $u_i(s) < 1 < \frac{\beta \ell}{\alpha}$ , - si  $\beta \ell < 1$ , alors la charge initiale  $u_i(s)$  doit vérifier :  $u_i(s) < \frac{\beta \ell}{\alpha} < 1$ .
- C'est-à-dire que la charge  $u_i(s)$  est tirée dans la distribution uniforme entre 0 et min  $(1, \frac{\beta \ell}{\alpha})$ .
- Elle subit une surcharge  $\beta \ell$ . La charge défaillante  $\ell_i = u_i(s) + \beta \ell$  causée par la défaillance de charge  $\ell$  est donc distribuée selon la densité de probabilité conditionnelle :

$$f_{L_i|L}(\ell_i|\ell) = \frac{1}{\min\left(1,\frac{\beta\ell}{\alpha}\right)}, \text{ si } \beta\ell < \ell_i < \beta\ell + \min\left(1,\frac{\beta\ell}{\alpha}\right).$$
(VI.23)

Cette densité de probabilité ne dépend donc que de la distribution initiale  $f_{L(0)}$  des charges et de la valeur de la charge défaillante précédente. L'histoire complète des défaillances n'intervient pas dans cette densité.

Sous ces approximations, le processus est entièrement déterminé par des règles de passage simples d'une génération à la suivante. Le nombre de descendants est défini par une loi de Poisson (VI.20), et les charges des descendants ne dépendent que de la charge de l'ancêtre et sont distribuées selon la loi de probabilité (VI.23).

## VI.4 Processus de branchement généralisé de Harris

Harris [92] propose un formalisme qui décrit l'évolution d'une famille dont les membres possèdent une certaine caractéristique qui influence leur descendance (voir le détail en appendice A). Sous les approximations détaillées section VI.3, le modèle de Lehmann & Bernasconi est un processus de branchement généralisé de Harris. On définit le modèle H comme un processus de branchement généralisé avec les caractéristiques issues du modèle de Lehmann & Bernasconi.

Un élément défaille à l'étape s = 0. Sa charge  $L(0) = \lambda$  est tirée dans la distribution  $f_{L(0)}$ uniforme sur (0, 1). Cette défaillance initiale engendre un nombre de défaillances subséquentes  $N_1(\lambda)$  distribué selon une loi de Poisson (VI.20) de paramètre  $\mu(\lambda)$ , qui dépend de la valeur  $\lambda$  de la charge initiale. Les défaillances produites à l'étape s = 1 engendrent à leur tour une génération de défaillances à l'étape s = 2, pour chacune selon une loi de Poisson dont le paramètre dépend de sa charge.

La charge d'une défaillance-fille *i* quelconque  $L_i = \ell_i$  dépend de la charge  $L = \ell$  de la défaillance de l'étape précédente qui l'engendre. Elle est donnée par  $\ell_i = u_i + \beta \ell$  où  $u_i$  est une variable aléatoire distribuée selon la densité de probabilité uniforme  $f_{u_i|L}$  qui dépend de la valeur de la charge de la défaillance L à l'étape précédente. Ce modèle reproduit les caractéristiques du modèle de L&B sous les approximations détaillées auparavant, et en particulier le fait que les charges défaillantes sont construites à partir d'une variable uniforme  $u_i$  comprise dans (0, 1) et vérifiant les conditions de défaillance :  $\alpha u_i < \beta \ell$ . La densité de probabilité de  $u_i$  suite à une défaillance-mère de charge  $L = \ell$  est donc :

$$f_{u_i|L}(u_i|\ell) = \begin{cases} \frac{1}{\min(1,\beta\ell/\alpha)} & \text{si } 0 < u_i < \min\left(1,\frac{\beta\ell}{\alpha}\right) \\ 0 & \text{sinon} \end{cases}, \quad (VI.24)$$

où  $\beta$  est la fraction de la charge de la défaillance redistribuée à l'élément et causant sa défaillance, et  $\alpha$  le seuil de défaillance, défini tel que l'élément *i* défaille si l'augmentation de sa charge dépasse une fraction  $\alpha$  de sa charge initiale  $u_i$ .

La densité de probabilité conditionnelle  $f_{L_i|L}$  que l'élément soit de charge  $\ell_i = u_i + \beta \ell$  si la charge de l'élément qui l'a engendré est  $L = \ell$  est donc :

$$f_{L_i|L}(\ell_i|\ell) = \frac{1}{\min\left(1, \frac{\beta\ell}{\alpha}\right)}, \text{ si } \beta\ell < \ell_i < \beta\ell + \min\left(1, \frac{\beta\ell}{\alpha}\right).$$
(VI.25)

Le paramètre de la loi de Poisson  $\mu(\ell)$  pour le nombre de défaillances issues d'une charge  $L = \ell$  est défini comme :

$$\mu(\ell) = \min\left(\frac{1}{\beta}, \frac{\ell}{\alpha}\right). \tag{VI.26}$$

Ces hypothèses de construction de la cascade dérivent directement de la redistribution bimodale sur réseau infini proposée par Lehmann & Bernasconi, que l'on a détaillée dans la section précédente VI.3.

## VI.5 Validité des approximations

On vérifie par le biais de la simulation de l'algorithme décrit dans la section VI.1 la validité des approximations effectuées section VI.3.

## VI.5.1 Distribution conditionnelle de la charge défaillante

Lorsque le nombre de défaillances totales est très faible devant le nombre total d'éléments dans le système, on vérifie sur la courbe verte de la figure VI.9 que le résultat théorique (VI.23) obtenu pour la distribution de la charge défaillante, si on fixe la charge de l'ancêtre, est une bonne approximation. Pour un nombre quelconque de défaillances totales (courbe bleue), la distribution de la charge défaillante  $L_i = \ell_i$  de l'élément *i* engendrée par une défaillance de charge  $L = \ell$  présente la forme de marche attendue, mais déformée. Cette déformation s'explique par le fait que, lorsque le nombre d'éléments survivants devient petit devant la taille du sytème, ils reçoivent plusieurs surcharges simultanées avant de défaillir, et la charge des éléments défaillants est donc plus importante que celle attendue sous l'hypothèse d'une unique surcharge. Cet effet est particulièrement visible sur la distribution de la charge défaillante issue de défaillances de petite charge.

La distribution présente un saut à  $\ell_i = \beta \ell$ , quel que soit l'échantillon statistique choisi, car la variable  $\ell_i$  est définie comme  $\ell_i = u_i + \beta \ell$ , où  $u_i \in (0, u_{max})$ .

### VI.5.2 Moyenne et distribution du nombre de défaillances

On suppose, dans la section VI.3.2, qu'à la génération s le nombre de descendants directs d'un élément de charge  $L = \ell$  est distribué selon une loi binomiale de probabilité  $p = p(s) \min \left(1, \frac{\beta \ell}{\alpha}\right)$ , de moyenne :

$$\mu(\ell) = \min\left(\frac{1}{\beta}, \frac{\ell}{\alpha}\right). \tag{VI.27}$$

Cette probabilité est exacte si les descendants ne subissent qu'une unique surcharge avant de défaillir. Dans le cas du modèle de Lehmann & Bernasconi à taille finie, un élément i de charge  $u_i(s)$  à la génération s peut avoir reçu une surcharge :

- $\beta \ell_t < \alpha u_i(0)$  à une génération t antérieure. Si cette surcharge n'a pas suffi à le faire défaillir, elle peut cependant le rendre vulnérable à une surcharge  $\beta \ell < \alpha u_i(0)$  ultérieure. En effet, si  $\beta(\ell + \ell_t) > \alpha u_i(0)$ , l'élément *i* défaille.
- $\beta \ell'$  à la génération courante s, en plus de la surcharge due à la défaillance  $L = \ell$  dont on calcule ici le nombre de descendants. Dans ce cas, l'élément *i* peut avoir défailli aussi bien à cause de de la défaillance de charge  $\ell$  qu'à cause de celle de charge  $\ell'$  et il ne compte alors que pour moitié dans la statistique du nombre de descendants de la défaillance de charge  $L = \ell$ .



FIGURE VI.9 – Distribution de la charge défaillante  $L_i = \ell_i$  sachant que la charge qui l'engendre est  $L = \ell$ , simulé selon l'algorithme du modèle de Lehmann & Bernasconi. Les différentes courbes représentent les résultats de simulations effectuées jusqu'à un nombre total de défaillances Mquelconque (courbe bleue), ou très petit devant  $N : M < \frac{N}{20}$  (courbe verte). Les résultats de simulation sont mis en regard avec les résultats analytiques, sous les approximations détaillées en section VI.3.3.



FIGURE VI.10 – Nombre moyen de défaillances issues d'une défaillance de charge  $L = \ell$  en fonction de la valeur de  $\ell$ , simulé selon l'algorithme du modèle de Lehmann & Bernasconi. Les différentes données sont prises jusqu'à un nombre total de défaillances M plus ou moins petit devant N.

On vérifie sur la figure VI.10 que ces effets sont négligeables si on ne regarde la cascade que pour un nombre total de défaillances petit devant la taille totale du système (points verts). En effet, tant que les défaillances ne sont pas trop nombreuses, la probabilité que l'un des deux évenements listés plus haut survienne est très faible, et le nombre moyen de défaillances causées par une charge  $\ell$  est bien donné par l'équation (VI.19).

Le premier effet énoncé plus haut influe sur la distribution du nombre de défaillances issues d'une défaillance de charge faible, tandis que le second déforme la distribution pour des défaillances de charges importante. On peut voir sur la figure VI.11 que dans le cas où la charge défaillante est faible (a), lorsque le nombre d'éléments survivants n'est pas imposé comme étant de l'ordre de la taille totale du système, la distribution montre une queue plus élargie, *i.e.* une faible charge peut engendrer un nombre élevé de défaillances. Si la charge défaillante est importante (b), la distribution est déplacée vers l'axe des ordonnées, *i.e.* une forte charge engendre moins de défaillances que prévues par la théorie suite aux approximations.

Lorsqu'on limite les simulations aux générations pour lesquelles le nombre de survivants est de l'ordre de la taille totale du système (courbes rouges), les approximations de la section VI.3.2 sont bien vérifiées. Le modèle H de branchement généralisé doté des caractéristiques du modèle L&B offre un cadre purifié dans lequel on peut déterminer analytiquement un certain nombre de quantités présentes dans le processus des cascades de défaillances.

## VI.6 Conclusion

Nous avons présenté dans ce chapitre le modèle de Lehmann et Bernasconi qui propose une étude simplifiée de transmission de défaillances en cascade dans un réseau de type électrique, de taille finie [74]. Les deux auteurs de ce modèle se sont intéressés aux quantités limites du processus de cascade : la probabilité  $P_{nc}$  qu'une seule défaillance se soit produite et qu'elle ait pas engendré de cascade, et la probabilité  $P_b$  que la cascade se soit étendue à tous les éléments du réseau. Lors de notre étude de ce modèle, nous avons considéré d'autres quantités globales, comme la distribution de la taille M des cascades ou la détermination de la valeur critique pour laquelle des cascades globales apparaissent dans le réseau. Nous avons également examiné l'effet de la distribution des charges dans le réseau sur la dynamique de la cascade, par exemple la distribution du nombre de défaillances selon la charge de la défaillance qui les engendre. La valeur des charges défaillantes d'une génération donnée est une variable aléatoire composée de la somme de variables aléatoires dont le nombre augmente avec les générations, et l'étude de leur distribution montre une fonction limite dont la forme explique les propriétés globales de la cascade.

Nous avons également introduit dans ce chapitre l'approximation du modèle L&B en processus de branchement généralisé proposée par les auteurs [74], et dont le cadre mathématique a été formalisé par Harris [92]. Nous avons montré que cette approximation n'est valide que si le réseau est infini  $N = \infty$ , ou tant que le nombre de survivants  $\nu$  dans le réseau est suffisamment grand devant 1. Nous définissons dans le chapitre suivant un modèle de branchement généralisé, le modèle H, dont les caractéristiques de création de nouvelles générations sont données par les propriétés du modèle L&B. Dans le cadre de ce modèle, nous proposons une étude analytique des quantités citées dans ce chapitre.



(b) Grande charge défaillante

FIGURE VI.11 – Distribution du nombre de défaillances engendrées par une défaillance de charge  $\ell = 0.1$  (a) et  $\ell = 1.2$  (b). Les différentes courbes sont déterminées pour des échantillons avec nombre total de défaillances plus ou moins petit devant N. Les courbes continues représentent la distribution attendue par la théorie.

# Chapitre **VII**

# Etude du processus de branchement généralisé

Le modèle décrit par Lehmann & Bernasconi peut, sous certaines approximations valides pour un système de taille infinie, être étudié comme un processus de branchement généralisé. Ce type de processus décrit l'évolution d'une famille dont les membres sont caractérisés par un type, *i.e* une variable aléatoire, dont la valeur ne dépend que des caractéristiques de la génération précédente, et dont la connaissance suffit à décrire la génération suivante. Dans le processus de branchement généralisé appliqué au modèle de défaillances dans un réseau électrique, le modèle H, on considère la famille de défaillances successives, et le type considéré est la charge des défaillances.

Les règles de passage d'une génération à la suivante se limitent au nombre de descendants directs d'un individu, donné par la loi de Poisson (VI.20), de fonction génératrice (VI.21), et la densité de probabilité jointe pour un individu de charge L d'engendrer  $N_1$  individus de charges  $L_1, L_2, \ldots, L_{N_1}$ , qui est donnée par l'équation (VI.23).

Dans le cadre de ce modèle H, que l'on va décrire en détail dans un premier temps, on étudie analytiquement les grandeurs considérées dans le modèle de Lehmann & Bernasconi, à la section VI.2.

## VII.1 Définition du modèle H

Ce modèle décrit l'arbre généalogique des défaillances engendrées à la suite d'une défaillance initiale. Toutes ces défaillances, à l'exception de la première, portent une charge définie à partir de la charge de la défaillance qui les engendre. La première défaillance porte une charge aléatoire tirée directement d'une distribution  $f_{L(0)}$  donnée. L'ensemble des défaillances nées après le même nombre s d'étapes à la suite de la défaillance initiale forment la s-ième génération.

1. À la génération s = 0:

Cette génération est constituée d'une unique défaillance, de charge L(0) tirée dans la distribution  $f_{L(0)}$ . Le nombre de membres de cette génération est donc :  $N_0 = 1$ , et le nombre total d'individus dans la famille est  $M_0 = 1$ .

2. À la génération s = 1 :

La défaillance de la génération précédente, de charge L(0) engendre  $N_1$  descendants en une génération. Le nombre de descendants directs est distribué selon une loi de Poisson, dont le paramètre  $\mu$  dépend de la charge de la défaillance-mère  $L(0) = \lambda$ :

$$P(N_1(\lambda) = n) = \frac{\mu(\lambda)^n}{n!} e^{-\mu(\lambda)},$$
 (VII.1)

où le paramètre est donné par :

$$\mu(\lambda) = \min\left(\frac{1}{\beta}, \frac{\lambda}{\alpha}\right).$$
(VII.2)

Cette loi de Poisson a pour fonction génératrice :

$$G(\lambda, t) = e^{\mu(\lambda)(t-1)}.$$
 (VII.3)

Les charges  $L_1(1), \ldots, L_n(1)$  des  $N_1(\lambda) = n$  défaillances de cette génération sont tirées dans la distribution uniforme  $f_{L_i(1)|L(0)}$  des charges  $L_i(1) = \ell$  issues d'une défaillance de charge  $L(0) = \lambda$ :

$$f_{L_i(1)|L(0)}(\ell|\lambda) = \frac{1}{\min\left(1,\frac{\beta\lambda}{\alpha}\right)}, \text{ si } \beta\lambda < \ell < \beta\lambda + \min\left(1,\frac{\beta\lambda}{\alpha}\right), \quad (\text{VII.4})$$

ou encore, en fonction du nombre moyen de descendants  $\mu(\lambda)$  de la défaillance de charge  $\lambda$ :

$$f_{L_i(1)|L(0)}(\ell|\lambda) = \frac{1}{\beta\mu(\lambda)}, \text{ si } \beta\lambda < \ell < \beta\left(\lambda + \mu(\lambda)\right), \quad (\text{VII.5})$$

Cette distribution est issue de la définition des charges défaillantes comme somme d'une variable aléatoire tirée de la distribution uniforme  $f_{L(0)}$  et d'une fraction  $\beta L(0)$  de la charge de la défaillance-mère.

Cette génération est donc composée de  $N_1(\lambda)$  défaillances, de charges  $L_i(1) = \ell_i$  tirées dans la distribution  $f_{L(1)}(\ell_i)$ . Le nombre de membres de cette génération est :  $N_1(\lambda)$ , et le nombre total d'individus dans la famille est  $M_1(\lambda) = M_0 + N_1(\lambda)$ .

3. À la génération s = 2 :

Chacune des  $N_1(\lambda) = n$  défaillances de la génération précédente, de charge

$$L_1(1) = \ell_1, \dots, L_n(1) = \ell_n,$$

engendre  $N_1(\ell_1), \ldots, N_1(\ell_n)$  descendants en une génération. Le nombre de descendants de chaque défaillance est distribué selon la loi de Poisson (VII.8) avec paramètre  $\mu(\ell_1), \ldots, \mu(\ell_n)$ . Les charges  $L_1(2), \ldots, L_{N_2(\lambda)}(2)$  des  $N_2(\lambda) = N_1(\ell_1) + \cdots + N_1(\ell_n)$  défaillances de la deuxième génération sont tirées comme suit. Les descendants de la défaillance de charge  $L_1(1) = \ell_1$  portent des charges  $L_{j,1}(2) = \ell_{j,1}, \ j = 1, \ldots, N_1(\ell_1)$ . Ces charges sont tirées dans une même distribution  $f_{L_{j,1}(2)|L_1(1)}$ . Les descendants de la défaillance de charge  $L_i(1) = \ell_i, \ i = 1, \dots, N_1(\lambda)$  portent des charges  $L_{j,i}(2) = \ell_{j,i}, \ j = 1, \dots, N_1(\ell_i)$  qui sont tirées dans la distribution  $f_{L_{j,i}(2)|L_i(1)}$ , définie par :

$$f_{L_{j,i}(2)|L_i(1)}(\ell'|\ell) = \frac{1}{\beta\mu(\ell)}, \text{ si } \beta\ell < \ell' < \beta\left(\ell + \mu(\ell)\right).$$
(VII.6)

Cette génération est donc composée de  $N_2(\lambda)$  défaillances. Ce sont les descendants de deuxième génération (ou petits-enfants) de la charge initiale  $L(0) = \lambda$ . La charge d'une défaillance  $L_i(2)$ ,  $i = 1, \ldots, N_2(\lambda)$  de cette génération prise au hasard est distribuée selon la densité  $f_{L(2)}$ . Le nombre total d'individus dans la famille à la deuxième génération est

$$M_2(\lambda) = M_1(\lambda) + N_2(\lambda).$$
(VII.7)

4. À la génération s :

Chaque défaillance de la génération précédente, de charge  $L_i(s-1) = \ell_i$ ,  $i = 1, \ldots, N_{s-1}$ engendre un nombre  $N_1(\ell_i)$  de descendants à la génération suivante, distribué selon la loi de Poisson de paramètre  $\mu(\ell_i)$ :

$$P(N_1(\ell_i) = n) = \frac{\mu(\ell_i)^n}{n!} e^{-\mu(\ell_i)},$$
 (VII.8)

où le paramètre est défini par :

$$\mu(\ell_i) = \min\left(\frac{1}{\beta}, \frac{\ell_i}{\alpha}\right),\tag{VII.9}$$

et qui a pour fonction génératrice :

$$G(\ell_i, t) = \left\langle t^{N_1(\ell_i)} \right\rangle = e^{\mu(\ell_i)(t-1)}.$$
 (VII.10)

Les  $N_1(\ell_i)$  défaillances issues de *i* portent les charges

$$L_{j,i}(s) = \ell_{j,i}, \ j = 1, \dots, N_1(\ell_i),$$

distribuées selon  $f_{L_{j,i}(s)|L_i(s-1)}$ :

$$f_{L_{j,i}(s)|L_i(s-1)}(\ell_{j,i}|\ell_i) = \frac{1}{\min\left(1,\frac{\beta\ell_i}{\alpha}\right)}, \text{ si } \beta\ell_i < \ell_{j,i} < \beta\ell_i + \min\left(1,\frac{\beta\ell_i}{\alpha}\right).$$
(VII.11)

Cette génération comprend

$$N_s(\lambda) = \sum_{i=0}^{N_{s-1}} N_1(L_i(s-1))$$
(VII.12)

membres, c'est-à-dire que la défaillance initiale  $L(0) = \lambda$  a engendré  $N_s(\lambda)$  descendants de s-ième ordre. La cascade contient

$$M_s(\lambda) = M_{s-1}(\lambda) + N_s(\lambda) \tag{VII.13}$$

individus au total après s générations.

$$s = 0$$

$$s = 1$$

$$s = 1$$

$$L_{1}(1)$$

$$L_{2}(1) = \ell_{1}$$

$$L_{j,2}(2) = \ell_{2} \in f_{L_{j,2}(2)|L_{2}}(1)$$

$$L_{j,2}(2) = \ell_{1} \in f_{L_{j,2}(2)|L_{2}}(1)$$

$$K_{j,2}(1) = \ell_{1} \in f_{L_{j,2}(2)|L_{2}}(1)$$

FIGURE VII.1 – Schéma du processus de branchement généralisé.

La cascade se poursuit jusqu'à ce qu'une certaine génération  $s_a$  ne contienne aucune défaillance :  $N_{s_a}(\lambda) = 0$ . Toutes les génération suivantes seront également vides, et on appelle cette génération  $s_a$  le temps d'arrêt de la cascade.

Le processus peut se résumer par le schéma d'un processus de branchement généralisé, figure VII.1.

Un résumé des notations se trouve en appendice B.

## VII.2 Probabilité de non-cascade

Le nombre de défaillances engendrées par la défaillance initiale d'un élément de charge  $\lambda$  distribuée sur la distribution initiale  $f_{L(0)}$  est donné par la loi de Poisson de paramètre  $\mu(\lambda)$ . La cascade s'interrompt sans défaillance ultérieure si la première défaillance de charge  $\lambda$  n'engendre aucun descendant :

$$P_{nc}(\lambda) = P(N_1(\lambda) = 0) = e^{-\mu(\lambda)}, \qquad (\text{VII.14})$$

où l'on rappelle que  $\mu(\lambda) = \min\left(\frac{1}{\beta}, \frac{\lambda}{\alpha}\right)$ . On obtient la probabilité de non-cascade pour une charge initiale quelconque en prenant la moyenne de la probabilité de non-cascade suite à une défaillance initiale sur la distribution uniforme  $f_{L(0)}$  des charges de la défaillance initiale :

$$P_{nc} = \begin{cases} \alpha - \alpha e^{-1/\alpha} & \text{si } \alpha \ge \beta \\ \alpha + \left(1 - \alpha - \frac{\alpha}{\beta}\right) e^{-1/\beta} & \text{si } \alpha < \beta \end{cases}.$$
 (VII.15)

## VII.3 Nombre total de défaillances

Si la cascade se poursuit au-delà de la première génération, on s'intéresse au nombre d'éléments qui ont défailli au cours de son histoire. On introduit le modèle simple dans lequel les éléments ne portent pas de charge.

## VII.3.1 Pour un processus de branchement simple

Le nombre total de défaillances dans un processus de cascades simple (*i.e.* sans charge) est donné par la somme du nombre de défaillances  $N_s$  à chaque génération s, jusqu'à l'extinction de la cascade :

$$M = N_0 + N_1 + \dots = \sum_{s=0}^{\infty} N_s.$$
 (VII.16)

Le nombre moyen de descendants directs d'une défaillance est donné par  $\mu$ , quelle que soit la génération considérée ou le nombre de défaillances à chaque étape. La distribution du nombre de descendants directs  $N_1$  d'une défaillance est donnée par une loi de Poisson de paramètre  $\mu$ :

$$P(N_1 = k) = \frac{\mu^k}{k!} e^{-\mu}.$$
 (VII.17)

Cette distribution a pour fonction génératrice :

$$G(t) = \left\langle t^{N_1} \right\rangle = \sum_{k=0}^{\infty} P(N_1 = k) t^k = e^{\mu(t-1)}.$$
 (VII.18)

Le nombre total de descendants issus d'un unique individu, étiqueté m, après s étapes est donné par la somme du nombre de membres  $N_i$  de chaque génération i:

$$M_s^{(m)} = \sum_{i=0}^s N_i,$$
 (VII.19)

avec  $N_0 = 1$ , et de fonction génératrice :

$$H_{s}^{(m)}(t) = \left\langle t^{M_{s}^{(m)}} \right\rangle = \sum_{n=1}^{\infty} P\left(M_{s}^{(m)} = n\right) t^{n}.$$
 (VII.20)

Et le nombre total de descendants après s + 1 étapes a pour fonction génératrice :

$$H_{s+1}^{(m)}(t) = \sum_{n=1}^{\infty} P\left(M_{s+1}^{(m)} = n\right) t^n,$$
 (VII.21)

où le nombre total de descendants  $M_{s+1}^{(m)}$  d'un élément m après (s+1) générations, si la défaillance m a  $N_1 = k$  descendants directs, est donné par :

$$M_{s+1}^{(m)} = M_s^{(1)} + M_s^{(2)} + \dots + M_s^{(k)}, \qquad (\text{VII.22})$$

où  $M_s^{(i)}$  est la taille totale de la famille issue de la défaillance d'indice i après s générations.

La fonction génératrice (VII.21) du nombre total de descendants d'un unique élément quelconque sur s + 1 générations devient donc :

$$H_{s+1}^{(m)}(t) = \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} P\left(N_1^{(m)} = k\right) P\left(M_s^{(1)} + \dots + M_s^{(k)} = n-1\right) t^n.$$
 (VII.23)

Or toutes les branches de l'arbre des défaillances sont indépendantes et identiquement distribuées. La dernière probabilité est donc une convolution de n termes, et on reconnait n fonctions génératrices du nombre total de descendants sur s générations :

$$H_{s+1}^{(m)}(t) = t \sum_{k=0}^{\infty} P(N_1^{(m)} = k) \left( H_s^{(.)}(t) \right)^k, \qquad (\text{VII.24})$$

où le point représente un indice d'une défaillance quelconque issue de la première défaillance, d'indice m. Toutes les défaillances ont, dans le modèle simple, la même distribution du nombre total de descendants, et on peut négliger l'indice du noeud dans les équations :

$$H_{s+1}(t) = t \sum_{k=0}^{\infty} P(N_1 = k) \left( H_s(t) \right)^k.$$
 (VII.25)

On reconnaît dans cette formule la fonction génératrice du nombre de descendants directs d'un unique individu :

$$G(t) = \sum_{k=0}^{\infty} P(N_1 = k)t^k,$$
 (VII.26)

et l'équation de récurrence pour la fonction génératrice du nombre total de défaillances en s + 1 générations après une unique défaillance initiale est donnée par :

$$H_{s+1}(t) = tG(H_s(t)).$$
(VII.27)

Le nombre total de défaillances survenues lors de la cascade est défini à partir du nombre total de défaillances jusqu'à la génération s comme :

$$M = \lim_{s \to \infty} M_s, \tag{VII.28}$$

et a donc pour fonction génératrice;

$$H(t) = tG(H(t)).$$
(VII.29)

Cette fonction vérifie donc l'équation :

$$H(t) = t \exp\left(\mu \left(H(t) - 1\right)\right)$$
(VII.30)

La solution de cette relation est donnée par introduction d'une fonction de Lambert W [94], définie comme la fonction inverse de :

$$g(W) = W e^W, (VII.31)$$

et dont une forme explicite peut être donnée par son développement en série de Taylor :

$$W(x) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} x^n.$$
 (VII.32)



FIGURE VII.2 – Branche principale de la fonction de Lambert.

On trace la fonction de Lambert sur son intervalle de définition dans  $\mathbb{R}$  sur la figure VII.2

La fonction génératrice du nombre total de défaillance M pour un processus de cascade sans charge, et gouverné par une loi de Poisson de paramètre  $\mu$ , est donc :

$$H(t) = -\frac{W(-\mu e^{-\mu}t)}{\mu},$$
 (VII.33)

et la probabilité d'avoir M = n défaillances dans la cascade est donnée par :

$$P(M=n) = \frac{1}{n!} \frac{\mathrm{d}^n H(0)}{\mathrm{d}t^n} = -\frac{1}{\mu} \frac{1}{n!} \left. \frac{\mathrm{d}^n}{\mathrm{d}t^n} W\left(-\mu \mathrm{e}^{-\mu} t\right) \right|_{t=0}.$$
 (VII.34)

La formule en série de Taylor de la fonction de Lambert permet d'obtenir facilement la dérivée n-ième de la fonction, et la distribution du nombre total de défaillances suit une loi de Borel-Tanner [95] :

$$P(M = n) = \frac{e^{-n\mu}}{n!} (n\mu)^{n-1}.$$
 (VII.35)

On représente, figure VII.3, la distribution de probabilité du nombre total de défaillances dans la cascade pour différentes valeurs du paramètre  $\mu$  de la loi de Poisson. La distribution décroît exponentiellement vite vers 0 pour  $\mu < 1$ . On observe un comportement critique lorsque  $\mu = 1$ , et la distribution suit alors une loi de puissance :

$$P(M=n) = \frac{e^{-n}}{n!} n^{n-1} \approx \frac{n^{-3/2}}{\sqrt{2\pi}}.$$
 (VII.36)



FIGURE VII.3 – Distribution du nombre total de défaillances dans le modèle de branchement simple pour différentes valeurs du nombre moyen  $\mu$  de descendants par individu. La formule de P(M = n) est donnée par l'équation (VII.35).

Pour  $\mu > 1$ , la distribution de probabilité à n fini décroît exponentiellement vite vers 0, mais la distribution n'est clairement pas normalisée sur n fini. La partie résiduelle du poids de la distribution est reporté à l'infini et correspond à la probabilité que la famille de défaillances ait une taille infinie :

$$P(M = \infty) = 1 - \sum_{n=1}^{\infty} P(M = n).$$
 (VII.37)

On reporte ce poids, sur la figure VII.3, à la valeur d'arrêt de la simulation N = 1000.

Ces caractéristiques peuvent s'interpréter aisement dans le cas du modèle simple. En effet, le paramètre  $\mu$  est le nombre moyen de descendants engendrés par individu, et pour  $\mu < 1$ , chaque génération est plus petite que la précédente avec probabilité finie. La cascade de défaillances s'éteindra donc en un temps fini avec probabilité 1 et le nombre total de défaillances est fini. Pour  $\mu > 1$  la probabilité que la cascade ne s'éteigne jamais est non nulle. À la transition entre ces deux régimes, la cascade s'arrête en général après un faible nombre de défaillances mais, avec probabilité non nulle, elle peut aussi être de taille  $M \gg 1$ .

## VII.3.2 Dans le modèle H

Un élément *i* de charge  $L_i = \ell_i$  engendre un nombre total de descendants  $M_s(\ell_i)$  en *s* générations, de fonction génératrice

$$H_s(\ell_i, t) = \left\langle t^{M_s(\ell_i)} \right\rangle = \sum_{n=1}^{\infty} P\left(M_s(\ell_i) = n\right) t^n.$$
(VII.38)

L'ensemble des défaillances issues d'un unique élément de charge  $L = \ell$  après (s + 1) générations est composé des branches de défaillances sur s générations issues de tous les  $N_1(\ell) = k$ descendants directs de l'élément de charge L, de charges  $L_1, \ldots, L_k$ :

$$M_{s+1}(\ell) = 1 + M_s(L_1) + \dots + M_s(L_k).$$
(VII.39)

La fonction génératrice du nombre total de défaillances  $M_{s+1}(\ell)$  issues d'un unique élément de charge  $\ell$  en (s + 1) générations est déterminée de la même façon que pour le modèle simple en VII.3.1. Les membres de la famille de défaillances se distinguent cependant ici les uns des autres par leurs charges respectives. Le nombre de descendants issus d'un individu i de la première génération de défaillances issue de la défaillance de charge  $\ell$  dépend de sa charge  $L_i = \ell_i$ , tirée dans la distribution uniforme  $U(\beta \ell, \beta \ell + \beta \mu(\ell))$ . La fonction génératrice du nombre total de défaillances en s + 1 générations après une première défaillance de charge  $L = \ell$  vérifie l'équation :

$$H_{s+1}(\ell, t) = t \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} P(N_1(\ell) = k)$$

$$\int d\ell_1 \dots d\ell_k f_{L_1|L}(\ell_1|\ell) \dots f_{L_k|L}(\ell_k|\ell)$$

$$t^{n-1} P(M_s(\ell_1) + \dots + M_s(\ell_k) = n-1).$$
(VII.40)

Le dernier terme est une convolution de k termes indépendants et identiquement distribués en fonction de la charge  $\ell_i$ . On introduit ici la définition de la moyenne d'une fonction f quelconque par rapport à la charge :

$$\overline{f(\ell)} = \int \mathrm{d}\ell' f_{L'|L}(\ell'|\ell) f(\ell').$$
(VII.41)

L'équation (VII.40) précédente s'écrit alors :

$$H_{s+1}(\ell,t) = t \sum_{k=0}^{\infty} P\left(N_1(\ell) = k\right) \overline{H_s(\ell,t)}^k, \qquad (\text{VII.42})$$

dans laquelle on reconnaît la fonction génératrice du nombre de descendants directs d'une défaillance de charge  $L = \ell$ :

$$G(\ell, t) = \left\langle t^{N_1(\ell)} \right\rangle = \sum_{k=0}^{\infty} P(N_1(\ell) = n) t^k = e^{\mu(\ell)(t-1)}, \quad (\text{VII.43})$$

soit :

$$H_{s+1}(\ell, t) = t \ G\left(\ell, \overline{H_s(\ell, t)}\right).$$
(VII.44)

La fonction génératrice du nombre total de défaillances engendrées en s générations par une unique défaillance de charge  $L = \ell$  vérifie donc l'équation de récurrence :

$$H_{s+1}(\ell, t) = t \exp\left[\mu(\ell) \left(\overline{H_s(\ell, t)} - 1\right)\right].$$
 (VII.45)

La taille finale de la cascade initialisée par la défaillance de charge  $L(0) = \lambda$  est notée  $M(\lambda)$ , sa fonction génératrice est notée  $H(\lambda, t)$ . Cette fonction vérifie l'équation suivante :

$$H(\lambda, t) = t \, \exp\left[\mu(\lambda) \left(\overline{H(\lambda, t)} - 1\right)\right], \qquad (\text{VII.46})$$

où  $\lambda$  est tirée aléatoirement de la distribution  $f_{L(0)}$ .

Pour résoudre cette équation, on distingue deux cas, selon la valeur de la charge de la défaillance initiale  $L(0) = \lambda$ , et donc selon la valeur de son nombre moyen de descendants  $\mu(\lambda)$ :

$$\mu(\lambda) = \min\left(\frac{1}{\beta}, \frac{\lambda}{\alpha}\right).$$
(VII.47)

a Cas  $\lambda > rac{lpha}{eta}$  :

Dans ce cas, le nombre moyen de descendants directs de la défaillance initiale prend la valeur :

$$\mu(\lambda) = \frac{1}{\beta},\tag{VII.48}$$

et les charges de ces descendants sont distribuées selon :

$$f_{L|L(0)}(\ell|\lambda) = 1, \quad \text{if } \beta\lambda < \ell < \beta\lambda + 1.$$
 (VII.49)

L'équation de la fonction génératrice de la taille  $M(\lambda)$  de la cascade se met sous la forme :

$$H(\lambda, t) = t \exp\left[\frac{1}{\beta} \int_0^1 H(y + \beta\lambda, t) dy - \frac{1}{\beta}\right].$$
 (VII.50)

Pour résoudre cette équation, on suppose que cette fonction ne dépend pas de la valeur de la charge initiale L(0):

$$H(\lambda, t) = h(t), \tag{VII.51}$$

et l'équation (VII.50) s'écrit, en termes de h:

$$h(t) = t \exp\left[\frac{1}{\beta} \left(h(t) - 1\right)\right].$$
 (VII.52)

Cette équation est semblable à l'équation obtenue pour le processus de branchement simple (VII.30), et la fonction génératrice, pour  $\lambda > \frac{\alpha}{\beta}$ , a pour équation :

$$H(\lambda, t) = h(t) = -\beta W\left(-\frac{te^{-1/\beta}}{\beta}\right).$$
 (VII.53)

Par définition de la fonction génératrice du nombre total de défaillances issues d'un élément de charge  $L(0) = \lambda$ ,

$$P\left(M(\lambda) = n\right) = \frac{1}{n!} \left. \frac{\mathrm{d}^n H(\lambda, t)}{\mathrm{d}t^n} \right|_{t=0},\tag{VII.54}$$

et donc, dans le cas où la charge initiale  $\lambda > \frac{\alpha}{\beta}$  :

$$P(M(\lambda) = n) = \frac{\mathrm{e}^{-\frac{n}{\beta}}}{n!} \left(\frac{n}{\beta}\right)^{n-1}.$$
 (VII.55)

Cette distribution ne dépend donc pas de la valeur  $\lambda$  de la charge de la défaillance initiale, tant que  $\lambda > \alpha/\beta$ , et on retrouve le résultat (VII.35) du cas simple pour  $\mu = 1/\beta$ . Pour  $\beta < 1$ , cette distribution n'est pas normalisée, et il existe donc un poids non nul à la limite :

$$P(M(\lambda) = \infty) = 1 - \sum_{n=1}^{\infty} P(M(\lambda) = n).$$
 (VII.56)

Pour  $\beta \geq 1$ , la distribution est normalisée, et la probabilité que la taille de la cascade initialisée par une charge  $L = \lambda > \alpha/\beta$  soit infinie est nulle.

On représente sur la figure VII.4 la distribution du nombre total de défaillance pour des cascades initiées par des défaillances de charges  $\lambda > \alpha/\beta$ .

b Cas  $\lambda < \frac{\alpha}{\beta}$ 

Dans ce cas, le paramètre de la distribution de Poisson devient :

$$\mu(\lambda) = \frac{\lambda}{\alpha},\tag{VII.57}$$

et la distribution de la charge issue de la première défaillance :

$$f_{L|L(0)}(\ell|\lambda) = \frac{\alpha}{\beta\lambda}, \quad \text{if } \beta\lambda < \ell < \beta\lambda \left(1 + \frac{1}{\alpha}\right).$$
 (VII.58)

Pour résoudre l'équation intégrale (VII.46), on définit la primitive de la fonction génératrice H comme :

$$\mathcal{H}(\lambda, t) = \int_0^\lambda \mathrm{d}\ell H(\ell, t), \qquad (\text{VII.59})$$

qui vérifie alors l'équation différentielle :

$$\frac{\mathrm{d}\mathcal{H}(\lambda,t)}{\mathrm{d}\lambda} = t \exp\left[\frac{1}{\beta}\left(\mathcal{H}\left(\beta\lambda + \frac{\beta\lambda}{\alpha},t\right) - \mathcal{H}(\beta\lambda,t)\right) - \frac{\lambda}{\alpha}\right].$$
 (VII.60)

Ici, on ne peut pas séparer la fonction  $\mathcal{H}$  en un produit de fonctions indépendantes des deux paramètres. On peut faire une approximation grossière en linéarisant l'équation intégrale. On appelle  $\tilde{\mathcal{H}}(\lambda, t)$  la solution de l'équation linéarisée :

$$\frac{\partial}{\partial\lambda}\tilde{\mathcal{H}}(\lambda,t) = t\left[1 + \frac{1}{\beta}\left(\tilde{\mathcal{H}}\left(\beta\lambda + \frac{\beta\lambda}{\alpha},t\right) - \tilde{\mathcal{H}}(\beta\lambda,t)\right) - \frac{\lambda}{\alpha}\right].$$
 (VII.61)



FIGURE VII.4 – Distribution du nombre total de défaillances, obtenue par simulation du modèle H, initiée par une défaillance de charge  $\lambda = 0, 4 > \alpha/\beta$  et  $\lambda = 0.8 > \alpha/\beta$ . La formule de  $P(M(\lambda) = n)$  est donnée par l'équation VII.56.

Et on résout cette équation par itérations successives. On pose  $\mathcal{H}_0(\lambda, t) = t\lambda$ . La première itérée est :

$$\tilde{\mathcal{H}}_1(\lambda, t) = \int_0^\lambda \mathrm{d}\ell \ t \left( 1 + \frac{\ell}{\alpha} (t-1) \right).$$
(VII.62)

Les solutions itérées convergent vers :

$$\tilde{\mathcal{H}}(\lambda,t) = t\lambda + \sum_{k=2}^{\infty} \lambda^k \frac{t^{k-1}(t-1)\prod_{i=1}^{k-1} \left((1+\alpha)^i - \alpha^i\right)\beta^{\frac{(k-1)(k-2)}{2}}}{k!\alpha^{\frac{k(k-1)}{2}}}.$$
 (VII.63)

La fonction génératrice linéarisée  $\tilde{H}(\lambda, t)$  du nombre total de défaillances est donc :

$$\widetilde{H}(\lambda,t) = \frac{\mathrm{d}\widetilde{\mathcal{H}}(\lambda,t)}{\mathrm{d}\lambda} \\
= t\left(1-\frac{\lambda}{\alpha}\right) + \sum_{k=2}^{\infty} \frac{t^k \lambda^{k-1}}{(k-1)!} \qquad (\text{VII.64}) \\
\prod_{i=1}^{k-1} \left((1+\alpha)^i - \alpha^i\right) \frac{\beta^{\frac{(k-1)(k-2)}{2}}}{\alpha^{\frac{k(k-1)}{2}}} \left(1 - \frac{\lambda\beta^{k-1}\left((1+\alpha)^k - \alpha^k\right)}{k\alpha^k}\right).$$

Cette approximation donne des résultats cohérents au premier ordre uniquement. En particulier, la probabilité que la cascade ne provoque qu'une unique défaillance est donnée par :

$$P(M=1) = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{H}(\lambda,0) = \left(1 - \frac{\lambda}{\alpha}\right) \approx \mathrm{e}^{-\frac{\lambda}{\alpha}}.$$
 (VII.65)

Aux ordres supérieurs, l'approximation de linarité de l'équation intégrale n'est pas vérifiée, et la probabilité que la cascade ait une taille  $M \ge 2$  ne peut pas être calculée à partir de la fonction génératrice linéarisée  $\tilde{H}(\lambda, t)$ .

#### c Charge initiale quelconque

La distribution du nombre total de défaillances dans une cascade issue d'un unique individu dont la charge est tirée aléatoirement dans  $f_{L(0)}$  est donnée par l'intégration de la distribution pour une charge initiale  $L(0) = \lambda$  sur la distribution  $f_{L(0)}$ :

$$P(M=n) = \int d\lambda f_{L(0)}(\lambda) P(M(\lambda)=n).$$
(VII.66)

On représente sur la figure VII.5 les résultats de la simulation de l'algorithme du modèle H et les valeurs obtenues par intégration numérique de l'équation (VII.46). Les courbes sont obtenues pour une valeur donnée du paramètre  $\alpha$  et différentes valeurs de  $\beta$ . Pour certaines valeurs des paramètres, la distribution de probabilité décroît exponentiellement vite vers 0. Pour de plus grandes valeurs de  $\beta$ , la distribution n'est normalisée que si on considère le poids à  $M = \infty$ . Ce poids est tracé, sur la figure VII.5, au temps d'arrêt de la simulation N = 2000. Entre ces deux régimes, il existe une valeur  $\beta_c$  du paramètre  $\beta$  pour laquelle la distribution de probabilité suit une loi de puissance, dont l'exposant dépend des valeurs des paramètres  $\alpha$  et  $\beta$ . La valeur



FIGURE VII.5 – Distribution de probabilité du nombre total M de défaillances suite à une défaillance initiale de charge quelconque, pour différentes valeurs du paramètre  $\beta$ , le paramètre  $\alpha$  étant fixé. Ces courbes sont obtenues par simulation du modèle H sur 10000 échantillons. La simulation est stoppée lorsque la taille de la cascade devient supérieure à N = 2000, et le poids de la distribution restant est cumulé en n = N. Les résultats théoriques sont obtenus par le calcul numérique de l'équation intégrale (VII.46) pour la fonction génératrice H, dérivée sur les premières valeurs pour obtenir le début de la courbe théorique de la distribution. Pour  $\alpha = 0.2$ , la valeur critique vaut  $\beta_c \approx 0.1792$ .

de transition  $\beta_c$  peut être déterminée numériquement comme la valeur de  $\beta$  pour laquelle la probabilité que le système présente des cascades infinies devient non nulle. On peut remarquer, par analogie avec le processus de branchement simple, que pour cette même valeur  $\beta_c$ , le nombre moyen de descendants par individu  $\bar{\mu}(s)$ , à une génération s donnée, converge asymptotiquement vers 1. Les propriétés de  $\bar{\mu}(s)$  sont détaillées section VII.9.2.

#### d Probabilité de non-cascade

La probabilité que la cascade s'éteigne dès la première défaillance est donnée par la probabilité de non-cascade  $P_{nc} = P(M = 1)$ . Cette probabilité correspond à la moyenne sur la distribution  $f_{L(0)}$  de la charge initiale L(0) de la probabilité qu'une cascade initiée par la charge L engendre une et une seule défaillance :

$$P_{nc} = P(M = 1) = \int d\lambda P(M(\lambda) = 1) f_{L(0)}(\lambda) = \int_0^1 d\lambda P(M(\lambda) = 1), \qquad (VII.67)$$

où  $M(\lambda)$  est la valeur asymptotique du nombre de défaillances totales engendrées par un individu de charge  $\lambda$ . Selon la valeur des paramètres  $\alpha$  et  $\beta$ , on distingue deux cas :

Si α ≥ β, alors α/β ≥ 1 et la charge initiale λ ≤ 1 ≤ α/β. La probabilité de non-cascade est donc donnée par :

$$P_{nc} = \int_0^1 d\lambda e^{-\frac{\lambda}{\alpha}} = \alpha - \alpha e^{-\frac{1}{\alpha}}.$$
 (VII.68)

• Si  $\alpha < \beta$ , alors  $\frac{\alpha}{\beta} < 1$  et l'intégrale sur la charge initiale se décompose en deux parties :

$$P_{nc} = \int_{0}^{\frac{\alpha}{\beta}} \mathrm{d}\lambda \mathrm{e}^{-\frac{\lambda}{\alpha}} + \int_{\frac{\alpha}{\beta}}^{1} \mathrm{d}\lambda \mathrm{e}^{-\frac{1}{\beta}} = \alpha + \left(1 - \alpha - \frac{\alpha}{\beta}\right) \mathrm{e}^{-\frac{1}{\beta}}.$$
 (VII.69)

On retrouve bien les résultats obtenus par l'approche directe de la section précédente VII.2.

#### e Premières valeurs

On confirme ici que l'approximation de linéarisation de la fonction génératrice dans le cas  $\lambda < \alpha/\beta$  sur les valeurs suivantes de la distribution de probabilité du nombre total de défaillances n'est pas vérifiée. La probabilité pour que la cascade ne contienne que deux défaillances suite à une défaillance initiale de charge  $\lambda$  est donnée par :

$$\int_{0}^{1} \mathrm{d}\lambda P\left(M(\lambda)=2\right) = \begin{cases} \int_{0}^{1} \mathrm{d}\lambda \frac{\lambda^{2}(2\alpha^{2}-\lambda\beta-2\lambda\alpha\beta)}{\alpha^{3}} & \mathrm{si} \ \alpha \ge \beta\\ \int_{0}^{\beta} \mathrm{d}\lambda \frac{\lambda^{2}(2\alpha^{2}-\lambda\beta-2\lambda\alpha\beta)}{\alpha^{3}} + \int_{\frac{\alpha}{\beta}}^{1} \mathrm{d}\lambda \frac{\mathrm{e}^{-2/\beta}}{\beta} & \mathrm{si} \ \alpha < \beta \end{cases}$$
(VII.70)

Les valeurs calculées ne correspondent pas aux valeurs obtenues par la simulation. L'approximation consistant à linéariser l'équation intégrale n'est donc pas valide pour les ordres supérieurs à 2. Cependant, le calcul menant à  $P(M(\lambda) = n)$  pour  $\lambda > \alpha/\beta$  est bien vérifié, quelque soit la valeur de n.

## VII.3.3 Calcul exact pour les cas simples

Le problème se simplifie pour les deux valeurs extrêmes des paramètres :  $\alpha = 0$  et  $\beta = 1$ .

a Cas  $\alpha = 0$ 

Pour  $\alpha = 0$ , un noeud surchargé défaille quelle que soit la charge de la défaillance qui l'engendre. Le nombre typique  $\mu(\ell)$  de descendants d'un individu de charge  $\ell$  ne dépend pas de la valeur  $\ell$  de sa charge :

$$\mu(\ell) = \frac{1}{\beta}.$$
 (VII.71)

On est ici toujours dans le cas  $\lambda > \alpha/\beta$  de la résolution de l'équation (VII.46), et la distribution de probabilité du nombre total de défaillances, intégrée sur toutes les valeurs possibles de la charge initiale  $L(0) = \lambda$  est donc donnée par la distribution de Borel-Tanner :

$$P(M=n) = \frac{\mathrm{e}^{-\frac{n}{\beta}}}{n!} \left(\frac{n}{\beta}\right)^{n-1}.$$
 (VII.72)

Cette distribution n'est normalisée que pour  $\beta = 1$ , pour les autres valeurs de  $\beta < 1$ , elle comporte une partie défective qui correspond à la probabilité que la cascade soit de taille infinie.

Dans ce cas simple, la résistance intrinsèque du réseau est nulle, et le nombre d'individus touchés ne dépend que des propriétés de connectivité du réseau : le nombre moyen de voisins d'un noeud, et donc du paramètre  $\beta$ .

b Cas  $\beta = 1$ 

Pour  $\beta = 1$ , le nombre typique de descendants d'un individu de charge  $L = \ell$  dépend de sa charge :

$$\mu(\ell) = \min\left(1, \frac{\ell}{\alpha}\right). \tag{VII.73}$$

La fonction génératrice du nombre total de défaillances doit donc être résolue selon les valeurs de la charge défaillante initiale  $\lambda$ .

La distribution de probabilité intégrée sur la distribution des charges initiales décroît cependant selon une loi de puissance d'exposant 3/2, c'est-à-dire qu'on retrouve le régime critique  $\mu = 1$  d'un processus de branchement simple. En effet, on montrera par la suite que le nombre moyen  $\bar{\mu}(s)$  de descendants par individus à la génération *s* converge très rapidement vers  $\bar{\mu}(s) = 1$ .

On vérifie sur la figure VII.6 que cette loi de puissance s'observe pour toutes les valeurs du paramètre  $\alpha$  pour  $n \gg 1$ , au préfacteur près. Pour de grandes valeurs de  $\alpha$ , la décroissance est plus rapide sur les premières valeurs de n. En effet, dans ce cas les quelques défaillances sont survenues dans les premières générations de la cascade, les charges défaillantes sont encore de l'ordre de 1, et le système sent encore la dépendance en  $\alpha$  de  $\mu(\ell)$ . Pour des cascades dont la taille est grande devant 1, les charges défaillantes se concentrent autour de valeurs de  $L > \alpha$ , et on retrouve un processus de paramètre  $\mu = 1$ . On détaille les propriétés de la distribution de la charge défaillante dans la suite de notre étude.



FIGURE VII.6 – Distribution de probabilité du nombre total M de défaillances suite à une défaillance initiale de charge quelconque, pour  $\beta = 1$ , et différentes valeurs du paramètre  $\alpha$ . Ces courbes sont obtenues par simulation du modèle H sur 10000 échantillons. La simulation est stoppée lorsque la taille de la cascade devient supérieure à N = 2000.

## VII.4 Nombre de défaillances par génération

L'incrément naturel du nombre total de défaillances dans la cascade après s générations est le nombre de défaillances  $N_s$  qui se produisent à la génération s donnée. On présente les quantités dans le cas simple du processus sans charges.

### VII.4.1 Processus de branchement simple

Le nombre de défaillances de la génération s + 1 de la cascade est la somme du nombre de descendants directs de toutes les défaillances de la génération s ou encore la somme du nombre de descendants à la *s*-ième génération de toutes les défaillances de la première génération :

$$N_{s+1} = N_s^{(1)} + \dots + N_s^{(k)}, \tag{VII.74}$$

où  $N_s^{(i)}$  est le nombre de descendants de la *s*-ième génération de la famille issue de la défaillance d'indice i = 1, ..., k, avec k le nombre de descendants directs de la défaillance initiale  $N_1 = k$ , tiré de la distribution de Poisson (VII.17). Les  $N_s^{(i)}$  sont des variables aléatoires indépendantes et identiquement distribuées. La fonction génératrice du nombre  $N_s$  de défaillances à la génération s suite à une unique défaillance initiale est notée :

$$F_s(t) = \left\langle t^{N_s} \right\rangle = \sum_{n=0}^{\infty} P(N_s = n) t^n, \qquad (\text{VII.75})$$

et vérifie l'équation de récurrence suivante :

$$F_{s+1}(t) = \sum_{n=0}^{\infty} P(N_{s+1} = n)t^n$$
  
=  $\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} P(N_1 = k)P(N_s^{(1)} + \dots + N_s^{(k)} = n)t^n$   
=  $\sum_{k=0}^{\infty} P(N_1 = k)(F_s(t))^k$   
=  $G(F_s(t)),$  (VII.76)

où G est la fonction génératrice de la distribution de Poisson du nombre de descendants directs  ${\cal N}_1.$ 

La fonction génératrice  $F_{\infty}$  de la valeur asymptotique du nombre de membres d'une génération  $N_{\infty}$  vérifie donc l'équation :

$$F_{\infty}(t) = \exp\left(\mu\left(F_{\infty}(t) - 1\right)\right),\tag{VII.77}$$

dont la solution ne dépend pas de l'argument t:

$$F_{\infty}(t) = -\frac{W(-e^{-\mu}\mu)}{\mu},$$
 (VII.78)

La distribution de probabilité est donc nulle partout, sauf en  $N_{\infty} = 0$ , où elle prend la valeur :

$$P(N_{\infty} = 0) = -\frac{W(-e^{-\mu}\mu)}{\mu}.$$
 (VII.79)

Cette probabilité possède deux régimes :

$$P(N_{\infty} = 0) = 1, \quad \text{si } \mu \le 1,$$
 (VII.80)

 $\operatorname{et}$ 

$$P(N_{\infty} = 0) < 1, \quad \text{si } \mu > 1.$$
 (VII.81)

Dans ce dernier cas, la probabilité complémentaire correspond à la probabilité que la taille de l'ultime génération soit infinie :

$$P(N_{\infty} = \infty) = 1 - P(N_{\infty} = 0) = 1 + \frac{W(-e^{-\mu}\mu)}{\mu}.$$
 (VII.82)

#### VII.4.2 Processus de branchement généralisé

La probabilité qu'un individu de charge  $L = \lambda$  ait  $N_2(\lambda)$  descendants après deux générations (*i.e.*  $N_2(\lambda)$  petits-enfants) dépend du nombre  $N_1(\lambda)$  de descendants directs qu'il a engendrés :

$$P(N_{2}(\lambda) = n) = \sum_{k=0}^{\infty} P(N_{1}(\lambda) = k) \int d\ell_{1} \dots d\ell_{k} f_{L_{1}|L}(\ell_{1}|\lambda) \dots f_{L_{k}|L}(\ell_{k}|\lambda)$$
  

$$P(N_{1}(\ell_{1}) + \dots + N_{1}(\ell_{k}) = n),$$
(VII.83)

où  $f_{L_i|L}$  est la distribution de la charge de la *i*-ème défaillance issue de  $L = \lambda$ . Le dernier terme de l'équation donne la probabilité que k charges indépendantes engendrent ensemble n descendants directs, quelle que soit la valeur que prennent ces charges.

La fonction génératrice du nombre de descendants à deux générations d'une charge initiale  $\lambda$  est notée  $F_2(\lambda, t)$ , et définie par :

$$F_2(\lambda, t) = \sum_n P(N_2(\lambda) = n)t^n.$$
 (VII.84)

On rappelle que la fonction génératrice du nombre de descendants directs d'une charge initiale  $\lambda$  est notée  $G(\lambda, t)$  (VII.10). On peut ici remarquer que la fonction génératrice du nombre de descendants à la première génération correspond à la fonction génératrice du nombre de descendants directs :  $G(\lambda, t) = F_1(\lambda)$ 

À partir de la relation (VII.83), on obtient l'équation de récurrence suivante pour la fonction génératrice du nombre de descendants à deux générations d'intervalle :

$$F_{2}(\lambda, t) = \sum_{k} P(N_{1}(\lambda) = k) \left( \int d\ell f_{L'|L}(\ell|\lambda) P(N_{1}(\ell) = \nu) t^{\nu} \right)^{k}$$
$$= \sum_{k} P(N_{1}(\lambda) = k) \overline{F_{1}(\lambda, t)}^{k}, \qquad (VII.85)$$

où la valeur moyenne est prise par rapport à la distribution des charges L(1) issues de la charge initiale L(0). On reconnait la fonction génératrice  $G(\lambda, t)$  du nombre de descendants directs de la charge initiale  $L(0) = \lambda$ :

$$F_2(\lambda, t) = G\left(\lambda, \overline{F_1(\lambda, t)}\right).$$
(VII.86)

On reproduit cette relation de récurrence pour le nombre de descendants  $N_s(\lambda)$  de la s-ième génération après défaillance d'un élément de charge  $L = \lambda$ . La fonction génératrice de cette quantité est notée :

$$F_s(\lambda, t) = \left\langle t^{N_s(\lambda)} \right\rangle.$$
 (VII.87)

La probabilité qu'un élément de charge  $\lambda$  ait n descendants à la s + 1-ième génération peut être déterminée par la probabilité que cet individu initial ait k descendants directs, et que la somme des descendants après s générations de ces k charges égale n:

$$P(N_{s+1}(\lambda) = n) = \sum_{k} P(N_1(\lambda) = k) P(N_s(L_1) + \dots + N_s(L_k) = n), \quad (\text{VII.88})$$

où les  $L_i$ , i = 1, ..., k sont les charges des k descendants directs de l'élément de charge  $\lambda$ . Et la fonction génératrice du nombre de descendants à la (s + 1)-ième génération est obtenue par analogie avec le raisonnement précédent :

$$F_{s+1}(\lambda, t) = \sum_{n=0}^{\infty} P(N_{s+1}(\lambda) = n)t^n$$
(VII.89)

$$= \sum_{k=0}^{\infty} P(N_1(\lambda) = k) \langle F_s(\lambda, t) \rangle^k$$
(VII.90)

$$= G\left(\lambda, \overline{F_s(\lambda, t)}\right).$$
(VII.91)

On obtient donc la relation de récurrence sur les fonctions génératrices du nombre de descendants d'un individu de charge  $L(0) = \lambda$  à la génération s + 1 par rapport à la génération s:

$$F_{s+1}(\lambda, t) = \exp\left[\mu(\lambda)\left(\overline{F_s(\lambda, t)} - 1\right)\right].$$
 (VII.92)

Dans la limite des grands temps  $s \to \infty$ ,  $F_{\infty}(\lambda, t) = \lim_{s \to \infty} F_s(\lambda, t)$ , l'équation de récurrence devient une équation intégrale :

$$F_{\infty}(\lambda, t) = \exp\left[\mu(\lambda)\left(\overline{F_{\infty}(\lambda, t)} - 1\right)\right], \qquad (\text{VII.93})$$

où l'intégrale est contenue dans la valeur moyenne de la fonction génératrice. La solution de cette équation ne dépend pas de l'argument t, et le nombre de membres de l'ultime génération ne peut donc prendre que deux valeurs :  $N_{\infty} = 0$  ou  $N_{\infty} = \infty$ . On étudie les probabilités de ces événements dans la section suivante.

## VII.4.3 Cas simple $\alpha = 0$

Dans le cas où  $\alpha = 0$ , on a déjà montré que le nombre moyen de descendants ne dépend pas de la charge de la défaillance. On se retrouve donc dans le cas du processus simple, avec  $\mu = 1/\beta$ . La fonction génératrice de la taille de l'ultime génération  $F_{\infty}$  et donc la probabilité que la taille de la dernière génération soit nulle, vérifie :

$$F_{\infty}(t) = P(N_{\infty} = 0) = -\beta W \left( -e^{-1/\beta}/\beta \right).$$
 (VII.94)

Cette probabilité n'est égale à 1 que pour  $\beta = 1$ , et sa complémentaire correspond à la probabilité que la taille de la génération soit asymptotiquement infinie :

$$P(N_{\infty} = \infty) = 1 + \beta W \left( -e^{-1/\beta}/\beta \right).$$
(VII.95)

On trace, sur la figure VII.7, la formule analytique de la probabilité que l'ultime génération soit de taille infinie, en fonction du paramètre  $\beta$ .

## VII.5 Probabilité de rupture

Dans le formalisme du modèle L&B, la rupture du système correspond à une cascade de défaillances qui affecte tous les noeuds du réseau. Le système étant, dans le formalisme du modèle H, infini, la probabilité de rupture correspond à la probabilité que la famille de défaillances ne s'éteigne jamais.



FIGURE VII.7 – Probabilité que la dernière génération  $s \to \infty$  soit de taille infinie pour  $\alpha = 0$  en fonction de  $\beta$ , à partir de l'équation (VII.95).

#### VII.5.1 Probabilité d'extinction

On définit la probabilité d'extinction  $d_s$  de la cascade à une étape s comme la probabilité que le nombre de membres  $N_s$  de cette génération soit nul. Pour une cascade issue d'un unique individu de charge  $L(0) = \lambda$ , le nombre de membres de la génération s est égal au nombre de descendants  $N_s(\lambda)$  de l'individu de charge  $L = \lambda$  à la s-ième génération, on note :

$$d_s(\lambda) = P(N_s(\lambda) = 0). \tag{VII.96}$$

La fonction génératrice de  $N_s(\lambda)$  est  $F_s(\lambda, t)$ , qui est donnée par la relation de récurrence (VII.92). La probabilité d'extinction  $d_s(\lambda)$  à la génération s d'une cascade issue d'un unique individu de charge  $\lambda$  est liée à la fonction génératrice  $F_s(\lambda, t)$  par la relation suivante :

$$d_s(\lambda) = F_s(\lambda, 0). \tag{VII.97}$$

On pose t = 0 dans la relation (VII.92), d'où la récurrence sur la probabilité d'extinction à la génération s d'une cascade issue de  $L(0) = \lambda$ :

$$d_{s+1}(\lambda) = \exp\left[\mu(\lambda)\left(\overline{d_s(\lambda)} - 1\right)\right].$$
 (VII.98)

Avec les caractéristiques du problème, la valeur moyenne de la probabilité d'extinction s'écrit encore :

$$\overline{d_s(\lambda)} = \int_{I(\lambda)} \frac{\mathrm{d}\ell}{\beta\mu(\lambda)} d_s(\ell) \tag{VII.99}$$

où  $I(\lambda)$  est l'intervalle de définition des charges des défaillances issues de la défaillance de charge  $L(0) = \lambda$ .

La probabilité d'extinction  $d_s$  pour une charge quelconque, :

$$d_s = \int d\lambda \ f_{L(0)}(\lambda) d_s(\lambda), \qquad (\text{VII.100})$$

à une génération donnée est tracée figure VII.8.

Dans le cas simple de  $\beta = 1$ , la probabilité d'extinction converge en 1/s vers 1, comme on peut l'observer sur la figure VII.9, quelle que soit la valeur de  $\alpha$ . Seul le préfacteur change, et correspond à la probabilité de non-cascade associée aux paramètre  $\beta = 1$  et  $\alpha$ .

Si à une étape s la cascade est éteinte, *i.e*  $N_s = 0$ , alors elle est éteinte à toutes les générations ultérieures. La probabilité que la cascade initiée par un élément de charge  $L = \lambda$  soit éteinte à l'ultime étape est la limite asymptotique de la probabilité  $d_s(\lambda)$  que la cascade issue de L soit éteinte à la génération  $s : d(\lambda) = \lim_{s\to\infty} d_s(\lambda)$ . On prend donc la limite  $s \to \infty$  dans les deux membres de l'équation de récurrence (VII.98), et la probabilité d'extinction  $d(\lambda)$  d'une cascade initiée par une charge  $\lambda$  vérifie l'équation intégrale :

$$d(\lambda) = \exp\left[\mu(\lambda)\left(\overline{d(\lambda)} - 1\right)\right], \qquad (\text{VII.101})$$

avec  $d(\lambda)$  la moyenne de la probabilité d'extinction sur la distribution des charges issues d'une défaillance  $L = \lambda$ . On retrouve ce résultat si on pose t = 0 dans l'équation intégrale (VII.93), car  $d(\lambda) = P(N_{\infty}(\lambda) = 0) = F_{\infty}(\lambda, 0)$ .


FIGURE VII.8 – Probabilité d'extinction en fonction de la génération pour différentes valeurs du paramètre  $\beta$ , pour  $\alpha$  fixé, par calcul de l'équation de récurrence (VII.98).



FIGURE VII.9 – Complémentaire de la probabilité d'extinction en fonction de la génération à  $\beta = 1$ , pour  $\alpha = 0.2$  et  $\alpha = 1$ .

#### VII.5.2 Probabilité de rupture

Si la cascade ne s'éteint jamais, un nombre infini de charges défaillent : il y a rupture. La probabilité de rupture  $P_b$  est donc la probabilité complémentaire de la probabilité asymptotique d'extinction d, et pour une cascade initiée par un élément de charge  $\lambda$ :

$$P_b(\lambda) = 1 - d(\lambda). \tag{VII.102}$$

On obtient donc l'équation intégrale suivante pour la probabilité de rupture suite à une défaillance initiale de charge  $\lambda$ :

$$1 - P_b(\lambda) = \exp\left[\mu(\lambda) \left(\int d\ell f_{L_i|L(0)}(\ell|\lambda) \left(1 - P_b(\ell)\right) - 1\right)\right],$$
 (VII.103)

où la densité de probabilité de la charge  $L_i$  des descendantes, sachant que la défaillance-mère a pour charge L(0), est normalisée, et l'équation (VII.103) se simplifie en :

$$1 - P_b(\lambda) = \exp\left[-\mu(\lambda) \int d\ell f_{L_i|L(0)}(\ell|\lambda) P_b(\ell)\right].$$
 (VII.104)

Le nombre moyen de défaillances  $\mu(\lambda)$  engendrées par une défaillance de charge  $L(0) = \lambda$  et la distribution conditionnelle  $f_{L_i|L(0)}(\ell|\lambda)$  de la charge  $L_i = \ell$  d'une défaillance engendrée par une défaillance de charge  $L(0) = \lambda$  sont données respectivement par les équations (VII.9) et (VII.11). La probabilité de rupture  $P_b(\lambda)$  suite à une défaillance initiale de charge  $L(0) = \lambda$  est donc :

$$1 - P_b(\lambda) = \exp\left[-\frac{1}{\beta} \int_{I(\lambda)} d\ell P_b(\ell)\right],$$
 (VII.105)

où  $I(\lambda)$  est l'intervalle de définition de la distribution conditionnelle de la charge-fille  $f_{L_i|L(0)}(\ell|\lambda)$  sachant qu'elle est engendrée par une défaillance de charge  $L = \lambda$ :

$$I(\lambda) = (\beta\lambda, \beta\lambda + \beta\mu(\lambda)).$$
(VII.106)

On retrouve bien la formule intégrale de Lehmann et Bernasconi [74].

On trace sur la figure VII.10 la probabilité de rupture obtenue par intégration numérique de la relation (VII.105) moyennée sur l'ensemble des valeurs possibles de la défaillance de la génération s = 0:

$$P_b = \int_0^1 \mathrm{d}\lambda P_b(\lambda). \qquad (\text{VII.107})$$

On observe une transition de  $P_b = 0$  à  $P_b \neq 0$  pour une valeur finie  $\beta_c$  du paramètre  $\beta$ , qui dépend de la valeur du paramètre  $\alpha$ .

La probabilité  $P_b(\lambda)$  de rupture après une défaillance initiale  $L(0) = \lambda$  est la probabilité complémentaire de la probabilité asymptotique d'extinction  $d_{\infty}$ , et donc la probabilité complémentaire de  $P(N_{\infty}(\lambda) = 0)$ , soit encore la partie défective de la normalisation de la distribution de probabilité du nombre de défaillances de l'ultime génération :

$$P_b(\lambda) = 1 - \sum_{n=0}^{\infty} P(N_{\infty}(\lambda) = n).$$
(VII.108)



FIGURE VII.10 – Probabilité de rupture après une défaillance initiale de charge quelconque en fonction du paramètre  $\beta$ , pour différentes valeurs du paramètre  $\alpha$ .

On retrouve bien, pour  $\alpha = 0$  la fonction :

$$P_b = 1 - \beta W \left( -\mathrm{e}^{-1/\beta}/\beta \right), \qquad (\text{VII.109})$$

décroissante de 1 à 0. Dans ce cas, la résistance intrinsèque du réseau est nulle, et la probabilité d'une cascade globale dépend de la connectivité du réseau sous-jacent.

#### a Comparaison avec l'algorithme L& B

On compare sur la figure VII.11 les résultats du calcul numérique de l'équation intégrale (VII.105) (ligne continue) et de la probabilité de rupture obtenue par simulation de l'algorithme L&B (points). On trouve un bon accord entre les valeurs calculées et les valeurs obtenues par simulation, sauf au niveau de la transition à gauche. Lehmann et Bernasconi indiquent dans leur article que cette différence est due à des effets de taille finie, et que pour un système de taille  $N \to \infty$ , cette différence disparaît.

On a constaté section VI.2 que la taille du système n'influe pas sur la validité des approximations, mais que c'est plutôt le rapport entre le nombre de survivants et la taille du système. En particulier pour les valeurs repérées par l'axe vertical rouge sur la figure VII.11, qui correspondent aux valeurs de  $\alpha$ ,  $\beta$  et N pour lesquelles tous les graphes de la section VI.3.1 ont été tracés, la valeur de la probabilité de rupture déterminée analytiquement correspond à la valeur obtenue par simulation. On a cependant remarqué précédemment que les approximations ne



FIGURE VII.11 – Probabilité de rupture obtenue par calcul numérique de l'équation (VII.105) (ligne) et par simulation de l'algorithme L&B (points). La ligne rouge pointe les valeurs utilisées dans les figures précédentes :  $\alpha = 0.2$  et  $\beta = 0.4$ . La probabilité de rupture pour la simulation correspond aux histoires dans les quelles le nombre total de défaillances égale le nombre total d'éléments dans le système.

sont pas vérifiées pour la statistique complète de la simulation, mais seulement pour un nombre de défaillances petit devant la taille du système, *i.e.* pour les premières générations.

Hors de la zone de transition, il semble donc que la probabilité que le système subisse une rupture est entièrement déterminée par les premières générations de la cascade, lorsque  $M \ll N$ . Pour les valeurs proches de la transition, la stabilité du système semble dépendre de toute l'histoire de la cascade.

#### VII.5.3 Intervalle de définition de P<sub>b</sub>

La probabilité de rupture  $P_b(\lambda)$  pour une cascade issue d'une défaillance de charge  $\lambda$  est obtenue à partir de la probabilité d'extinction asymptotique  $d(\lambda)$ . La probabilité  $d_s(\ell)$  d'extinction à la génération s est définie par récurrence, et prend ses valeurs dans l'intervalle  $(\ell_{min}(s), \ell_{max}(s))$ . Cet intervalle contient toutes les valeurs possibles des charges défaillantes. Selon l'équation intégrale (VII.98), pour déterminer la probabilité d'extinction après deux générations, on doit connaître les valeurs de la probabilité d'extinction sur l'intervalle  $I(\lambda) = (\ell_{min}(1), \ell_{max}(1))$ . On définit les bornes de cet intervalle par un raisonnement issu du modèle de Lehamnn et Bernasconi. On ajoute à une charge  $u \in (0, 1)$  une fraction  $\beta$  d'une charge  $\lambda \in (0, 1)$ . Les charges u et  $\lambda$  sont telles que le noeud défaille, *i.e.*  $\ell(1) = u + \beta \lambda > u^{max} = (1 + \alpha)u$ .

Un noeud de charge u = 0 défaille pour toute valeur de la surcharge incidente et la borne inférieure de l'intervalle de définition  $I(\lambda)$  est donc  $\ell_{min} = 0$ . On peut trouver une borne supérieure  $\ell_{max}$  de l'intervalle en fonction de deux cas distincts. La valeur de la charge défaillante initiale est toujours  $\lambda \in (0, 1)$ .

• Si  $\alpha > \frac{\beta}{1-\beta}$ , l'intégrale de la première itération est définie sur des valeurs de  $\ell(1)$  telles que :

$$\begin{aligned} \alpha u &< \beta \lambda \\ \Rightarrow \frac{\beta}{1-\beta}u &< \beta \lambda < \beta \\ \Rightarrow u &< 1-\beta \\ \Rightarrow \ell(1) &< 1-\beta(1-\lambda) < 1 \end{aligned}$$
(VII.110)

Le calcul des valeurs de  $d_s(\ell)$  pour les valeurs de  $\ell \in I(\lambda)$  dépend des valeurs de  $d_{s-1}$  dans l'intervalle  $(\ell_{min}(2), \ell_{max}(2))$  et ainsi de suite. Par récurrence, si lors de *s*-ième itération la charge défaillante  $\ell(s)$  est inférieure à 1, la charge défaillante  $\ell(s+1)$  de l'itération suivante vérifie :

$$\begin{aligned} \alpha u &< \beta \ell(s) \\ \Rightarrow u &< 1 - \beta \\ \Rightarrow \ell(s+1) &< 1 - \beta(1 - \ell(s)) < 1 \end{aligned}$$
(VII.111)

Et la borne supérieure de l'intervalle est  $\ell_{max} = 1$ .

• Si  $\alpha < \frac{\beta}{1-\beta}$ , pour toute valeur de la charge  $u \in (0,1)$  de l'élément surchargé, on a pour

les itérations successsives :

$$0 < \ell(1) = u + \beta\lambda < 1 + \beta$$
  

$$0 < \ell(2) = u + \beta\ell(1) < 1 + \beta(1 + \beta)$$
  

$$\vdots$$
  

$$0 < \ell(s+1) = u + \beta\ell(s) < \sum_{i=0}^{s+1} \beta^i \le \frac{1}{1-\beta} \forall s,$$
 (VII.112)

et la borne supérieure de la valeur de la charge défaillante est  $\ell_{max} = \frac{1}{1-\beta}$ .

Notons que cet intervalle  $(0, \ell_{max})$  est l'intervalle sur lequel on doit résoudre l'équation intégrale pour d et donc pour  $P_b$ . Cependant, la probabilité de rupture  $P_b$  quelle que soit la valeur de la défaillance initiale L(0) est la moyenne sur la distribution  $f_{L(0)}$  des charges initiales :

$$P_b = \int_0^1 P_b(\lambda) \mathrm{d}\lambda, \qquad (\text{VII.113})$$

et ce, quelle que soit la valeur de  $\ell_{max}$ .

#### VII.5.4 Résolution numérique de l'équation intégrale

On discrétise l'intervalle de définition  $(\ell_{min}, \ell_{max})$  de la fonction  $P_b$  avec une résolution K:

$$\ell_i = \ell_{min} + \frac{i(\ell_{max} - \ell_{min})}{K}, \ i = 0, \dots, K.$$
 (VII.114)

La valeur prise par la fonction  $P_b$  aux points discrets  $\ell_i$  est notée  $P_i = P_b(\ell_i)$ . On initialise la résolution par itérations successives en posant :

$$P_b^0(\ell_i) = P_i^0 = \frac{i}{K}.$$
 (VII.115)

On injecte ces valeurs dans l'intégrale de l'équation (VII.105) :

$$P_b(\ell) = 1 - \exp\left\{-\frac{1}{\beta} \int_{\beta\ell}^{\beta\ell + \min(1,\beta\ell/\alpha)} \mathrm{d}y P_b^0(y)\right\}.$$
 (VII.116)

On discrétise l'intégrale par la méthode du trapèze. Pour une intégrale que lconque, avec résolution K, on a :

$$\int_{a}^{b} f(x) dx = \frac{b-a}{2K} \sum_{i=0}^{K-1} \left( f_i + f_{i+1} \right), \qquad (\text{VII.117})$$

où  $f_i = f(a + i\frac{b-a}{K}).$ 

Dans notre cas, la résolution m de la discrétisation de l'intégrale est imposée par la résolution K de la discrétisation de l'intervalle de définition de  $P_b$ . Pour une valeur  $\ell_i = \frac{i\ell_{max}}{K}$  donnée, l'intégrale est définie entre les bornes :

$$\beta \frac{i\ell_{max}}{K} \text{ et } \beta \frac{i\ell_{max}}{K} + \min\left(1, \frac{\beta}{\alpha} \frac{i\ell_{max}}{K}\right).$$
 (VII.118)

Dans le contexte du calcul discret, on approxime ces bornes par les valeurs discrètes  $\ell_{i^-}$  et  $\ell_{i^+}$  les plus proches, qui correspondent aux parties entières suivantes :

$$i^{-} = \lfloor i\beta \rfloor, \ i^{+} = \lfloor i\beta + \min\left(\frac{K}{\ell_{max}}, \frac{i\beta}{\alpha}\right) \rfloor,$$
 (VII.119)

c'est-à-dire que la résolution de la discrétisation de l'intégrale est :

$$m = \left\lfloor i\beta + \min\left(\frac{K}{\ell_{max}}, \frac{i\beta}{\alpha}\right) \right\rfloor - \lfloor i\beta \rfloor.$$
 (VII.120)

Si  $m \neq 0$ , l'équation intégrale discrétisée s'écrit donc :

$$P_i = P_b(\ell_i) = 1 - \exp\left\{-\frac{1}{\beta}\min\left(1, \frac{\beta\ell_i}{\alpha}\right)\frac{1}{2m}\sum_{j=0}^{m-1}\left(P^0_{\lfloor i\beta \rfloor + j} + P^0_{\lfloor i\beta \rfloor + j+1}\right)\right\},\qquad(\text{VII.121})$$

et si  $m = 0, P_i = 0.$ 

Puis on itère le processus, avec  $P_i^0 = P_i$  à chaque étape, jusqu'à ce que les solutions aient convergé. Les courbes figure VII.10 sont obtenues grâce à ce calcul numérique.

## VII.6 Valeurs de transition

On peut donc implémenter l'équation discrétisée, et obtenir les valeurs numériques de la probabilité de rupture. On s'intéresse en particulier au seuil de transition de  $P_b$ :  $\beta_c$ , et on en détermine numériquement les valeurs, en fonction du paramètre  $\alpha$ . On représente  $\beta_c(\alpha)$  sur la figure VII.12.

Sous cette valeur de transition  $\beta_c$ , il existe un intervalle de valeurs du paramètre  $\beta$  dans lequel on peut voir apparaître de grandes cascades qui survivent durant un grand nombre de générations. Sous une certaine valeur  $\beta_{\star}$ , la probabilité de survie de la cascade au-delà d'un nombre de générations de l'ordre de 1 est nulle. Cette autre valeur de transition  $\beta_{\star}$  du paramètre de redistribution dépend de la valeur du paramètre  $\alpha$ . Elle est définie comme la valeur de  $\beta$  telle que pour un  $\alpha$  donné, la probabilité que la taille d'une génération à un temps asymptotiquement long soit finie devient strictement nulle. La fonction génératrice  $F_s(\lambda, t)$  de la taille de la *s*-ième génération après une défaillance initiale  $\lambda$  est donnée par l'équation de récurrence (VII.92), dans laquelle la moyenne est prise sur l'intervalle  $I(\lambda)$  (VII.106), dont les bornes dépendent de l'argument  $\lambda$ . La fonction génératrice du nombre de défaillances d'une génération quelle que soit la valeur de la défaillance initiale  $0 < \lambda < 1$ . Dans la limite des temps longs, la fonction génératrice vérifie l'équation (VII.93), et ne dépend plus du paramètre t. La probabilité  $P(N_s = 0)$  que la génération s, pour  $s \to \infty$ , soit vide est donnée par :

$$P(N_s = 0) = \int_0^1 F_s(\lambda, t = 0) d\lambda.$$
 (VII.122)

La valeur de transition vers un régime de grandes cascades correspond donc au jeu de paramètres  $(\alpha, \beta_{\star})$  pour lequel  $P(N_s(\lambda) = 0)$  devient inférieur à 1  $\forall \lambda \in (0, 1)$ , ou encore  $1 - P(N_s(\lambda) = 0) = 1 - F_s(\lambda, 0)$  devient supérieur à 0.



FIGURE VII.12 – Valeur critique  $\beta_c$  du paramètre  $\beta$  en fonction de  $\alpha$ , déterminé numériquement comme la valeur de  $\beta$  pour lequel, pour un  $\alpha$  donné,  $P_b$  devient non nul, et  $\bar{\mu}(s) \xrightarrow[s \to \infty]{} 1$ . Les valeurs de  $\beta_c$  sont mesurées à une précision de  $10^{-4}$ .

L'équation intégrale (VII.93) de la fonction génératrice du nombre de descendants de la génération  $s \to \infty$  à la suite d'une défaillance initiale de charge  $\lambda$  peut être écrite sous la forme d'une équation différentielle pour la primitive :

$$\mathcal{F}_s(\lambda, t) = \int_0^\lambda F_s(x, t) \mathrm{d}x.$$
 (VII.123)

Cette primitive vérifie l'équation différentielle :

$$\frac{\mathrm{d}\mathcal{F}_s(\lambda,t)}{\mathrm{d}\lambda} = \exp\left[\frac{-\mathcal{F}_s\left(\beta\lambda + \min(1,\beta\lambda/\alpha,t)\right) + \mathcal{F}(\beta\lambda,t)}{\beta} - \mu(\lambda)\right].$$
 (VII.124)

On définit une fonction d'étude  $\mathcal{G}(\lambda, t) = \lambda - \mathcal{F}_s(\lambda, t)$ , primitive de  $1 - F_s(\lambda, t)$ , et dont l'équation différentielle est donnée par :

$$-\frac{\mathrm{d}\mathcal{G}(\lambda,t)}{\mathrm{d}\lambda} + 1 = \exp\left[\frac{-\mathcal{G}\left(\beta\lambda + \min(1,\beta\lambda/\alpha),t\right) + \mathcal{G}(\beta\lambda)}{\beta},t\right].$$
 (VII.125)

Cette équation différentielle a une solution constante en  $\lambda$  si  $\frac{\beta\lambda}{\alpha} < 1$ , quelle que soit la fonction  $\mathcal{G}$ . En première approximation pour la résolution de l'équation (VII.125), on pose donc :

$$\mathcal{G}(\lambda, t) = \mathcal{G}_1(\lambda, t) = a(t), \text{ si } \frac{\beta\lambda}{\alpha} < 1 \Leftrightarrow \lambda < \frac{\alpha}{\beta}.$$
 (VII.126)

L'argument de la seconde occurrence de la fonction  $\mathcal{G}$  dans le terme de droite de l'équation (VII.125) est toujours inférieur à l'argument de la fonction du terme de gauche :  $\beta\lambda < \lambda$  car  $0 < \beta < 1$ . L'argument de la première fonction du terme de droite peut être plus grand que l'argument de gauche, mais surtout :

$$\beta\lambda\left(1+\frac{1}{\alpha}\right) > \beta\lambda,$$
 (VII.127)

et la première fonction  $\mathcal G$  du terme de droite est définie par la fonction  $\mathcal G_1$  si et seulement si :

$$\beta\lambda\left(1+\frac{1}{\alpha}\right) < \frac{\alpha}{\beta}.$$
 (VII.128)

La solution de l'équation différentielle est donc une fonction constante si :

$$\lambda < \frac{\alpha^2}{\beta^2} \frac{1}{1+\alpha}.$$
 (VII.129)

On définit la probabilité que la génération  $s \to \infty$  soit vide à partir de la fonction d'étude par :

$$P(N_s(\lambda) = 0) = 1 - \frac{\mathrm{d}\mathcal{G}(\lambda, 0)}{\mathrm{d}\lambda}, \qquad (\text{VII.130})$$

donc pour  $\mathcal{G}(\lambda, t) = a(t)$ ,  $P(N_s(\lambda) = 0) = 1$ . La probabilité que la génération ne soit pas vide pour une défaillance initiale quelconque est donnée par la probabilité que la génération ne soit pas vide pour une charge initiale  $L(0) = \lambda$  intégrée sur l'intervalle (0, 1). La probabilité intégrée est donc nulle si :

$$1 - P(N_s(\lambda) = 0) = 0 \text{ pour tout } \lambda < 1, \qquad (\text{VII.131})$$

c'est-à-dire que si :

$$\lambda < 1 < \frac{\alpha^2}{\beta^2} \frac{1}{1+\alpha},\tag{VII.132}$$

alors  $1 - P(N_s(\lambda) = 0) = 0, \forall \lambda < 1$ , et

$$\frac{\alpha^2}{\beta^2} \frac{1}{1+\alpha} > 1 \Leftrightarrow \beta < \frac{\alpha}{\sqrt{1+\alpha}},\tag{VII.133}$$

en première itération.

Cependant, pour  $\lambda < \frac{\alpha^2}{\beta^2} \frac{1}{1+\alpha}$ , l'argument  $\beta \lambda \left(1 + \frac{1}{\alpha}\right)$  doit aussi être inférieur à la borne pour que la fonction  $\mathcal{G}$  soit constante et  $1 - P\left(N_s(\lambda) = 0\right) = 0$ . On doit donc avoir :

$$\beta\lambda\left(1+\frac{1}{\alpha}\right) < \frac{\alpha^2}{\beta^2} \frac{1}{1+\alpha}$$
(VII.134)

$$\Leftrightarrow \lambda < \frac{\alpha^3}{\beta^3} \frac{1}{(1+\alpha)^2}.$$
 (VII.135)

Et la probabilité que la génération ne soit pas vide est nulle pour une défaillance initiale quelconque si :

$$\frac{\alpha^3}{\beta^3} \frac{1}{(1+\alpha)^2} > 1,$$
 (VII.136)

 $\operatorname{soit}$ 

$$\beta < \frac{\alpha}{(1+\alpha)^{2/3}}.\tag{VII.137}$$

On itère ce raisonnement, jusqu'à la limite asymptotique de la valeur de transition du paramètre  $\beta$  : pour

$$\beta < \beta_{\star} = \frac{\alpha}{1+\alpha},\tag{VII.138}$$

la probabilité que la s-ième génération de la cascade ne soit pas vide est nulle, pour  $s \to \infty$ . Cette valeur de transition est cependant différente de la valeur de transition  $\beta_c$ . En effet, pour  $\beta_{\star} < \beta < \beta_c$ , la taille de chaque génération s peut être non nulle tant que  $s < \infty$ . Lors du passage à l'infini, seuls deux cas peuvent se produire :  $N_{\infty} = 0$  ou  $N_{\infty} = \infty$ . Sous  $\beta_c$ , la distribution de la taille de la famille décroît plus vite qu'une loi de puissance, et n'a donc pas de poids à l'infini. La cascade s'éteint donc *in fine* lorsque les valeurs du paramètre  $\beta < \beta_c$ . On montre dans la section VII.9.2 ci-dessous que pour  $\beta_{\star} < \beta < \beta_c$ , le nombre asymptotique moyen de descendants par défaillance prend une valeur finie, inférieure à 1.

Le résultat obtenu pour la valeur de  $\beta_{\star}$  est semblable au résultat que Lehmann et Bernasconi ont obtenu pour la valeur de transition d'un modèle de faisceau de fibres [96], dont les propriétés sont proches du modèle étudié ici.

On représente sur la figure VII.13 les trois premières itérations du calcul de la valeur de transition  $\beta_{\star}$  ainsi que la valeur asymptotique. On observe que la valeur asymptotique est très



FIGURE VII.13 – Valeur de transition  $\beta_{\star}$  du paramètre  $\beta$  en fonction du paramètre  $\alpha$ , pour laquelle  $P_b = 0$ . Les différentes courbes représentent les trois premières approximations pour le calcul de la valeur de transition et la valeur asymptotique.

proche des valeurs obtenues pour les premières itérations pour de faibles valeurs de  $\alpha$ , tandis que pour  $\alpha \sim 1$ , la valeur de transition est étalée pour les itérations successives.

À cette transition, la propagation des cascades dans le système est principalement limitée par la stabilité locale des éléments du réseau. Le nombre de défaillances totales dans la cascade décroît plus vite qu'une loi de puissance, comme on peut le voir sur la figure VII.14. Dans certains rares cas, une cascade de très grande taille peut se produire. Watts [20] décrit dans un modèle de cascades sur réseaux aléatoires une telle transition entre un régime de cascades globales et un régime dans lequel les cascades sont petites devant la taille du système.

## VII.7 Temps d'arrêt de la cascade

Le temps d'arrêt de la cascade est défini comme la génération  $s_a$  pour laquelle aucune défaillance ne se produit, *i.e.* la génération d'extinction de la cascade. La probabilité d'extinction en fonction de la génération est donnée par l'équation (VII.99). La probabilité que la cascade s'éteigne à la génération  $s_a$  est donnée par la dérivée discrète de la probabilité d'extinction d(s):

$$f_{s_a}(s) = \operatorname{Prob}(s_a = s) = d(s+1) - d(s).$$
 (VII.139)



FIGURE VII.14 – Distribution du nombre total de défaillances dans la cascade, pour un système juste au-dessus de la transition  $\beta_{\star}$ .

On compare sur la figure VII.15 les résultats de la simulation et les résultats de la résolution numérique de l'équation intégrale (VII.98) pour la probabilité d'extinction en fonction de la génération.

Pour les valeurs des paramètres tels que la probabilité de rupture est non nulle, la génération d'arrêt de la cascade est distribuée selon une loi exponentielle. À la transition  $\beta = 1$ , on vérifie bien que la distribution du temps d'arrêt est la dérivée discrète de la probabilité d'extinction en fonction du temps. En effet, la probabilité d'exctinction pour  $\beta = 1$  se comporte en 1/s, et on vérifie sur la figure VII.15(e) que la distribution du temps d'arrêt de la cascade converge vers une droite d'équation  $1/s^2$ .

## VII.8 Distribution des charges défaillantes

Dans le modèle H, une charge défaillante L(s) à la génération s est une variable aléatoire somme de deux variables aléatoires indépendantes : la charge initiale  $u_i$  de l'élément i et la surcharge  $\beta L(s-1)$ . La charge initiale est tirée de la distribution  $f_{L(0)}$ , la surcharge est déterminée à partir de la charge défaillante L(s-1), tirée de la distribution des charges défaillantes de la génération s-1.

Le nombre de descendants d'une défaillance de la (s-1)-ième s génération dépend de



(e) À la transition  $\beta = 1$ 

FIGURE VII.15 – Distribution du temps d'arrêt de la cascade pour une même valeur du paramètre  $\alpha$  et différentes valeurs de  $\beta$  telles que le jeu de paramètres soit avant (a), après (c)-(d) et à la transition à gauche (b) ou à droite pour  $\beta = 1$  (e). Les simulations ont été réalisées sur des cascades de taille maximale N = 2000 188

la valeur de sa charge  $L(s-1) = \ell_{s-1}$ , et est distribué selon une loi de Poisson (VII.8) de paramètre  $\mu(\ell_{s-1})$ . Le poids de la charge des descendants de l'élément de charge L(s-1) dans la distribution des charges défaillantes de la génération s dépend donc de la distribution du nombre de défaillances issues de cet élément.

#### VII.8.1 Distribution à la première génération

À la première génération s = 1, les charges défaillantes sont issues d'une unique défaillance dont la charge  $L(0) = \lambda$  est distribuée selon  $f_{L(0)}$ . La densité de probabilité non normalisée du nombre de descendants de charge  $L(1) = \ell_1$  engendrés par un individu de charge  $\lambda$  est donnée par :

$$\tilde{\rho}_{L(1),N_1}(\ell_1, n | \lambda) = P(N_1(\lambda) = n) n f_{L(1)|L(0)}(\ell_1 | \lambda), \qquad (\text{VII.140})$$

où  $f_{L(1)|L(0)}$  est la densité de probabilité conditionnelle (VII.4) de la charge d'une défaillance L(1) issue d'une défaillance L(0). Le facteur de normalisation de la densité du nombre d'éléments de charge  $\ell_1$  à la première génération est donc donné par :

$$\bar{\rho}_{L(1)}(\ell_1) = \int d\lambda f_{L(0)}(\lambda) \sum_n P(N_1(\lambda) = n) n f_{L(1)|L(0)}(\ell_1|\lambda), \qquad (\text{VII.141})$$

et le nombre typique de descendants d'une défaillance initiale  $L(0) = \lambda$  est :

$$\bar{\mu}(0) = \int d\lambda f_{L(0)}(\lambda)\mu(\lambda) = \int_0^1 d\lambda \sum_n P(N_1(\lambda) = n)n.$$
(VII.142)

La distribution d'une charge défaillante à la génération s = 1 est calculée quels que soient la charge de la défaillance initiale et le nombre de défaillances engendrées à la première génération, et est donc donnée par :

$$f_{L(1)}(\ell) = \frac{\bar{\rho}_{L(1)}(\ell)}{\bar{\mu}(0)} = \frac{1}{\bar{\mu}(0)} \int_0^1 d\lambda \mu(\lambda) f_{L(1)|L(0)}(\ell|\lambda), \qquad (\text{VII.143})$$

le nombre moyen de défaillances directes d'une défaillance de charge  $\lambda$  étant  $\mu(\lambda)$ . On remplace la distribution conditionnelle de la charge défaillante par son expression explicite :

$$f_{L(1)}(\ell_1) = \frac{1}{\bar{\mu}(0)} \int_0^1 d\lambda \frac{\mu(\lambda)}{\min\left(1, \frac{\beta\lambda}{\alpha}\right)} \Theta(\ell_1 - \beta\lambda) \Theta\left(\beta\lambda + \min\left(1, \frac{\beta\lambda}{\alpha}\right) - \ell_1\right), \qquad (\text{VII.144})$$

où  $\mu(\lambda) = \min(1/\beta, \lambda/\alpha)$ . La fonction  $\Theta$  est la fonction de Heaviside, utilisée ici pour exprimer les bornes de la densité conditionnelle  $f_{L(1)|L(0)}$  de la charge L(1) sachant que  $L(0) = \lambda$ . La densité de probabilité de la charge d'une défaillance de la première génération est donc :

$$f_{L(1)}(\ell_1) = \frac{1}{\bar{\mu}(0)\beta} \int_0^1 d\lambda \Theta(\ell_1 - \beta\lambda) \Theta\left(\beta\lambda + \min\left(1, \frac{\beta\lambda}{\alpha}\right) - \ell_1\right).$$
(VII.145)

On distingue deux cas pour le facteur de normalisation  $\bar{\mu}(0)$ :

• Si  $\alpha > \beta$ , alors  $\alpha/\beta > 1$  et le nombre moyen de défaillances  $\mu(\lambda) = \lambda/\alpha$ ,  $\forall \lambda \in (0, 1)$ . Le facteur de normalisation vaut alors :

$$\bar{\mu}(0) = \int_0^1 \mathrm{d}\lambda \frac{\lambda}{\alpha} = \frac{1}{2\alpha}.$$
 (VII.146)

• Si  $\alpha < \beta$ , alors  $\alpha/\beta < 1$  et le nombre moyen de défaillances  $\mu(\lambda)$  prend deux formes différentes selon la valeur de  $\lambda$ . Le facteur de normalisation vaut alors :

$$\bar{\mu}(0) = \int_0^{\frac{\alpha}{\beta}} \mathrm{d}\lambda \frac{\lambda}{\alpha} + \int_{\frac{\alpha}{\beta}}^1 \frac{\mathrm{d}\lambda}{\beta} = \frac{2\beta - \alpha}{2\beta^2}.$$
 (VII.147)

#### VII.8.2 Distribution à la *s*-ième génération

Le raisonnement suivi dans la section précédente est valable pour toute génération ultérieure. A la génération s, une défaillance est issue de l'une des défaillances de la génération s - 1, de charge L(s-1), distribuée selon la densité de probabilité  $f_{L(s-1)}$ , et la distribution de la charge d'une défaillance de la génération s est donc donnée par :

$$f_{L(s)}(\ell_s) = \frac{\sum_n \int d\ell_{s-1} f_{L(s-1)}(\ell_{s-1}) \tilde{\rho}_{L(s),N_1}(\ell_s | \ell_{s-1})}{\sum_n \int d\ell f_{L(s-1)}(\ell_{s-1}) n P(N_1(\ell_{s-1}) = n)},$$
(VII.148)

où  $\tilde{\rho}_{L(s),N_1}(\ell_s|\ell_{s-1}) = P(N_1(\ell_{s-1}) = n)nf_{L(s)|L(s-1)}(\ell_s|\ell_{s-1})$  comprend la distribution conditionnelle  $f_{L(s)|L(s-1)}(\ell_s|\ell_{s-1})$  qu'une défaillance soit de charge  $L(s) = \ell_s$  à la génération s sachant qu'elle est engendrée par une défaillance de charge  $L(s-1) = \ell_{s-1}$ . Cette distribution est la distribution conditionnelle (VII.11), qui ne dépend pas de la génération dans le modèle H. La valeur moyenne du nombre de défaillances (VII.9) engendrées par une défaillance est aussi indépendante de la génération considérée, et ne dépend que de la charge de la défaillance-mère. Le facteur de normalisation se note :

$$\bar{\mu}(s-1) = \int d\ell_{s-1} f_{L(s-1)}(\ell_{s-1}) \mu(\ell_{s-1}), \qquad (\text{VII.149})$$

et dépend donc aussi de la distribution des charges défaillantes de la génération précédente. La distribution de la charge d'un élément défaillant à l'étape s est donc :

$$f_{L(s)}(\ell_s) = \frac{1}{\bar{\mu}(s-1)\beta} \int d\ell f_{L(s-1)}(\ell) \Theta(\ell_s - \beta \ell) \Theta\left(\beta \ell + \min\left(1, \frac{\beta \ell}{\alpha}\right) - \ell_s\right).$$
(VII.150)

On représente, sur la figure VII.16, l'évolution des distributions des charges défaillantes pour un paramètre  $\alpha$  donné, selon que le paramètre  $\beta$  prend des valeurs telles que le système est endessous ou au-dessus de la transition de rupture. On vérifie sur la figure VII.18 que pour un système au-dessus de la transition  $\beta_{\star}$ , la distribution des charges converge vers une distribution de charges de valeurs finies. Pour un système sous la transition  $\beta_{\star}$ , la distribution de charges tend vers la fonction delta  $\delta(\ell)$ . On verra dans la section suivante que cette remarque est liée à la convergence, pour  $\beta > \beta_{\star}$ , du nombre moyen de descendants par individu d'une génération donnée *s* vers une valeur asymptotique non nulle :  $\bar{\mu} > 0$ .



(b) Au-dessus de la transition

FIGURE VII.16 – Distribution de la charge des défaillances pour quelques unes des premières générations. La valeur de  $\alpha$  est fixée pour les deux figures, et  $\beta$  est choisi tel que le système est sous la transition (a) ou au-dessus (b).



(b) À la transition  $\beta_c$ 

FIGURE VII.17 – Distribution de la charge des défaillances pour quelques unes des premières générations. La valeur de  $\alpha$  est fixée pour les deux figures, et  $\beta$  est choisi tel que le système est à la transition  $\beta_{\star} = \alpha/(1+\alpha)$  (b) ou au point critique  $\beta_c \approx 0,1792$  (b).



FIGURE VII.18 – Distribution de la charge des défaillances pour la cinquième génération, pour une valeur fixée de  $\alpha$  et des valeurs de  $\beta$ , telles que le système est sous, au-dessus et à la transition.



FIGURE VII.19 – Distribution asymptotique de la charge défaillante au-dessus du seuil de transition  $\beta_{\star} = \alpha/(1+\alpha)$  pour  $\alpha = 0.6$  et  $\alpha = 0.2$ .

### VII.8.3 Distribution limite

La distribution des charges défaillantes converge vers une distribution limite  $f_L$  pour  $s \to \infty$ dont la forme dépend de la phase du système. Cette distribution limite est solution de l'équation intégrale :

$$f_L(\ell) \int d\ell' f_L(\ell') \mu(\ell') = \frac{1}{\beta} \int d\ell' f_L(\ell') \Theta(\ell - \beta \ell') \Theta\left(\beta \ell' + \min\left(1, \frac{\beta \ell'}{\alpha}\right) - \ell\right). \quad (\text{VII.151})$$

Pour des valeurs des paramètres  $\alpha$  et  $\beta$  telles que la borne supérieure de la charge défaillante à toute étape est  $\ell_{max} = 1$ , la distribution limite est une fonction  $\delta : f_L(\ell) = \delta(\ell)$ ,  $\forall \ell$ . Si  $\beta > \beta_\star = \alpha/(1+\alpha)$ , la distribution converge vers une courbe non triviale. Juste au-dessus du point  $\beta_\star = \alpha/(1+\alpha)$ , la courbe présente un maximum proche de l'axe des ordonnées et décroit ensuite exponentiellement vers 0, comme on peut l'observer sur la figure VII.19. Ces courbes ont été obtenues pour 4000 itérations de l'équation VII.151, pour des valeurs de  $\beta$  proches du point critique  $\beta_\star$ . Lorsqu'on approche de la valeur de transition, la distribution converge infiniment lentement vers une distribution  $\delta$ .

#### VII.8.4 Exemple de cas simples

#### a Cas $\alpha = 0$

L'étude de la distribution des charges défaillantes ne présente pas d'intérêt dans ce cas, car le processus est indépendant de la valeur des charges défaillantes.

b Cas  $\beta = 1$ 

Les charges défaillantes, pour  $\beta = 1$ , sont définies à une génération s donnée sur l'intervalle (0, s + 1). Leur distribution converge très vite vers une distribution normale, qui émerge de leur construction en somme de variables aléatoires identiquement distribuées. En effet, après les premières générations, presque toutes les charges défaillantes sont grandes devant  $\alpha$ , et tout élément surchargé défaille. La distribution des charges défaillantes de la génération s converge vers :

$$f_{L(s)}(\ell) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\ell-m)^2}{2\sigma^2}},$$
 (VII.152)

où m = (s+1)/2 et  $\sigma^2 = \sqrt{(s+1)/2}$ . On observe bien sur la figure VII.20 un comportement en loi normale, et en redimensionnant les distributions selon m(s) et  $\sigma(s)$ , on vérifie sur la figure VII.21 que la distribution limite redimensionnée des charges défaillantes est la distribution normale.

L'itération du calcul de la distribution des charges défaillantes d'une génération s à la génération s + 1 confirme l'estimation d'une distribution normale, où pour  $s \gg 1$ , on considère que  $\min(1, \ell/\alpha) = 1$ , car le poids de la distribution des charges est concentré en  $\ell > 1$ :

$$f_{L(s+1)}(\ell) = \int_{0}^{s+1} \mathrm{d}x \frac{1}{\sqrt{2\pi(s/2)(1/2)}} e^{-\frac{(x-s/2)^{2}}{2\sqrt{s/2}}} \Theta(\ell-x)\Theta(x+1-\ell),$$

$$= \frac{1}{2} \left( \mathrm{erf}\left(\frac{2+s-2\ell}{2(2s)^{1/4}}\right) - \mathrm{erf}\left(\frac{s-2\ell}{2(2s)^{1/4}}\right) \right) \text{ si } 1 < \ell < s.$$
(VII.153)

et pour  $\ell < 1$  et  $s < \ell < s + 1$ ,  $f_{L(s+1)}(\ell) \approx 0$ . On vérifie que :

$$\frac{1}{2} \left( \operatorname{erf} \left( \frac{2+s-2\ell}{2(2s)^{1/4}} \right) - \operatorname{erf} \left( \frac{s-2\ell}{2(2s)^{1/4}} \right) \right) =$$

$$\frac{1}{\sqrt{2\pi} ((s+1)/2)^{1/4}} \exp \left( -\frac{(\ell-(s+1)/2)^2}{2\sqrt{(s+1)/2}} \right) = f_{L(s+1)}(\ell).$$
(VII.154)

Et la distribution des charges défaillantes dans ce cas simple converge bien vers une distribution normale.

## VII.9 Nombre typique de descendants par individu

### VII.9.1 Processus de Galton-Watson

Dans un processus de branchement simple, tous les éléments composant la famille sont identiques et, en particulier, leur nombre de descendants directs est identiquement distribué. On



FIGURE VII.20 – Distribution de la charge défaillante pour  $\beta=1,\,\alpha=0.2.$ 



FIGURE VII.21 – Distribution redimensionnée de la charge défaillante pour  $\beta = 1, \alpha = 0.2$ .

note G(t) la fonction génératrice de la distribution du nombre de descendants directs. On suppose que le premier moment de cette distribution existe, et il est noté  $\mu = G'(1)$ . La probabilité d'extinction d d'une famille définie par cette distribution est la probabilité qu'il existe une génération finie s pour laquelle le nombre d'individus  $N_s$  appartenant à cette génération soit égal à 0. La fonction génératrice  $F_s(t)$  du nombre d'individus de la s-ième génération est la composée s fois de la fonction génératrice du nombre de descendants directs, et on a en particulier :

$$F_{s+1}(t) = G(F_s(t)).$$
 (VII.155)

La probabilité d'extinction de la famille est donnée par :

$$d = \lim_{s \to \infty} F_s(0), \tag{VII.156}$$

ce qui est aussi valable pour :

$$d = \lim_{s \to \infty} F_{s+1}(0) = \lim_{s \to \infty} G(F_s(0)).$$
 (VII.157)

La probabilité d'extinction vérifie donc l'équation :

$$d = G(d). \tag{VII.158}$$

La probabilité d'extinction d prend par définition des valeurs dans (0,1), et G est une fonction croissante et strictement convexe sur cet intervalle.

Si  $\mu \leq 1$ , on a  $G'(1) \leq 1$ , et G étant strictement convexe, G'(x) < 1 pour  $0 \leq x < 1$ , et donc G(b) - G(a) < b - a pour a < b dans (0, 1). On pose b = 1, et on a donc :

$$G(t) > x, \text{ si } 0 \le x < 1,$$
 (VII.159)

et donc on doit avoir d = 1.

Si  $\mu > 1$ , il existe une solution dans (0,1), et la probabilité d'extinction prend une valeur finie d < 1.

On distingue donc trois régimes selon la valeur de  $\mu$  :

- Si  $\mu < 1$ , la probabilité d'extinction de la cascade est d = 1, et la probabilité de rupture  $P_b = 0$ .
- Si  $\mu > 1$ , la probabilité d'extinction de la cascade prend une valeur finie d < 1, et la probabilité de rupture est non nulle.
- Si  $\mu = 1$ , la probabilité d'extinction de la cascade converge lentement vers d = 1, et in fine la probabilité de rupture  $P_b$  est nulle.

Le nombre moyen de descendants par individu est donc lié à la probabilité d'extinction de la famille, et on cherche à généraliser cette propriété pour le processus de branchement généralisé.

#### VII.9.2 Processus de branchement généralisé

Dans le processus de branchement généralisé, la distribution du nombre de descendants par individu est une loi de Poisson de paramètre  $\mu(L)$ . Ce paramètre ne dépend que de la charge Lde l'individu défaillant. La distribution  $f_{L(s)}$  de la charge défaillante est fonction de la génération s de la défaillance. Le nombre typique de descendants d'une défaillance prise au hasard dans la s-ième génération est donc donné par :

$$\bar{\mu}(s) = \int \mathrm{d}\ell f_{L(s)}(\ell)\mu(\ell) \tag{VII.160}$$

La distribution de la charge défaillante  $f_{L(s)}$  converge vers une distribution limite  $f_L$  pour  $s \gg 1$ , et le nombre typique de descendants d'un individu pris au hasard dans une génération quelconque  $s \gg 1$  converge donc vers une valeur asymptotique :

$$\bar{\mu} = \int \mathrm{d}\ell f_L(\ell)\mu(\ell). \tag{VII.161}$$

Dans notre modèle, la valeur asymptotique  $\bar{\mu}$  distingue quatre régimes différents.

- Pour  $\bar{\mu} = 0$ , *i.e*  $\beta < \beta_{\star}$ , la probabilité de rupture est nulle :  $P_b = 0$ .
- Pour 0 < μ
   <li>1, *i.e* β<sub>\*</sub> < β < β<sub>c</sub>, le nombre total de défaillances est distribué selon une loi de puissance pour les petites valeurs de M, puis la distribution décroît exponentiellement vite vers zéro. La taille de la cascade M peut donc prendre de grandes valeurs, mais reste toujours finie, et la probabilité de rupture est nulle.
- Pour  $\bar{\mu} = 1$ , *i.e*  $\beta = \beta_c$ , la taille de la cascade est distribuée selon une loi de puissance sur les valeurs de  $M \ll \infty$ . La probabilité de rupture subit pour cette valeur de  $\beta$  une transition, d'identiquement nulle à une valeur finie.
- Pour  $\bar{\mu} > 1$ , *i.e*  $\beta > \beta_c$ , la taille de la cascade est distribuée selon une loi de Borel-Tanner dont le paramètre est supérieur à 1. La cascade a donc une probabilité non nulle d'avoir une taille infinie, c'est-à-dire que la probabilité de rupture est finie.

On représente ces différents cas sur la figure VII.22.

On rappelle que l'étude analytique de la transition  $\beta_{\star}$  donne la formule suivante :

$$\beta_{\star}(\alpha) = \frac{\alpha}{\alpha+1}.$$
 (VII.162)

Pour ce jeu de paramètres à la transition, le nombre typique de descendants décroît vers 0 selon une loi proche de la loi de puissance avec le temps, et on observe sur la figure VII.23 que cette valeur de transition sépare le régime dans lequels le nombre typique de descendants d'une défaillance à la génération s décroit exponentiellement vite, et le régime dans lequel  $\bar{\mu}(s)$  converge vers une valeur asymptotique finie.

#### VII.9.3 Cas simples

a Cas 
$$\alpha = 0$$

Dans ce cas, le nombre moyen de descendants par individu ne dépend pas de la charge de la défaillance-mère, et la distribution des charges étant normalisée, le nombre de descendants typique des défaillances d'une génération quelconque s est donné par :

$$\bar{\mu}(s) = \frac{1}{\beta}.$$
 (VII.163)

On a vu que dans ce cas, la probabilité de rupture n'est nulle que pour  $\beta = 1$ , soit  $\bar{\mu}(s) = \bar{\mu} = 1$ , ce qui confirme l'analyse précédente.



FIGURE VII.22 – Nombre moyen de descendants par individu défaillant à la génération s en fonction de la génération. On représente cette quantité pour différentes valeurs du paramètre  $\beta$  pour un  $\alpha$  donné.



FIGURE VII.23 – Graphe log-log du nombre typique de descendants par défaillance en fonction de l'étape, pour des valeurs des paramètres en-dessous, au-dessus et très proche de la transition  $\beta \gtrsim \beta \star = \alpha/(1+\alpha)$ . La droite a pour équation :  $f(x) = 4,289(\pm 0,008)x^{-0.9224(\pm 0,0006)}$ .

#### b Cas $\beta = 1$

On a vu lors de l'étude de la distribution des charges défaillantes que le poids de cette distribution se déplace très vite vers de grandes valeurs de  $\ell$ . Pour  $s \gg 1$ , on a donc en bonne approximation :

$$\bar{\mu}(s) = \int_0^{s+1} \mathrm{d}\ell \min\left(1, \frac{\ell}{\alpha}\right) \frac{1}{\sqrt{2\pi}(s/2)^{1/4}} \mathrm{e}^{-\frac{(\ell-s/2)^2}{2\sqrt{s/2}}} \approx 1.$$
(VII.164)

On atteint dans ce cas particulier très vite le régime stationnaire de la cascade, qui se comporte alors comme un processus de branchement simple avec  $\mu = 1$ . On retrouve en effet bien une distribution en loi de puissance d'exposant 3/2 pour la taille de la cascade.

## VII.10 Conclusion

Nous avons présenté dans ce chapitre un modèle de branchement généralisé dont les propriétés de passage d'une génération à la suivante imitent le modèle de cascade de défaillances dans un réseau électrique de Lehmann et Bernasconi [74]. Nous avons montré que la dynamique de ce modèle dépend du régime considéré, défini par le jeu de paramètres ( $\alpha, \beta$ ).

- Pour des valeurs du paramètre β inférieures à la valeur-seuil β<sub>\*</sub>(α), le système résiste aux cascades globales. La taille des cascades de ce régime est distribuée selon une loi exponentielle et au bout d'un nombre fini de générations, la cascade s'éteint avec probabilité 1. La distribution de la génération d'arrêt de la cascade est centrée sur les premières valeurs de s. La probabilité de rupture d'un système avec ce jeu de paramètres est nulle, et la distribution des charges défaillantes d'une génération s converge très vite vers un pic δ.
- Pour β = β<sub>\*</sub>(α), la distribution de la taille des cascades décroît vers 0 comme une loi de puissance avec une queue de décroissance exponentielle. Cette valeur correspond donc à la transition entre un régime dans lequel les cascades touchent un petit nombre d'éléments et un régime dans lequel la taille des cascades peut être grande devant 1. À cette valeur précise, la distribution des charges défaillantes converge infiniment lentement vers une fonction δ de Dirac.
- Pour  $\beta = \beta_c(\alpha)$ , la distribution de la taille des cascades est distribuée selon une loi de puissance. En effet, pour ces valeurs des paramètres, la probabilité que la cascade soit globale devient non nulle. La distribution de la charge des défaillances d'une génération converge très vite vers une distribution limite, à partir de laquelle on calcule le nombre moyen de descendants par individu. À  $\beta_c$  ce nombre converge vers une valeur asymptotique  $\bar{\mu} = 1$ . La distribution de la taille de la cascade en loi de puissance s'explique par analogie avec les processus de branchement simples, pour lesquels  $\mu = 1$  correspond au régime critique.
- Pour  $\beta_{\star}(\alpha) < \beta < \beta_c(\alpha)$ , la distribution du nombre total de défaillances converge petit à petit vers la loi de puissance de  $\beta_c$ . La probabilité de rupture est donc nulle, et la

probabilité d'extinction converge vers 1, cependant de grandes cascades peuvent toucher un nombre d'éléments grand devant 1. La distribution de la charge défaillante converge vers une distribution limite. La moyenne de cette distribution est inférieure à 1, et la valeur asymptotique du nombre moyen de descendants par individu d'une génération donnée converge vers une valeur finie  $0 < \bar{\mu} < 1$ .

- Pour  $\beta_c(\alpha) < \beta < 1$ , la probabilité de rupture prend une valeur finie qui est donnée par la partie défective de la normalisation du nombre total de défaillances, distribué selon une loi proche de la loi de Borel-Tanner de paramètre  $\bar{\mu} > 1$ . La distribution des charges défaillantes converge vers une loi limite non triviale.
- Pour  $\beta = 1$ , quelque soit la valeur de  $\alpha$ , le nombre total de défaillances est distribué selon une loi de puissance d'exposant 3/2. Cette distribution n'a pas de composante à l'infini, et la probabilité de rupture est donc nulle. La distribution des charges défaillantes est la distribution d'une somme de variables aléatoires identiquement distribuées, et converge donc vers une loi normale de moyenne (s+1)/2 et de variance  $\sqrt{(s+1)/2}$ . Le nombre moyen de défaillances par individu d'une génération donnée converge en quelques générations vers  $\bar{\mu} = 1$ .

On trouve dans ce modèle simple un recouvrement entre un comportement en loi de puissance et un comportement bimodal de la propagation de cascades, selon la valeur des paramètres. Watts a souligné dans une étude sur la propagation de cascades [20] que ces deux types de comportements correspondent aux propriétés de rareté de cascades globales. On peut remarquer ici que la transition de droite  $\beta = 1$  apparaît à cause de la contrainte topologique imposée par le nombre moyen de descendants  $\bar{\mu} = 1$ , tandis que celle de gauche  $\beta_c$  marque l'apparition de grandes cascades dues à la vulnérabilité locale des éléments. On a pu, dans le modèle H, déterminer les paramètres qui influent sur ces deux types de transition, et présenter donc les justifications analytiques de chacune de ces transitions. Partie D Conclusion

# CHAPITRE VIII Résultats et perspectives

Cette thèse a porté sur l'étude des réseaux hors d'équilibre selon deux axes différents. Dans la première partie, nous nous sommes intéressés aux conséquences sur la distribution du degré et sur la statistique des extrêmes de la taille finie des réseaux. Dans la seconde partie nous avons étudié quels sont les mécanismes de propagation d'une défaillance sur un réseau.

## VIII.1 Effets de taille finie

Nous nous sommes focalisés sur l'étude de trois modèles simples et stochastiques de réseaux en croissance :

- Le modèle d'attachement uniforme (AU), dont les règles de construction sont les plus simples possibles pour un réseau en croissance : le nouveau noeud qui arrive à chaque étape choisit au hasard un des noeuds présents et s'y connecte.
- Le modèle de Barbási-Albert (BA), qui est le réseau en croissance invariant d'échelle le plus simple : la probabilité pour un noeud de recevoir une nouvelle connexion à chaque étape ne dépend que du nombre de connexions qu'il possède déjà.
- Le modèle d'attachement préférentiel généralisé (APG), qui est le réseau le plus général pour lequel la fonction de partition ne dépend pas de l'histoire du réseau, et qui permet une interpolation entre le modèle uniforme et les modèles préférentiels.

Pour chacun de ces réseaux, nous avons établi la dépendance en la taille n de la distribution de leurs degrés. Nous avons montré que pour chacun d'entre eux, il existe une valeur de transition  $k_{\star}(n)$  qui distingue les régimes stationnaires et de grande déviation de la distribution du degré. Cette valeur de transition prend les valeurs suivantes :

- Dans le modèle AU :  $k_{\star} \sim \ln n$ ,
- Dans le modèle BA :  $k_{\star} \sim n^{1/2}$ ,
- Dans le modèle APG :  $k_{\star} \sim n^{1/(c+2)}$ .

À temps fini, la distribution du degré est donnée par la distribution stationnaire, corrigée par une fonction d'échelle multiplicative  $\Phi$ , qui dépend des conditions initiales considérées. On a en effet distingué dans notre étude deux conditions initiales qui diffèrent par la symétrie du degré des premiers noeuds. La différence entre ces deux symétries apparaît clairement dans la fonction d'échelle des modèles d'attachement préférentiels. On reproduit sur la figure VIII.1 les fonctions d'échelle de la distribution du degré pour les trois modèles, à chaque fois pour les deux choix



FIGURE VIII.1 – Fonctions d'échelle de temps fini pour les trois modèles considérés, en fonction du degré redimensionné par la valeur de transition  $k_{\star}$ .

de conditions initiales. Ces fonctions d'échelles précisent l'effet de taille finie sur la distribution du degré. Pour  $k < k_{\star}$ , la fonction d'échelle est proche de 1, et les noeuds de ce degré ne sentent pas les effets de taille finie du réseau et les degrés sont distribués selon la distribution stationnaire. Pour  $k > k_{\star}$ , la fonction d'échelle est proche de 0. La probabilité de trouver un noeud de cet intervalle de degré est très faible, et donnée par la fonction de grande déviation, calculée pour chaque modèle. Pour  $k = k_{\star}$ , la distribution du degré décroît exponentiellement vite. On retrouve la coupure de la queue de la distribution observée dans les réseaux réels.

Nous avons montré par la suite que le degré maximum d'un réseau de taille n est typiquement de l'ordre de ce degré de coupure. Pour chaque modèle, nous avons étudié des quantités typiques de la statistique des extrêmes, et explicité la dynamique en jeu à la tête du réseau. Dans le modèle AU, le meneur à changé d'identité  $\ln n$  fois au cours de l'histoire du réseau. Le noeud qui devient le nouveau meneur n'est pas forcément l'un des premiers noeuds qui sont entrés dans le système, car l'indice du meneur croît continuement au cours de l'histoire du réseau. Le nouveau meneur est même avec une probabilité de l'ordre de 0.65 un noeud qui n'a jamais été meneur dans toute l'histoire du réseau. La probabilité de persistance du premier meneur converge donc asymptotiquement vers 0, selon une loi inspirée de la dynamique d'un marcheur aléatoire avec des murs absorbants croissants.

Dans les modèles d'attachement préférentiel, l'histoire du réseau connaît autant de changements de meneur que dans le modèle AU, cependant la prépondérance sur la connectivité du réseau se partage entre un petit nombre de noeuds qui monopolisent les nouveaux noeuds entrants. En effet, dans ces modèles, la probabilité pour les noeuds de plus grand degré d'attirer une nouvelle connexion est bien plus grande que pour les autres noeuds, et eux seuls croissent. Dans ces modèles, la probabilité de survie du premier meneur converge vers une valeur asymptotique finie.

Pour illustrer la statistique des extrêmes dans les réseaux en croissance, on représente sur la figure VIII.2 les valeurs asymptotiques atteintes par les différentes quantités observées en fonction du paramètre  $\nu$  du modèle d'attachement préférentiel généralisé, qui permettent de retrouver le comportement du modèle BA pour  $\nu = 1/2$  et celui du modèle AU pour  $\nu \to 0$ .



(a) Indice moyen asymptotique (b) Nombre moyen asymptotique (c) Probabilité de survie asympdu meneur de meneurs distincts totique du premier meneur

FIGURE VIII.2 – Valeurs asymptotique des quantités d'extrêmes dans le modèle APG en fonction du paramètre  $\nu$ .

## VIII.2 Propagation de défaillance

Dans un second temps, nous avons voulu comprendre comment des réseaux construits selon les règles de la partie précédente se comportent si l'un des éléments vient à défaillir. Nous avons étudié un modèle très simplifié de propagation de défaillances pour en tirer des évaluations analytiques et en extraire les fondementaux du développement d'une cascade dans un système complexe.

Nous avons prolongé le raisonnement de Lehmann et Bernasconi, et montré qu'une cascade de défaillance peut s'approximer comme un processus de branchement généralisé dans lequel les éléments sont les défaillances, regroupées en arbres hiérarchiques. Nous avons vérifié l'existence de deux valeurs de transition pour les paramètres du problème. Ces deux valeurs  $\beta_c$  et  $\beta = 1$  présentent une transition différente vers un domaine de l'espace des paramètres pour lequel des cascades globales peuvent apparaître avec une probabilité non nulle.

La transition  $\beta = 1$  correspond à une transition de connectivité : au-delà de  $\beta = 1$ , la taille du réseau connecté est petite par rapport à la taille du système, et les cascades ne touchent qu'une fraction finie de la population de noeuds. La taille des cascades dans ce régime est distribuée en loi de puissance.

La transition  $\beta_c$  correspond à la limite de la résistance locale de chaque noeud. La distribution du nombre total de défaillances dans la cascade décroit exponentiellement vite vers 0, avec une composante renvoyée à l'infinie, qui correspond à de rares cascades globales. Un système dans ce régime semblera en général très stable, ne présentant que quelques cascades de faible taille, avant qu'un choc initial indiscernable des précédents conduise le système à un effondrement total. On parle alors de comportement robuste mais fragile.

Une variante possible de ce modèle est de le contraindre spatialement (voir appendice C). Dans un tel modèle, les défaillances ne se propagent que de proches en proches, et on pourrait s'intéresser à l'influence d'une telle contrainte sur les propriétés des transitions. Une autre perspective consiste à modéliser le réseau de défaillances comme un réseau invariant d'échelle, et de vérifier la validité des résultats précédents. Appendices
## Annexe **A** Équations intégrales

## A.1 Relation de récurrence dans le processus de branchement généralisé

#### A 1.1 Distribution de points et intégrales aléatoires

Un ensemble fini de variables aléatoires définies sur X peut être représenté comme une distribution ponctuelle  $\omega$ , *i.e.* une combinaison finie de  $x_i \in X$  et d'entiers  $n_i$ . Le nombre d'éléments de type  $x_i$  (ou dans  $[x_i, x_i + dx]$ ) est  $n_i$ , et la distribution ponctuelle s'écrit :

$$\omega = (x_1, n_1; \dots; x_K, n_K), \tag{A.1}$$

où le nombre de types représentés K est aussi une variable aléatoire. L'ensemble des combinaisons de  $x_i, n_i$  est donc un ensemble aléatoire dans lequel chaque distribution de points  $\omega$  a une mesure de probabilité  $P(\omega)$ .

Une intégrale aléatoire  $I_h$  est une variable aléatoire réelle déterminée à partir d'une fonction h définie positive appliquée à une distribution ponctuelle aléatoire. Elle est donnée par :

$$I_h(\omega) = \int_X h(x)\rho(x)dx = \sum_i n_i h(x_i), \qquad (A.2)$$

où  $\rho(x)$  est la densité de masse de la distribution ponctuelle  $\omega$ .

On définit la fonction  $\chi$  caractéristique de l'ensemble X par :

$$\chi(x) = \begin{cases} 1 & \text{si } x \in X \\ 0 & \text{sinon} \end{cases}$$
(A.3)

Si la distribution ponctuelle considérée représente les défaillances d'une étape quelconque, l'intégrale aléatoire appliquée à  $\chi$  est la variable aléatoire du nombre de défaillances  $N_s$  de cette génération. La fonction génératrice des moments  $\phi$  de cette variable aléatoire est définie par :

$$\phi(t) = \left\langle e^{-tI_{\chi}(\omega)} \right\rangle. \tag{A.4}$$

On appelle fonctionnelle génératrice des moments (FGM) d'une distribution ponctuelle l'expression généralisée pour une fonction quelconque h:

$$\Phi[h] = \left\langle e^{-I_h(\omega)} \right\rangle, \tag{A.5}$$

et on retrouve la fonction génératrice des moments du nombre de défaillances d'une étape par :

$$\phi(t) = \Phi[t\chi]. \tag{A.6}$$

La moyenne est prise sur l'espace de probabilité de  $I_h$ , donc sur l'espace de probabilité des distributions de points considérées.

Si une distribution ponctuelle  $\omega$  est la somme d'un nombre fini de distributions ponctuelles indépendantes  $\omega_1, \ldots, \omega_k$ , et que chaque distribution de point  $\omega_i$  a pour fonctionnelle génératrice  $\Phi_i$ , alors la fonctionnelle génératrice des moments  $\Phi$  de la distribution  $\omega$  est donnée par :

$$\Phi[h] = \Phi_1[h]\Phi_2[h]\dots\Phi_k[h].$$
(A.7)

#### A.1.2 Relation de récurrence sur les fonctionnelles génératrices des moments

On considère une suite de distributions ponctuelles  $\omega_0, \omega_1, \ldots$  représentant les membres des générations d'une famille, issue d'une unique défaillance de paramètre  $L = \lambda_0$  à l'étape 0 :  $\omega_0 = (\lambda_0, 1)$ . La distribution des types des membres d'une génération ne dépend que de la distribution des types de la génération immédiatemment précédente. Chaque individu de type  $\lambda$  engendre à l'étape suivante une distribution de points  $\omega_{\lambda}^{(1)}$  qui ne dépend que de la valeur du paramètre  $\lambda$ , *i.e.* les branches sont indépendantes.

La probabilité qu'une distribution ponctuelle  $\omega_s$  engendre une distribution ponctuelle  $\omega_{s+1}$  à la génération suivante est notée  $P^{(1)}(\omega_s, \omega_{s+1})$ . La probabilité qu'un individu de type  $\lambda$  engendre une distribution de points  $\omega$  est donnée par :

$$p_{\lambda}(\omega) = P^{(1)}\left((\lambda, 1), \omega\right), \qquad (A.8)$$

et la FGM de cette distribution ponctuelle, notée  $\omega_{\lambda}^{(1)}$ , est donnée par :

$$\Phi_{\lambda}^{(1)}[h] = \left\langle \mathrm{e}^{I_h(\omega_{\lambda}^{(1)})} \right\rangle = \int_{\Omega} \mathrm{e}^{I_h(\omega_{\lambda}^{(1)})} \mathrm{d}P^{(1)}\left((\lambda, 1), \omega_{\lambda}^{(1)}\right), \tag{A.9}$$

où  $\Omega$  est l'ensemble des distributions de points.

Si la distribution ponctuelle représentant la *s*-ième génération est donnée par  $\omega_s = (\lambda_1, n_1; \ldots; \lambda_k, n_k)$ , la distribution de points de la s + 1-ième génération est la somme des distributions ponctuelles  $\omega_{\lambda_1}^{(1)}, \ldots, \omega_{\lambda_k}^{(1)}$  issues des membres de la génération *s*. Ces distributions de points étant indépendantes, la FGM  $\Phi_{\omega_{s+1}}$  de la distribution de la génération s + 1 est donnée par :

$$\Phi_{\omega_{s+1}}[h] = \left(\Phi_{\lambda_1}^{(1)}[h]\right)^{n_1} \dots \left(\Phi_{\lambda_k}^{(1)}[h]\right)^{n_k}.$$
(A.10)

Si l'on prend le logarithme de cette fonctionnelle, on retrouve l'intégrale aléatoire appliquée à  $\ln \left( \Phi^{(1)}_{\cdot}[h] \right)$  pour la distribution ponctuelle  $\omega_s$ :

$$\Phi_{\omega_{s+1}}[h] = \exp\left\{\int_X \ln\left(\Phi_{\lambda}^{(1)}[h]\right) \mathrm{d}\omega_s(\lambda)\right\}.$$
(A.11)

Lorsque l'indice d'une FGM est une distribution ponctuelle  $\omega$ , il est sous-entendu qu'on ne moyenne pas sur l'ensemble des distributions ponctuelles, mais qu'on se limite à la seule distribution ponctuelle  $\omega$ . On peut donc réécrire l'équation précédente sous la forme :

$$\Phi_{\omega_{s+1}}[h] = \Phi_{\omega_s} \left[ -\ln\left(\Phi_{\cdot}^{(1)}[h]\right) \right].$$
(A.12)

Le point figure le type des membres de la distribution  $\omega_s$ , variable muette dans cette notation.

Si on itère cette relation de récurrence pour un processus initié par un unique ancêtre de de type  $\lambda_0$ , on retrouve au bout de s étapes :

$$\Phi_{\omega_{s+1}(\lambda_0)}[h] = \Phi_{\omega_0}^{(1)} \underbrace{\left[ -\ln\left(\Phi_{\cdot}^{(1)}\left[ -\ln\left(\Phi_{\cdot}^{(1)}[\dots]\right)\right] \right) \right]}_{s \text{ fois}}.$$
(A.13)

Les termes entre crochets représentent une succession d'individus sur s générations, initiée par chaque membre la la première génération. On peut donc réécrire cette équation comme :

$$\Phi_{\omega_{s+1}(\lambda_0)}[h] = \Phi_{\lambda_0}^{(1)} \left[ -\ln \Phi_{\omega_s(.)}[h] \right]$$
(A.14)

La notation  $\Phi_{\lambda}^{(1)}$  sous-entend la moyenne (A.9) sur l'ensemble des distributions qui peuvent être engendrées par un individu de type  $\lambda$ . On cherche donc la probabilité de la distribution ponctuelle aléatoire  $\omega$  après une défaillance de charge  $\lambda : p_{\lambda}(\omega)$ . Cette probabilité dépend des caractéristiques du processus considéré.

#### A.2 Modèle de cascade de Harris

#### A.2.1 Relation de récurrence pour la FGM

Le processus de branchement s'applique à la famille formée par l'ensemble des défaillances dans un système. Un ensemble de charges défaillantes peut être représenté comme une distribution ponctuelle  $\omega$ . Les charges défaillantes prennent une valeur réelle positive distribuée aléatoirement, *i.e.* l'espace de définition des types  $\lambda_i$  de la distribution des charges est  $X = \mathbb{R}^+$ . Les règles de passage d'une défaillance à la distribution de charge qu'elle engendre à la génération suivantes sont définies comme suit.

Le nombre de descendants d'une charge  $L = \lambda$  est tiré d'une loi de Poisson de paramètre  $\mu(\lambda)$ , la valeur moyenne du nombre de défaillances :

$$P(N(\lambda) = k) = \frac{\mu(\lambda)^k}{k!} e^{-\mu(\lambda)}, \qquad (A.15)$$

où le nombre moyen de défaillances  $\mu(\lambda)$  induites par la défaillance d'une charge  $L = \lambda$  à l'étape précédente ne dépend que de la valeur de la charge  $\lambda$ . La fonction génératrice du nombre  $N(\lambda)$ de descendants directs d'une défaillance de charge  $\lambda$  est :

$$F^{(\lambda)}(t) = e^{\mu(\lambda)(t-1)}.$$
(A.16)

La charge des défaillances-filles L dépend de la charge de la défaillance L à l'étape précédente. On note  $f_{L_i|L}$  la densité de probabilité de la charge d'une descendante L si la défaillance qui l'engendre est L.

Avec ces définitions, la densité de probabilité jointe qu'un individu de charge  $L = \lambda$  engendre N descendants de charges  $L_1, L_2, \ldots, L_N$  est  $f_{N;L'_1,\ldots,L_N|L}$ . La moyenne dans la définition de la FGM de la distribution ponctuelle issue d'une défaillance de charge  $\lambda$  (A.9) correspond à l'intégration sur cette densité de probabilité jointe. Comme les charges issues de la défaillance de charge  $L = \lambda$  sont des variables aléatoires indépendantes, la densité de probabilité jointe s'écrit encore :

$$f_{N;L_1,...,L_N|L}(n;y_1,y_2,...|\lambda) = P(N(\lambda) = n) \left( f_{L_i|L}(y|\lambda) \right)^n.$$
(A.17)

La FGM de la distribution de charges issue de la défaillance de charge  $L = \lambda$  s'obtient par la moyenne de l'intégrale aléatoire  $I_h$  sur l'ensemble des  $\omega_{\lambda}^{(1)}$  (A.9), soit :

$$\Phi_{\lambda}^{(1)}[h] = \sum_{n} \int dy_1 \dots dy_n f_{N;L_1,\dots,L_N|L}(n;y_1,\dots,y_n|\lambda) e^{I_h(\omega_{\lambda}^{(1)})}.$$
 (A.18)

On peut décomposer l'intégrale aléatoire en une somme de n termes :

$$I_h(\omega_\lambda^{(1)}) = \sum_{i=0}^n h(\lambda_i), \tag{A.19}$$

où les  $\{\lambda_i\}$  sont les types de chaque individu de la distribution ponctuelle issue de  $\lambda$ . De plus, la distribution de probabilité (A.17) est la densité de probabilité jointe de variables aléatoires indépendantes, et on a donc :

$$\Phi_{\lambda}^{(1)}[h] = \sum_{n=0}^{\infty} P(N(\lambda) = n) \left( \int f_{L_i|L}(y|\lambda) \mathrm{e}^{-h(y)} \mathrm{d}y \right)^n.$$
(A.20)

On reconnait la fonction génératrice (A.16) du nombre de descendants directs d'une défaillance de charge  $\lambda$  :

$$\Phi_{\lambda}^{(1)}[h] = F^{(\lambda)} \left( \int f_{L_i|L}(y|\lambda) \mathrm{e}^{-h(y)} \mathrm{d}y \right), \tag{A.21}$$

et donc :

$$\Phi_{\lambda}^{(1)}[h] = \exp\left[\mu(\lambda)\left\{\int f_{L_i|L}(y|\lambda)\mathrm{e}^{-h(y)}\mathrm{d}y - 1\right\}\right].$$
(A.22)

On injecte ce résultat dans (A.14) pour obtenir l'équation de récurrence sur la FGM d'une distribution de charges après s + 1 étapes à la suite d'une défaillance de charge  $L = \lambda_0$  par rapport à la FGM de la distribution ponctuelle après s étapes suite aux défaillances engendrées à la première étape par la défaillance de charge  $\lambda_0$ :

$$\Phi_{\omega_{s+1}(\lambda_0)}[h] = \exp\left[\mu(\lambda_0)\left\{\int \mathrm{d}y f_{L_i|L}(y|\lambda_0)\Phi_{\omega_s(y)}[h] - 1\right\}\right].$$
(A.23)

Cette récurrence construit l'arbre généalogique des défaillances "à l'envers", *i.e.* un arbre de s étapes issu d'une unique défaillance peut être considéré comme la branche de défaillances issue d'un descendant direct d'une unique défaillance à l'étape précédente, et donc une partie d'un arbre de s + 1 générations issu d'un unique individu.

#### A.2.2 Probabilité de rupture

Il existe une relation simple (A.6) entre la FGM d'une distribution de charges d'une génération  $\Phi_{\omega_s}$  et la fonction génératrice des moments du nombre de membres d'une génération  $\phi_s$ . Si la distribution de charge est issue d'une unique défaillance  $L = \lambda$  après *s* étapes, on note la FGM  $\Phi_{\omega_s(\lambda)}$  et la fonction génératrice  $\phi_{\lambda}^{(s)}$ . Si on pose  $h = t\chi$ , on obtient alors à partir de (A.23) la relation de récurrence de la fonction génératrice des moments du nombre de défaillances  $N_s(\lambda)$ à l'étape *s*:

$$\phi_{s+1}^{(\lambda)}(t) = \exp\left[\mu(\lambda)\left\{\int \mathrm{d}y f_{L_i|L}(y|\lambda)\phi_s^{(y)}(t) - 1\right\}\right].$$
(A.24)

# ANNEXE **B Notations**

#### B.1 Notations générales

- s : indice de la génération de la cascade.
- $\alpha$  : paramètre de résistance du réseau.
- $\beta$  : paramètre correspondant à la fraction de la charge défaillante redistribuée à un descendant.
- $\beta_c$ : valeur critique du paramètre  $\beta$  pour laquelle le nombre total de défaillances est distribué selon une loi de puissance.
- $\beta_{\star}$ : valeur critique du paramètre  $\beta$  pour laquelle  $P_b = 0$ .
- L(0) : charge de la défaillance initiale.
- L(s) : charge d'une défaillance quelconque de la génération s.
- $f_{L(0)}$ : distribution des charges initiales
- $\lambda$  : valeur prise par la charge de la défaillance initiale.
- $f_{L(s)}$ : distribution des charges des défaillances de la génération s.
- $\ell$  : valeur prise par la charge d'une défaillance quelconque
- $f_{L'|L}$ : distribution de la charge L' d'une défaillance engendrée par une défaillance de charge L.
- $N_0$ : nombre de défaillances initiales, ici  $N_0 = 1$ .
- $N_s$  : nombre de défaillances de la génération s.
- $F_s$ : fonction génératrice du nombre de défaillances de la génération s.
- $N_1(\ell)$  : nombre de défaillances directement engendrées par une défaillance de charge  $\ell$ .
- G : fonction génératrice du nombre de descendants directs.
- $M_0$ : nombre total de défaillances à la fin de la génération initiale, ici  $M_0 = 1$ .
- $M_s$ : nombre total de défaillances à la fin de la génération s.
- $H_s$ : fonction génératrice du nombre total de défaillances à la fin de la génération s.
- M : nombre total de défaillances dans une cascade.
- H : fonction génératrice du nombre total de défaillances dans une cascade.
- M(λ) : nombre total de défaillances dans une cascade initiée par une défaillance de charge λ.
- $P_b$  : probabilité de rupture lors d'une cascade quelconque.
- $P_b(\lambda)$  : probabilité de rupture lors d'une cascade initiée par une défaillance de charge  $\lambda$ .
- $d_s$ : probabilité d'extinction de la cascade à la génération s.

- d : probabilité d'extinction asymptotique de la cascade.
- $P_{nc}$  : probabilité de non-cascade.
- $\tilde{\rho}_{L',N_1|L}$ : densité de probabilité non normalisée du nombre de descendants de charge L' d'une défaillance de charge L.
- $\bar{\rho}_{L'|L}$  : facteur de normalisation de la densité ci-dessus.
- $\bar{\mu}(s)$  : nombre typique de descendants d'une défaillance quelconque de la génération s.
- $s_a$  : génération d'arrêt due à l'extinction de la cascade.

#### B.2 Notations spécifiques au modèle LB

- N : taille du système.
- i = 1, ..., N : indice de l'un des éléments initiaux du système.
- $u_i(0)$  : charge initiale de l'élément *i*.
- $u_i^{max}$  : charge seuil de l'élément i.
- $u_i(s)$  : charge de l'élément i à la génération s
- $\Delta_{i,j}$  : fraction aléatoire de la défaillance j ajoutée à l'élément i.
- p(s) : probabilité de surcharge pour un élément quelconque à la génération s.
- $s_r$  : génération d'arrêt de la cascade due à la défaillance de tous les éléments du système.
- $n_{anc}$  : nombre d'ancêtres d'une défaillance.

## ANNEXE C Cascade avec extension spatiale sur l'axe 1D

On s'intéresse dans ce chapitre à la propagation spatiale d'une cascade de défaillances. On répartit donc les éléments du réseau dans l'espace à une ou deux dimensions. On considère qu'un élément défaillant peut distribuer sa charge dans un rayon limité. On se demande alors quelles sont les conséquences de ce nouveau paramètre sur les propriétés de la cascade de défaillances.

#### C.1 Modèle Lehmann & Bernasconi

On considère pour ce modèle la construction décrite précédemment, section VI.1, pour un réseau composé de N éléments.

1. À la génération s = 0:

On tire indépendamment la charge  $u_i(0)$ , i = 1, ..., N de chaque élément dans la distribution  $f_{L(0)}$ , ainsi que chaque position  $x_i$ , i = 1, ..., N dans la distribution uniforme sur (0, 1). La charge initiale et la position sont tirées une fois et une seule pour toute l'histoire du réseau. La charge peut évoluer au cours de la cascade, tandis que la position de chaque élément reste fixe tout au long de la cascade.

La charge-seuil de chaque élément est déterminée par le paramètre  $\alpha$  et la charge initiale  $u_i(0)$ :

$$u_i^{max} = (1+\alpha)u_i(0). \tag{C.1}$$

Le noeud défaille lorsque la charge  $u_i(s)$  du noeud *i* à l'étape *s* devient supérieure à ce seuil :  $u_i(s) > u_i^{max}$ .

La cascade est initialisée par l'élément 1, de charge  $L(0) = u_1(0)$  tirée dans  $f_{L(0)}$  et placé au milieu de l'intervalle en  $x_1 = 0.5$ , qui défaille.

Le nombre de défaillances  $N_0$  de la génération s = 0 est donc :  $N_0 = 1$ , et le nombre total d'éléments qui ont défailli jusqu'à l'étape s = 0 est :  $M_0 = N_0 = 1$ .

2. À la génération s = 1 :

La portée de la transmission est limitée dans ce modèle, c'est-à-dire que l'élément défaillant ne peut ici redistribuer sa charge qu'aux éléments situés à une distance inférieure à la portée  $r_0$  des défaillances. Chaque élément *i* à portée  $r_0$  de l'élément défaillant 1 reçoit une fraction aléatoire  $\Delta_{i,1}(0)$  de la charge défaillante L(0), et sa charge à la première génération est donc :

$$u_i(1) = u_i(0) + \Delta_{i,1}(0)L(0).$$
(C.2)

On se place dans le cas de conditions aux limites périodiques, c'est-à-dire que les éléments susceptibles de subir une surcharge autour de la défaillance à la position  $x_1$  sont les éléments dont la position x' vérifie :

$$|x' - x_1| < r_0 \text{ ou } |x' - x_1| > 1 - r_0.$$
 (C.3)

Les variables aléatoires  $\Delta_{i,1}(0)$  sont définies de telle façon que, en moyenne, la charge défaillante L(0) est entièrement redistribuée sur les éléments survivants à portée. On fait l'approximation que la densité d'éléments est spatialement constante, et donc qu'en tout point de l'intervalle (0, 1), la densité est égale au nombre  $\nu(1)$  d'éléments survivants à l'étape 1. Pour que la charge de l'élément défaillant soit en moyenne redistribuée aux éléments à portée, :  $2r_0\nu(1) \langle \Delta(0) \rangle = 1$ , soit, pour  $\nu(1) \gg 1$ :

$$\sum_{i \in C_1} \Delta_{i,1}(0) \approx 1, \tag{C.4}$$

où on définit  $C_1 = \{\{j\}\}, |x_1 - x_j| < r_0$  l'ensemble des éléments à portée de la défaillance initiale.

Parmi les éléments à portée de la défaillance au début de la génération s = 1, un nombre aléatoire  $N_1 = n$  d'éléments de charges  $L_1(1), \ldots, L_n(1)$  et de positions  $x_{I(1)}, \ldots, x_{I(n)}$ défaillent suite à la surcharge provoquée par la redistribution de la charge L(0). Le nombre total de défaillances à la fin de l'étape s = 1 est donc  $M_1 = M_0 + N_1$ .

3. À la génération s = 2 :

Chaque élément *i* survivant à la génération s = 2 reçoit une fraction aléatoire  $\Delta_{i,1}(1), \ldots, \Delta_{i,j}(1)$ de la charge des *j* défaillances dont il est à portée, parmi les  $N_1$  défaillances de la génération s = 1. La charge  $u_i(2)$  d'un élément *i* parmi les  $\nu(2) = N - M_1$  éléments survivants à l'étape s = 2 est alors :

$$u_i(2) = u_i(1) + \Delta_{i,1}(1)L_1(1) + \dots + \Delta_{i,j}(1)L_j(1), \qquad (C.5)$$

où les variables aléatoires  $\Delta_{i,k}(1)$  vérifient la conservation de chaque charge défaillante associée  $L_k(1)$  sur les éléments à portée de chaque défaillance :

$$2r_0\nu(2)\langle\Delta_{i,k}(1)\rangle = 1, \ j = 1,\dots,N_1,$$
 (C.6)

soit pour  $\nu(2) \gg 1$ :  $\sum_{i \in C_k} \Delta_{i,k}(1) \approx 1$ ,  $\forall k$ , où  $C_k$  est le "cercle" de survivants autour de la défaillance k.

Un nombre aléatoire  $N_2$  d'éléments défaillent suite à cette redistribution de charge et engendreront à leur tour des défaillances subséquentes. Le nombre total de défaillances à la fin de l'étape s = 2 est donc  $M_2 = M_1 + N_2$ .

#### 4. À la génération s :

Les défaillances de charges  $L_1(s-1), \ldots, L_{N_{s-1}}(s-1)$  de la génération précédente s-1 distribuent leur charges à ceux des  $u_1(s), \ldots, u_{\nu(s)}(s)$  éléments survivants au début de la *s*-ième génération qui sont à leur portée. Le nombre total de survivants à la génération *s* est donné par :

$$\nu(s) = N - M_{s-1} = N - \sum_{t=0}^{s-1} N_t = \nu(s-1) - N_{s-1}$$
(C.7)

avec  $M_{s-1}$  le nombre total de défaillances à la fin de la génération (s-1) et  $N_t$  le nombre de défaillances survenues pendant la génération t.

Chaque élément *i* survivant reçoit une fraction aléatoire  $\Delta_{i,k}(s-1)$  de la charge  $L_k(s-1)$  de chacune des  $\mathcal{N}_{r_0}$  défaillances de la génération précédente (s-1) qui peuvent l'atteindre, *i.e.* qui sont à une distance  $r < r_0$ :

$$u_i(s) = u_i(s-1) + \sum_{k=0}^{N_{r_0}} \Delta_{i,k}(s-1)L_k(s-1).$$
 (C.8)

Pour assurer la conservation en moyenne des charges défaillantes, la relation suivante sur les  $\Delta(s-1)$  doit être vérifiée à toute génération s:

$$2r_0\nu(s) \langle \Delta(s-1) \rangle = 1,$$
 (C.9)  
 $j = 1, \dots, \nu(s-1).$ 

où la moyenne est prise sur l'ensemble des valeurs possibles pour j donné, et où l'on fait l'approximation que la densité de charges survivantes est constante sur l'intervalle (0,1). Ces variables aléatoires  $\Delta(s)$  dépendent du nombre total de défaillances, et donc de l'histoire de la cascade depuis la génération d'origine s = 0.

Suite à cette redistribution,  $N_s$  éléments défaillent à la génération s et redistribueront leur charge aux survivants à portée lors de la génération suivante de la cascade.

La cascade continue jusqu'à ce que tous les éléments aient défailli ou jusqu'à ce qu'une génération ne contienne aucune défaillance. On définit la rupture du réseau dans ce modèle lorsque le nombre d'éléments survivants à la fin de la cascade est inférieur à un dizième de la taille du réseau.

Les variables aléatoires  $\Delta(s)$  sont données par la distribution binomale à la génération s :

$$\Delta(s-1) = \begin{cases} \beta & \text{avec probabilité } p(s) \\ 0 & \text{avec probabilité } 1-p(s) \end{cases},$$
(C.10)

où  $\beta \in [0,1]$  est la fraction de charge redistribuée avec probabilité p(s) aux éléments de la génération s. La probabilité p(s) est définie de telle façon que les équations sur la moyenne de  $\Delta(s)$  sont vérifiées. La valeur moyenne de la variable aléatoire distribuée selon cette loi bimodale est donc :

$$\langle \Delta(s-1) \rangle = \beta p(s) = \frac{1}{2r_0\nu(s)},\tag{C.11}$$

où on fait l'approximation que la densité de charges défaillantes est uniforme à toute étape sur l'ensemble de l'intervalle (0, 1).

On représente sur la figure C.1 le temps de défaillance d'un élément situé à la position x pour des histoires caractéristiques des jeux de paramètres  $\alpha$ ,  $\beta$  et  $r_0$ .

La probabilité de rupture est, dans ce modèle, la probabilité que tous les éléments ont défailli, sauf éventuellement une petite fraction de l'ensemble. Le processus de cascade se propage en effet à un nombre exponentiel d'éléments, et les quelques éléments restants le sont à la suite des nouvelles contraintes topologiques appliquées au système, *i.e.* si ces éléments ne sont pas à portée des dernières défaillances, ils ne peuvent défaillir. On définit donc arbitrairement la probabilité de rupture comme :

$$P_b = P\left(M_\infty < \frac{N}{10}\right). \tag{C.12}$$

Dans le cas des conditions aux limites périodiques, on vérifie sur la figure C.2(a) que pour une portée  $r_0$  de transmission de la charge défaillante couvrant l'ensemble de l'intervalle, on retrouve bien la même courbe pour la probabilité de défaillance que dans le cas à portée infinie. Si on diminue la portée  $r_0$  de transmission de la charge défaillante, la courbe de la probabilité de rupture en fonction du paramètre  $\beta$  est modifiée. En particulier, les deux seuils de probabilité de rupture non nulle sont décalés vers de plus petites valeurs de  $\beta$ .

#### C.2 Modèle Recuit

Les approximations qui nous permettent de décrire analytiquement les cascades du modèle de Lehmann & Bernasconi peuvent être résumées dans le modèle recuit : un élément ne subit qu'une unique surcharge au cours de son histoire. Dans le modèle recuit, les charges de tous les éléments survivants sont retirées dans la distribution initiale  $f_{L(0)}$  avant chaque surcharge. Dans le cas de la cascade sur un réseau spatial, la position des éléments est fixée à la génération s = 0.

1. À la génération s = 0:

On distribue uniformément N - 1 éléments sur l'intervalle (0, 1), aux positions  $x_i$ , i = 2, ..., N. On tire une charge L(0) dans la distribution  $f_{L(0)}$ , qui est la charge défaillante initiale. On fixe cet élément au milieu de l'intervalle, à la position  $x_0 = 0.5$ . La génération s = 0 contient donc une unique défaillance,  $N_0 = 1$ .

2. À la génération s = 1 :

Chacun des  $\nu(1) = N - N_0 = N - 1$  éléments survivants subit une surcharge de la défaillance de la première génération avec probabilité p(1), si il est à portée  $r_0$  de la défaillance initiale. La probabilité p(1) assure que, en moyenne, toute la charge L(0) est redistribuée au réseau :

$$p(1) = \frac{1}{2r_0\nu(1)},\tag{C.13}$$

où on considère que la densité d'éléments survivants est uniforme sur l'intervalle (0, 1). Si un élément *i* subit une surcharge, on tire sa charge  $u_i$  dans la distribution  $f_{L(0)}$ , à laquelle on ajoute la surcharge  $\beta L(0)$ . Si cette surcharge ajoutée à la charge  $u_i$  dépasse la



(c) Même portée,  $\beta$  différent

FIGURE C.1 – Temps de défaillance en fonction de la position de l'élément au cours d'une unique cascade, pour un même paramètre  $\alpha$  et différentes valeurs des paramètre  $\beta$  et  $r_0$ . On observe l'apparition de deux branches de défaillances suite à la diminution de la portée pour une même valeur de  $\beta$  (b), et pour des valeurs différentes de  $\beta$ , l'apparition de branches de direction opposée à la direction générale de la cascade (c).



(a) Comparaison entre le modèle sans portée et le modèle avec portée entendue sur tout l'intervalle



(b) Comparaison entre deux portées

FIGURE C.2 – Probabilité de rupture en fonction de  $\beta$  pour un paramètre  $\alpha$  donné, obtenue par simulations du modèle L&B et du modèle L&B avec introduction de la portée, lorsque la portée couvre l'ensemble de l'intervalle de définition des positions des éléments du système.

valeur-seuil  $u_i^{max} = (1 + \alpha)u_i$  associée à l'élément *i*, alors l'élément *i* situé en  $x_i$  défaille, avec une charge défaillante  $L(1) = u_i + \beta L(0)$ . À la fin de la génération  $s = 1, N_1$  éléments ont défailli, de charges  $L_1, \ldots, L_{N_1}$  et de positions  $x_1, \ldots, x_{N_1} \in C_{r_0}(x_0)$ , où  $C_{r_0}(x_0)$  est l'intervalle de longueur  $2r_0$  centré en  $x_0$ . Le nombre total de défaillances à la fin de la génération s = 1 est donc  $N_f(1) = 1 + N_1$ 

3. À la génération s = 2 :

Pour chacune des  $N_1$  défaillances de la génération s = 1, chaque élément parmi les  $\nu(2) = N - N_1 - N_0$  survivants subit une surcharge avec probabilité p(2), s'il est à portée de la défaillance considérée. Pour la défaillance de charge  $L_j(1)$  à la position  $x_j$ , par exemple, en moyenne  $2r_0\nu(2)$  éléments sont à portée de transmission, et k éléments parmi ceux-ci subissent une surcharge  $\beta L_j(1)$ . On tire alors k charges  $u_1, \ldots, u_k$  indépendemment dans la distribution  $f_{L(0)}$ . Les valeurs-seuil associées à ces charges sont  $u_i^{max} = (1+\alpha)u_i$ . Parmi ces k charges,  $N_1(L_j(1))$  défaillent, *i.e.*  $u_i + \beta L_j(1) > u_i^{max}$ ,  $i = 1, \ldots, N_1(L_j(1))$ . La défaillance de charge  $L_{j'}(1)$  peut transmettre une surcharge  $\beta L_{j'}$  aux éléments survivants à portée, avec probabilité p(2). Une surcharge est subie par k' éléments, dont on tire la charge avant surcharge dans la distribution  $f_{L(0)}$ . Parmi ces k' éléments,  $N_1(L_{j'}(1))$  défaillent. On itère ce processus pour les  $N_1$  défaillances de la génération s = 1. À la fin de la génération s = 2, il y a donc  $N_2 = N_1(L_1(1)) + \cdots + N_1(L_{N_1(\lambda)}(1))$  éléments qui ont défailli, de charges  $L_{N_1+1}, \ldots, L_{N_1+N_2}$ , de positions  $x_{N_1+1}, \ldots, x_{N_1+N_2}$ , et  $M_2 = M_1 + N_2$ 

défaillances totales. On itère ainsi le processus, jusqu'à ce que tous les N éléments aient défailli, ou qu'à une certaine génération s, le nombre de défaillances  $N_s$  est nul. On dit qu'il y a rupture du système s'il reste à la fin de la cascade moins d'une petite fraction des éléments initiaux.

On représente sur la figure C.3 la probabilité de rupture pour les différents modèles, sans portée, avec portée, recuit ou non recuit.

La position des éléments engendrés par une défaillance i située à la position  $x_i$  n'est pas tirée d'une distribution, pour assurer la distribution uniforme des éléments sur l'axe. Les caractéristiques spatiales de ce modèle sont liées à un phénomène de dépeuplement de l'axe, qui déforme la distribution spatiale uniforme des défaillances issues d'une défaillance de la génération précente.

Dans le modèle recuit, les charges défaillantes à chaque étape sont distribuées comme dans le modèle  $\mathcal{H}$ : la charge défaillante est composée de la somme d'une variable aléatoire tirée selon  $f_{L(0)}$  et astreinte à défaillir suite à une surcharge  $\beta L_i$  à l'étape précédente, et de la surcharge  $\beta L_i$ .

On représente figure C.4 quelques histoires typiques du modèle recuit. On observe que les caractéristiques des histoires observées dans le modèle L&B spatial se retrouvent pour les mêmes valeurs des paramètres dans le modèle recuit.

On considérera donc dans la suite le modèle recuit spatial.

#### C.3 Propriétés spatiales du modèle recuit

Les styles caractéristiques de chaque jeu de paramètres sont clairement visibles sur la figure C.5 où on représente trois distances caractéristiques du modèle : la distance moyenne d entre les défaillances de l'étape s et la défaillance initiale, la distance maximale  $d_{max}$  atteinte par les



FIGURE C.3 – Probabilité de rupture en fonction de  $\beta$  pour une valeur donnée de  $\alpha$ . On compare sur cette figure les cas du modèle L&B à portée infinie et le modèle recuit avec portée, selon que les éléments défaillants peuvent transmettre leur charge à l'ensemble de éléments ou à une fraction seulement.



(c) Même portée,  $\beta$  différent

FIGURE C.4 – Temps de défaillance en fonction de la position de l'élément au cours d'une unique cascade, pour un même paramètre  $\alpha$  et différentes valeurs des paramètre  $\beta$  et  $r_0$  dans le modèle recuit spatial. On observe la même structure de l'extension spatiale des cascades pour les mêmes valeurs des paramètres  $\alpha$ ,  $\beta$  et  $r_0$  que dans le cas du modèle L&B C.1.

défaillances de l'étape s et la distance r entre le centre de gravité des charges défaillantes à l'étape s et la défaillance initiale.

Selon les valeurs des paramètres  $\alpha$ ,  $\beta$  et  $r_0$ , les caractéristiques de la cascade de défaillances apparaissent dans les trois distances d, r et  $d_{max}$ . Dans le cas C.5(a) d'une interaction à longue portée et pour des paramètres induisant une probabilité de rupture  $P_b \sim 1$  (ou un nombre typique de descendants M > 1), la cascade spatiale diffuse rapidement vers les bords du système,  $d_{max}(s) \approx 0.5$ , reste symétrique au cours de son histoire, *i.e.*  $r(s) \approx 0$  et après quelques étapes, les défaillances sont réparties uniformément sur l'ensemble de l'intervalle :  $d(s) \approx 0.25$ .

Dans le cas C.5(b) d'une interaction à faible portée et pour des paramètres induisant un grand nombre typique de descendants par défaillance, la cascade diffuse continuement vers les bords de l'intervalle avec une vitesse typique v et avec un étalement des défaillances  $(d_{max} - d)$  qui reste constant tant que la cascade progresse. Dans le cas d'un grand nombre de descendants par individu, les deux branches symétriques par rapport à l'origine des défaillances survivent tout au long de la phase de diffusion de la cascade, et le rayon moyen des défaillances de la s-ième génération reste nul : r(s) = 0.

Dans le cas C.5(c) d'une interaction à faible portée et pour un nombre typique de descendants de l'ordre de 1, le front de la cascade progresse à une vitesse v similaire à la vitesse obtenue pour une même portée  $r_0$  mais des valeurs des paramètres  $\alpha$  et  $\beta$  tels que le nombre typique de descendants par défaillance soit plus grand. La distance moyenne des défaillances de la génération s à l'origine des défaillances est cependant plus faible pour ce jeu de paramètres, et l'écart entre le front et le "bulk" augmente avec le temps. Le nombre typique de descendants étant de l'ordre de 1, le front de défaillances progresse trop vite pour que tous les éléments de l'intervalle visité aient défailli. De nouvelles branches, symétriques par rapport à la défaillance les engendrant peuvent apparaître au cours de l'histoire de la cascade. Le rayon moyen des défaillances de la génération s n'est pas nul et montre des oscillations correspondant au passage de la cascade secondaire de part et d'autre de l'origine des défaillances. Ce phénomène déforme également les autres distances typiques qui montrent également des oscillations.

L'introduction de la composante spatiale impose un réservoir de charge faillibles fini. Le nombre typique de descendants par défaillance est donc réduit, mais la définition de la probabilité de surcharge assure que les caractéristiques globales de la cascade sont semblables aux caractéristiques de la cascade sur réseau complet. On représente sur la figure C.6 le nombre typique de descendants d'une défaillance moyenné sur toutes les histoires, en fonction de  $\beta$  pour le modèle LB et le modèle spatial. Pour ce dernier modèle, on présente deux rayons de défaillance, l'un couvrant l'ensemble de l'axe, le second très restreint autour de la défaillance. On observe que  $\bar{\mu}$  est du même ordre pour chacun des rayon considéré, mais, conformément à la probabilité de rupture, pour  $r_0 = 0.03$ , le nombre typique de défaillances devient inférieur à 1 pour une valeur du paramètre  $\beta_c < \beta < 1$ .



(a) Grande portée, grande probabilité de rupture





(c) Portée faible, probabilité de rupture plus faible

FIGURE C.5 – Distances moyenne et maximale des défaillances de l'étape s à la défaillance initiale et distance du centre de gravité des défaillances de la génération s, moyenné sur 10000 réalisations.



FIGURE C.6 – Nombre typique de descendants par défaillance, moyenné sur 10000 histoires, en fonction de  $\beta$ , pour différents modèles.

### Bibliographie

- R. Albert, H. Jeong, and A.L. Barabási. Internet : Diameter of the world-wide web. Nature, 401 :130-131, 1999.
- [2] S.-H. Yook, H. Jeong, and A. L. Barabási. Modeling the internet's large scale topology.
- [3] R.V. Solé, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde. Robustness of the european power grids under intentional attack. 2007.
- [4] J.R. Banavar, A. Maritan, and A. Rinaldo. Size and form in efficient transportation networks. *Nature*, 99 :130, 1999.
- [5] S. Redner. How popular is your paper? an empirical study of citation distribution. *Eur. Phys. J. B*, 4 :131, 1998.
- [6] N. Barkai and S. Leibler. Robustness in simple biochemical networks. Nature, 387 :913, 1997.
- [7] L. Euler. Solutio problematis ad geometriam situs pertinentis. Commentarii academiae scientiarum Petropolitanae, 8.
- [8] P. Erdős and A. Rényi. On random graphs. Publicationes Mathematicae, 6, 1959.
- [9] P. Erdős and A. Rényi. On the evolution of random graphs. Zentralblatt, 103.
- [10] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Generic scale of "scale-free" networks. Phys. Rev. E, 63 :062101.
- [11] P. L. Krapivsky and S. Redner. Organization of growing random networks. Phys. Rev. E, 63 :066123, May 2001.
- [12] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. Advances in Physics, 51(4):1079-1187, 2002.
- [13] B. Bollobas. The evolution of random graphs. Trans. Am. Math. Soc., 286:257, 1984.
- [14] V. F. Kolchin. Random mapping. Theor. Probab. Appl., 31:439, 1986.
- [15] T. Luczak. Component behavior near the critical point of the random graph process. Random Struct. Algorithms, 1:287, 1990.
- [16] S.N.Dorogovtsev and J.F.F. Mendes. Evolution of Networks. Oxford University Press, 2003.
- [17] J. Travers and S. Milgram. An experimental study of the small world problem. Sociometriy, 1969.

- [18] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world' networks. Nature, 393 :440-442, 1998.
- [19] M. E. J. Newman. The structure and function of complex networks. SIAM Review, 45(2) :pp. 167–256, 2003.
- [20] D.J. Watts. A simple model of global cascades on random networks. PNAS, 99(9) :5766– 5771, 2002.
- [21] S.A. Pandit and R.A. Amritkar. Characterization and control of small-world networks. *Phys. Rev. E*, 60 :1119, 1999.
- [22] M. Barthelemy and L. A. N. Amaral. Small-world networks : Evidence for a crossover picture. Phys. Rev. Lett., 82 :3180 - 5180(E), 1999.
- [23] A. Barrat. Comment on "small-world networks : Evidence for a crossover picture". e-print cond-mat/9903323, 1999.
- [24] M. E. J. Newman and Watts D. J. Renormalization group analysis of the small-world network model. *Phys. Lett. A*, 263 :341, 1999.
- [25] M. Argollo de Menezes, Moukarzel C. F., and Penna T. J. P. First-order transition in small-world networks. *Europhys. Lett.*, 50:574, 2000.
- [26] A. Barrat and Weigt M. On the properties of small-world network models. Eur. Phys. J. B, 13:547, 2000.
- [27] D.J de Solla Price. Networks of scientific papers. *Science*, 149:510.
- [28] Kochen M. de Sola Pool, I. Contacts and influence. Social Networks, 1:5–51, 1978.
- [29] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. Rev. Mod. Phys., 74 :47–97, Jan 2002.
- [30] R. Albert, H. Jeong, and A.L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406 :378–381, 2000.
- [31] L. A. N. Amaral, A. Scala, M. Barthelemy, and Stanley H. E. Classes of small-world networks. Proc. Natl. Acad. Sci. U.S.A., 97 :11 149, 2000.
- [32] R. Albert, I. Albert, and G.L. Nakarado. Structure vulnerability of the borth american power grid. Phys. Rev. E, 69 :025103, 2004.
- [33] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. Science, 286(5439):509-512, 1999.
- [34] M. E. J. Newman. Scientific collaboration networks (i & ii). Phys. Rev. E, 64 :016131, 2001.
- [35] M. E. J. Newman. The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. U.S.A., 98 :404, 2001.
- [36] A.-L. Barabasi, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A*, 311, 2001.
- [37] R. Kumar, P. Raghavan, S. Rajalopagan, and A. Tomkins. Proceedings of the 9th ACM Symposium on Principles of Database Systems, page p. 1, 1999.

- [38] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structure in the web. *Computer Networks*, 33(1-6):309-320, 2000.
- [39] L.A. Adamic and B.A. Huberman. Power-law distribution of the world wide web. Science, 287(5461) :2115, 2000.
- [40] A. Capocci, V.D.P. Servedio, F. Colaiori, L.S. Buriol, D. Donato, S. Leonardi, and G. Caldarelli. Phys. Rev. E, 74 :036116, 2006.
- [41] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *Comput. Commun. Rev.*, 29 :251, 1999.
- [42] R. Govindan and H. Tangmunarunkit. Heuristics for internet map discovery. Proceedings of IEEE INFOCOM 2000, 3:1371, 2000.
- [43] H. Jeong, Z. Neda, and A.-L. Barabasi. Measuring preferential attachment for evolving networks. *Europhys.Lett.*, 61:567–572, 2003.
- [44] R. Pastor-Satorras, A. Vazquez, and A. Vespignani. Dynamical and correlation properties of the internet. *Phys. Rev. Lett*, 87 :258701, 2001.
- [45] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. J. Shenker, and W. Willinger. The origin of power laws in internet topologies revisited. *Proceedings of the 21st Annual Joint Conference* of the IEEE Computer and Communications Societies, 2002.
- [46] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. Proc. Natl. Acad. Sci. USA, 101.
- [47] H. Jeong, S. P. Mason, Z. N. Oltvai, and A.-L. Barabasi. Lethality and centrality in protein networks. *Nature*, 411 :41, 2001.
- [48] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Structure of growing networks with preferential linking. *Phys. Rev. Lett.*, 85(21):4633-4636, 2000.
- [49] L. Kullmann and J. Kertesz. Preferential growth : Exact solution of the time-dependent distributions. Phys. Rev. E, 63 :051112, 2001.
- [50] P.L. Krapivsky, S. Redner, and F. Leyvraz. Phys. Rev. Lett., 85:4629, 2000.
- [51] J. Kleinberg, R. Kumar, S. Raphavan, and A. Tomkins. Lecture Notes in Computer Science, volume 1627. Springer-Verlag Berlin, 1999.
- [52] P.L. Krapivsky and S. Redner. Finiteness and fluctuations in growing networks. J. Phys. A, 35 :9517, 2002.
- [53] M. Abramowitz and I.A. Stegun. Handbook of Mathematical Functions. Dover, 1965.
- [54] B. Wacklaw and I.M. Sokolov. Phys. Rev. E, 75 :056114, 2007.
- [55] S. Lawrence and C.L. Giles. Science, 280(5360) :98-100, 1998.
- [56] B.A. Huberman, P.L.T. Pirolli, J.E. Pitkow, and R.M. Lukose. Strong regularities in world wide web surfing. *Science*, 280(5360) :95–97, 1998.
- [57] S. Lawrence and C.L. Giles. Accessibility of information on the web. Nature, 400 :107, 1999.
- [58] G. Peach. Theory of the pressure broadening and shift of spectral lines. Advances in Physics, 30(3):367-474, 1981.

- [59] P. Erdös and T. Luczak. Publ. Math. Inst. Hungar. Acad. Sci., 5(17), 1960.
- [60] P. L. Krapivsky and S. Redner. Statistics of changes in lead node in connectivity-driven networks. *Phys. Rev. Lett.*, 89 :258703, Dec 2002.
- [61] A.A. Moreira, Jr. Andrade, J.S., and L.A.N. Amaral. Extremum statistics in scale-free network models.
- [62] M. Bauer, C. Godrèche, and J.-M. Luck. J. Stat. Phys., 96 :963, 1999.
- [63] G. Bianconi and A.L. Barabási. Phys. Rev. Lett., 86:5632, 2001.
- [64] C. Godrèche and Luck J.M. On leaders and condensates in a growing network. J. Stat. Mech., page P07031, 2010.
- [65] D.S. Callaway, M. E. J. Newman, S. H. Strogatz, and Watts D.J. Network robustness and fragility : Percolation on random graphs. *Phys. Rev. Lett.*, 85 :5468, 2000.
- [66] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin. Resilience of the internet to random breakdowns. *Phys. Rev. Lett.*, 85(21):4626, 2010.
- [67] Y. Moreno, J. B. Gomez, and A. F. Pacheco. Instability of scale-free networks under node-breaking avalanches. *Europhys. Lett.*, 58 :630–636, 2002.
- [68] R. Guimera, A. Arenas, A. Diaz-Guilera, F. Vega-Redondo, and A. Cabrales. Optimal network topologies for local search with congestion. *Phys. Rev. Lett.*, 89 :248701, 2002.
- [69] A.E. Motter and Y.-C. Lai. Cascade-based attacks on complex networks. Phys. Rev. E, 66 :065102, 2002.
- [70] P. Crucitti, V. Latora, and M. Marchiori. A model for cascading failures in complex networks. *Phys. Rev. E*, 69 :045104R, 2004.
- [71] M. L. Sachtjen, B. A. Carreras, and Lynch V. E. Disturbances in a power transmission system. *Phys. Rev. E*, 61 :4877, 2000.
- [72] S. Bikhchandani, D. Hirshleifer, and Welch I. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100:992–1026, 1992.
- [73] D. ben Avraham and S. Havlin. Diffusion and Reactions in Fractals and Disordered Systems. Cambridge University Press, 2000.
- [74] J. Lehmann and J. Bernasconi. Stochastic load-redistribution model for cascading failure propagation. *Phys. Rev. E*, 81:031129, 2010.
- [75] F. T. Peirce. J. Text. Ind., 17:355, 1926.
- [76] H. E. Daniels. Proc. Roy. Soc. London A, 183:405, 1945.
- [77] S. Pradhan, A. Hansen, and B.K. Chakrabarti. Failure processes in elastic fiber bundles. *Rev. Mod. Phys.*, 82 :499, 2010.
- [78] D. Sornette. J. Phys. A, 22 :243, 1989.
- [79] W. I. Newman and A. M. Gabrielov. Int. J. Fract., 50 :1, 1991.
- [80] P. C. Hemmer and A. Hansen. ASME J. Appl. Mech., 59 :909, 1992.
- [81] J. B. Gomez, D. Iniguesz, and A.F. Pacheco. Phys. Rev. Lett., 71:380, 1993.
- [82] J. V. Andersen, D. Sornette, and K. Leung. Phys. Rev. Lett., 78 :2140, 1997.

- [83] Zapperi S., P. Ray, H. E. Stanley, and A. Vespignani. Phys. Rev. Lett., 78 :1408, 1997.
- [84] M. Kloster, A. Hansen, and P. C. Hemmer. Phys. Rev. E, 56 :2615, 1997.
- [85] F. Kun, S. Zapperi, and H. J. Herrmann. Europhys. J. B, 17:269, 2000.
- [86] S. Pradhan and P. C. Hemmer. Phys. Rev. E, 75:056112, 2007.
- [87] Bhattacharyya P., S. Pradhan, and B. K. Chakrabarti. Phys. Rev. E, 67:046112, 2003.
- [88] D. G. Harlow and S. L. Phoenix. Int. J. Fracture, 17:601, 1981.
- [89] R. C. Hidalgo, Y. Moreno, F. Kun, and H. J. Herrmann. Phys. Rev. E, 65:046148, 2002.
- [90] D.N. Kosterev, C.W. Taylor, and W.A. Mittelstadt. Model validation for the august 10, 1996 wscc system outage.
- [91] I. Dobson, B.A. Carreras, and D.E. Newmann. A branching process approximation to load-dependent system failure. *Thirty-seventh Hawaii International Conference on System Sciences, IEEE*, 2004.
- [92] T.E. Harris. Theory of branching processes. Dover NY, 1989.
- [93] I. Simonsen, L. Buzna, K. Peters, S. Bornholdt, and D. Helbing. Transient dynamics increasing network vulnerability to cascading failures. *Phys. Rev. Lett.*, 100 :218701, 2008.
- [94] J. H. Lambert. Observations variae in mathesin puram. Acta Helvitica, physicomathematico-anatomico-botanico-medica 3 :128–168, 1758.
- [95] F.A. Haight and M.A. Breuer. The borel-tanner distribution. 47(1-2):143-150, 1960.
- [96] J. Lehmann and J. Bernasconi. Breakdown of fiber bundles with stochastic loadredistribution. Chemical Physics, 375(2-3):591-599, 2010.